Published in final edited form as:

J Educ Psychol. 2012 February ; 104(1): . doi:10.1037/a0025323.

Prevalence and Nature of Late-Emerging Poor Readers

Hugh W. Catts,
University of Kansas

Donald Compton,
Vanderbilt University

J. Bruce Tomblin, and
University of Iowa

Mindy Sittner Bridges
University of Kansas

Abstract

Some children demonstrate adequate or better reading achievement in early school grades, but fall significantly behind their peers in later grades. These children are often referred to as late-emerging poor readers. In this study, we investigated the prevalence and heterogeneity of these poor readers. We also examined the early language and nonverbal cognitive abilities of late-emerging poor readers. Participants were 493 children who were a subsample from an epidemiological study of language impairments in school-age children. In kindergarten, children were administered a battery of language, early literacy, and nonverbal cognitive measures. Word reading and reading comprehension achievement was assessed in second, fourth, eighth, and tenth grades. Latent transition analysis was used to model changes in reading classification (good vs. poor reader) across grades. Population estimates revealed that 13.4% percent of children could be classified as late-emerging poor readers. These children could be divided into those with problems in comprehension alone (52%), word reading alone (36%), or both (12%). Further results indicated that late-emerging poor readers often had a history of language and/or nonverbal cognitive impairments in kindergarten. Subtypes of poor readers also differed significantly in their profiles of language, early literacy, and nonverbal cognitive abilities in kindergarten. Results are discussed in terms of causal factors and implications for early identification.

Keywords

Late-emerging poor readers; reading disabilities; latent transition analysis; subtypes of poor readers; early identification

Considerable research attention has been devoted to the early identification and prevention of reading disabilities (Compton, Fuchs, Fuchs, & Bryant, 2006; Jenkins, 2003; Vellutino, Scanlon, Small, & Fanuele, 2006; Wood, Hill, Meyer, & Flowers, 2005). This work has led to the identification of children at risk for reading disabilities (RD) as early as kindergarten or first grade and to the development of effective intervention programs (Denton & Mathes, 2003; Foorman, Francis, Fletcher, Schatschneider & Mehta, 1998; Simmons, Coyne, Kwok, McDonagh, Harm, & Kame'enui, 2008). Despite progress, at least one group of children is routinely missed in this process. These are children whose reading problems are not apparent until the later school grades. These so-called "late-emerging" poor readers get off to a good start in learning to read but experience significant problems in later grades. Late-

emerging poor readers present a significant challenge for early identification and prevention efforts because only limited information is available about the prevalence and nature of their reading difficulties and the factors that underlie these problems.

Chall (1983) was among the first to bring attention to late-emerging poor readers. In her discussion of the so-called “fourth grade slump,” she noted that some children show adequate or better progress in beginning reading, but fall behind in the middle elementary school grades. She suggested that some of these children may develop sufficient word reading skills but lack the linguistic and/or conceptual skills and knowledge necessary to understand more demanding texts found in later grades. Alternatively, Chall noted that other children may fail to develop fluency in word reading, which in turn, could disrupt comprehension as texts become more challenging.

Despite the recognition of late-emerging RD more than 25 years ago, there have been relatively few empirical investigations of these reading problems. The most often cited study of late-emerging poor readers was conducted by Leach, Scarborough, and Rescorla (2003). They investigated 161 fourth or fifth grade children, including 65 children with an educational history of RD. All participants were administered a battery of reading, spelling, language, and cognitive measures at the time of the study. Information concerning history of RD was obtained from parental report and confirmed by school records. Using a combination of data from these direct assessments and parental/school records, 31 children were identified as having late-emerging reading disabilities. Further data showed that these children were heterogeneous in terms of the type of reading problems they had. Ten were found to have late-emerging deficits in reading comprehension, 11 in word reading, and 10 in both reading comprehension and word reading. Other analyses also suggested that late-emerging poor readers did not simply have less severe disorders than those with persistent reading disabilities, but rather had similar reading difficulties. Furthermore, retrospective analyses showed that these children indeed had late-emerging problems, not late-identified problems. That is, late-emerging poor readers did not appear to have early reading problems that had been missed as a result of flaws in the identification process (e.g., overlooked due to their high intelligence, good behavior, or compensatory strategies) as had sometimes been assumed.

Whereas Leach et al. (2003) provided an initial examination of the nature of late-emerging RD, more direct evidence of these difficulties can be obtained from longitudinal studies that investigated various reading and reading-related abilities across grades. For example, Badian (1999) conducted a longitudinal study in which a large group of children (N=1,008) were followed from pre-kindergarten through seventh or eighth grades. Each year, the children were administered measures of word reading, listening comprehension, and reading comprehension. The primary purpose of the study was the investigation of the prevalence and stability of RD based on a discrepancy between listening and reading comprehension. However, in secondary analyses, children were divided into those showing RD in Grades 1-4 (early poor readers), Grades 5-8 (late poor readers) or both (consistent poor readers). For these analyses, children were defined as having RD if their mean score in reading comprehension across the four-year grade span (Grade 1-4 or 5-8) was below the 25th percentile. Results showed that 1.9% of the population were early poor readers, 5.8% were late poor readers, and 6.8% were persistent poor readers. Other findings demonstrated that when compared to early poor readers, late poor readers had significantly higher word reading scores in the early grades and significantly lower listening comprehension scores in the later grades.

The above study only considered poor readers who were identified on the basis of deficits in reading comprehension. As a result, no information was available concerning the prevalence

of children with late-emerging word reading deficits. In another longitudinal study, Lipka, Lesaux, and Siegel (2006) examined reading and reading-related abilities of children with poor word-reading skills. From a sample of 1,100 children who had been followed from kindergarten through fourth grade, 22 children were identified with word-reading deficits in fourth grade. Seven of the poor readers had persistent problems across grades, 8 had late-emerging deficits (after 3rd grade) and 7 had borderline deficits at other grades. Additional results indicated that those with late-emerging word-reading problems had phonological processing deficits, especially after second grade. Such deficits were evident on tests of phonological awareness, phonological decoding, and spelling. The authors suggested that these children may have been able to compensate for their phonological deficits in the early grades, but as words became more complex, they showed reading and spelling difficulties.

In another longitudinal study, Parrila, Aunola, Leskinen, Nurmi, and Kirby (2005) used latent growth curve modeling to examine individual differences in reading growth across grades in samples of English and Finnish speaking children. Data from the Finnish sample, however, were only available for first and second grades, and thus, is less relevant for the present study. The English speaking children, however, were followed from first through fifth grade and administered tests of word identification, word attack, and passage comprehension at each grade. Growth mixture models identified latent classes based on homogeneous reading growth trajectories. Of particular interest was the identification of a class (for both word attack and passage comprehension) that included children who were relatively poor readers at the beginning of school but who had caught up to other good readers by fifth grade (i.e., early poor readers). The growth mixture models did not, however, uncover a class of late-emerging poor readers. The authors suggested that given the relatively small overall sample size (N=198) and the conservative nature of the statistical procedure employed, such a class may have been too small to be identified.

Compton, Fuchs, Fuchs, Elleman, and Gilbert (2008) also used a latent variable approach to examine late-emerging poor readers. They employed latent transition analysis (LTA), which is an extension of latent class analysis, to model movement across latent reading classes (e.g., RD) as a function of grade. Because LTA relies on multiple indicators of a reading class, it is less susceptible to measurement error that can give the appearance of children changing reading categories, when no true change has occurred. Compton and colleagues selected 177 children who showed initial risk for RD (based on teacher judgments) and followed them from first through fourth grade. They identified 5 late-emerging poor readers from the sample but did not disaggregate these children by deficit area (i.e., word reading vs. reading comprehension). Because the sample was not a representative one, no conclusions can be drawn concerning population prevalence. However, the researchers were primarily interested in whether or not late-emerging poor readers could be identified on the basis of their early reading or language abilities. Results showed that late-emerging poor readers had poorer listening comprehension and more limited word-reading growth than did typical children. However, these indices did not reliably distinguish children with late-emerging RD from typical readers, because of the high rate of false positives.

The present investigation was undertaken to examine late-emerging poor readers by estimating the prevalence of RD subtypes in the population of developing readers and exploring potentially important differences across subtypes that might be used in the early identification and treatment of late-emerging poor readers. Specifically, we examined the following questions: (1) What is the population prevalence of late-emerging poor readers? (2) What portion of this population show late-emerging word reading deficits, comprehension deficits, or combined deficits? (3) Do late-emerging poor reader subtypes display different profiles of kindergarten language and nonverbal cognitive abilities/disabilities when compared to each other and typical readers? To address these questions,

we utilized latent transition analysis to examine data from a longitudinal study of the language and reading abilities of children followed from kindergarten through tenth grade. Our sample was a subsample of children who participated in an epidemiologic investigation of language impairments in children (Tomblin et al., 1997). Whereas our sample was not completely representative of the larger sample, we were able to estimate how it differed from the epidemiologic sample and applied a weighting procedure to better assure its representativeness. An advantageous aspect of our study was the use of local sample-based norms for all measures. This allowed us to compare the prevalence of RD subgroups based on word reading or comprehension deficits without the confound of relying on tests that have different sample- or population-based norms. Also, because we had multiple indicators of reading across a wide grade range (2nd-10th grades), we were able to use a latent variable approach. Finally, we had measures of language, pre-literacy, and non-verbal cognitive abilities in kindergarten that could potentially provide insight into factors that preview late-emerging RD.

Method

Participants

This study involved 493 participants who were followed from kindergarten through tenth grade. These participants were initially recruited as part of an epidemiologic study of language impairment in kindergarten, involving a stratified cluster sample of 7,218 children. This sample was stratified by residential setting (i.e. rural, urban, suburban) and cluster sampled by building. (The sample was 33% rural, 37% urban, and 30% suburban; 51% male and 49% female; and 83% White, 12.7% Black, and 4% Other). Following the completion of the epidemiologic study, a subsample of children was recruited to participate in a longitudinal investigation. Because the primary focus of the latter study was developmental language impairments, all children who displayed a language impairment (including some with a nonverbal cognitive deficit) in kindergarten were asked to participate. In addition, a sample of children with a nonverbal cognitive deficit alone and a random sample of children without a language or nonverbal cognitive impairment were also recruited. The initial longitudinal sample included 604 children, but complete data on all reading measures were available for 493 participants through the tenth grade. These participants included 100 children with a specific language impairment (SLI), 72 children with a specific nonverbal cognitive deficit (SCD), 75 children with both a language impairment and a nonverbal cognitive deficit (referred to as having a non-specific language impairment - NLI), and 246 children without a language impairment or a nonverbal cognitive deficit in kindergarten¹. This sample was comparable to the larger epidemiologic sample in terms of their demographic characteristics (32% rural, 32% urban, and 36% suburban; 51% male and 49% female; and 86% White, 11% Black, and 3% Other).

Because the sample included more children with a language impairment and/or a nonverbal cognitive deficit than would be expected in the normal population, a weighting procedure was used to provide population estimates for the prevalence of late-emerging poor readers. This weighting procedure took advantage of information from the epidemiologic study and estimated the likelihood that a subject with a given language, nonverbal cognitive, and gender profile would appear in a representative sample, and weighted his or her score accordingly; children with a language and nonverbal cognitive impairment received

¹Children were identified as having a language impairment if they performed at least 1.25 SDs below the mean on 2 of 5 language composite scores derived from the TOLD-2:P and the Narrative Skills task described in Table 1. This criterion is approximately equal to having an overall language composite score which is at least 1.14 SD below the mean. Children were identified as having a nonverbal cognitive deficit if they performed at least 1 SD below the mean on a composite measure that included the Block Design and Picture Completion subtests from the WPPSI-R.

proportionally less weight in the analyses than children who showed typical language and cognitive development.² As a result, the weighting procedure allowed us to provide results that were more representative of those expected from the epidemiologic sample, and by extension, the general population.

Measures

The purpose of this study was to examine late-emerging RD. Previous research suggested that these reading problems could involve difficulties in word recognition and/or reading comprehension. Participants in our sample were assessed on multiple measures of each of these aspects of reading in second, fourth, eighth, and tenth grades. To assess word recognition, the Word Identification and Word Attack subtests of the *Woodcock Reading Mastery Test-Revised* (WRMT-R; Woodcock, 1987) were administered across grades. Subjects' performances on these measures were converted to z-scores using local sample-based weighted means and standard deviations. Measure specific z-scores were then combined and converted to a composite z-score again using local sample-based means and standard deviations. The use of local sample-based norms for these and all other measures allowed us to take advantage of our large sample size and to have a common reference across all tests. The latter is particularly important because it helps avoid error that is often introduced by using published norms from different samples. Specifically, the use of different norming samples could lead to apparent profile differences in reading/language subskills (e.g., word reading vs. comprehension) when none are present. Reading comprehension was measured by several instruments. In all grades, the Passage Comprehension subtest from the WRMT-R and the comprehension portion of the *Gray Oral Reading Test-3* (Wiederholt & Bryant, 1994) were administered. In second and fourth grades, the Reading Comprehension subtest from the *Diagnostic Achievement Battery-2* (DAB-2; Newcomer, 1990) was also administered. Because this measure was not appropriate for eighth and tenth grade, it was replaced at these grades with the reading comprehension component of the *Qualitative Reading Inventory-2* (Leslie & Caldwell, 1995), which was very similar in format (silent reading of passages followed by open-ended questions) to the DAB-2. In each grade, the three measures of reading comprehension were converted to z-scores using local weighted means and standard deviations. Measure specific z-scores were then combined and converted to a composite z-score for reading comprehension.

Another aim of the study, was the examination of the early language and/or nonverbal cognitive deficits associated with late-emerging RD. To pursue this aim, we examined participants' performances on a battery of tests administered in kindergarten that included measures of vocabulary, grammar, narration, letter knowledge, phonological awareness, and nonverbal IQ. These measures are briefly described in Table 1. In each case, scores on these measures were converted to z-scores using local sample-based weighted means and standard deviations. Finally, we had available to us data on participants' classification in terms of whether or not they had a language and/or nonverbal cognitive impairment.

As noted above, it was advantageous for us to use local-based norms for all measures in our analyses. However, to provide more information concerning our sample, we also include in Table 2 weighted means and standard deviations of selected measures that are based on published norms. These data include prorated standard scores from the Block Design and Picture Completion subtests of the WPPSI-R (kindergarten only) and standard scores from

²For example, the epidemiologic study found that boys with a specific language impairment (who also failed an earlier language screen) composed 2.8% of the general population. In our sample, however, these children composed 9.3%. To ensure that the children from this group did not contribute disproportionately to our results, their scores were adjusted by weighting them by a constant that was equal to the expected prevalence of these children (2.8%) divided by their actual prevalence (9.3%; constant = .30).

the WRMT-R and GORT-3. Also included were standard scores from the PPVT-R and the full Performance Scale of the WISC-R (Grade 2 only), which were not included in other analyses for the present study. These data show that our sample was very representative of that used for norming the PPVT-R and the WPPSI-R/WISC-III. Our sample did, however, have somewhat lower reading scores at most grades than the samples used to norm the WRMT-R and GORT-3 at most grades.

Procedures

Fourteen examiners participated in the administration of the test battery in kindergarten and another three in the administration of the test batteries in the other grades. Seven of the examiners were certified speech-language pathologists and the remaining had undergraduate degrees in speech and hearing ($n=3$) or education ($n=7$). All language measures were administered by examiners certified in speech-language pathology. In addition, all examiners received approximately one week of training by the investigators on the administration of the testing protocols. Testing was conducted in specially designed vans parked at the participants' schools or homes.

Data Analysis

In this study, we explore intra-individual change over time in latent variables representing word reading and reading comprehension skill. A useful methodological framework for studying intra-individual change over time is latent growth curve modeling using continuous variables (e.g., Bollen & Curran, 2006; Duncan, Duncan, & Strycker, 2006). In contrast to continuous growth, we focus on the movement of children across RD and normal reader categories defined by the latent variables using a form of latent transition analysis (see Collins, 2006; Lanza, Flaherty, & Collins, 2003; Kaplan, 2008). Latent transition analysis (LTA) addresses questions concerning prevalence of discrete states and incidence of transitions between states (see Collins & Wugalter, 1992; Graham, Collins, Wugalter, Chung, & Hansen, 1991; Martin, Velicer, & Fava, 1996). LTA is well suited to the study of change where there are numerous states, individuals can transition relatively freely among the states, and the states are measured with multiple indicators. Because LTA relies on latent variable analyses, classes are not directly observable, but rather are inferred from a combination of manifest indicators. LTA provides a way of statistically modeling movement between latent states, including estimating the prevalence of each discrete state and the incidence of transitions between states, adjusted for measurement error. In addition, LTA models provide a measure of classification reliability via the estimated posterior probabilities of class membership for each individual (Muthen, 2000). Reliable classification of an individual into a latent class is indicated by a high posterior probability indicating membership in a single class compared to all other classes.

In order to address the first two research questions about the prevalence and heterogeneity of late-emerging poor readers, we used LTA to identify six different subtypes of readers across time. These were typically developing across grades (TD), early identified (and persistent) RD (EIRD), late emerging RD with deficits in word reading (LERD-W), late-emerging RD with deficits in comprehension (LERD-C), late-emerging RD with deficits in comprehension and word reading (LERD-CW) and late-emerging typical developing (LETD). In some of the analyses, we further delineate the EIRD class into those with deficits in word reading, comprehension, or both. At any given time, the array of latent class membership defines an individual's latent status. The use of latent class models to identify subtype membership helps alleviate problems associated with the use of single indicators to identify subtypes. Single-measure cut-scores allow small changes in scores, often due to measurement error, to move children near the cut-score across category boundaries (e.g., Fergusson et al., 1996;

Wright, Field, & Newman, 1996). This causes instability in classification across repeated measures with the magnitude of fluctuation varying across tests and cut-scores (see Francis et al., 2005). Francis et al. argued for the use of a broader set of measures to better infer RD class membership (i.e., latent classes). The use of numerous measures to form latent classes increases the stability of RD classification and the accuracy with which we identify children with various subtypes of RD.

LTA models produce parameter estimates corresponding to the proportion of individuals in each latent class at a point in time, as well as a transition probability matrix, consisting of estimates of the probability of latent class membership at the next time point, contingent on latent class membership at the previous point. In this study we employed mixture LTA (i.e., a mover-stayer model). Mixture LTA modeling takes advantage of the benefits in reliability and validity afforded latent variables and the flexibility of mixture modeling that allows for multiple models within the population. The mover-stayer model is a form of LTA in which a subpopulation exists that does not transition over time and another subpopulation exists that can transition over time. Detailed descriptions of LTA modeling (see Collins & Wugalter, 1992; Graham et al., 1991; Kaplan, 2008; Kaplan & Walpole, 2005; Martin, Velicer, & Fava, 1996) along with statistical coding systems (Kaplan, 2008) are available in the literature.

To identify the subtypes, we used the five reading outcomes measured at second, fourth, eighth, and tenth grades. Two measures of word-level reading and three measures of reading comprehension were used to form latent classes representing word reading and reading comprehension skill at each grade. We used a cut-score of -1 SD below the weighted sample mean on each measure to indicate RD at each point in time. This cut-score is similar to that used in other studies to identify RD children (e.g., Catts, Adlof, Hogan, & Weismer, 2005; Lambrecht Smith, Roberts, Locke, & Tozer, 2010; McArthur, Hogben, Edwards, Heath, & Mengler, 2000). Accordingly, the score on each measure at each point in time was coded as 1 (above) or 0 (below) for individuals above or below the -1 SD cut-off. This coding system allowed us to represent whether an individual child was performing above or below the RD threshold on the reading measures at each point in time. A child could score below our normalcy threshold on an outcome measure and still not qualify as RD due to performance on the other outcome measure(s). At each point in time we allowed 4 latent classes in the model: normal reader (NR), RD-word reading only, RD-reading comprehension only, and RD-word reading and reading comprehension. Over time, this permits identification of 2 stable classes (TD and EIRD) and 3 transitioning LERD classes (LERD-W, LERD-C, & LERD-CW) and 1 transitioning TD class (LETD). The latter group of children had RD in the early grades but typical reading in the later grades. Models were generated using mixture modeling routines contained in Mplus 5.0 (Muthén & Muthén, 2007).

Our third research question asked if LERD subtypes differed in kindergarten language and cognitive abilities/disabilities. To answer this question, we carried out three analyses. In the first of these analyses, we compared subtypes in terms of the percentage of children who had been classified as having typical language (TL) or a language and/or nonverbal cognitive deficit in kindergarten (i.e., SLI, NLI, SCD). We expected that both EIRD and LERD subtypes to have a high prevalence of deficits in kindergarten. In a second analysis, we examined profile differences in the LERD subtypes on kindergarten measures of vocabulary (picture vocabulary, oral vocabulary), grammar (grammatical completion, grammatical understanding, sentence imitation), narrative skill (narrative comprehension, narrative recall), phonological awareness, letter identification, and nonverbal IQ. Some of these measures were the same as those used to classify children in language and/or cognitive deficit groups. However, the profile analysis allowed us to examine subtype differences in

language and cognitive abilities from another perspective. In the profile analysis, we assessed whether the profiles of LERD subtypes were significantly different from those of TD, LETD, and EIRD subtypes. We hypothesized that the kindergarten profiles of the LERD subtypes would include deficits in language abilities but strengths in phonological awareness and letter identification, whereas EIRD subtypes would have deficits in phonological awareness, letter identification, and perhaps language abilities. Profiles were compared using a 2-way repeated measures ANOVA with the between-subjects factor being subtype and the within-subjects factor being kindergarten measures (see Tabachnick & Fidell, 2007). The ten kindergarten measures were scaled to have means of 0 and standard deviations of 1. The main effect for subtype, referred to as the *elevation* effect, represents differences between subtypes averaged across the kindergarten measures. The main effect for kindergarten measure, referred to as the *flatness* effect, represents differences among the dimensions averaged across the subtypes. Within profile analysis, the interaction between subtype and kindergarten measure, referred to the *shape* effect, represents differences in the shape of the profile across subtypes. Significant elevation, flatness, and shape effects were probed using post-hoc follow-up analyses corrected for multiple comparisons as described by Tabachnick and Fidell (2007).

In the final set of analyses, we sought to determine if membership in reading disability subtypes (i.e., LERD & EIRD) and typical reader group could be accurately predicted on the basis of kindergarten measures. To accomplish this, we collapsed the three LERD subtypes into a single late-emerging group and used the 10 kindergarten measures to predict membership in the LERD and EIRD groups using multinomial logistic regression. This allowed us to not only identify important kindergarten predictors of LERD and EIRD but also examine the classification accuracy of an early identification model of LERD. Multinomial logistic regression allows the dependent measure to have more than two categories, in this case three, and compares each level of the dependent class with a specified reference class. From an early identification standpoint, our primary question was whether the predictors vary across the LERD and EIRD subtypes compared to the TD group. Therefore we designated the TD group as the reference group for comparisons. Our hypothesis was that language measures would predict LERD membership whereas phonological awareness and letter identification measures would predict EIRD membership.

Results

Intercorrelations among word reading and reading comprehension composite scores across the four assessment waves are displayed in Table 3. All correlations were statistically significant ($p < .001$) and ranged from .46 to .92. Consistent with previous studies (Hulslander, Olson, Willcutt, & Wadsworth, 2010; Foorman, Francis, Shaywitz, Shaywitz, & Fletcher, 1997; Wagner et al., 1997), rank order of word reading and comprehension performance as a function of time was relatively steady. The magnitudes of correlations within reading skill (i.e., word reading with word reading and reading comprehension with reading comprehension) across time were larger than the correlations between reading skills (i.e., word reading with reading comprehension). Overall, the high correlations between word reading and reading comprehension imply reasonably stable reading scores across time, suggesting that transitions between reading classes as a function of time should be somewhat uncommon.

As described in the data analysis section, we used LTA to derive four latent classes (NR, RD-W, RD-C, RD-CW) at each grade. The results of this analysis are shown in Table 4. These data indicate relatively high percentages of children with RD across grades, which was expected given our sample included more children with language and/or nonverbal cognitive deficits than would be predicted in a more representative sample. Because of this,

in the analyses that follow that concern prevalence estimates, we use a weighting procedure (described in the Methods section) to correct for this overrepresentation. The unweighted data in Table 4 indicate that in second grade 71% of the sample was classified as NR (i.e., normal readers) and 29% RD. Despite the high correlations between reading skills across time, the percentage of children classified as NR dropped to 54% and RD increased to 46%, indicating the transition of children from the NR class to the various RD classes. Note that the majority of the change took place between second and fourth grades.

The upper portion of Table 5 provides the unweighted proportions of movers and stayers in the sample. (For this study stayers were defined as children with zero probability of moving classes, whereas movers had at least a small probability of moving classes as a function of time.) Across the sample, only 2% of the children were classified as stayers and these children all occupied the RD-CW class. Even though only 2% of the sample was classified as stayers, the Bayesian information criterion for the mover-stayer model was lower than the estimate for a model assuming a homogeneous population, suggesting that the mover-stayer model better represented the fit of the manifest response frequencies. The movers were initially distributed across NR (71%), RD-W (9%), and RD-CW (20%) classes.

The lower portion of Table 5 provides the transition probabilities of movers across grade boundaries. In terms of transitions from second to fourth grades, 77% of children classified as NR in second grade were predicted to remain NR in fourth grade, 8% were predicted to transition to RD-W, and 15% to RD-C. Children in the RD-W class only transitioned to the RD-CW class, whereas children in the RD-CW class transitioned to NR, RD-W, and RD-C classes. The diagonal of the transition probability matrix represents the stability of each class across the two time points. In general, the stability of classes increased as children progressed to higher grades with diagonal transition probabilities exceeding .90 for all four classes during the eighth to tenth grade transition. The magnitude of transition probabilities between the RD-W and RD-CW classes (in both directions) were relatively large indicating movement between the two classes between second and eighth grades. Model fit was estimated using the likelihood ratio chi-square (LR^2). The likelihood ratio compares the observed response proportions against the response proportions predicted by the model. The degrees of freedom are calculated by subtracting the number of free parameters from the total number of response pattern possible (Kaplan, 2008). As with most SEM-based models the null hypothesis for chi-square model tests is that the specified model holds for the given population, and therefore accepting the null hypothesis implies that the model is plausible. Overall, the model was found to fit the data ($LR^2(1,048,172, N=493) = 726.01, p > .05$).

A total of 512 different latent transition cases were possible with the LTA model specified [$2 \text{ (mover-stayers)} \times 4^4 \text{ (4 possible classes at each of 4 time points)}$]. Table 6 provides frequency counts and percentages for the identified latent classes. Also, included are the population estimates based on the weighting procedure derived from the epidemiologic study. Only 26 of the 512 possible classes (5%) were represented in the data, with 2 TD, 2 LERD-W, 4 LERD-CW, 2 LERD-C, 3 LETD, 5 EIRD-W, 5 EIRD-CW, and 3 EIRD-C classes. The data responded in an orderly fashion with transitions generally followed by stable class membership, meaning that once a child transitioned from say NR to RD-C (i.e., LERD-C), she or he tended to stay in that class. Predicted probabilities, the estimated probability that a child has been assigned to the correct class, for the 26 classes ranged from .623 to .918, with 22 of the classes having probabilities above .80. Population estimates for the classes are as follows: 67.9% TD, 4.8% LERD-W, 1.6% LERD-CW, 7.0% LERD-C, 1.9% LETD, and 16.8% EIRD. These estimates were weighted using procedures described above to better assure their representativeness. Overall, 84.7% of the sample remained in the same class (i.e., stable) across grades 2-10 (67.9% TD; 16.8% EIRD). The remaining 15.3 % of the sample made a class transition across time. Approximately 13% of

the population transitioned from the NR to one of the RD classes with comprehension deficits (LERD-C & LERD-CW) making up the largest transitioning group. Further, the LERD population represents nearly 50% of the total sample identified with RD [LERD/(LERD + early identified RD)]. It is noteworthy that 4.8% of the population showed LERD in word reading only, indicating a class of children who develop word reading problems later in school. Our results further showed that about 2% of the population transitioned from early RD to normal readers. The vast majority of transitions took place between grades 2-4, with far fewer occurring between grades 4-8, and a negligible number occurring between grades 8-10. To further clarify subtype performance, Table 7 displays norm referenced composite *z*-scores for word reading and reading comprehension by subtype as a function of time. Contrary to the TD and EIRD latent classes who exhibited relatively stable performance as referenced against sample weighted norms, LERD classes continued to lose ground with time in deficit areas, indicating that the deficit areas were becoming more severe as a function of time in the LERD classes.

We also examined gender differences among LERD subgroups. Although studies often show a higher male:female ratio among poor readers identified on the basis of word reading (Shaywitz, Shaywitz, Fletcher, & Escobar, 1990), and the reverse for poor readers identified on the basis of comprehension (Nation & Snowling, 1998; Yuill & Oakhill, 1991), a weighted frequency analysis indicated no significant difference between males and females in the 3 LERD subtypes ($\chi^2(2, N=93) = 2.24, p = .326$). Ratios were in the expected direction with late emerging comprehension problems occurring more often in females (LERD-C 57%, LERD-CW 75%) and late emerging word reading problems less often (LERD-W 45%), however due to the low power none of the values were found to be significantly different from 50%.

Further analyses were undertaken to explore subgroup differences in kindergarten language and nonverbal cognitive abilities/disabilities. As described in the participant section, the sample for this study was drawn from an epidemiologic investigation of language impairments in children. At the time of kindergarten testing, children were classified into four different groups (TL, SLI, NLI, SCD). Table 8 presents class counts, percentages, and weighted population estimates for the subtypes as a function of kindergarten language/cognitive classification. As might be expected all of the LERD classes contained many children with SLI, ranging from roughly 13% to 17% of the class compared to only about 3% in the TD class. In addition, LERD classes had a substantial percentage of children from the SCD and NLI categories. The one exception was that the LERD-W class had approximately the same percentage of children from the NLI category as did the TD class. Results also showed that the majority of children with EIRD had a history of language and/or nonverbal cognitive deficits. In addition, children with a history of deficits in both areas (NLI) were most frequently classified as children with EIRD.

To further explore the relationships between kindergarten measures (i.e., vocabulary, grammar, narration, phonological awareness letter identification, and nonverbal IQ) and latent classes membership, a profile analysis was conducted. Table 9 lists correlation coefficients between word reading and reading comprehension composite scores across the assessment waves and kindergarten measures. All kindergarten measures were significantly correlated ($p < .001$) with word reading and reading comprehension skill across the four assessment waves. No univariate or multivariate outliers were identified ($p = .001$). Three measures showed some limited skew (picture vocabulary, sentence imitation, and phonological awareness) with none reaching statistical significance. Profile analysis using a 2-way repeated measures ANOVA was performed to explore elevation, flatness, and shape effects on the 10 kindergarten measures across the 6 subtypes (TD, LERD-C, LERD-W, LERD-CW, LETD, EIRD), with profiles displayed in Figure 1. When performance was

averaged across kindergarten measures subtypes differed significantly on elevation, $F(5, 487) = 56.07, p < .001$, with TD performance $>$ LERD-W = LETD = LERD-C = LERD-CW $>$ EIRD (adjusting $\alpha = .008$ to control for multiple pairwise comparisons among groups). When performance was aggregated across subtypes, the effect of flatness was found to be nonsignificant, $F(9, 479) = 1.72, p = .065$. The lack of a flatness effect is likely due to the mirror image-like performance across various groups on the measures. The test of shape, defined by the subtype by kindergarten measures interaction, was significant, $F(45, 2415) = 2.37, p < .001$, indicating deviation from parallelism across the subtypes on the kindergarten measures. These shape effects indicate the possibilities of unique strengths and weaknesses across the kindergarten measures as a function subtype. Relative strengths and weaknesses may be helpful in considering early predictors of LERD subtypes. To identify relative strengths and weakness on kindergarten measures in the subtypes, confidence intervals were calculated around the mean of the profile for the six groups combined. Since the kindergarten measures were converted to z-scores prior to analysis, the grand mean for each measure was 0 and the SD was 1. Alpha for each of the 60 comparisons was set at .0008 to achieve an adjusted error rate of 5%. Therefore, 99.92% confidence limits were evaluated for the pooled profile. The horizontal lines at $z = .16$ and $z = -.16$ on Figure 1 represent the 99.92% confidence intervals for the combined sample with subtype performance above the upper limit considered a strength and performance below a weakness. Scores within the confidence interval was considered average performance. As evident from Figure 1, the TD subtype exhibited a relative strength on all kindergarten measures whereas relative weaknesses were present across all LERD subtypes on measures of oral vocabulary, grammatic understanding, sentence imitation, and phonological awareness.

Finally, a multinomial logistic regression analysis was performed to explore the predictive utility of the ten kindergarten measures in distinguishing EIRD and LERD groups from the TD group. The kindergarten measures were entered as a single block with the TD group serving as referent. Comparing the log-likelihood ratios of the base model (null model) to the model with the ten predictors revealed a significant improvement in the final model over the base model $\chi^2(20, N = 493) = 275.91, p < .001$. Overall the goodness-of-fit statistic was excellent with the deviance statistic $p = .993$ and Nagelkerke $R^2 = .50$. Despite the fit of the model, the accuracy of classification of the groups was quite poor. As seen in Table 10, accuracies ranged from 15.1% to 85.9% and an overall classification accuracy of 67.0%. In general, the model was most accurate in correctly classifying TD, followed by EIRD, and LERD. Over 50% of the LERD group was misclassified as TD using the kindergarten measures, suggesting that while the groups differed on kindergarten measures, these differences were not adequate to accurately distinguish the two groups in kindergarten. Since the profile analysis showed that the LERD-W group had a quite different pattern of strengths and weaknesses than the other LERD subtypes, we reran the multinomial regression analysis with the four groups (TD, EIRD, LERD-W, and LERD-C/CW). Again prediction accuracy was good for TD (87%), and fair for EIRD (70%) and quite poor for LERD-W (0%) and LERD-C/CW (16%).

Discussion

Prevalence of Late-Emerging Poor Readers

Our first research question concerned the prevalence of late-emerging poor readers. To address this question, we used latent transition analysis to identify good and poor readers at second, fourth, eighth, and tenth grades and to subsequently classify these readers into stable or transitioning latent classes across grades. Because our sample included more children with language and/or nonverbal cognitive deficits than would be found in a representative sample, we used weighted scores to estimate the prevalence of latent classes. Whereas the use of weighted scores might be considered a limitation, the fact that these scores were

based on the results of an epidemiologic study better assures the representativeness of our results. These results estimated the prevalence of late-emerging poor readers to be 13.4 % of school-aged children. This value and that of the other RD subtypes are influenced, of course, by the criteria we used for designating a poor reader. According to our criteria, a child could potentially be identified as a poor reader if he/she scored 1 SD below the weighted sample mean on a measure of word reading and/or reading comprehension in second, fourth, eighth, or tenth grade. Thus, for a given measure, approximately 16% of our sample (after weighting) would be classified as a poor reader on that measure at each grade. Because a child could be identified as RD on the basis of word reading measures, or comprehension measures, or both at any grade, the expected prevalence of RD in our sample was considerably higher than 16%. In fact, our LTA analysis indicated that 32.1% of the sample could be classified as RD at one or more grades. Late-emerging poor readers, thus, represented about 42% of all those classified as a poor readers across grades (13.4/32.1%). Persistent poor readers constituted about 52% of the poor readers and 6% of poor readers were those with early reading problems only.

Our estimate of the prevalence of late-emerging poor readers can be compared to that of others who have examined this group of poor readers. However, given the variability in the criteria used to identify poor readers across studies (e.g., cut-scores of 16th vs. 25th percentile and the use of population vs. sample-based norms), comparisons are best made not in terms of the overall prevalence of late-emerging poor readers in a sample but rather in terms of the proportion of poor readers who are late-emerging. In this regard, the results of our epidemiological-based study are quite similar to those obtained in several of the other investigations. As reported above, we found that 42% of children who were RD at any of the grades we examined were late-emerging poor readers. Leach et al. (2003) found 47% of fourth or fifth graders with RD could be classified as late-emerging whereas Badian (1999) showed that 46% of poor readers in grades 5-8 did not have a RD in grades 1-4. Finally, Lipka et al. (2006) reported a lower prevalence of late-emerging deficits (36%) within fourth grade poor readers. However, in their study, RD was defined on the basis of word reading problems only. Taken together, these results indicate that while many children may experience difficulties in beginning reading, a sizable percentage of poor readers may not manifest their difficulties until fourth grade or later.

Heterogeneity of Late-Emerging Poor Readers

Previous research has demonstrated that late-emerging poor readers represent a heterogeneous group (Leach et al., 2003; Lipka et al., 2006). This work suggests that some children in this group have problems in reading comprehension alone, while other have difficulties in word reading alone or in combination with those in comprehension. Our results also showed heterogeneity among late-emerging poor readers. Approximately half (52%) of the late-emerging poor readers in our sample had problems in reading comprehension alone. Such a finding is not surprising given what is known about reading development and the changing nature of reading curricula. In the initial phases of reading development, children must learn to decode and recognize printed words (Adams, 1990). As a result, most reading curricula place limited demands on comprehension processes and focus primarily on helping children become fluent readers. However, as children progress through the school grades, reading texts (and reading assessments) change to include a greater percentage of informational passages or more complex narratives that place higher demands on the language and cognitive processing needed for comprehension. As a result, some children who have had no problems in the initial phases of reading development may begin to encounter difficulties. Our data suggest that most of these children will start to show reading problems by fourth grade. However, in a few cases, students may encounter

these difficulties between fourth and eighth grades. As discussed in a later section, these problems may result from deficits in language and/or other aspects of cognitive processing.

A cautionary note is necessary in the interpretation of the results of late emerging comprehension deficits. Because one of the reading comprehension tests (i.e., DAB-2) administered in second and fourth grades was not appropriate for older children, another similar measure (i.e., QRI-2) was substituted for it in eighth and tenth grades. Although this measure represented only one of the three assessments in the comprehension composite and was scaled on the same sample, it could be argued that this change in reading tests contributed to transitions in latent classes. However, such a contribution is likely to be small at best since the majority of transitions involving children with late-emerging comprehension deficits took place between second and fourth grades and not between fourth and eighth grades where the switch in reading tests could have impacted results.

Our results also showed that a portion of late-emerging poor readers had problems in word reading alone (36%) or in combination with problems in reading comprehension (12%). In almost all cases these difficulties were first documented in the fourth grade. Leach et al. (2003) and Lipka et al. (2006) also observed late-emerging poor readers who had problems specific to word reading. Consistent with our findings, Leach et al. (2003) reported about a third of late-emerging poor readers demonstrated this profile. These problems may arise as a result of the changes in the nature of the words encountered in the middle elementary school grades. At these grades, children begin to encounter more complex multisyllabic words that require advanced skills in phonological decoding, orthographic processing, and deviational morphology. Thus, children who have gotten off to a good start in word reading may begin to show difficulties if they have deficiencies in these skills (Leach et al. 2003). Alternatively, these children may fall behind not because of inherent limitations but because of a lack of experience and/or practice with reading. Children with limited experience/practice might get off to a good start but may not have sufficient learning opportunities to develop the orthographic knowledge needed to be successful word readers at later elementary grades. Finally, Juel (1991) has also suggested that some children may rely heavily on memorization of words and appear to be successful in beginning word reading, but struggle when such memorization becomes inefficient in the later grades.

History of Language and/or Nonverbal Cognitive Deficits

Previous research on late-emerging poor readers or children with related problems (e.g., poor comprehenders) suggests that deficits in language and/or other cognitive abilities may underlie their reading difficulties (Catts, Adlof, & Weismer, 2006; Compton et al., 2008; Locascio, Mahone, Eason, & Cutting, 2010; Nation et al., 2004). Our data allowed us to examine factors that might underlie late-emerging RD in several ways. First, our participants had been classified in kindergarten as having a specific language impairment (SLI), a specific nonverbal cognitive deficit (SCD), a combined language and cognitive deficit (NLI) or typical language abilities (TL). This allowed us to examine the prevalence of children with a history of each of these classifications among each of the poor reader subtypes. Our results showed that after sample adjusted weighting, all poor reader subtypes contained a significant proportion of children with a history of language and/or nonverbal cognitive deficits. Children with a history of language impairments (i.e., SLI, NLI) were quite prevalent among those with persistent reading problems (EIRD) and late-emerging comprehension difficulties (i.e., LERD-C, LERD-CW). Furthermore, all subtypes had a sizable proportion of children with a history of a specific nonverbal cognitive deficit. In fact, when this proportion was combined with that involving children with a history of both a nonverbal cognitive deficit and a language impairment (i.e., NLI), our findings showed that all late-emerging subtypes had a higher proportion of children with a nonverbal cognitive deficit than with those with a language impairment (SLI and NLI combined).

In another set of analyses, we examined profile differences between reader subtypes on kindergarten measures. These measures included the language and nonverbal cognitive measures that were used in the kindergarten classification discussed above as well as measures of phonological awareness and letter identification. Our results revealed that subtypes differed significantly in their patterns of strengths and weaknesses on these measures. Most children in the TD subtype performed well across all assessments, whereas those in the EIRD subtype had weaknesses on all measures. Both groups of late-emerging poor readers with comprehension problems (i.e., LERD-C, LERD-CW) demonstrated deficiencies in oral language. This result and the findings above that these groups had a high proportion of children with a history of a language impairment is consistent with previous research showing a relationship between oral language and reading comprehension (Kendeou, van den Broek, White, & Lynch, 2009; Storch & Whitehurst, 2002). These results further suggest that developmental language problems may underlie, in part, late-emerging comprehension problems. Deficits in vocabulary, grammar, and/or narrative abilities likely contribute to the difficulties that at least some of these children experience when faced with the linguistically more demanding text encountered in middle elementary school.

The factors that underlie late-emerging comprehension problems appear to go beyond oral language deficits. We also found that these children had weaknesses in nonverbal IQ and were frequently classified as having a nonverbal cognitive deficit in kindergarten. The role of IQ (especially nonverbal IQ) in reading development has been played down for some time (Siegel, 1989; Stanovich, 1991). This has resulted for the most part from a focus on the development of word reading skills that have been shown to be relatively independent of IQ (Stuebing, Fletcher, LeDoux, Lyon, Shaywitz & Shaywitz, 2002; Vellutino, Scanlon, & Lyon, 2000). However, few would question the impact of IQ on reading comprehension. Clearly verbal IQ, which is often assessed by subtests measuring vocabulary or other aspects of oral language, impacts the ability to understand written text. Measures of nonverbal IQ would also seem to be related to comprehension in that these measures assess children's ability to plan, organize, and critically analyze stimuli. In fact, recent research has shown that children with specific reading comprehension problems may have deficits in strategic planning and organizing (Cutting, Materek, Cole, Levine, & Mahone, 2009; Locascio et al., 2010). Thus, it may be that problems in the latter areas contributed to some of our participants' difficulties on nonverbal cognitive measures (e.g., Block Design) and reading comprehension assessments.

The finding that children in the LERD-C and LERD-CW also had deficits in phonological awareness is somewhat surprising in light of the results from studies of poor comprehenders. Poor comprehenders, who have a specific deficit in reading comprehension, often overlap with late-emerging poor readers. Research has typically found that poor comprehenders have poor oral language abilities but good phonological awareness skills (Cain, Oakhill, & Bryant, 2000; Nation et al, 2004). One possible explanation for the inconsistencies in the results may be the time at which phonological awareness abilities have been measured. Studies of poor comprehenders have typically examined phonological awareness in well after children have learned to read. It may be that many poor comprehenders (and children with LERD-CW or LERD-C) have deficits in phonological awareness during preschool/ kindergarten which improve as the result of their developing orthographic knowledge. Bishop McDonald, Bird, and Hayiou-Thomas (2009) have offered a similar account to explain why children with SLI but good word reading (i.e., a group that should overlap with poor comprehenders and LERD) had deficits in phonological awareness in preschool but not in the primary grades.

It is also important to note that children with late-emerging comprehension problems did show some relative strengths in their kindergarten abilities. Both the LERD-C and LERD-CW subtypes performed in the average or above range in letter identification in kindergarten. It is unclear from our data whether this performance was influenced by early literacy experiences or from some inherent cognitive ability. Regardless, their early alphabetic knowledge may have served as a protective factor during the initial phases of learning to read words.

Late-emerging poor readers with word reading problems alone (i.e., LERD-W) demonstrated both strengths and weaknesses in kindergarten measures. Specifically, these children typically scored above average in receptive vocabulary, narrative skills, and letter knowledge. They showed below average performance in expressive vocabulary, some measures of grammar, and phonological awareness. However, the latter deficits were generally rather mild and don't seem sufficient, by themselves, to account for the problems these children experienced in later word reading. Some of children in the LERD-W subtype did have a history of a specific language or nonverbal cognitive impairment. However, it is unclear how such problems might cause specific word reading difficulties in some children and specific comprehension deficits in others. Further research is needed to better understand the factors that cause children to have late-emerging RD both in word reading and comprehension. This work should also go beyond the examination of factors intrinsic to children and consider other variables known to influence reading performance. One of the limitations of this investigation was that we were unable to consider the impact of variables such as reading instruction, experience, and/or motivation on reading outcomes.

The need to explore both intrinsic and extrinsic causal factors in children with LERD is further highlighted by our results concerning the difficulties predicting these reading problems based on kindergarten measures. Multinomial regression analysis indicated that kindergarten measures could accurately identify children with TD, and to a lesser extent, those with EIRD. However, children with LERD were seldom identified accurately. Many were misclassified as children with EIRD but even more were misidentified as children with TD. Thus, despite the fact that children with LERD often have language and/or nonverbal cognitive deficits, these deficits alone (or at least the way we measured them) were not sufficient to distinguish them from children with TD or EIRD. Enough children with TD also had a history of deficits in language and/or cognitive abilities (or at least mild problems in these areas), which made it difficult to distinguish them from children with LERD. As predicted, language and/or cognitive deficits were also quite prevalent in children with EIRD; in fact, the vast majority of these children had deficits in one or both of these areas. We had predicted that children with LERD would perform well in phonological awareness and letter identification and that this performance might distinguish them from children with EIRD. However, whereas many children with LERD had strengths in letter identification, most had at least mild deficits in phonological awareness. The resultant pattern of performance was not sufficiently different to distinguish them from children with EIRD.

Late-Emerging versus Late-Identified

In this paper, we have referred to the children of primary interest as late-emerging poor readers. However, given that most of these children were shown to have a history of language and/or nonverbal cognitive impairments, is it a misnomer to refer to them in this way? Presumably, many of these children evidenced difficulties in language and/or nonverbal cognitive abilities prior to beginning school. Thus, in this sense, their problems might be considered early rather than late to emerge. Nevertheless, in the current paper and other related papers (Leach et al., 2003; Lipka et al., 2006), the term late-emerging is used specifically to refer to reading difficulties. Our results indicate that these children do not have reading problems in the early school grades that are somehow missed and only

identified in the latter grades. As a group, these children score well within the normal range in reading in the early grades and only fall behind in fourth grade and beyond. The early to emerge language and cognitive deficits may partly explain these reading difficulties but are not the defining characteristics of this group. In fact, as noted above, both children with TD and EIRD may have “early emerging” language and nonverbal cognitive deficits as well. Perhaps with further investigation we will find early emerging developmental difficulties that distinguish children with LERD from those with TD and EIRD. When and if such a profile is uncovered, the term late-emerging could be revised to better reflect the developmental nature of the disorder.

References

- Adams, M. *Beginning to read: Thinking and learning about print*. MIT Press; Cambridge, MA: 1990.
- Badian N. Reading disability defined as a discrepancy between listening and reading comprehension. *Journal of Learning Disabilities*. 1999; 32:138–148. [PubMed: 15499714]
- Bishop DVM, McDonald D, Bird S, Hayiou-Thomas ME. Children who read words accurately despite language impairments: Who are they and how do they do it. *Child Development*. 2009; 80:593–605. [PubMed: 19467013]
- Bollen, KA.; Curran, PJ. *Latent curve models: A structural equation perspective*. John Wiley & Sons Inc.; Hoboken, NJ: 2006.
- Cain K, Oakhill J, Bryant P. Phonological skills and comprehension failure: A test of the phonological processing deficit hypothesis. *Reading and Writing: An Interdisciplinary Journal*. 2000; 13:31–56.
- Catts HW. The relationship between speech-language impairments and reading disabilities. *Journal of Speech and Hearing Research*. 1993; 36:948–958. [PubMed: 8246483]
- Catts H, Adolf SM, Hogan TP, Ellis Weismer S. Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language and Hearing Research*. 2005; 48:1378–1396.
- Catts HW, Adlof SM, Ellis Weismer SE. Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*. 2006; 49:278–293.
- Chall, JS. *Stages of reading development*. McGraw-Hill; New York: 1983.
- Collins LM. Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*. 2006; 57:505–528.
- Collins LM, Wugalter SE. Latent class models for stage-sequential dynamic latent variables. *Multivariate Behavioral Research*. 1992; 27:131–157.
- Compton DL, Fuchs D, Fuchs LS, Bryant JD. Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology*. 2006; 98(2):394–409.
- Compton DL, Fuchs D, Fuchs LS, Elleman AM, Gilbert JK. Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. *Learning and Individual Differences*. 2008; 18:329–337.
- Culatta B, Page J, Ellis J. Story retelling as a communicative performance screening tool. *Language, Speech, and Hearing Services in Schools*. 1983; 14:66–74.
- Cutting LE, Materek A, Cole CA, Levine TM, Mahone EM. Effects of fluency, oral language, and executive function on reading comprehension performance. *Annals of Dyslexia*. 2009; 59:34–54. [PubMed: 19396550]
- Denton, CA.; Mathes, PG. Intervention for struggling readers: Possibilities and challenges.. In: Foorman, BR., editor. *Preventing and remediating reading difficulties: Bringing science to scale*. York Press; Timonium, MD: 2003. p. 229-251.
- Duncan, TE.; Duncan, SC.; Strycker, LA. *An introduction to latent variable growth curve modeling: Concepts, issues, and application*. Erlbaum; Mahwah, NJ: 2006.
- Dunn, LM.; Dunn, LM. *Peabody Picture Vocabulary Test - Revised*. American Guidance Service; Circle Pines, MN: 1981.

- Fergusson DM, Horwood LJ, Caspi A, Moffitt TE, Silva PA. The (artefactual) remission of reading disability: Psychometric lessons in the study of stability and change in behavioral development. *Developmental Psychology*. 1996; 32:132–140.
- Foorman, BF.; Fletcher, JM.; Francis, DJ.; Carlson, CD.; Chen, D.; Mouzaki, A.; Schatschneider, C.; Wristers, K.; Taylor, RH. Technical Report Texas Primary Reading Inventory Technical. 1998 Edition. Center for Academic and Reading Skills and University of Houston; Houston, TX: 1998.
- Foorman BR, Francis DJ, Fletcher JM, Schatschneider C, Mehta P. The role of instruction in learning to read: Preventing reading disabilities in at-risk children. *Journal of Educational Psychology*. 1998; 90:37–55.
- Foorman, BR.; Francis, DJ.; Shaywitz, SE.; Shaywitz, BA.; Fletcher, JM. The case for early reading intervention.. In: Blachman, B., editor. *Foundations of reading acquisition and dyslexia: Implications for early intervention*. Lawrence Erlbaum Associates, Inc.; Mahwah, NJ: 1997. p. 243-264.
- Francis DJ, Fletcher JM, Stuebing KK, Lyon GR, Shaywitz BA, Shaywitz SE. Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities*. 2005; 38:98–108. [PubMed: 15813593]
- Graham JW, Collins LM, Wugalter SE, Chung NK, Hansen WB. Modeling transitions in latent stage-sequential processes: A substance use prevention example. *Journal of Consulting and Clinical Psychology*. 1991; 59:48–57. [PubMed: 2002142]
- Hulstlander J, Olson RK, Willcutt EG, Wadsworth SJ. Longitudinal stability of reading-related skills and their prediction of reading development. *Scientific Studies of Reading*. 2010; 14(2):111–136. [PubMed: 20563241]
- Jenkins, JR. Candidate measures for screening at-risk students.. Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention symposium; Kansas City, MO. Dec. 2003
- Juel, C. Beginning reading.. In: Barr, R.; Kamil, ML.; Mosenthal, PB.; Pearson, PD., editors. *Handbook of reading research*. Vol. 2. Longman; New York: 1991. p. 759-788.
- Kaplan D. An overview of Markov chain methods for the study of stage-sequential developmental processes. *Developmental Psychology*. 2008; 44:457–467. [PubMed: 18331136]
- Kaplan D, Walpole S. A stage-sequential model of reading transitions: Evidence from the early childhood longitudinal study. *Journal of Educational Psychology*. 2005; 97(4):551–563.
- Kendeou P, van den Broek P, White MJ, Lynch JS. Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology*. 2009; 101:765–778.
- Lambrecht Smith S, Roberts JA, Locke JL, Tozer R. An exploratory study of the development of early syllable structure in reading-impaired children. *Journal of Learning Disabilities*. 2010; 43:294–307. [PubMed: 20581371]
- Lanza, ST.; Flaherty, BP.; Collins, LM. Latent class and latent transition analysis.. In: Schinka, JA.; Velicer, WF., editors. *Handbook of psychology: Research methods in psychology*. Vol. 2. John Wiley & Sons Inc.; Hoboken, NJ: 2003. p. 663-685.
- Leach JM, Scarborough HS, Rescorla L. Late-emerging reading disabilities. *Journal of Educational Psychology*. 2003; 95(2):211–224.
- Leslie, L.; Caldwell, J. *Qualitative reading inventory-2*. Harper Collins; New York: 1995.
- Lipka O, Lesaux N, Siegel L. Retrospective analyses of the reading development of Grade 4 students with reading disabilities: Risk status and profiles over 5 years. *Journal of Learning Disabilities*. 2006; 39:364–378. [PubMed: 16895160]
- Locascio G, Mahone EM, Eason S, Cutting L. Executive dysfunction among children with reading comprehension deficits. *Journal of Learning Disabilities*. 2010; 20(5):1–14.
- Martin RA, Velicer WF, Fava JL. Latent transition analysis to the stages of change for smoking cessation. *Addictive Behaviors*. 1996; 21:67–80. [PubMed: 8729709]
- McArthur GM, Hogben JH, Edwards VT, Heath SM, Mengler ED. On the “specifics” of specific reading disability and specific language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 2000; 41:869–874.

- Muthen, B. Latent variable mixture modeling.. In: Marcoulides, GA.; Schumacker, RE., editors. Advanced structural equation modeling: New developments and techniques. Erlbaum; Mahwah, NJ: 2000. p. 1-33.
- Muthen, LK.; Muthen, BO. Mplus user's guide and software. 5th ed.. Muthén & Muthén; Los Angeles, CA: 2007.
- Nation K, Clarke P, Marshall CM, Durand M. Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2004; 47:199–211.
- Nation K, Snowling MJ. Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child Development*. 1998; 69:996–1011. [PubMed: 9768483]
- Newcomer, P. Diagnostic achievement battery-2. Pro-Ed.; Austin, TX: 1990.
- Newcomer, P.; Hammill, D. Test of Language Development-2 Primary. Pro-Ed.; Austin, TX: 1988.
- Oakhill J, Cain KE, Bryant PE. The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes*. 2003; 18:443–468.
- Parrila R, Aunola K, Leskinen E, Nurmi J, Kirby JR. Development of individual differences in reading: Results from longitudinal studies in English and Finnish. *Journal of Educational Psychology*. 2005; 97:299–319.
- Shaywitz SE, Shaywitz BA, Fletcher JM, Escobar MD. Prevalence of reading disability in boys and girls: Results of the Connecticut Longitudinal Study. *Journal of the American Medical Association*. 1990; 264(8):998–1002. [PubMed: 2376893]
- Siegel LS. IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities*. 1989; 22:469–479. [PubMed: 2794763]
- Simmons DC, Coyne MD, Kwok O, McDonagh S, Harn BA, Kame'enui EJ. Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities*. 2008; 41:158–173. [PubMed: 18354935]
- Stanovich KE. Discrepancy definitions of reading disability: Has intelligence led us astray? *Reading Research Quarterly*. 1991; 26:7–29.
- Storch S, Whitehurst G. Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*. 2002; 38:934–947. [PubMed: 12428705]
- Stuebing KK, Fletcher JM, LeDoux JM, Lyon GR, Shaywitz SE, Shaywitz BA. Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal*. 2002; 39:469–518.
- Tabachnick, BG.; Fidell, LS. Using multivariate statistics. 5th ed.. Pearson; New York, NY: 2007.
- Tomblin JB, Records N, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*. 1997; 40:1245–1260.
- Vellutino FR, Scanlon DM, Lyon GR. Differentiating between difficult-to-remediate and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities*. 2000; 33:223–238. [PubMed: 15505962]
- Vellutino FR, Scanlon DM, Small S, Fanuele DP. Response to intervention as a vehicle for distinguishing between reading disabled and non-reading disabled children: Evidence for the role of kindergarten and first grade intervention. *Journal of Learning Disabilities*. 2006; 39(2):157–169. [PubMed: 16583795]
- Wagner RK, Torgesen JK, Rashotte CA, Hecht SA, Barker TA, Burgess SR, Donahue J, Garon T. Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology*. 1997; 33:468–479. [PubMed: 9149925]
- Wechsler, D. Wechsler Preschool and Primary Scale of Intelligence-Revised. Psychological Corporation; New York: 1989.
- Wechsler, D. Wechsler Intelligence Scale for Children-III. Psychological Corporation; San Antonio, TX: 1991.
- Wiederholt, JL.; Bryant, BR. Gray oral reading tests. 3rd ed.. Pro-Ed.; Austin, TX: 1994.

- Wood FB, Hill DF, Meyer MS, Flowers DL. Predictive Assessment of Reading. *Annals of Dyslexia*. 2005; 55:193–216. [PubMed: 17849193]
- Woodcock, R.; W.. Woodcock reading mastery test-revised. AGS; Circle Pines, MN: 1987.
- Wright SF, Fields H, Newman SP. Dyslexia: Stability of definition over a five year period. *Journal of Research in Reading*. 1996; 19:46–60.
- Yuill, N.; Oakhill, J. Children's problems in text comprehension: An experimental investigation. Cambridge University Press; New York: 1991.

Figure 1.

Profiles (with 99.92% confidence intervals) of the kindergarten measures for the subtypes. *Note.* PV = picture vocabulary; OV = oral vocabulary; GC = grammatic completion; GU = grammatic understanding; SI = sentence imitation; NC = narrative comprehension; NR = narrative recall; PA = phonological awareness; LID = letter identification; IQ = nonverbal IQ. TD = typically developing; LERD = late emerging RD; EIRD = early identified RD; W = word reading only; C = comprehension only; CW = comprehension + word reading.; LETD = late emerging typically developing

Table 1**Kindergarten Measures**

Vocabulary and Grammar	
<i>Test of Language Development-2: Primary (TOLD-2:P; Newcomer & Hammill, 1988)</i>	
Picture Vocabulary	The child points to pictures named by the examiner.
Oral Vocabulary	The child names pictures presented by the examiner.
Grammatical Understanding	The child points to the picture that illustrates the sentence spoken by the examiner. Tests understanding of grammatical concepts such as past tense, plurals, and negation.
Sentence Imitation	The child repeats sentences spoken by the examiner. Sentences increase in grammatical complexity.
Grammatical Completion	The child supplies correct word to complete sentence spoken by examiner. Tests use of plurals, past tense, comparatives and superlatives, etc.
Narrative Skills (Culatta, Page, & Ellis, 1983)	
Narrative Comprehension	The child answers open-ended questions after hearing a story read aloud by the examiner.
Narrative Expression	The child retells a story read aloud by the examiner.
Phonological Awareness (Catts, 1993)	
Syllable/ Phoneme Deletion	The child repeats a word spoken by the examiner, then deletes specified sound/syllable, and says remaining word. Example, "Say <i>cowboy</i> . Now say cowboy without the <i>cow</i> ."
Letter Knowledge	
<i>Woodcock Reading Mastery Test-Revised (WRMT-R; Woodcock, 1987)</i>	
Letter Identification	The child names printed letters of the alphabet.
Nonverbal IQ	
<i>Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R; Weschler, 1989)</i>	
Block Design	The child reproduces geometric designs using colored blocks.
Picture Completion	The child identifies the missing component of a picture. For example, the child is shown a picture of a face without an ear and responds, "ear."

Table 2
Weighted Means and SDs for the Standard Scores of Selected Measures using National Norms

Grade	K	2	4	8	10
PPVT-R		100.6 (15.9)	102.6 (15.4)	100.7 (15.1)	101.6 (15.9)
NVIQ	100.7 (14.5)	100.3 (15.1)			
WRMT-R: WI		104.5 (18.8)	97.6 (14.9)	96.0 (15.7)	93.1 (9.0)
WRMT-R: WA		94.8 (16.4)	94.2 (15.7)	93.2 (12.6)	91.5 (8.0)
WRMT-R: PC		99.9 (14.9)	96.3 (15.4)	96.0 (15.7)	96.2 (10.3)
GORT-3		93.5 (16.7)	95.6 (17.2)	95.4 (21.5)	99.1 (21.5)

Note. PPVT-R=Peabody Picture Vocabulary Test-Revised; NVIQ=prorated from the Block Design and Picture Completion subtests of the Wechsler Preschool and Primary Scale of Intelligence-Revised (K) or the Performance Scale of the Wechsler Intelligence Scale for Children-III (Grade 2); WRMT-R:WI= Woodcock Reading Mastery Tests-Revised; Word Identification; WRMT-R:WA= Woodcock Reading Mastery Tests-R: Word Attack; WRMT-R:PC= Woodcock Reading Mastery Tests-R: Passage Comprehension; GORT-3=Gray Oral Reading Test-3.

Table 3
Correlation Coefficients for Word Reading and Comprehension Composite Scores Across Grades

Composite Score	1	2	3	4	5	6	7	8
1. Word Reading Grade 2	-							
2. Comprehension Grade 2	.80 [*]	-						
3. Word Reading Grade 4	.90 [*]	.76 [*]	-					
4. Comprehension Grade 4	.66 [*]	.79 [*]	.69 [*]	-				
5. Word Reading Grade 8	.85 [*]	.72 [*]	.93 [*]	.65 [*]	-			
6. Comprehension Grade 8	.58 [*]	.69 [*]	.60 [*]	.80 [*]	.59 [*]	-		
7. Word Reading Grade 10	.82 [*]	.70 [*]	.90 [*]	.64 [*]	.92 [*]	.60 [*]	-	
8. Comprehension Grade 10	.46 [*]	.62 [*]	.50 [*]	.75 [*]	.49 [*]	.84 [*]	.51 [*]	-

* $p < .001$.

Table 4

Class Counts for the NR and RD Latent Classes by Assessment Wave

Wave	Latent Class			
	NR	RD-W	RD-C	RD-CW
Grade 2	349	35	0	109
Grade 4	277	83	55	78
Grade 8	267	78	59	89
Grade 10	267	78	60	88

Note. NR=normal readers; RD-W = reading disabled word reading only; RD-C = reading disabled comprehension only; RD-CW = reading disabled in both comprehension and word reading.

Table 5

Transition Probabilities for the Movers and Stayers

Grade 2 Latent Classes					
Variable	NR	RD-W	RD-C	RD-CW	Proportion of Sample
Movers	.71	.09	.00	.20	.98
Stayers	.00	.00	.00	1.00	.02
Results for Movers					
Grade 2 (rows) by Grade 4 (column)					
NR	.77	.08	.15	.00	
RD-W	.00	.79	.00	.21	
RD-C	.00	.00	.00	.00	
RD-CW	.07	.22	.06	.65	
Grade 4 (rows) by Grade 8 (column)					
NR	.91	.00	.09	.00	
RD-W	.06	.73	.00	.21	
RD-C	.07	.00	.75	.19	
RD-CW	.00	.20	.00	.80	
Grade 8 (rows) by Grade 10 (column)					
NR	.99	.00	.01	.00	
RD-W	.00	.91	.00	.09	
RD-C	.00	.00	.95	.05	
RD-CW	.00	.06	.00	.94	
Fit Indices					
$LR^2(1,048,172, N=493) = 726.01, p > .05$					
BIC = 9062.36					

Note. NR=normal readers; RD-W = reading disabled word reading only; RD-C = reading disabled comprehension only; RD-CW = reading disabled in both comprehension and word reading; BIC = Bayesian information criterion.

Table 6

Frequency Counts, Percentages, and Population Estimates for the Identified Latent Classes

Class	Latent Class Designation				Frequency	Percent	Population Estimate
	Grade 2	Grade 4	Grade 8	Grade 10			
TD	1	1	1	1	255	51.7	67.8
TD	1	2	1	1	1	0.2	0.1
Total					256	51.9	67.9
LERD-W	1	2	2	2	26	5.2	4.5
LERD-W	1	3	4	2	3	0.6	0.3
Total					29	5.8	4.8
LERD-CW	1	3	4	4	8	1.6	1.0
LERD-CW	1	3	3	4	3	0.6	0.3
LERD-CW	1	1	3	4	1	0.2	0.1
LERD-CW	1	2	4	4	1	0.2	0.2
Total					13	2.6	1.6
LERD-C	1	3	3	3	37	7.5	5.1
LERD-C	1	1	3	3	14	2.8	1.9
Total					51	10.3	7.0
LETD	4	1	1	1	7	1.4	1.0
LETD	4	2	1	1	3	0.6	0.5
LETD	2	2	1	1	1	0.2	0.4
Total					11	2.2	1.9
EIRD-W	2	2	2	2	19	3.8	3.0
EIRD-W	4	4	2	2	15	3.0	2.0
EIRD-W	4	2	2	2	13	2.6	2.9
EIRD-W	2	4	4	2	1	0.2	0.1
EIRD-W	4	4	4	2	1	0.2	0.1
EIRD-CW	4	4	4	4	52	10.5	5.2
EIRD-CW	4	2	4	4	11	2.2	0.9
EIRD-CW	2	2	4	4	4	0.8	0.3
EIRD-CW	2	4	4	4	7	1.4	0.8
EIRD-CW	4	2	4	4	1	0.2	0.4
EIRD-C	2	2	2	3	3	0.6	0.4
EIRD-C	4	4	2	3	2	0.4	0.4
EIRD-C	4	3	3	3	4	0.8	0.3
Total					133	26.9	16.8

Note. TD = typically developing; LERD = late emerging RD; LETD = late emerging typically developing; EIRD = early identified RD; W = word reading only; C = comprehension only; CW = comprehension + word reading; 1 = NR; 2 = RD-W; 3 = RD-C; 4 = RD-CW.

Table 7
Reading Profiles, Expressed as Norm-References Composite Z-scores, for the Latent Classes as a Function of Grade

Measure	TD	Latent Class					LETD
		LERD-W	LERD-C	LERD-CW	EIRD		
Grade 2							
Word Reading	.39	-.60	-.07	-.50	-1.35	-.87	
Comprehension	.29	-.49	-.55	-.73	-1.33	-.89	
Grade 4							
Word Reading	.43	-.88	-.09	-.69	-1.39	.02	
Comprehension	.25	-.47	-.91	-1.28	-1.15	-.16	
Grade 8							
Word Reading	.42	-1.06	-.02	-.99	-1.38	.05	
Comprehension	.24	-.59	-1.15	-1.56	-1.03	-.18	
Grade 10							
Word Reading	.40	-1.21	-.08	-1.03	-1.25	.26	
Comprehension	.21	-.49	-1.31	-1.55	-.91	-.01	

Note. TD = typically developing; LERD = late emerging RD; EIRD = early identified RD; LETD = late emerging typically developing; W = word reading only; C= comprehension only; CW = comprehension + word reading.

Table 8

Class Count, Percentage, and Population Estimate for Latent Classes as a Function of Kindergarten Language Classification

Class	Kindergarten Language Classification			
	TL (n=246)	SLI (n=100)	NLI (n=72)	SCD (n=75)
TD				
Count	185	30	12	29
Percentage	72.3	11.7	4.7	11.3
Population Estimate	87.8	3.3	2.1	6.9
LERD-W				
Count	11	9	3	6
Percentage	37.9	31.0	10.3	20.7
Population Estimate	62.5	16.7	4.2	16.7
LERD-C				
Count	13	16	8	14
Percentage	25.5	31.4	15.7	27.5
Population Estimate	42.9	17.1	14.3	25.7
LERD-CW				
Count	3	4	3	3
Percentage	23.1	30.3	23.1	23.1
Population Estimate	50.0	12.5	12.5	25.0
EIRD				
Count	30	39	44	20
Percentage	22.6	29.3	33.1	15.0
Population Estimate	43.9	17.1	23.2	15.9
LETD				
Count	4	2	2	3
Percentage	36.4	18.2	18.2	27.3
Population Estimate	55.6	11.1	11.1	22.2

TD = typically developing; LERD = late emerging RD; EIRD = early identified RD; LETD = late emerging typically developing; TL=typical language; SLI = specific language impairment; NLI = nonspecific language impairment; SCD = specific nonverbal cognitive deficit; W = word reading only; C= comprehension only; CW = comprehension + word reading;

Table 9

Correlation Coefficients Between Word Reading and Comprehension Composite Scores and Kindergarten Early Predictors

Composite Score	Kindergarten Measures									
	PV	OV	GC	GU	SI	NC	NR	PA	LID	IQ
1. WR Grade 2	.36 [*]	.48 [*]	.45 [*]	.49 [*]	.54 [*]	.35 [*]	.25 [*]	.58 [*]	.57 [*]	.37 [*]
2. Comp Grade 2	.50 [*]	.58 [*]	.58 [*]	.56 [*]	.61 [*]	.42 [*]	.36 [*]	.59 [*]	.59 [*]	.52 [*]
3. WR Grade 4	.32 [*]	.44 [*]	.44 [*]	.43 [*]	.52 [*]	.31 [*]	.23 [*]	.54 [*]	.50 [*]	.36 [*]
4. Comp Grade 4	.50 [*]	.57 [*]	.59 [*]	.56 [*]	.62 [*]	.43 [*]	.35 [*]	.55 [*]	.45 [*]	.51 [*]
5. WR Grade 8	.33 [*]	.45 [*]	.45 [*]	.44 [*]	.53 [*]	.28 [*]	.18 [*]	.56 [*]	.48 [*]	.35 [*]
6. Comp Grade 8	.49 [*]	.56 [*]	.57 [*]	.55 [*]	.58 [*]	.42 [*]	.29 [*]	.49 [*]	.36 [*]	.49 [*]
7. WR Grade 10	.33 [*]	.47 [*]	.49 [*]	.46 [*]	.52 [*]	.29 [*]	.20 [*]	.55 [*]	.46 [*]	.34 [*]
8. Comp Grade 10	.47 [*]	.51 [*]	.54 [*]	.53 [*]	.55 [*]	.40 [*]	.32 [*]	.47 [*]	.32 [*]	.45 [*]

Note. WR = word reading; Comp = comprehension; PV = picture vocabulary; OV = oral vocabulary; GC = grammatic completion; GU = grammatic understanding; SI = sentence imitation; NC = narrative comprehension; NR = narrative recall; PA = phoneme awareness; LID = letter identification; IQ = nonverbal IQ.

* $p < .001$.

Table 10

Classification Accuracy of the Multinomial Logistic Regression Model

Observed	Predicted			Percent Correct
	TD	LERD	EIRD	
TD	220	11	25	85.9%
LERD	47	14	32	15.1%
EIRD	39	5	89	66.9%
Overall Percent	63.5%	6.2%	30.3%	67.0%