Article

Hammill Institute on Disabilities

Journal of Learning Disabilities
2016, Vol. 49(1) 21–35

DOI: 10.1177/0022219414522706
journaloflearningdisabilities.sagepub.com
SAGE

A Longitudinal Examination of the Persistence of Late Emerging Reading Disabilities

Jill M. Etmanskie, PhD[1], Marita Partanen, MSc, MEd[1], and Linda S. Siegel, PhD[1]

Abstract

There are some children who encounter unexpected reading difficulties in the fourth grade. This phenomenon has been described as late emerging reading disabilities (LERD). Using Grade 4 as a starting point, this study examined the reading development of 964 children between kindergarten and Grade 7. The results showed that 72.0% of children had typical reading performance in Grade 4, whereas there was 0.7% with poor word reading, 12.6% with poor reading comprehension, 2.5% with poor word reading and comprehension, and 12.2% with borderline performance. We also showed that there were similar proportions of children who had early versus late emerging reading difficulties; however, most of the late emerging poor readers recovered by Grade 7. Furthermore, our study showed that poor comprehenders showed poorer performance than typical readers on word reading, pseudoword decoding, and spelling between Grade 1 and Grade 7 and poorer performance on a working memory task in kindergarten. Overall, this study showed that most children recover from late emerging reading problems and that working memory may be an early indicator for reading comprehension difficulties.

Keywords

reading disabilities, reading comprehension, late emerging, longitudinal, children

The fourth grade has long been considered a critical point in reading development during which time some children experience an unexpected drop in reading achievement. This trend is thought to be related to the pedagogical shift that occurs in the school system between the third and fourth grades. From kindergarten to the third grade, the predominant focus is on teaching foundational reading skills, such as letter knowledge and word reading. In fourth grade, however, the focus changes from learning to read to reading to learn (Chall, 1983). Grade 4 reading material becomes significantly more complex as textbooks serve as the mechanism to gain access to all other curriculum, content, and knowledge. Storybooks and narrative texts are replaced with math and social studies textbooks. Reading instruction is not continued past the primary grades, and given the complexities of the advanced literacy demands, a portion of skilled readers can suddenly present with significant reading problems.

Research has shown that 5% to 10% of students acquire reading problems for the first time after Grade 3 (Catts & Hogan, 2002; Juel, 1988; Scarborough, 1998). The "fourth grade slump" was described as the sudden drop in reading performance between the third and fourth grades (Chall & Jacobs, 2003). This phenomenon has been linked to visible minorities and children from low-income families based on the assumption that the language skills of children from less affluent environments would decelerate over time (Chall, Jacobs, & Baldwin, 1990; Kieffer, 2010). However, others have found this phenomenon to exist despite income level (Leach, Scarborough, & Rescorla, 2003) or language status (Lipka, Lesaux, & Siegel, 2006). Leach and colleagues referred to this phenomenon as late emerging reading disabilities (LERD). Studies have estimated the proportion of LERD among older elementary and middle school students to be as high as 41% to 46% (Leach et al., 2003; Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992). These prevalence rates have become an increasing concern because children who are found to have literacy difficulties beyond the third grade are not usually attended to with the same urgency as their younger peers (DeFord, Lyons, & Pinnell,

[1]University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Marita Partanen, University of British Columbia, 2125 Main Mall, Vancouver, BC V6T1Z4, Canada.
Email: partanen@alumni.ubc.ca

1991). In addition, intervention efforts beyond Grade 3 do not attract the same level of school funding, further limiting the amount of help available to these children (Worthy, Patterson, Salas, Prater, & Turner, 2002). The longer students with LERD continue without adequate assistance and support, the more pervasive their deficits become. Consequently, this progression will likely affect other areas of learning, behavior, and self-motivation (Tunmer, Chapman, Greaney, & Prochnow, 2002).

In Chall's theory of reading, Grade 4 represents a key transitional point in schooling where children are thought to have mastered the basic reading skills required for integrating vocabulary knowledge with their ability to think critically about the broader text (Chall, 1983; Chall & Jacobs, 2003). The abundance of more advanced text is thought to cause reading performance to temporarily drop until children adjust to the higher cognitive demands. For instance, higher-level schoolwork requires the use of certain cognitive abilities that, up until Grade 4, may have been underutilized or unknowingly weak all along (Snow, Burns, & Griffin, 1998). Given that each subsequent year of school requires more qualitatively and quantitatively advanced readings, children with less efficient cognitive processing abilities (e.g., poor working memory) will struggle to sustain the same level of reading as before. Moreover, children with poor reading skills may read less and avoid challenging texts, which in turn has negative consequences for the development of vocabulary knowledge (Chall, 1983; Stanovich, 1986).

In some cases, the occurrence of LERD may be the result of flaws in the identification process. For example, children with good behavior, high ability, and good compensation strategies are less likely to appear as struggling readers (Leach et al., 2003). Another possibility is that the early screening tools used to identify at-risk readers are ineffective for detecting those at-risk for LERD (Leach et al., 2003). If the latter assumption is correct, then it is unclear whether the absence of good early indicators of LERD is a result of the fact that they are truly indistinguishable from their typical reader counterparts, or if the situation is more reflective of inadequacies within the current assessment tools for identifying at-risk children.

To gain a broader understanding of LERD, it is important to consider these findings within the broader domain of reading development. Reading is a multifaceted process in which individuals must use and integrate both lower-level and higher-level word reading skills to achieve comprehension. Research has shown that deficits in phonological awareness, which is the ability to detect and manipulate sounds in oral language (e.g., rhyming, phoneme blending, phoneme deletion), represent a core deficit in both children and adults with reading disabilities (reviewed in Adams, 1990). However, difficulties in working memory, oral language skills, and rapid naming are also important for reading development (reviewed in Lesaux & Siegel, 2003; Norton & Wolf, 2012). Working memory is conceptualized as a limited resource system responsible for the brief storage and processing of visual or auditory information (Baddeley & Hitch, 1994). Oral language skills, such as the knowledge of syntax, morphology, and vocabulary, are important factors in reading development (Elbro & Scarborough, 2004). Finally, rapid naming is the ability to automatically perceive what is seen (letters, numbers, objects) with what is said; it has been suggested that rapid naming and reading involve some of the same processes such as automaticity and working memory, and this is why rapid naming is related to reading proficiency (Norton & Wolf, 2012).

Children with LERD have been classified into three reading subtypes: children with primary problems in word recognition, children with primary problems in comprehension, and children with both word recognition and comprehension difficulties (Leach et al., 2003; Shankweiler et al., 1999). Young children with poor word recognition demonstrate weaknesses on phonological processing tasks (e.g., Siegel, 1986; Stanovich & Siegel, 1994), and they may also perform poorly on verbal short-term memory tasks (Adams, 1990). Children with poor reading comprehension may have difficulty with oral language measures, especially those that tap vocabulary knowledge, listening comprehension, semantics, and morphology (Cain, Oakhill, & Bryant, 2004). Other sources of reading comprehension difficulties include working memory and knowledge of comprehension strategies (Swanson, Kehler, & Jerman, 2010). Children with both poor word recognition and reading comprehension may have difficulty with tasks of phonological awareness, rapid naming, listening comprehension, vocabulary knowledge, and nonverbal reasoning skills (Leach et al., 2003). In summary, these studies indicate that children with poor word recognition and/or reading comprehension may have different underlying deficits.

Previous longitudinal research has examined the composition of both early and LERD. Leach et al. (2003) found that of the 31 students identified as LERD in Grades 4 and 5, 36% were poor decoders, 32% had a specific comprehension deficit, and 32% had difficulty with both decoding and comprehension. Since more than two thirds of the late emergers also had at least some difficulty with word identification and decoding, their study provided support that many children who exhibit LERD have difficulty at the word level. More recently, Catts, Compton, Tomblin, and Bridges (2012) showed that of the 93 students with LERD, 36% had late emerging word reading difficulties, 52% had late emerging comprehension difficulties, and 12% had difficulties in both word reading and comprehension. These latter results suggest that children with LERD may be more likely to have difficulties with reading comprehension than with word recognition, but further evidence is needed because of these conflicting results.

The purpose of the current study was to explore LERD with a longitudinal study, allowing us to determine the permanence or lack of the LERD. Using a longitudinal cohort of children between kindergarten and Grade 7, we examined the prevalence and persistence of early and late emerging reading difficulties. This cohort was evaluated previously when children were in Grade 4 (Lipka et al., 2006). These authors showed that there were three subgroups of children with reading disabilities: those who had consistently below average reading scores between kindergarten and Grade 4, those who fluctuated between below average and average scores between kindergarten and Grade 4, and those who had late emerging reading difficulties in Grade 4. The purpose of the current study was to expand these previous results by evaluating the children's performance up to Grade 7. Specifically, we evaluated the number of children who had reading problems for the first time in Grade 4 and examined the persistence of reading and reading-related cognitive abilities. We also evaluated whether there were any differences between early and late emerging poor readers' performance on reading and cognitive measures in kindergarten. Our results expand previous longitudinal research in LERD (e.g., Catts et al., 2012; Compton, Fuchs, Fuchs, Elleman, & Gilbert, 2008; Leach et al., 2003; Lipka et al., 2006) by examining children's performance each year from kindergarten to Grade 7 on a variety of reading and cognitive measures.

Method

Participants

The children in the present study were part of a district-wide longitudinal study of literacy development that began in their kindergarten year. This school district implemented an early literacy program called Firm Foundations (North Vancouver School District, 2001). All children received phonological awareness training, and those who were identified as at risk for reading difficulty received small group intervention in kindergarten. The school district also implemented the Reading 44 program in Grades 1 through 7 (North Vancouver School District, 1999). This program included reading and instructional activities that placed less emphasis on phonological awareness and more emphasis on strategies for reading comprehension. Struggling readers continued to receive intervention from teachers across grades. Annual assessments were conducted from kindergarten to Grade 7 in all 30 of the participating elementary schools.

Of particular interest to this study were children's Grade 4 reading scores. Using this grade as a focal point, there were a total of 1,212 children who completed the reading and cognitive measures in Grade 4 ("overall sample"). The mean age was 9.76 years old (SD = 0.41). Of these 1,212 children, there were 590 females and 622 males, 943 children with English as a first language (L1) and 269 children with English as a second language (ESL). In all, 18 languages were represented in the overall sample, including languages from Europe, the Middle East, South Asia, East Asia, and Southeast Asia. School catchment area is a reliable measure of socioeconomic status (SES; D'Angiulli & Siegel, 2004; Tremblay, Ross, & Berthelot, 2001); we included all catchments from a school district, and thus a variety of SES levels were represented.

Because one of the purposes of this study was to examine reading performance over time, it was necessary to exclude children who did not have at least one reading score before Grade 4 and at least one reading score after Grade 4. From the overall sample (n = 1,212), there were 964 children who met the criterion of having data before and after Grade 4 ("Grade 4 sample"). These children had a mean age of 9.77 years (SD = 0.40). Of these 964 children, there were 495 females and 469 males, 783 L1 and 181 ESL children.

Our main analyses were conducted with children with typical reading ability and children with poor reading comprehension (as described further below, there was a small number of children with poor word reading skills in Grade 4). This included 815 children with a mean age of 9.76 years (SD = 0.39) in Grade 4 ("longitudinal sample"). Of these 815 children, there were 430 females and 385 males, 671 L1 and 144 ESL children. However, our longitudinal sample included different sample sizes in each grade due to attrition. Rather than using listwise deletion for missing cases, we retained all cases that met our criteria for the Grade 4 sample and the longitudinal sample. Data were missing at random and were not related to student's academic performance, gender, or language. The sample sizes for each grade were as follows: kindergarten (n = 665), Grade 1 (n = 662), Grade 2 (n = 729), Grade 3 (n = 801), Grade 4 (n = 815), Grade 5 (n = 716), Grade 6 (n = 749), and Grade 7 (n = 639). The proportion of language status (L1/ESL) was similar between Grade 1 and Grade 7, $\chi^2(1)$, $p > .15$, but there were slightly more L1 children than ESL children in kindergarten, $\chi^2(1) = 6.12$, $p = .01$.

Measures

The following measures were administered according to the child's grade. All of the measures were individually administered by trained graduate students, except for the reading comprehension and spelling tests, which were administered to a group. The total testing time was approximately 30 min per child and 40 min per group. Estimates of reliability are presented with Cronbach's alpha coefficient (α), using our longitudinal sample (n = 815).

Reading

Letter identification. Each child was shown 26 lowercase letters in random order, and he or she was asked to name the letter. This task has a maximum score of 26 (α = .94), and it was administered in kindergarten.

Wide Range Achievement Test–3 (WRAT-3): Reading. On WRAT-3 Reading (Wilkinson, 1993), each child was asked to read a list of words of increasing difficulty (e.g., *cat, stretch, itinerary*). Age level percentile scores are reported (α = .86-.92). This task was administered from kindergarten to Grade 7.

Stanford Diagnostic Reading Test–4 (SDRT-4): Reading comprehension. This test (Karlsen & Gardner, 1995) was administered in groups. Each child was required to read several short passages and answer the corresponding multiple-choice questions within a 20-min time limit. Age level percentile scores are reported (α = .79-.94). This task was administered from Grade 2 to Grade 7.

Spelling

Simple spelling. Children were asked to print their names and five simple words (i.e., *no, mom, dad, I, cat*). Raw scores are reported with a maximum score of 6 (α = .76). This task was administered in kindergarten.

WRAT-3: Spelling. This subtest (Wilkinson, 1993) was administered in groups. The children were required to produce the correct spelling of verbally presented words of increasing difficulty. Age level percentile scores are reported (α = .85-.95). This task was administered from Grade 2 to Grade 7.

Phonological processing

Rhyme detection. This task is from the *Phonological Abilities Test* (PAT; Muter, Hulme, & Snowling, 1997). In this task, the child was shown four pictures. A picture of the target word appeared above three pictures. Each child was asked which of the three words rhymed with the target word (e.g., "What rhymes with boat? Foot, bike, or coat?"). Raw scores are reported with a maximum score of 8 (α = .85). This task was administered in kindergarten.

Syllable identification and phoneme identification. These tasks were from the PAT (Muter et al., 1997). In the syllable identification task, the examiner presented a picture (e.g., rabbit) to the child. The examiner said the first part of the word (i.e., *ra*) and asked the child to finish the word (i.e., *bit*). In the phoneme identification task, the examiner presented a picture (e.g., *fish*) and said the first part of the word (i.e., *fi*) and asked the child to complete the word (i.e., *sh*). There was a maximum score of 8 on the syllable task (α = .85) and a maximum score of 8 on the phoneme task (α = .90). This task was administered in kindergarten.

Phoneme deletion. This task is from the PAT (Muter et al., 1997). The examiner presented the children with a picture of a word, and the child was asked to delete a phoneme (initial or final) from the word. There were eight initial phoneme deletion items and eight final phoneme deletion items. Examples of initial phoneme deletion include "Seat without /s/ says ____" and "Cake without /c/ says ____." Examples of the final phoneme deletion include "Sad without the /d/ says ____" and "Cup without the /p/ says ____." Raw scores are reported with a maximum score of 16 (α = .93). This task was administered in kindergarten.

Woodcock Reading Mastery Tests–Revised: Word Attack. This subtest (Woodcock, 1987) consists of a list of pseudowords of increasing difficulty, which were all pronounceable (e.g., *zurk, plib, straced*). Children were asked to decode as many pseudowords as possible. Age level percentile scores are reported (α = .95-.99). This task was administered from Grade 1 to 7.

Working memory and language

Memory for sentences. This task came from the *Stanford–Binet Intelligence Scales, Fourth Edition* (Thorndike, Hagen, & Sattler, 1986). Each child was asked to repeat sentences verbatim from simple two word sentences (e.g., "Tall girl") to complex sentences (e.g., "It was raining this morning, so the children carried umbrellas to school"). Raw scores are reported with a maximum score of 42 (α = .86). This task was administered in kindergarten.

Working memory for numbers. In this task (Siegel & Ryan, 1989), children were asked to count yellow dots that were presented randomly among blue dots. In each set, the task became more difficult by increasing the number of cards. Children were then asked to recall the number of yellow dots on each card as well as the order in which they were presented. The task was discontinued when all answers at one level were incorrect. The raw score was computed as the number of correct responses (α = .70-.73). This task was administered from Grade 2 to Grade 7.

Rapid naming. This measure was a variation of the rapid automatized naming task from Denckla and Rudel (1976). This task was used to assess lexical access, or word retrieval. In kindergarten, each child was asked to name 40 items on a page consisting of line drawings of 5 items (i.e., tree, chair, bird, pear, car). The same items were repeated in random order eight times. The score was the time required to name all 40 pictures (seconds). In Grades 2 to 7, children were asked to name 9 numbers (1 to 9) that were presented on a page in random order. The numbers formed 5 rows and 5 columns. The score was the time required to name all 25 numbers (seconds).

Table 1. Characteristics of the Grade 4 Sample.

Group[a]	*n* (% of Total)	Age in Grade 4 (*M, SD*)	Gender (% Female, Male)	Language (% L1, ESL)
Typical reader	694 (72.0)	9.77, 0.39	56.2, 43.8	82.7, 17.3
Poor word reader	7 (0.7)	9.90, 0.47	71.4, 28.6	85.7, 14.3
Poor comprehender	121 (12.6)	9.74, 0.39	33.1, 66.9	80.2, 19.8
Poor word reader and comprehender	24 (2.5)	9.77, 0.41	41.7, 58.3	79.2, 20.8
Borderline scores	118 (12.2)	9.80, 0.42	42.4, 57.6	73.7, 26.3

Note. ESL = English as a second language; L1 = first language.
[a]Groups are defined by Grade 4 performance.

Oral cloze. Syntactic awareness was assessed using an oral cloze task (Siegel & Ryan, 1989; Willows & Ryan, 1986). Sentences were read to the child, and then the child provided the missing word in each sentence. An example of this task includes "The moon shines bright in the ____." Raw scores are reported with a maximum score of 11 to 20 items depending on the grade ($\alpha = .84$, from the test's development). In our study the oral cloze task is used as an operational definition of grammatical knowledge and specifically, a measure of syntax.

Group Definition

Children's performances on the Grade 4 WRAT-3 Reading and SDRT-4 Reading Comprehension tasks were used to determine group status. Children who scored at or below the 25th percentile were considered to be poor readers, whereas children who scored at or above the 35th percentile were considered to be typical readers. The decision for this cutoff was guided by our earlier study, which used the same criteria (Lipka et al., 2006). Children who scored at or above the 35th percentile on both tasks were considered "typical readers." Children who scored at or below the 25th percentile on the word reading task and at or above the 35th percentile on the reading comprehension task were considered "poor word readers." Children who scored at or below the 25th percentile on the reading comprehension task and at or above the 35th percentile on the word reading task were considered "poor comprehenders." Children who scored at or below the 25th percentile on both word reading and comprehension tasks were considered "poor word readers/comprehenders." Children who obtained scores between the 26th and 34th percentile on either task were considered "borderline readers." Using the Grade 4 sample ($n = 964$), there were 694 typical readers, 7 poor readers, 121 poor comprehenders, 24 poor word readers/comprehenders, and 118 borderline readers. Table 1 shows the descriptive statistics for each group.

We then grouped children with poor word reading and/or reading comprehension into early emerging, late emerging, or inconsistent readers, according to their performance in Grades 1, 2, and 3 (or Grades 2 and 3 for reading comprehension). Children who had below average scores in all grades (i.e., ≤ 25th percentile) were considered "early emerging" poor readers, those who had average scores in all grades (i.e., ≥ 35th percentile) were considered "late emerging" poor readers, and those who fluctuated between the average and below average ranges were considered "inconsistent" poor readers.

Because of the small number of children with poor reading skills, formal analyses were conducted with children with typical reading ability and children with poor reading comprehension only. We have included descriptives for the poor word reader and poor word reader/comprehender groups for illustrative purpose in Tables 1 and 2.

Results

Prevalence of LERD

Our first analysis was to determine the number of children who developed poor reading skills for the first time in Grade 4. Using the criteria described above, there were a total of 7 children with poor word reading skills, of whom 3 were early poor readers, 2 were late poor readers, and 2 were inconsistent poor readers. There were also 121 children with poor reading comprehension, of whom 32 were early poor comprehenders, 32 were late poor comprehenders, and 57 were inconsistent comprehenders. Finally, there were 24 children with both poor word reading and comprehension, of whom 5 were early poor readers and 19 were inconsistent poor readers. These results indicate that there were similar proportions of children with early versus late emerging word reading or reading comprehension difficulties, whereas there were more children who had fluctuating scores between grades. Table 2 shows the descriptive statistics for these groups.

Differences in Age, Gender, and Language Status

An ANOVA showed that mean age was similar between the typical readers and poor comprehender groups, $F(3, 811) = 0.90$, $p > .40$. Chi-square analyses were completed to evaluate whether proportions of gender and language status in

Table 2. Characteristics of Early Emerging, Late Emerging, and Inconsistent Readers.

Poor Reader Group	Group[a]	*n* (% Within Poor Reader Group)	Age in Grade 4 (*M, SD*)	Gender (% Female, Male)	Language (% L1, ESL)
Poor word reader	Early	3 (42.9)	9.93, 0.12	66.7, 33.3	66.7, 33.3
	Late	2 (28.6)	9.85, 0.07	50.0, 50.0	100.0, 0.00
	Inconsistent	2 (28.6)	9.90, 1.13	100.0, 0.00	100.0, 0.00
Poor comprehender	Early	32 (26.4)	9.66, 0.38	31.3, 68.8	71.9, 28.1
	Late	32 (26.4)	9.81, 0.44	28.1, 71.9	87.5, 12.5
	Inconsistent	57 (47.1)	9.75, 0.36	36.8, 63.2	80.7, 19.3
Poor word reader and comprehender	Early	5 (20.8)	9.78, 0.39	60.0, 40.0	80.0, 20.0
	Late	0 (0.00)	—	—	—
	Inconsistent	19 (79.2)	9.77, 0.42	36.8, 63.2	78.9, 21.1

Note. ESL = English as a second language; L1 = first language.
[a]Groups are defined by Grade 1, 2, and 3 performance.

each group was similar to the total sample (51 female:49 male; 81 LI:19 ESL). The results showed that the four groups had similar proportions of L1 and ESL children compared to the total sample, typical readers $\chi^2(1) = 1.32$, $p > .25$; early poor comprehenders $\chi^2(1) = 1.73$, $p > .15$; late poor comprehenders $\chi^2(1) = 0.88$, $p > .30$; and inconsistent comprehenders $\chi^2(1) = 0.01$, $p > .95$. However, all groups had different proportions of males and females. The typical reader group had more females than males compared to the total sample, $\chi^2(1) = 7.50$, $p < .01$, whereas the poor comprehender groups had more males than females compared to the total sample, early poor comprehenders $\chi^2(1) = 5.00$, $p < .05$; late poor comprehenders $\chi^2(1) = 6.70$, $p < .01$; and inconsistent comprehenders $\chi^2(1) = 4.57$, $p < .05$. These results indicate that although language status is similar between groups, there are significant gender differences. Finally, analyses were not conducted with the poor word reader groups because of small sample size; however, mean age was similar to the typical reader group, whereas the proportions of gender and language status fluctuated between groups.

Persistence of LERD

The persistence of reading ability and differences between typical reader and poor comprehender groups was explored using multilevel modeling (MLM; Shek & Ma, 2011; Tabachnick & Fidell, 2006). This analysis allowed us to examine the persistence of abilities in our longitudinal design, without having to use listwise deletion for missing cases. Analyses for the reading comprehension, word reading, spelling, and word attack measures were conducted on percentile ranks due to more desirable statistical properties on that scale (for a discussion, see Zimmerman & Zumbo, 1993, 2005), and analyses for the working memory and rapid naming tasks were conducted with raw scores. Finally, because the oral cloze task had a different maximum value across grades, *z* score transformations were conducted on these data to make them suitable for comparisons across grades.

Our MLM design was conducted as follows. Test scores were grand-mean centered and groups were dummy-coded to compare the poor comprehender groups to the typical reader group. We included time as a repeated measures factor (i.e., Grades 1–7 or Grades 2–7) and gender and group as between-subjects factors. We also included quadratic effects of time because we expected that some groups would not show a linear relationship between time and test scores; for example, late poor comprehenders might show a quadratic effect (U shape) because they had a drop in their reading performance in Grade 4. We were interested in the fixed effects of the model, that is, the average differences between groups and time points. The results were corrected for multiple comparisons at $p < .001$.

A summary of the MLM results is shown in Table 3, and group performance is illustrated in Figures 1 through 7. The summary table shows that males performed more poorly than females on the reading comprehension test only. Across groups, there was a linear effect of time on word reading, spelling, word attack, working memory, and rapid naming tests. This linear effect demonstrates that scores increased significantly over time (and decreased on the rapid naming test). The positive quadratic effect of time (U shape) on the reading comprehension and rapid naming tests indicates that there was a depression in scores in the middle grades, but these improved slightly over time. The negative quadratic effect of time (inverted U shape) on word reading, spelling, word attack, and working memory tests indicates that there was an increase in scores in the middle grades in comparison to earlier and later grades. Overall, these results demonstrate that test scores changed across time in a linear or quadratic fashion.

Table 3. Fixed Effects (*t*-Value) From Longitudinal Analyses for Typical Reader and Poor Comprehender Groups.

Parameter	Reading Comprehension	Word Reading	Spelling	Word Attack	Working Memory	Rapid Naming	Oral Cloze
Gender[a]	–3.33*	2.41	0.07	1.29	1.33	–0.37	–1.36
Time[b]	–2.36	8.79*	13.40*	16.09*	11.16*	–31.22*	–0.38
Time (quadratic)[c]	7.94*	–8.23*	–7.15*	–12.36*	–5.03*	15.00*	0.13
Early poor comprehender[d]	–13.18*	–7.25*	–6.21*	–6.52*	–1.73	2.75	–6.75*
Late poor comprehender[d]	–3.08	–4.53*	–3.31*	–3.45*	–0.38	0.65	–3.04
Inconsistent comprehender[d]	–11.05*	–5.59*	–3.51*	–5.52*	–5.08*	3.93*	–6.68*
Time × early poor	2.49	0.63	–1.02	1.82	–0.26	0.17	1.43
Time × late poor	–6.85*	0.57	0.63	0.59	–1.09	–0.95	–0.05
Time × inconsistent	–5.28*	–0.04	–5.85*	0.64	1.14	–2.40	0.85
Time (quadratic) × early poor	–2.37	–0.56	1.38	–1.46	0.48	–0.55	–0.90
Time (quadratic) × late poor	5.95*	–0.09	–1.14	0.19	0.20	0.72	–0.51
Time (quadratic) × inconsistent	4.15*	0.36	5.28*	0.12	–0.56	1.63	–0.42

[a]Negative values indicate males scored lower than females. [b]Positive values indicate linear increase over time. [c]Positive values indicate quadratic growth over time (i.e., U shape), negative values indicate inverted U shape. [d]Negative values indicate worse performance compared to typical readers.
*Significant at $p < .001$.

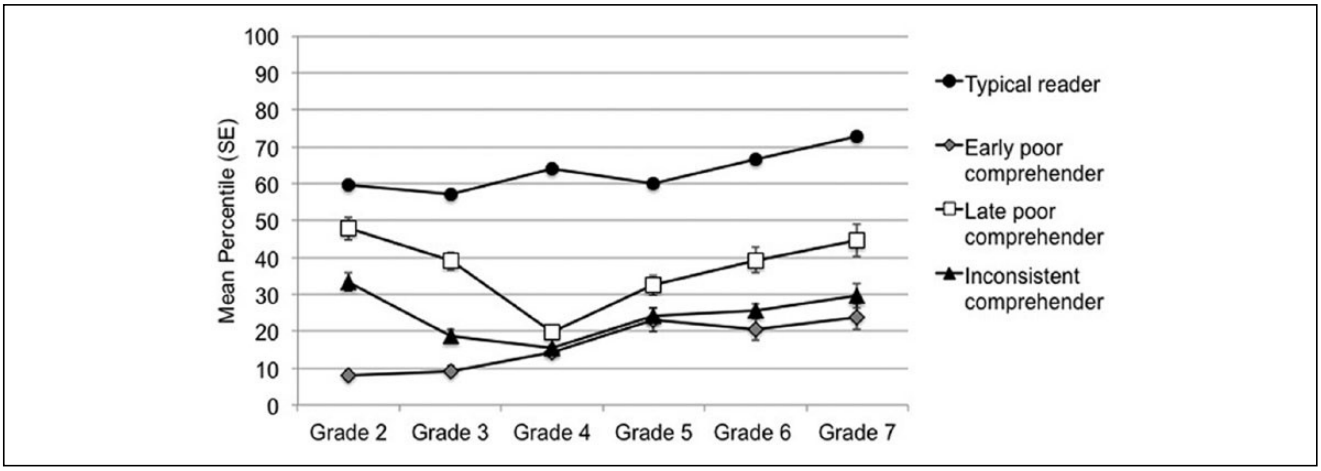

Figure 1. Reading comprehension scores between Grade 2 and Grade 7.

In the group comparisons, the three poor comprehender groups were contrasted to the typical reader group across grades. These results showed that the early poor comprehender group performed more poorly than the typical group on reading comprehension, word reading, spelling, word attack, and oral cloze measures, whereas the late poor comprehender group performed more poorly on the word reading, spelling, and word attack measures. Finally, the inconsistent comprehender group showed significantly poorer performance than the typical readers on all measures. Across time, these results indicate that the poor comprehender groups performed at a lower level than the typical readers on several tests.

Finally, we examined the interactions between time and group on all measures. A significant interaction would indicate that the poor comprehender group had a different progression in their scores when compared to the typical reader group, whereas a nonsignificant interaction would indicate that the poor comprehender and typical reader groups had a similar progression in their scores over time. The results showed that there was no interaction between time and the early poor comprehender group, which implies that this group developed at a similar pace to the typical readers (but they had lower scores on the tests described above). There was an interaction between time and the late poor comprehender group on the reading

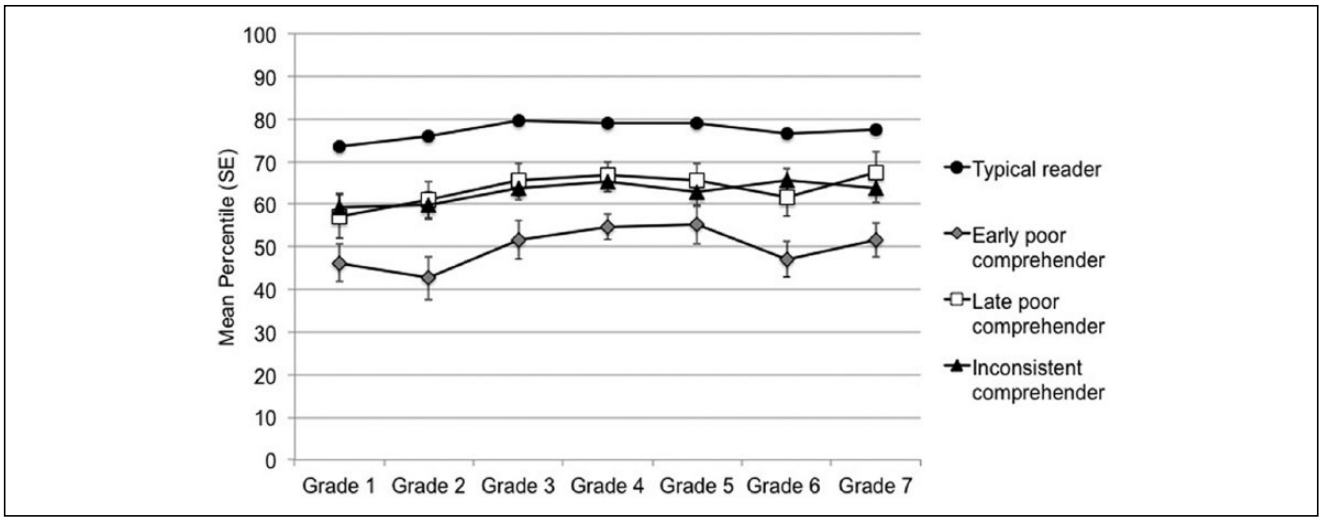

Figure 2. Word reading scores between Grade 1 and Grade 7.

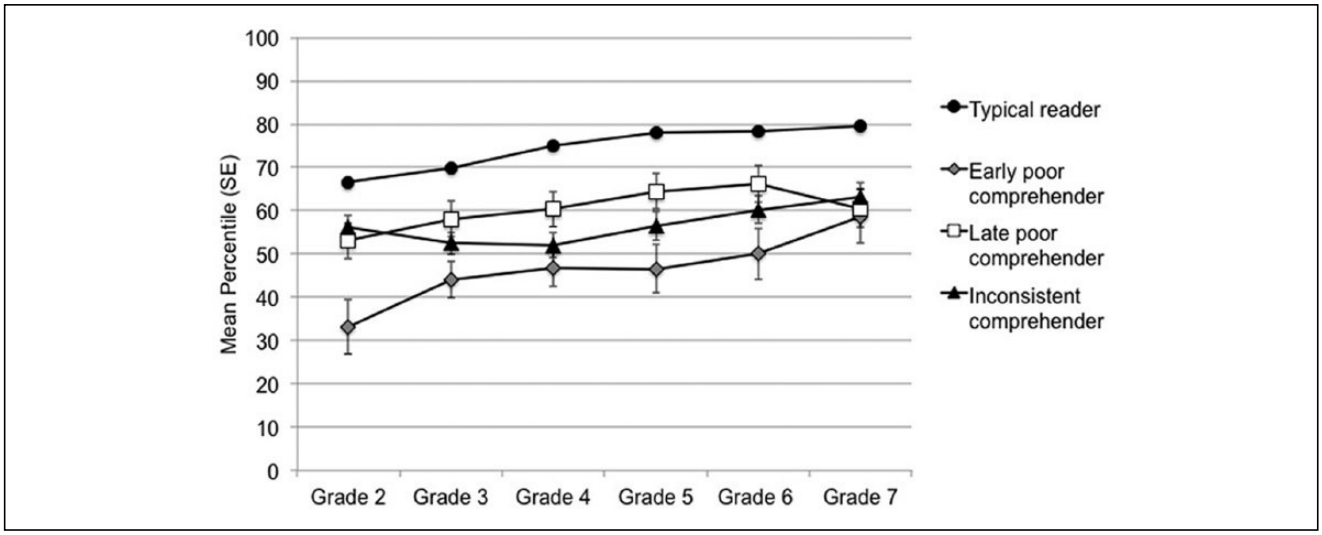

Figure 3. Spelling scores between Grade 2 and Grade 7.

comprehension measure, which was expected based on the group definition. However, there were no other significant interactions with this group, which suggests that there were not any longitudinal measures that could describe why the late poor comprehenders showed a drop in their reading comprehension scores in Grade 4. Finally, there was an interaction between time and the inconsistent poor comprehender group on the reading comprehension and spelling measures. This result indicates that the inconsistent poor comprehender group showed a linear decrease and quadratic increase (U shape) in comparison to the typical reader group, which implies a larger gap between the two groups in the middle grades on the comprehension and spelling measures.

Although the poor comprehender groups scored lower than the typical readers across time, we were also interested in how many children performed within the average range (≥ 35th percentile) or below average range (≤ 25th percentile) in Grade 7. Table 4 shows that the majority of early poor comprehenders (65%) and inconsistent poor comprehenders (57%) scored in the below average range, whereas the majority of late poor comprehenders scored in the average range (67%). This latter group of late poor comprehenders represents 2.2% of the total sample in Grade 7 (sample size was 639). These results suggest that although there are some children who experience a "slump" in reading comprehension in Grade 4, most score in the average range by Grade 7. Furthermore, most of the

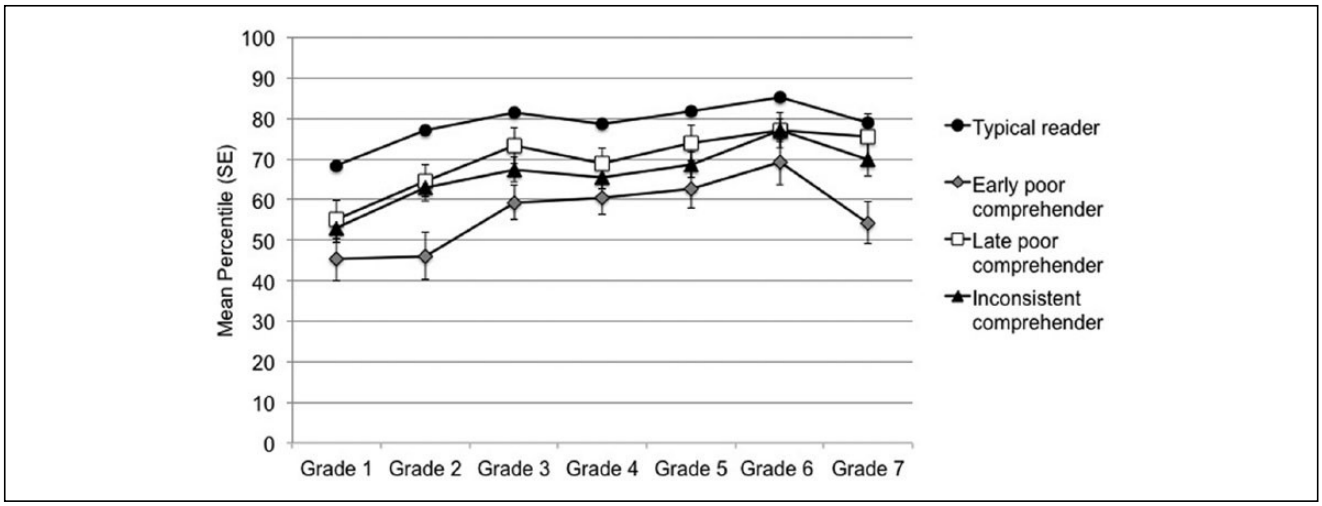

Figure 4. Word attack scores between Grade 1 and Grade 7.

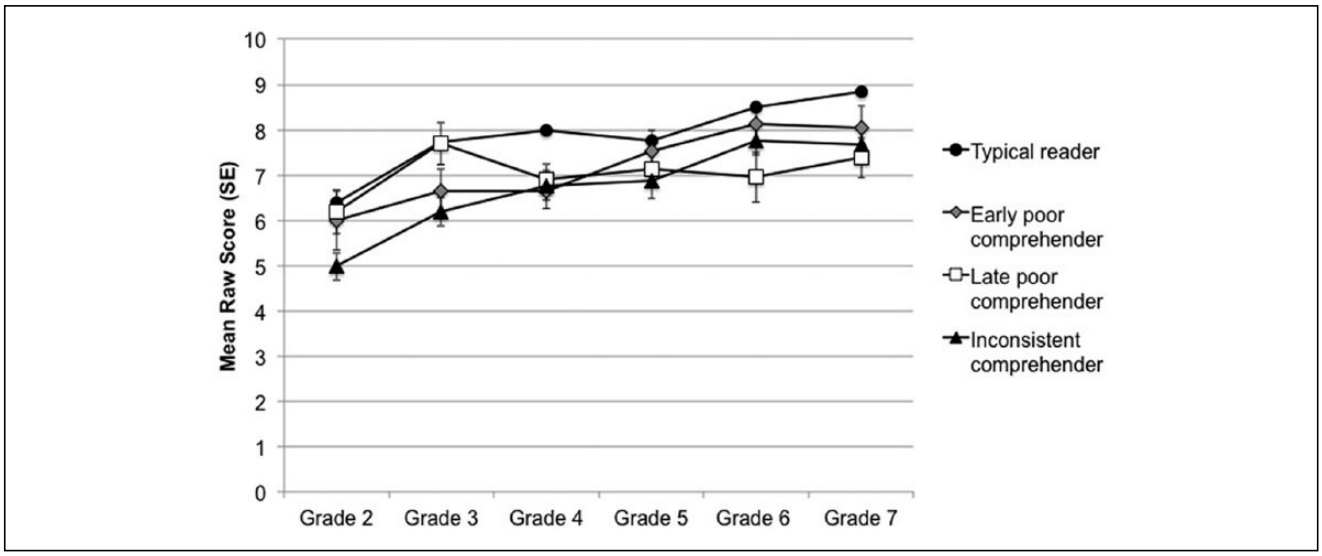

Figure 5. Working memory scores between Grade 2 and Grade 7.

children who experienced early or inconsistent reading difficulties continued to score within the below average range in Grade 7.

Early Indicators of Poor Reading Skills

We also sought to determine whether there were any early indicators of poor reading comprehension by examining group differences on reading and cognitive skills in kindergarten. We included the typical reader and three poor comprehender groups that were defined above; this included 566 typical readers, 26 early poor comprehenders, 29 late poor comprehenders, and 46 inconsistent comprehenders in kindergarten. An ANOVA was conducted for the 10 kindergarten measures separately, as we were interested in the univariate effects only. The measures included letter identification, word reading, simple spelling, rhyme detection, syllable identification, phoneme identification, phoneme deletion, working memory, rapid naming, and oral cloze (see Table 5 for group mean performance). The results were corrected for multiple comparisons at $p < .001$, and Welch's F test was used when homogeneity of variance was violated (Levene's test, $p < .15$). Post hoc comparisons included Tukey's test (homogeneity of variance assumed) and Games–Howell test (homogeneity of variance not assumed).

The ANOVA results showed significant differences between groups on all measures at the $p < .001$ level, letter

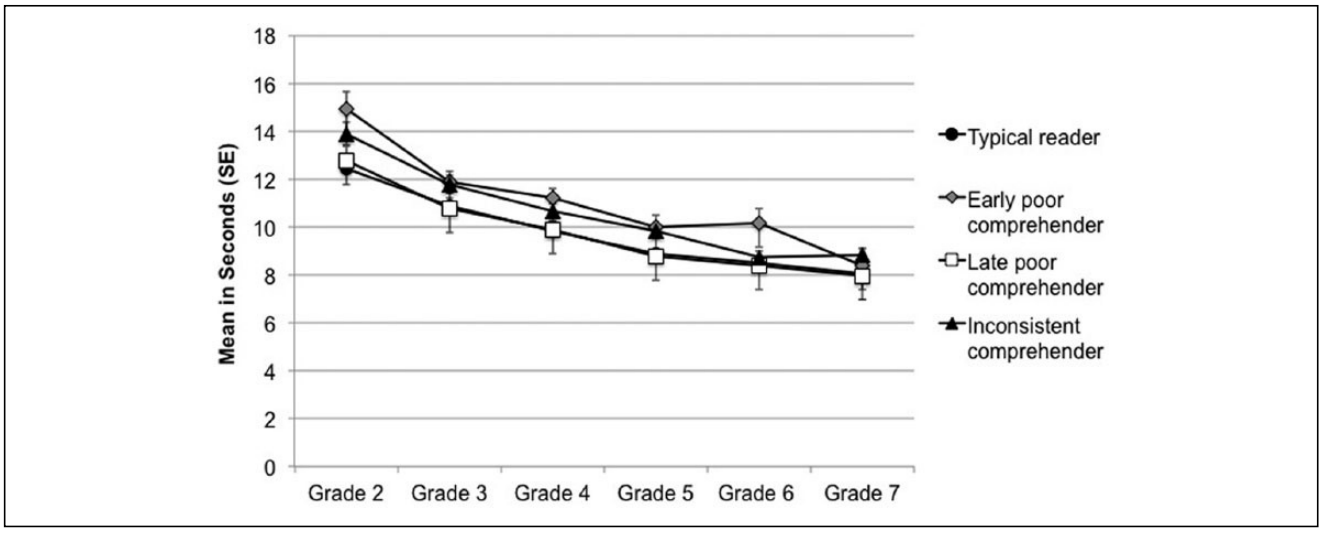

Figure 6. Rapid naming scores between Grade 2 and Grade 7.

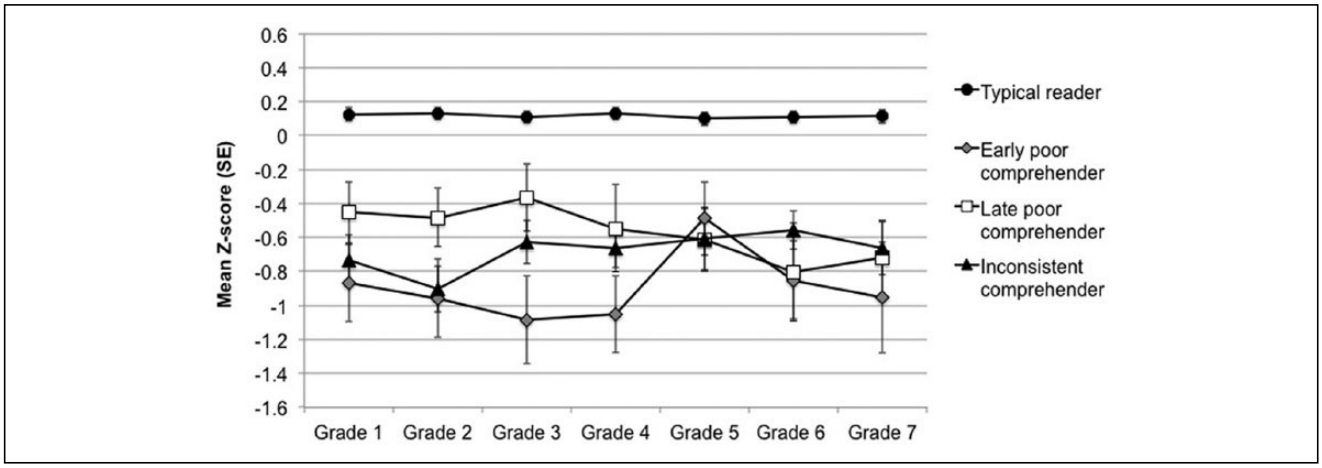

Figure 7. Oral cloze scores between Grade 1 and Grade 7.

Table 4. Frequencies of Poor Comprehender Groups in Grade 7.

	Grade 7 Performance (% Within Group)		
Group	Below Average	Borderline	Average
Early poor comprehender	13 (65.0)	2 (10.0)	5 (25.0)
Late poor comprehender	3 (14.3)	4 (19.0)	14 (66.7)
Inconsistent comprehender	25 (56.8)	5 (11.4)	14 (31.8)

identification $F(3, 661) = 15.47$; word reading Welch's $F(3, 54.20) = 11.29$; simple spelling Welch's $F(3, 54.09) = 16.75$; rhyme detection Welch's $F(3, 54.87) = 10.39$; syllable identification Welch's $F(3, 55.01) = 8.69$; phoneme identification Welch's $F(3, 58.58) = 9.95$; phoneme deletion Welch's $F(3, 63.62) = 15.01$; working memory Welch's $F(3, 59.98) = 19.46$; rapid naming $F(3, 649) = 7.28$; oral cloze Welch's $F(3, 62.42) = 15.81$. Post hoc comparisons with an a priori contrast between the typical reader and 3 poor comprehender groups were conducted (see Table 5). These results showed that the early poor comprehender group showed poorer performance on letter identification,

Table 5. Descriptives of Kindergarten Measures for All Groups.

Group	*n*	Letter Identification (Raw)	Word Reading (Percentile)	Simple Spelling (Raw)	Rhyme Detection (Raw)	Syllable Identification (Raw)	Phoneme Identification (Raw)	Phoneme Deletion (Raw)	Working Memory (Raw)	Rapid Naming (Sec)	Oral Cloze (Raw)
Typical reader[a]	566	16.92 (7.10)	59.53 (27.41)	2.88 (1.85)	7.26 (2.81)	4.93 (2.45)	3.12 (2.99)	3.96 (4.81)	16.99 (3.76)	66.43 (20.93)	2.57 (2.74)
Poor word reader											
Early	0	—	—	—	—	—	—	—	—	—	—
Late	2	5.50†	69.00	2.00	7.50	6.00	2.50	2.50	21.50	74.60	0.00†
Inconsistent	1	6.00†	16.00†	0.00†	0.00†	0.00†	0.00†	0.00†	16.00†	95.00†	0.00†
Poor comprehender											
Early	26	9.31*	32.64*	1.59*	5.04	3.77	1.77	1.46*	13.65*	75.89	1.15*
Late	29	14.17	42.71	1.90	6.52	4.03	3.10	2.66	14.76*	71.78	1.86
Inconsistent	46	12.24*	45.87	1.71*	4.93*	2.93*	1.46*	1.53*	15.11*	80.13*	1.05*
Poor word reader and comprehender											
Early	4	4.50†	9.00†	0.50†	6.25	1.50†	0.00†	0.25	13.25	104.50†	0.25
Late	0	—	—	—	—	—	—	—	—	—	—
Inconsistent	13	7.08†	21.08†	1.23†	4.77	4.00	0.92	1.85	14.92	85.73	1.31

[a]Includes mean (*SD*) for typical readers.
*Significant difference at $p < .001$ for a priori comparisons of typical reader group vs. poor comprehender groups; statistics were not conducted with poor word reader groups because of small sample sizes.
†1 *SD* below typical reader mean (1 *SD* above mean for rapid naming).

word reading, spelling, phoneme deletion, working memory, and oral cloze. The late poor comprehender group had worse performance on the working memory measure only. Finally, the inconsistent poor comprehender group had worse performance on all measures in comparison to the typical reader group, except for the word reading task. The most consistent finding is that the poor comprehender groups had lower working memory scores than the typical reader group in kindergarten.

Finally, we examined how the poor word reader groups performed in comparison to the typical reader group. Descriptive statistics are presented in Table 5, with an emphasis on scores 1 *SD* below the typical reader mean (or 1 *SD* above the mean for rapid naming). The late poor word readers obtained lower scores on the letter identification and oral cloze tests, the inconsistent poor word reader obtained lower scores on all tests, the early poor word reader/comprehenders obtained lower scores on letter identification, word reading, simple spelling, syllable identification, phoneme identification, and rapid naming, and the inconsistent poor word reader/comprehenders obtained lower scores on the letter identification, and word reading, spelling. These results indicate that although groups were defined by performance in later grades, these children performed more poorly on some reading and cognitive tests in kindergarten. Furthermore, all groups had lower scores on the letter identification test, which indicates that it may be an early indicator of later reading achievement.

Discussion

How Many Children Have LERD?

The initial research question established that there were 0.2% of children who had poor word reading skills and 3.3% of children who had poor reading comprehension skills for the first time in Grade 4. Interestingly, these numbers represent a relatively small proportion of the entire student body, a finding that contrasts with the results of the Leach et al. (2003) study, where they found a higher proportion of late emerging poor word readers than late emerging poor reading comprehenders. Our results are more consistent with those of Catts et al. (2012), who found that there were a larger proportion of children with reading comprehension than word recognition difficulties. These inconsistencies raise some questions: Primarily, what are the main reasons for these different outcomes, and furthermore, are there certain conditions or educational practices that would make subgroups of LERD more likely to surface? It may be that the findings differed, at least in part, due to differences in group selection criteria among the studies (e.g., in terms of measures used and cutoff scores chosen). This in itself gives rise to challenges when comparing studies, which naturally limit the direct comparison of findings. Another possibility is likely related to the differing early intervention practices and pedagogical approaches adopted by schools. In our study, all participating schools had adopted a consistent pedagogical approach to early intervention. All children identified as poor readers in our study participated in two early intervention programs developed by teachers. One was called Firm Foundations, which used activities to focus on vocabulary, syntax, phonological awareness, and phonological skills. The other program, Reading 44, used a number of strategies to develop text comprehension and vocabulary. There was a high level of commitment to these programs and all poor readers received the same early intervention instruction. The benefit of such early intervention could potentially lend support to why a smaller proportion of students were found to have reading difficulties than previous reports.

Finally, because there were relatively few children with poor word reading skills, this suggests that poor word recognition and poor reading comprehension may not necessarily occur together. Evidence is growing that word recognition and reading comprehension skills represent related, but different, constructs (de Jong & van der Leij, 2002; Oakhill & Cain, 2012). For example, poor word readers have difficulty with working memory, phonological awareness, and rapid naming, whereas poor comprehenders have difficulty with various aspects of higher-level language (Geva & Massey-Garrison, 2013). The current results suggest that although most children had adequate word recognition skills, some continued to struggle with reading comprehension. Our results likely represent a group of children that were "unexpected poor comprehenders," such that poor comprehension was unexpected given their adequate word reading skills (e.g., Tong, Deacon, Kirby, Cain, & Parrila, 2011). Tong et al. (2011) showed that a group of unexpected poor comprehenders performed more poorly on a measure of morphological awareness, suggesting a role of higher-level language in LERD.

Does LERD Persist?

Our study was also interested in comparing the performance of the early and late poor readers to the performance of other students, as well as addressing the longevity of LERD. The results indicated that all poor comprehender groups performed more poorly than typical readers on tasks of word reading, pseudoword decoding, and spelling between Grade 1 and Grade 7. Furthermore, the early poor comprehenders progressed at a similar pace to the typical readers across grades, whereas the late poor comprehenders showed a depression in their scores in the middle grades. This depression in scores may reflect changes in the reading comprehension test itself; that is, word reading accuracy is typically measured in the early grades, whereas comprehension and language skills are emphasized in the

later grades. Furthermore, the depression in scores was shown only on the reading comprehension task, and therefore we do not have any longitudinal measures that coincided with the onset of the LERD. These results indicate that although the late emerging poor comprehenders showed below average performance for the first time in Grade 4, they had consistently lower performance on measures of word reading, nonword reading, and spelling across time. Finally, it is possible that higher-level language skills are related to the onset of LERD, such as vocabulary knowledge or making inferences (Geva & Massey-Garrison, 2013). However, these skills were not directly measured on our oral cloze task, and thus further studies may be able to elucidate the complex relationship between language skills and LERD.

Second, 33% of children with LERD were found to have persistent reading comprehension problems. Thus, the majority, or 67%, of children with newly identified reading comprehension problems in Grade 4 were able to recover by the end of Grade 7. From this, it was determined that the late emergers represented 2% of the sample, contrasting earlier studies reporting 41% to 46% (e.g., Badian, 1999; Leach et al., 2003; Shaywitz et al., 1992). One reason that differences in percentages may exist among studies is due to the different methodological procedures used. For example, Leach et al. (2003) used Grade 4 samples and matched them to kindergarten typical readers before conducting their analyses retrospectively. The current study used a longitudinal sample of children that were not matched based on reading ability. Also, our previous analysis of this cohort found a higher percentage of children with LERD in Grade 4 (Lipka et al., 2006). Our current results suggest that most of the "late" emergers were able to recover from their reading deficits by the time they were in Grade 7.

Our findings are supported by other research, which has shown LERD prevalence rates between 2.8% and 13.4% (Catts et al., 2012; Compton et al., 2008; Kieffer, 2010). Together, these results suggest that earlier reports may have overestimated the number of children with late emerging reading problems and that the fourth grade slump is just that—a slump. As a caveat and as previously mentioned, these low numbers may also be due to the two early intervention programs implemented across the school district, in which reading difficulties were remediated on an ongoing and consistent basis. This model of early intervention is comparable to the response to intervention (RTI) approaches currently being implemented and studied in districts across North America. Specifically, RTI approaches are designed to minimize differences among lower- and higher-achieving children in their reading development (Vellutino et al., 1996). Consistent with previous reports, our results indicate that reading skills are generally stable over time and a small percentage of children show persistent reading difficulties (Compton et al., 2008).

Are There Any Early Indicators of Poor Reading Skills?

Finally, this study addressed whether or not there were early indicators of LERD. Our results showed that, in comparison to the typical reader group, the early emerging and inconsistent comprehender groups performed more poorly on a variety of reading, spelling, phonological processing, working memory, and language tasks, but that the late emerging poor comprehender group had poorer performance on the working memory task only. Furthermore, the poor word reader groups had consistently lower scores on a letter identification task. These results indicate that working memory and letter naming skills may be early indicators for reading difficulties, and that children with early and late emerging reading comprehension difficulties may have weaker working memory skills in kindergarten. Consistent with these results, others have shown that children with poor reading comprehension also struggle with working memory (e.g., Cain, 2006). Longitudinal studies have shown that children with LERD may have poorer phonological awareness, decoding efficiency, listening comprehension, vocabulary, grammar, and sentence imitation on kindergarten or Grade 1 measures (Catts et al., 2012; Compton et al., 2008). Our results indicate that working memory skills may be another indicator of late emerging difficulties in reading comprehension.

Other Findings

There were more males than females in our poor comprehender groups, and results showed that males performed worse on reading comprehension but similarly to females on all other tasks measured. These results are in accordance with previous findings suggesting that more males than females have reading difficulties (Shaywitz, Shaywitz, Fletcher, & Escobar, 1990), but are in contrast to the results of Catts et al. (2012), who found more females than males with late emerging comprehension problems. These differences in results may reflect how reading difficulties were identified or the differing intervention practices, as discussed above.

Our study also showed that there were a larger proportion of children with fluctuating scores than early or late emerging difficulties. This result could reflect the arbitrary cutoff scores used in this study; however, it was also shown that this group performed more poorly on a variety of measures than the typical readers. We suggest that some children may indeed show varying performance across time, and this may be indicative of their working memory or attention difficulties. Because of the high co-occurrence between reading and attention difficulties (Sexton, Gelhorn, Bell, & Classi, 2012), it would not be surprising if some poor readers also showed attention problems. This latter

statement is speculative, as our study did not collect information on co-occurring disorders, but future research should evaluate whether children with fluctuating performance in reading over time are more likely to have attention difficulties.

Conclusions

Overall, the results of this study suggest that very few children actually have LERD, with the majority steadily improving from what has been traditionally referred to as the "fourth grade slump." However, although these children improved in their reading comprehension, they performed more poorly than typical readers on word reading, pseudoword decoding, and spelling measures between the first and seventh grades. Furthermore, in kindergarten, the poor comprehender groups obtained lower scores than typical readers on a variety of literacy, phonological awareness, working memory, rapid naming, and language tests, and all groups performed more poorly on a working memory measure. Our results suggest that children who develop poor reading comprehension skills may also have poor working memory that is evident from an early age. This result implies that the assessment of working memory skills may be important for identifying children who may develop subsequent reading comprehension difficulties.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support with respect to the research, authorship, and/or publication of this article: This research was supported by funding from the Natural Sciences and Engineering Research Council of Canada.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Baddeley, A. D., & Hitch, G. J. (1994). Development in the concept of working memory. *Neuropsychology*, *8*, 485–493.

Badian, N. A. (1999). Reading disability defined as a discrepancy between listening and reading comprehension: A longitudinal study of stability, gender differences, and prevalence. *Journal of Learning Disabilities*, *32*, 138–148.

Cain, K. (2006). Individual differences in children's memory and reading comprehension: An investigation of semantic and inhibitory deficits. *Memory*, *14*, 553–569.

Cain, K., Oakhill, J., & Bryant, P. E. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*, 31–42.

Catts, H. W., Compton, D., Tomblin, J. B., & Bridges, M. S. (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology*, *104*, 166–181.

Catts, H. W., & Hogan, T. P. (2002, November). *At what grades should we assess phonological awareness?* Poster presented at the annual convention of the American Speech-Language-Hearing Association, Atlanta, GA.

Chall, J. S. (1983). *Stages of reading development*. New York, NY: McGraw-Hill.

Chall, J. S., & Jacobs, V. A. (2003). Poor children's fourth grade slump. *American Educator*, *27*, 14–15.

Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.

Compton, D. L., Fuchs, D., Fuchs, L. S., Elleman, A. M., & Gilbert, J. K. (2008). Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. *Learning and Individual Differences*, *18*, 329–337.

D'Angiulli, A., & Siegel, L. S. (2004). Early literacy instruction, SES, and reading development in English language learners and children with English as a first language. *Learning Disabilities: Research & Practice*, *19*, 202–213.

DeFord, D., Lyons, C., & Pinnell, G. (1991). *Bridges to literacy: Learning from Reading Recovery*. Portsmouth, NH: Heinemann.

de Jong, P. F., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading*, *6*, 51–77.

Denckla, M., & Rudel, R. G. (1976). Rapid "automatized" naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, *14*, 471–479.

Elbro, C., & Scarborough, H. S. (2004). Early identification. In P. Bryant & T. Nunes (Eds.), *Handbook of children's literacy* (pp. 339–359). Dordrecht, Netherlands: Kluwer.

Geva, E., & Massey-Garrison, A. (2013). A comparison of the language skills of ELLs and monolinguals who are poor decoders, poor comprehenders, or normal readers. *Journal of Learning Disabilities*, *46*, 387–401.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, *78*, 243–255.

Karlsen, B., & Gardner, E. (1995). *Stanford Diagnostic Reading Test* (4th ed.). San Francisco, CA: Harcourt Brace.

Kieffer, M. J. (2010). Socioeconomic status, English proficiency, and late-emerging reading difficulties. *Educational Researcher*, *39*, 484–486.

Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology*, *95*, 211–224.

Lesaux, N. K., & Siegel, L. S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology*, *39*, 1005–1019.

Lipka, O., Lesaux, N. K., & Siegel, L. S. (2006). Retrospective analyses of the reading development of Grade 4 students with reading disabilities: Risk status and profiles over 5 years. *Journal of Learning Disabilities, 39*, 364–378.

Muter, V., Hulme, C., & Snowling, M. (1997). *The Phonological Abilities Test*. London, UK: Psychological Corporation.

North Vancouver School District. (1999). *Reading 44: A core reading framework*. North Vancouver, BC, Canada: Author.

North Vancouver School District. (2001). *Firm foundations: Early literacy teaching and training*. North Vancouver, BC, Canada: Author.

Norton, E. S., & Wolf, M. (2012). Rapid automatized naming fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology*, *63*, 427–452.

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, *16*, 91–121.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75–119). Timonium, MD: York Press.

Sexton, C. C., Gelhorn, H. L., Bell, J. A., & Classi, P. M. (2012). The co-occurrence of reading disorder and ADHD: Epidemiology, treatment, psychosocial impact, and economic burden. *Journal of Learning Disabilities*, *45*, 538–564.

Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., Brady, S., . . . Shaywitz, B. A. (1999). Comprehension and decoding: Pattern of association in children with reading difficulties. *Scientific Studies of Reading*, *3*, 69–94.

Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine*, *326*, 145–150.

Shaywitz, S. E., Shaywitz, B. A., Fletcher, J. M., & Escobar, M. D. (1990). Prevalence of reading disability in boys and girls: Results of the Connecticut longitudinal study. *Journal of the American Medical Association*, *264*, 998–1002.

Shek, D. T. L., & Ma, C. M. S. (2011). Longitudinal data analysis using linear mixed models in SPSS: Concepts, procedures and illustrations. *Scientific World Journal*, *11*, 42–76.

Siegel, L. S. (1986). Phonological deficits in children with a reading disability. *Canadian Journal of Special Education*, *2*, 45–54.

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development*, *60*, 973–980.

Snow, C., Burns, S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academies Press.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407.

Stanovich, K. E., & Siegel, L. S. (1994). The phenotypic performance profile of reading-disabled children: A regression-based test of phonological-core variable difference model. *Journal of Educational Psychology*, *86*, 24–53.

Swanson, H. L., Kehler, P., & Jerman, O. (2010). Working memory, strategy knowledge, and strategy instruction in children with reading disabilities. *Journal of Learning Disabilities*, *43*, 24–47.

Tabachnick, B. G., & Fidell, L. S. (2006). *Using multivariate statistics* (6th ed.). Boston, MA: Allyn & Bacon.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford–Binet Intelligence Scales* (4th ed.). Chicago, IL: Riverside.

Tong, X., Deacon, S. H., Kirby, J. R., Cain, K., & Parrila, R. (2011). Morphological awareness: A key to understanding poor reading comprehension in English. *Journal of Educational Psychology*, *103*, 523–534.

Tremblay, S., Ross, N., & Berthelot, J.-M. (2001). Factors affecting Grade 3 student performance in Ontario: A multilevel analysis. *Education Quarterly Review*, *7*, 25–36.

Tunmer, W. E., Chapman, J. W., Greaney, K. T., & Prochnow, J. E. (2002). The contribution of educational psychology to intervention research and practice. *International Journal of Disability, Development and Education*, *49*, 11–29.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficulty-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, *88*, 601–638.

Wilkinson, G. S. (1993). *The Wide Range Achievement Test–3*. Wilmington, DE: Jastak Associates.

Willows, D. M., & Ryan, E. B. (1986). The development of grammatical sensitivity and its relation to early reading achievement. *Reading Research Quarterly*, *21*, 253–266.

Woodcock, R. W. (1987). *Woodcock Reading Mastery Tests–Revised*. Circle Pines, MN: American Guidance Service.

Worthy, J., Patterson, E., Salas, R., Prater, S., & Turner, M. (2002). "More than just reading": The human factor in reaching resistant readers. *Reading Research and Instruction*, *41*, 177–202.

Zimmerman, D. W., & Zumbo, B. D. (1993). Relative power of parametric and nonparametric statistical methods. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences, Volume 1: Methodological issues* (pp. 481–517). Hillsdale, NJ: Lawrence Erlbaum.

Zimmerman, D. W., & Zumbo, B. D. (2005). Can percentiles replace raw scores in statistical analysis of test data? *Educational and Psychological Measurement*, *65*, 616–638.