Learning and Individual Differences 21 (2011) 438–452

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Adolescent struggling readers in urban schools: Results of a Latent Class Analysis

Irma F. Brasseur-Hock [a,*], Michael F. Hock [a], Michael J. Kieffer [b], Gina Biancarosa [c], Donald D. Deshler [a]

[a] The University of Kansas, United States
[b] Teachers College, Columbia University, United States
[c] The University of Oregon, United States

ARTICLE INFO

Article history:
Received 18 March 2010
Received in revised form 3 December 2010
Accepted 17 January 2011

Keywords:
Adolescent readers
Latent Class Analysis
Adolescent reading profiles
Subgroups of adolescent readers
Reading skill profiles

ABSTRACT

In this study, our goal was to identify unique clusters of Adolescent Struggling Readers (ASRs) and examine the reading skill profiles each cluster presented. We assessed 319 students attending urban schools on three standardized measures of reading comprehension and eight standardized measures of component skills, including vocabulary, listening comprehension, word- and text-level reading accuracy and fluency. The mean age of the students was 14.9 years, and they were entering 9th grade students. Using Latent Class Analysis (LCA), we identified four empirically distinct levels of reading comprehension and classified a subsample of students as below-average comprehenders ($n = 195$). Within this subsample, we performed a second LCA that identified five subgroups with distinct profiles of component skills. These results suggest a great deal of heterogeneity in strengths and weaknesses of component reading skills within the ASR group, implying the need for diagnostic assessment and differentiated intervention.

Recently, a great deal of attention has been focused on adolescent literacy, particularly as it relates to reading proficiency (Berman & Biancarosa, 2005; Biancarosa & Snow, 2004; Carnegie Council on Advancing Adolescent Literacy, 2010; Cassidy & Cassidy, 2009; Christenbury, Bomer, & Smagorinsky, 2009; McCardle & Chhabra, 2004; National Center for Educational Statistics, 2004; Phelps, 2005; Snow, 2002). While this trend is welcomed by researchers, practitioners, and policy makers, many claims about the nature of the reading challenges faced by Adolescent Struggling Readers (ASRs), especially those in urban schools, are largely lacking scientific foundation. For example, The National Governors Association suggested that only 10% of adolescent struggling readers might need word-level interventions to improve reading comprehension based upon a study done with a limited number of students in grade 5 (e.g., Berman & Biancarosa, 2005; Buly & Valencia, 2003). Others suggest that this figure might be higher and change developmentally (e.g., Catts, Hogan, & Adlof, 2005). Thus, evidence that describes the reading skill profile of ASRs and that further identifies statistically unique subgroups of struggling readers is limited. This limitation presents a serious challenge for those designing interventions intended to have a significant impact on the achievement performance of ASRs. Because of the continued magnitude of the achievement gap and the limited amount of instructional time available to close the gap with ASRs, it is important to increase the precision with which we can identify subgroups of ASRs so interventions can be better tailored to the unique needs of students and optimize the return on investment of our intervention work. In this study, we use Latent Class Analysis to identify and describe the reading component skill profile of subgroups of ASRs attending urban schools to help inform such efforts.

1. The challenge

The magnitude of challenges facing adolescents, especially those who attend urban high schools, is striking. For example, in some of the largest urban school districts, nearly 65% of all adolescents read below the "satisfactory" level on state reading assessments (Council of Great City Schools, 2001). Consequently, these students are unable to understand and respond to the complex literacy demands of secondary school subject-matter courses (Hock et al., 2009; Lee, Grigg, & Donahue, 2007). The reading achievement scores for adolescents who struggle with learning have remained virtually unchanged for the last 30 years with 70% of the students at the basic or below basic level and unable to understand complex material written at grade level (Lee et al., 2007). Collectively, these data are being characterized as one of the most pressing crises facing the educational system in the U.S. (Deshler, 2006; Kamil et al., 2008; Scammacca et al., 2007). This is especially true for students attending schools in urban settings. The educational and life outcomes for urban students are lower than other students in all achievement areas (e.g., Lippman, Burns, & McArthur, 1996; NCES, 2009; Schultz, 1993). Generalizing findings from studies of other struggling readers may not address unique learning needs of the urban student population.

* Corresponding author at: The University of Kansas, Center for Research on Learning, 1122 West Campus Road, JRP Hall, Room 500, Lawrence, KS 66045, United States. Tel.: +1 785 760 4069(Office); fax: +1 785 864 5728.
E-mail address: ibrasser@ku.edu (I.F. Brasseur-Hock).

1041-6080/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.lindif.2011.01.008

A variety of factors have contributed to the literacy crisis confronting large portions of our high school populations. One of the most important is the ineffectiveness of instruction that is provided to many adolescents who struggle with reading (Kamil, 2003). While there are some adolescents who leave elementary grades as virtual non-readers or who are severely deficient in word recognition, the largest group of ASRs is hypothesized to be those who have acquired some, but not sufficient, reading skills to enable them to escape the "fourth-grade slump" (Biancarosa, Palincsar, Deshler, & Nair, 2007; Chall, 1983). Specifically, nearly 65% of ASRs in poor urban settings fall between the 5th and the 30th percentile in reading performance (Council of Great City Schools, 2001; Curtis, 2002), raising the hypothesis that these students have some decoding skills, but perhaps not at a level that enables them to deal fluently with subject-matter reading demands (Biancarosa et al., 2007). In addition, there is reason to believe that they lack the required skills and strategies to meet comprehension expectations (Biancarosa et al., 2007). However, in the absence of data on their component skills using multiple, independent measures, it is unclear whether these claims are justified.

A limited understanding of the component reading skill profiles of this heterogeneous population and the consequent lack of targeted intervention may be at the root of the continued poor literacy performance of ASRs. That is, if literacy instruction is not aligned with the unique profiles of ASRs, closing the academic literacy gap experienced by ASRs may be impossible. Thus, knowledge of the unique instructional needs of subgroups within the general category of ASR and the need to accurately and regularly measure their component skills may help inform instruction that addresses the specific needs of struggling readers (e.g., Connor, Morrison, Fishman, Ponitz, et al., 2007; Connor, Morrison, Fishman, Schatschneider, & Underwood, 2007; Mellard, Fall, & Mark, 2008).

2. Theoretical underpinnings

The ultimate goal of our study was to identify specific clusters of ASRs and the unique reading component skill profiles they present. Our primary motivation was to inform instruction and policy by adding to the empirical knowledge base about ASR's differences in reading. Theoretically, we framed our analysis of the data using an overarching view of reading comprehension as described in the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990).

The Simple View of Reading proposes that reading comprehension is a product of word recognition and linguistic comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990). This view does not deny the complexities of reading, but rather divides them into two components. The word recognition component is responsible for translating print into language, and the comprehension component makes sense of this linguistic information. The word recognition component encompasses efficient decoding, accurate sight word recognition, fluent word reading and access to appropriate words in the reader's mental lexicon that provides semantic information at the word-level. Thus, efficient word recognition allows the reader to quickly pronounce a word and triggers recognition of words acquired through language experiences (e.g., prior knowledge). Linguistic comprehension is defined as knowledge of facts and concepts, vocabulary, language and text structures, and verbal reasoning structures and strategies. Some researchers refer to linguistic comprehension as a language comprehension since measures of language comprehension seem to capture that domain (e.g., Catts et al., 2005). The interaction of these two components results in reading comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990). Both of these theoretical components are essential to developing fluent and effective reading comprehension, but neither is sufficient on its own.

In the Simple View of Reading, skilled reading requires both bottom–up and top–down processes that are not always sequential but rather occur simultaneously and in relationship to each other. If one component is weak, then efficient and effective comprehension is difficult. For example, if word recognition accuracy is poor and effortful, comprehension of text will be limited, as most cognitive energy will be required to decode words in text (e.g., Adams, 1990; Ehri, 1997, 1998; LaBerge & Samuels, 1974; Perfetti, Marron, & Foltz, 1996; Torgesen, 1999; Torgesen, Wagner, & Rashotte, 1999). Similarly, if linguistic comprehension is limited, even though words can be read accurately and with acceptable pace, understanding of text may be constrained due to deficiencies in vocabulary, semantics, prior knowledge, or text structure knowledge (e.g., Catts, Fey, Tomblin, & Zhang, 2002; Catts et al., 2005; Kintsch, 1998; Kintsch, 2005; McCardle, Scarborough, & Catts, 2001). Numerous studies have been conducted to demonstrate the effects of the two components described above and their relationship to the Simple View of Reading (e.g., Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Rashotte, MacPhee, & Torgesen, 2001; Torgesen, Rashotte, & Alexander, 2001). The results of such studies show that each component is necessary for proficient reading comprehension and that neither is independently sufficient for comprehension. However, little empirical evidence exists showing whether the component reading skills identified in the Simple View hold for ASRs entering high school or for older struggling readers.

One implication of this componential theory of reading is that there may be subgroups of struggling readers who have weaknesses in specific component skills, leading researchers to investigate whether there is empirical evidence for heterogeneity in the sources of reading difficulties. If specific subgroups of readers are empirically identified among the population of ASRs that both support the Simple View and expand theoretical knowledge of ASRs, our overall understanding of reading development may be enhanced. For example, we may be able to answer questions like, "Have most ASRs acquired at least a basic command of word-level skills as described in the Simple View, and therefore, struggle with comprehension primarily because of language comprehension deficits or will they cluster into unique classes with distinct component skill profiles that may include both word-level and language comprehension deficits?"

In a previous study, the Hock et al. (2009) conducted a descriptive analysis in which 345 adolescents attending urban schools were assessed using eleven reading component skills measures. The sample included equal numbers of students classified as either unsatisfactory, below basic, satisfactory, above average, or exemplary on the Kansas Reading Assessment, a measure of Adequate Yearly Progress (AYP) (Kansas State Department of Education, 2005). A principal-components analysis (PCA) was conducted to determine if the 11 reading measures administered formed distinct reading components. The results of the PCA indicated that the 11 variables could be summarized by four components: word-level, fluency, vocabulary, and comprehension. In this study, struggling readers were identified on the basis of their scores on a comprehension composite score, the mean of the Woodcock Language Proficiency Battery-Revised (WLPB-R) passage comprehension subtest and the Gray Oral Reading Test-4 (GORT-4) comprehension subtest score. Those who scored at or below the 40th percentile (standard score of 96) on the composite score were defined as struggling readers; those who scored above the 40th percentile were defined as proficient readers. Using this criterion, the sample used in the final analysis consisted of 202 adolescent "struggling" readers and 143 "proficient" readers.

An examination of 202 struggling readers showed that 123 (61%) scored low on every component, including comprehension. Another 26 scored low on every component except word-level. Aside from comprehension, the component on which the largest number of students scored below the mean standard score of 96 was fluency (177 or 88%). The overall reading profile of proficient and non-

proficient readers was identified with marked differences in reading component skills. Specifically, non-proficient readers lacked all of the critical reading skills identified by researchers as necessary for proficient reading by a full standard deviation. In contrast, proficient readers demonstrated a high level of competence in all the reading component skills assessed (Hock et al., 2009).

There are limitations to this study. First, the study was conducted in only two school districts in the Midwest. Thus, the prevalence rates described are not representative of a national sample. Second, the type and number of measures used were limited. Additional measures in listening comprehension, additional student characteristics, and contextual factors might help further define the nature of adolescent struggling readers. Finally, while the overall profile of reading component skills for the ASRs was identified, the specific nature and instructional needs of ASR subgroups was not investigated.

While the identification of non-proficient readers and their skill profile is important, such general findings do little if anything to identify unique subgroups of struggling readers and thus provide targeted instructional information. In the present study, we analyzed the subset of ASR data from the larger data set described above. In both studies the battery of reading component measures was aligned with the two main components of the Simple View of Reading in an effort to determine if the model holds for ASRs or if it might differ. The measures were also aligned with the essential reading components identified by the National Institute of Child Health and Human Development (2000) and included measures of phonics, fluency, vocabulary, and comprehension. For example, we sought first to determine if ASRs who were poor comprehenders lacked word-level skill proficiency (i.e., phonics, decoding, word recognition, and fluency skills) or if they lacked skills related to language comprehension (i.e., vocabulary, semantics, lack of prior knowledge, or text structure knowledge deficits). Second, we wanted to determine whether ASRs fit empirically unique groups that were statistically distinct within the structure of the Simple View of Reading. Limited research on reading skill profiles has been conducted with the targeted age group and no studies have been conducted to determine ASR profiles linked to the Simple View of Reading.

Previous descriptive studies have been conducted in order to describe the reading skill profiles of struggling readers. Although these studies have added to our understanding of the reading skills of struggling readers, the data are limited and do not include high school readers nor do they focus primarily on struggling readers in urban schools. Moreover, findings are often contradictory. Therefore, the field would benefit from additional research in this area.

2.1. Foundational studies with children

Research indicates that a significant number of children have reading disabilities, but little research has been conducted that describes the characteristics of these children and their reading skill profiles (Badian, 1999; Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992). One study that did look at student characteristics and reading subgroups with a population of young readers was conducted by Konold, Juel, McKinnon, and Deffes (2003). In this study, a multivariate analysis of early reading skill acquisition of students 5 to 10 years of age was conducted. The researchers analyzed a large subset ($n = 1604$) of the standardization sample for the Woodcock Diagnostic Reading Battery (Woodcock, 1997). In a descriptive analysis, they identified six distinct and homogeneous core profiles. They found that there was a developmental shift in skills and abilities that influence reading proficiency with strength in phonological processing outweighing other strengths at age 5. They also found that strengths in comprehension knowledge and short-term memory were of greater importance in reading proficiency for 10-year-olds. Additionally, while overall reading skill profiles were stable and somewhat flat for all readers in all six groups, overall ability mediated reading success and distinguished reading subtypes. An important finding was that there seem to be multiple paths to reading proficiency.

Vellutino, Tunmer, Jaccard, and Chen (2007) studied children in grades 2 and 3. They administered a large battery of tests to the children that measured specific reading subskills and reading-related cognitive skills. The same tests were also administered to an older group of students in grades 6 and 7, allowing for age-group comparisons. The students were from middle and upper middle class schools. Students included in the study had to score at the >90 standard score level on at least one of the subscales of the WISC-R (Wechsler, 1991). Key findings for younger students indicate that word identification and related phonological skills on one hand and reading comprehension on the other were strongly related. In contrast, the relationship between language comprehension and reading comprehension was stronger for the older group. Importantly, the researchers found that language comprehension becomes the dominant process in reading comprehension when the reader has become proficient in word identification. Thus, the link between word-level and language-level skills and processes was affirmed. Both of these studies help establish a foundation for research with older populations.

2.2. Younger adolescents

Buly and Valencia (2003) examined the reading skills of 108 fifth-grade students who scored poorly on the Washington Assessment of Student Learning (WASL). Students who participated in the study scored at levels 1 and 2 (on a scale of 1 to 4) on the reading portion of the assessment. At the district level, 57% of the students were Caucasian and 43% were students of color. Students in the sample performed poorly on all reading measures, including word identification, phonemic awareness, comprehension, vocabulary, rate, and expression. However, three factors accounted for 78% of the variance on the WASL scale scores: word identification, meaning, and fluency.

The authors found that poor student performance on the state reading assessment was due primarily to issues related to reading fluency and comprehension. Further, word-level problems contributed minimally to poor reading performance, and only about 9% of the students in the sample were poor readers in terms of word recognition, fluency, and meaning. Thus, most struggling readers in the study needed instruction primarily in comprehension and fluency, with very few needing instruction in all three areas.

In another descriptive study, Leach, Scarborough, and Rescorla (2003) studied late-identified reading disability (RD) in a sample of 161 fourth and fifth graders. The sample consisted primarily of Caucasian students; only 5% of the students were students of color. Ninety-five of these young adolescents were considered typically achieving readers, and 66 were identified as having some type of RD based on a standard score of 86 or less on reading comprehension tests.

On the basis of reading skill component scores, the authors assigned all 161 students to one of four groups: (a) the reading comprehension (RC) group, which included students with good word-level skills but poor comprehension; (b) the word-level (WL) group, which consisted of students with deficits in word-level skills but good comprehension; (c) the WL-RC, group in which students had deficits in both word-level and comprehension skills; and (d) the no reading disability NRD group, in which deficits were not detected in either word-level or comprehension skills.

In the three groups with reading deficits, 35% of the students had word-level processing deficits with adequate comprehension (WL), 32% had deficits in comprehension with adequate word-level skills (RC), and 32% had deficits in both word-level and comprehension skills (WL-RC). Thus, about two-thirds of the 66 poor readers had comprehension deficits, and about two-thirds had word-level deficits.

Additionally, 41% to 47% of the 66 poor readers were late-developing RD in that they had adequate reading skills before the fourth grade.

2.3. Older adolescents

In a longitudinal study of students followed from second grade through eighth grade, Catts et al. (2005) examined the word recognition and listening comprehension skills of poor readers over time. They assessed students at grades 2, 4, and 8. Within the first portion of the study, the authors used data from 527 students who participated in a longitudinal and epidemiological study through eighth grade. A regression analysis showed that word recognition and listening comprehension accounted for 76.6% (second grade), 71.8% (fourth grade), and 72.8% (eighth grade) of the composite variance in measures of reading comprehension across grade levels. Word recognition and listening comprehension varied in their unique contributions to reading comprehension across grade levels and across time. For example, word recognition played a large role in predicting reading comprehension in the early grades, whereas listening comprehension was significantly more predictive of overall reading comprehension as students grew older.

For the second portion of the study, the authors selected from the sample of 527 students who could be identified as poor readers ($n=154$). Eighth-grade readers in this analysis clustered into one of three skill categories: (a) dyslexic or students with deficits in word recognition but adequate listening comprehension (13.3%); (c) mixed RD or students with deficits in both word recognition and listening comprehension (36%); and (c) specific comprehension deficit or students with adequate word recognition but deficits in listening comprehension (30%). Thus, Catts et al. (2005) found that about 49% of the eighth-grade poor reader group had poor word recognition and about 66% had poor comprehension.

These findings clarified the influence that developmental stages have on student reading skill profiles. For example, in the second-grade analysis, listening comprehension accounted for only 9% of the unique variance in reading comprehension, whereas in the eighth-grade analysis, it accounted for 36% of the unique variance. Thus, these findings support the developmental nature of reading and highlight the shifting importance that word-level and language comprehension skills play in predicting reading comprehension. Also, according to Catts et al. (2005), by the eighth grade, word-level reading skills contribute minimally to reading comprehension, and the percentage of poor readers who struggle with comprehension nearly doubles by the fourth and eighth grades. Catts et al. (2005) also found evidence of a fourth-grade slump, whereby students considered to be satisfactory readers in second grade were identified as struggling readers by fourth grade.

2.4. Adults

Given that young adult populations often overlap with older adolescents, studies conducted with adults may be helpful in understanding adolescent reading proficiency. Sabatini (2002) examined the impact of rate or speed of processing on reading proficiency in adult struggling readers. While this study focused on key reading and ability elements (rate or speed of processing) with a somewhat older population, the study also provided profiles of reading component skills by examining how comprehension, word-level skills, and basic processing speed vary across word recognition skill levels in adult proficient and struggling readers. In addition, Sabatini found that overall ability played a significant role in determining proficiency in both word-level skills (decoding and rate) and comprehension.

Adult readers in the study presented three significantly different profiles in terms of reading ability. Specifically, the proficient group possessed both accurate and efficient word-level skills, the average group demonstrated adequate word-level skills, but they were not automatic or fluent in their use, and the low group struggled with accuracy in word-level reading as well as with efficiency in reading rate. A key mediator of reading proficiency seemed to be processing speed and general ability, especially for the two lower reading groups. Thus, adult struggling readers present heterogeneous reading profiles and may require differentiated instruction in multiple reading skills to meet their needs.

Taken together across the age span, these studies bring clarity to the discussion about the reading skill profiles of struggling readers. However, none of these studies occurred with older adolescents in high school settings and the extent to which the findings can be generalized to them is unknown. Additionally, none of the studies focused on struggling readers in urban schools with the intent to capture the range of skills possessed by this specific population. Distinguishing reading skills among students in urban, suburban, and rural schools seems warranted as students in urban schools present different achievement and social skill profiles. For example, urban students compare less favorably on all education outcome measures and are more likely to live in poverty (Lippman et al., 1996).

Furthermore, the results of the studies reviewed here are somewhat mixed. For example, Buly and Valencia (2003) found only 9% of the population to have significant reading deficits in word identification, fluency, and comprehension. In contrast, Catts et al. (2005) and Leach et al. (2003) found between 49% and 67%, respectively, of the struggling reader group to demonstrate deficits in word identification and comprehension.

Some of the differences among these studies might be attributable to their focus on different parts of the age span, but they are just as likely due to important differences in their purposes. For example, while Catts and colleagues sought to investigate the contributions of language impairment to reading skills over time and Vellutino and colleagues sought to describe the etiology of reading disability over time, Buly and Valencia sought to describe the reading skills of students failing a state reading test, and Leach and colleagues sought to establish whether reading difficulties can truly emerge later in development. These differences in their purposes led to very different approaches to sampling, measures, and methods.

Perhaps most importantly, previous studies have differed considerably in whether and how they defined poor readers before investigating subtypes of struggling readers. For example, Konold and colleagues sampled the full range of achievement from a test-standardization sample and used cluster analysis to define their struggling readers based on their performance on measures of reading skills thought to underlie reading achievement: short-term memory, processing speed, two measures of phonological awareness, and two measures of language comprehension. Vellutino and colleagues stratified their sample based on four levels of performance on a single reading comprehension measure (severely impaired, moderately impaired, average, and good readers); then they employed confirmatory factor analysis using at least two measures of each theorized latent construct. In contrast, Buly and Valencia's sample was defined as “struggling” by virtue of students having failed the state reading test, and multiple measures were reduced to three factors (word identification, fluency, and meaning) using exploratory factor analysis. Then they used cluster analysis to define subtypes of struggling readers based on their performance on the three factors. Consistent with many other studies of struggling readers, Catts and colleagues used a cut-score to define poor readers: in this case performing 1 *SD* or more below the mean on a composite of three reading comprehension measures. They then defined subgroups of struggling readers again using 1 *SD* cut-scores, this time on composites of listening comprehension and of word recognition. However, the authors also noted that examination of a bivariate scatterplot of struggling readers' performance on the listening comprehension and word recognition composites revealed that readers did not fall into clear subgroups but

rather into one data cloud; thus, the choice of cut-scores seemed primarily responsible for their definition.

The differences among these four studies alone lay bare how influential the choice of method is in a study of subtypes of struggling readers. Although each of these methods has some strengths relative to the purposes of the individual studies, none of them utilizes modern model-based techniques for differentiating between struggling and average performing readers across several measures of reading comprehension. In particular, the use of cut-scores, while offering advantages for clinical diagnoses, may not necessarily represent the qualitative breaks in the distribution that empirically separate struggling readers from others, as suggested by the findings of Catts and colleagues. In addition, most researchers conducting studies in this area hypothesize that they will find multiple subtypes of readers in the population of interest; however, the descriptive analytic techniques used (i.e., cluster analysis and cut-score classifications) often assume a number of possible subtypes and do not allow researchers to test explicitly whether this hypothesis holds or whether apparent subgroups are due to sampling error. It should also be noted that most of the studies reviewed adhered strictly to a Simple View definition of component reading skills and did not assess oral reading fluency or processing speed, with the exception of Konold and colleagues and of Buly and Valencia. It remains unclear how findings might differ were a wider range of measures used and were measures not constrained to represent one or another latent construct prior to struggling reader subtype definition. The present study adds to this literature in four ways: by investigating subtypes of struggling adolescent readers, an understudied population; by using multiple measures and avoiding cut-scores to define struggling readers; by defining subtypes of struggling readers using multiple indicators of a range of component skills; and by employing Latent Class Analysis, the benefits of which are described in detail in the Methods section.

3. Research question

Given the limited research on overall reading skill profiles and the nonexistence of research that examines ASRs in urban schools, our goal was to identify and describe the reading skill profiles of unique groups of ASRs attending urban schools. According to the National Center for Educational Statistics, urban schools are located in central cities of the Metropolitan Statistical Areas with more than 50,000 people; suburban schools in suburbs of those cities; and rural schools are located in towns of less than 50,000 people that are not suburbs of a larger city (Lippman et al., 1996). Following an analysis to identify struggling adolescent readers using multiple measures, we asked: Do adolescents with below-average comprehension exhibit differentiated profiles of component reading skills including word reading accuracy, word-level and passage-level fluency, and oral language?

4. Methods

4.1. Overall sampling plan

The current study took place in the context of a larger descriptive study of adolescent readers (Hock et al., 2009). The overall sampling plan was to recruit approximately 60 students in each of the five categories of the Kansas Reading Assessment (KRA) (i.e., unsatisfactory, basic, proficient, advanced, and exemplary) so that adequate subgroups of students could be assessed and their reading skill component profiles analyzed (Kansas State Department of Education, 2005). This approach has the advantage of maximizing statistical power across the entire range of the achievement distribution and thereby increasing the precision with which we can estimate population means for and detect differences between profiles in the tails of distribution, where the data would otherwise be sparse. Without such a sampling plan, it is possible that we would have lacked sufficient power to detect profiles that are less common in the population of adolescent readers or to differentiate between somewhat similar skill profiles. The disadvantage of this approach is that it somewhat limits our ability to make generalizations about the prevalence of particular skill profiles beyond the current sample.

We were able to obtain data on a single cohort of 345 students entering their ninth-grade year. Students came from two small urban junior high schools and two urban middle schools, which fed into three urban high schools in two Midwestern cities. Participating students were recruited during the end of their eighth-grade year and were assessed at the commencement of the ninth-grade year. Due to summer attrition, 37 additional students were recruited at the beginning of the ninth-grade school year and assessed with the original recruits. Students ranged in age from 13.45 years to 17.5 years with an average age of 14.9 years. The urban community consisted of 145,004 residents; the smaller urban community consisted of 81,873 residents (U.S. Census Bureau, 2002).

4.2. Analytic sample for the Latent Class Analysis

We restricted the original sample of 345 students to the 319 students with complete data on all measures. That is, students with missing data on one or more of the assessments were excluded from this analysis. As a result, each achievement level had a minimum of 55 students. Excluded students were not significantly different from included students on a range of demographic indicators. A demographic summary of the analytic sample is in Table 1.

4.3. Measures and instruments

Instruments were selected and grouped within a reading component framework identified in the literature as essential to the reading success of younger and adolescent readers (Curtis, 2002; NICHD, 2000). This framework includes word-level skills, fluency, vocabulary, and comprehension. The measures consisted of a battery of language and literacy tasks selected to cover the framework described above.

We chose not to create composites of the measures for this analysis, because we hypothesized that we might be able to differentiate groups of students based on different aspects of a latent construct (e.g., students who demonstrate weaknesses in phonological decoding, tapped by the word attack task, but not sight word reading accuracy, tapped by the letter-word identification task). Unlike some previous studies, we approached this as an empirical question rather than as an a priori assumption that groups would (or would not) be differentiated along particular lines. The rationale for this approach is that only one prior investigation studied subtypes of struggling readers in this age range, and this study found no clear differentiation among subtypes based on two a priori constructs:

Table 1
Demographic summary of the participating students ($n = 319$).

Demographic characteristic	All students	KRA level 1	KRA level 2	KRA level 3	KRA level 4	KRA level 5
Female	45.8%	46.3%	36.8%	52.7%	38.1%	55.0%
Urban	81.5%	100.0%	98.5%	83.8%	61.9%	63.3%
Free/reduced-price lunch	51.1%	67.9%	71.6%	54.9%	38.1%	27.1%
Special education (IEP)	9.1%	16.7%	10.3%	5.4%	9.5%	5.0%
Race						
Asian	2.5%	0.0%	2.9%	6.8%	0.0%	1.7%
Black	51.1%	77.8%	60.3%	56.8%	33.3%	28.3%
Hispanic	15.7%	11.1%	25.0%	14.9%	11.1%	15.0%
White	29.2%	9.3%	10.3%	21.6%	55.6%	50.0%
Other	1.5%	1.9%	0.0%	0.0%	0.0%	1.7%
Language background						
ELL	5.3%	1.9%	4.4%	6.8%	7.9%	5.0%

listening comprehension and word recognition (Catts et al., 2005). Though widely theorized (e.g., Biancarosa & Snow, 2004; Kamil, 2003), it remains unproven that a wider range of skills reliably differentiates struggling readers when they are adolescent.

4.3.1. Reading comprehension

Reading comprehension was assessed with three measures, the Kansas Reading Assessment (Kansas DOE, 2005), the WLPB-R Passage Comprehension subtest (Woodcock, 1991), and the GORT-4 Comprehension subtest (Wiederholt & Bryant, 2001). The Kansas Reading Assessment (KRA) is a group-administered test given annually in the spring to students in grades 5, 6, 7, 8, and 11 to measure AYP as defined in the NCLB Act of 2000. By the end of eighth grade, students are assessed on their proficiency in comprehending narrative, expository, and technical text. Measures are also taken of such skills as identification of main ideas, details, and the author's purpose, comparing, contrasting, problem solving, and using text organizers. Additionally, students are assessed on fluency, decoding, and prior knowledge. The KRA scores used were proficiency levels from 1 to 5; although more finely graduated scale scores would have been preferable, these were not available from the participating schools.

The WLPB-R Passage Comprehension subtest requires the reader to silently read a sentence or a short passage and supply a word that fits the meaning and context of the passage. This modified cloze procedure measure is completed in about 6 min. The WLPB-R has internal consistency values ranging from .77 to .95 with a medium $r = .89$. Split-half reliability exceeds .90. The GORT-4 Comprehension subtest requires the reader to read graded passages orally and to respond to comprehension questions read by the examiner and presented in a multiple-choice format. Passages range from about 20 to 160 words in length. The task takes about 10 min; two forms (A and B) are available. The reliabilities of the GORT-4 all average internal consistency reliabilities of .90 or above. The GORT-4 was normed on a sample of more than 1600 students aged 6 through 18. The normative sample was stratified to correspond to key demographic variables including race, gender, ethnicity, and geographic region.

4.3.2. Word reading accuracy

Two measures of word-level reading skills were administered. Word decoding and word identification were measured using the Word Attack and Word Identification subtests of the Woodcock Language Proficiency Battery-Revised (WLPB-R; Woodcock, 1991). The Word Attack subtest requires individuals to apply phonics and structural analysis skills to pronounce nonsense words ordered in increasing difficulty. The publisher reports reliability estimates that are greater than .90. The Letter-Word Identification subtest uses real letters and words in isolation, graded in order of difficulty. Participants read the increasingly difficult letters and words until a ceiling score is attained. The reliability of this subtest also exceeded .90. Each subtest takes about 5 min to administer. The tests are administered individually.

4.3.3. Reading fluency

Fluency was assessed using three norm-referenced subtests. First, the Test of Word Reading Efficiency (TOWRE) Sight Word Efficiency subtest (Torgesen et al., 1999) measures the number of real printed words accurately decoded within 45 s. This subtest has two forms (A and B) of equivalent difficulty. The reported reliability is .84 for students age 10–18 years. Second, the TOWRE Phonemic Decoding Efficiency subtest measures the number of pronounceable nonwords that are accurately decoded within 45 s. Its reliability is .89 for students age 10–18 years. Overall testing time is 2–3 min for each of the subtests.

Finally, the Gray Oral Reading Test-4 (GORT-4) was administered to evaluate oral reading rate and accuracy (Wiederholt & Bryant, 2001). The GORT-4 is comprised of 12 passages. Participants are required to read aloud passages as quickly and as accurately as possible and then answer five comprehension questions. For each passage administered, and depending on basal and ceiling criteria, the examiner documented the time in seconds required to read the passage, the total number of reading errors, and responses to comprehension questions. The GORT-4 rate and accuracy subtest scores are summed to provide an overall reading fluency score. Reported reliability is .92 for the fluency measures.

4.3.4. Vocabulary and language comprehension

Receptive oral vocabulary was assessed using the Peabody Picture Vocabulary Test-Third Edition (PPVT-III; Dunn & Dunn, 1997). The PPVT-III requires the student to point to the one of the four pictures that represents a stimulus word pronounced by the examiner. The words become increasingly difficult. Test administration takes 10–12 min. Reading vocabulary was assessed using the Woodcock Language Proficiency Battery-Revised (WLPB-R; Woodcock, 1991) Reading Vocabulary subtest. The Reading Vocabulary subtest is comprised of two parts that assess a person's knowledge of Synonyms and Antonyms, respectively. The Synonym portion measures participants' ability to identify a word that has the same or nearly the same meaning as the test item presented by the examiner. The Antonym portion measures participants' ability to identify a word whose meaning is the opposite or nearly the opposite in meaning of the test item presented by the examiner. Performance on the Synonym and Antonym portions of the Reading Vocabulary subtest forms a single index of expressive vocabulary. Reliability exceeds .90.

Language comprehension was assessed using the WLPB-R Listening Comprehension subtest (Woodcock, 1991). The test focuses on a number of semantic operations beginning with simple verbal analogies and associations and moving to the ability to infer implications. This 38-item cloze procedure requires the participant to listen to a sentence and then supply a key word that completes the meaning of the sentence. The task requires about 10 min. Reliability exceeds .90.

4.4. Procedures

Participants were tested individually during one 2- to 2.5-hour testing session after school or on a Saturday at participants' schools in a quiet classroom or the library. A total of 16 examiners participated in administering the test battery. All examiners completed an extensive six-hour training conducted by the investigators regarding administration and scoring procedures for each test within the assessment. The first assessment was observed one-on-one for consistency in following the script, and the student record booklet was reviewed for recording/scoring accuracy, followed by immediate feedback. Reliability checks were completed for each measure that involved scorer judgment. Two scorers independently scored 10% of the student responses on the GORT, the WLPB-R word attack subtest, and the TOWRE subtests for sight word reading and phonemic word reading. The inter-scorer reliability was 96.5% on the GORT-4, 92% on the WLPB-R, and 95.5% on the TOWRE.

4.5. Data analysis

Latent Class Analysis (LCA) was used in a two-step analytical process: first, to identify below-average comprehenders and subsequently, to describe the component skill profiles of below-average comprehenders. Because LCA has yet to be used extensively in reading research, a discussion of its purposes and strengths is warranted here. LCA is conceptually related to cluster analysis, in that it is a multivariate method designed to identify unobserved (or latent) subpopulations of individuals based on multiple observed measures (Lazarsfeld & Henry, 1968; Lubke & Muthén, 2005). LCA uses maximum likelihood estimation to fit a hypothesized model in which membership in a specified number of latent classes is related to performance on the included measures and to produce fitted

probabilities of class membership for individuals. The observed measures can be categorical or continuous as in the present study.[1] Each latent class can be interpreted as a subpopulation with homogenous profiles on the multiple observed measures included in the analysis, while the differences between the latent classes indicate heterogeneity in the population studied.

LCA has several advantages over non-parametric cluster analysis approaches, such as mean centroid cluster analysis. First and foremost, as a model-based method, LCA provides statistical tests and goodness of fit indices for testing inferential hypotheses about the number of classes that exist in the population of interest. These tests are considered to be more rigorous than the arbitrary criteria used in traditional cluster analysis (Lubke & Muthén, 2005; Magidson & Vermunt, 2002). The model-fitting process begins with a one-class model (i.e., a model in which all readers are hypothesized to demonstrate a single, homogeneous profile) to which additional classes are added one at a time (i.e., hypothesizing each time that one additional subgroup exists), and statistical tests are conducted at each step to determine if the additional class significantly improves the goodness of fit of the model. Simulation studies in the statistical literature have found that these tests are robust and specific in determining when latent classes can and cannot be differentiated in the population (e.g., Lo, Mendell, & Rubin, 2001; Tofighi & Enders, 2006; Yang, 2006). Most recently, Nylund, Asparouhov, and Muthén (2007), using a Monte Carlo simulation, found that the bootstrapped Likelihood Ratio Test (bLRT) is a robust and consistent indicator for the existence of additional classes in the population, at various sample sizes and under various assumptions. This test builds on the likelihood ratio test commonly used in structural equation and multilevel modeling, in which the change in the −2 log-likelihood (Δ-2LL) statistic is compared to a chi-square distribution to determine if a hypothesized model (i.e., a model with n classes) has significantly better fit than a comparison model (a model with n–1 classes); the bLRT improves upon this test by comparing the Δ-2LL to a distribution estimated via the bootstrapping resampling method, thus not invoking assumptions about the distribution of Δ-2LL statistics that may not hold for latent class models. In addition to this primary test, secondary criterion include the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC); if these statistics decrease as additional classes are added, it provides further support for the conclusion that the additional subgroups found exist in the population.

In the present study, we combined the statistical evidence provided by the bLRT, AIC, and BIC with substantive concerns about the size and interpretability of the resulting subtypes of readers found, as recommended by Lubke and Muthén (2005) and others. Specifically, based on prior research, we hypothesized that we would find heterogeneity both in the reading comprehension levels of all adolescent readers and in the component skill profiles of struggling readers. Given this expectation, we were conservative in regarding additional subgroups as meaningful only if they both had consistent statistical support across the bLRT, BIC, and AIC, and if they did not yield additional small subgroups (i.e., making up fewer than 5% of the sample participants)[2] that were difficult to interpret as meaningfully distinct from the other groups found (i.e., only differ on one observed measure in a way that suggests noise more than signal). In this sense, our findings indicate a conservative lower bound for the extent of heterogeneity in the population.

Second, unlike traditional cluster analysis approaches, LCA provides standard errors associated with the parameter estimates for the means of each latent class on each of the multiple measures. This allows one to construct 95% confidence intervals around those means and to determine on which measures two similar classes converge and diverge by examining which intervals do and do not overlap. For instance, examining these confidence intervals might indicate that two classes of poor readers are differentiated from one another on word reading measures but do not perform significantly different on oral language measures. In the present study, we used this advantage of LCA to interpret the strengths and weaknesses of each profile relative to the others. It also allowed us to determine on which measures a given class was significantly below meaningful benchmarks on national norms (i.e., standard scores of 100 and 85).

Third, LCA has an advantage over non-parametric cluster analysis methods in that it allows for likelihood ratio tests to determine whether substantive variables predict class membership, a technique that was used in the current study to investigate whether comprehension level predicted component skill profile. In this test, a hypothesized model that included a regression path between comprehension level and component skill profile was compared to a model in which this regression path was set to 0 (i.e., indicating that comprehension level was not related to component skill profile). It is worth noting that although some LCA models invoke an assumption of conditional independence of indicators, such that the multiple measures are assumed to be uncorrelated within each latent class, this assumption can be relaxed (as in the models fitted below) when there are theoretical reasons to consider specific indicators to be correlated within class.

5. Results

Initial analyses were conducted to identify groups of students demonstrating empirically different levels of reading comprehension on three measures of reading comprehension. Subsequently, those groups demonstrating below-average reading comprehension were analyzed to identify empirically distinct profiles on a range of reading component skill measures.

5.1. Initial analyses to identify below-average comprehenders

Initial analyses indicated that adolescents presented at least four empirically different levels of reading comprehension. A sequence of LCA models fitted with increasing numbers of latent classes indicated several categorically different levels of reading comprehension achievement, based on students' performance on the three measures of reading comprehension. It should be noted that although the sample was stratified based on KRA performance, inclusion of this measure in the analysis to define struggling readers adds to the reliability of profile distinctions; the more measures of a latent construct included in an analysis, the more reliable the results. Given the stratified sampling plan and the five KRA achievement levels, it might be expected that five profiles will emerge from the analysis. Nonetheless, it may be that KRA levels do not fully determine profile membership, and this is indeed the case as indicated in the findings.

In each model, the three measures were allowed to correlate within class, thus relaxing the assumption of conditional independence. This decision has theoretical support in that the three reading comprehension measures would be expected to correlate even after partialing out class membership (as well as empirical support based on the residual covariances from models that did not include these

[1] Note that LCA with continuous observed variables is sometimes referred to as Latent Profile Analysis, although recent frameworks and software such as Mplus do not make this distinction.

[2] We were conservative in not selecting models with small subgroups for both methodological and practical reasons. In methodological terms, small latent classes will tend to have imprecisely estimated means, making it difficult to characterize such classes. For instance, a small component skill class might have a mean standard score of 85 on a picture vocabulary measure, but a large standard error and thus a 95% confidence interval that cover 100. It would be difficult to say whether such a class is below average or in the average range for picture vocabulary. In more practical terms, a small subgroup is more likely to be the result of idiosyncratic effects of instruction or background characteristics that are specific to our convenience sample of schools, or to idiosyncratic, construct-irrelevant effects of the measures used, and thus may not represent subtypes that are meaningful for informing intervention efforts more generally.

Table 2
Goodness-of-fit statistics, results of bootstrapped log-likelihood ratio test, and estimates of within-class correlations for Latent Class Analysis describing levels of reading comprehension achievement ($n = 319$).

	2-class solution	3-class solution	4-class solution	5-class solution
−2LL	6010.312	5961.574	5932.512	5914.52
AIC	6036.312	5995.574	5974.513	5964.520
BIC	6085.219	6059.529	6053.516	6058.572
Δ-2LL (Δ df = 4)	49.154***	48.738***	29.061***	17.993***
Within-class correlations				
GORT with WLPB-R	.57	.49	.49	.41
KRAS with WLPB-R	.53	.41	.24	.30
KRAS with GORT	.50	.25	.19	.19

*** p<.000.

correlations). Table 2 provides the goodness-of-fit statistics for the two-class, three-class, four-class, and five-class models.

The four-class solution indicating four levels of reading comprehension was chosen, based on the bLRT, BIC, and AIC that all indicated at least four significantly different classes and our substantive conclusion that four classes, but not more, were of practically important size and theoretically interpretable. In particular, the bLRT indicated that the addition of a third and fourth class each yielded a statistically significant improvement in the model fit (Δ-$2LL = 48.738$; $\Delta df = 4$; $p < .0001$ and Δ-$2LL = 29.061$; $\Delta df = 4$; $p < .0001$, respectively), a result which was also supported by changes in the BIC and AIC, which both declined with the addition of a third and fourth class. The addition of a fifth class also led to a statistically significant improvement in the model fit as indicated by the bLRT (Δ-$2LL = 17.993$; $\Delta df = 4$; $p < .0001$) and a decline in the AIC. However, the BIC increased with the additional class, suggesting worsening fit. More importantly, the additional fifth class was quite small (including fewer than 5 participants or less than 2% of the sample), and it was difficult to interpret as representing a meaningful group of students; this class demonstrated very low performance on the GORT and WLPB-R (standard scores near 75), but high performance on the KRA (performance level near 4.5), suggesting that these students may be outliers who happened to score particularly well on the KRA for idiosyncratic reasons.

The four classes generally displayed ordinal differences in reading comprehension achievement. Table 3 displays the predicted mean of each class on the observed measures, corresponding 95% confidence intervals, and the proportion[3] and number of students in the sample classified into each class. In order from lowest to highest performance on all three measures, the groups were designated as *Struggling Comprehenders* ($n = 121$), *Low Average Comprehenders* ($n = 74$), *Average Comprehenders* ($n = 72$), and *Advanced Comprehenders* ($n = 51$). As shown in Table 3, each successive class demonstrates mean scores that are higher than the previous class on all three measures with typical differences between adjacent classes of roughly half a standard deviation (7.5 standard score points on the GORT and WLPB-R or approximately .7 achievement levels on the KRA).

Despite this apparent ordinality, not all three measures differentiated between each pair of comprehender classes, as evidenced by the differences in predicted mean scores in adjacent classes on each measure. For instance, the WLPB-R scores differentiated the higher two classes from one another very well (as evidenced by their 29 standard score point difference and non-overlapping 95% confidence intervals) and differentiated the lower two classes well (as evidenced by their nearly 11 standard score point difference and non-overlapping confidence intervals). However, this measure did not differentiate between the low average and average comprehenders. The 2.34 standard score point fitted difference on the WLPB-R between these latter two classes is not statistically significant (i.e., the 95% confidence intervals for the latent means on WLPB-R for these two classes overlap). Similarly, the KRA differentiated the lower three comprehender classes well (as evidenced by differences of more than one achievement level and non-overlapping confidence intervals), but did not differentiate between the average and advanced comprehenders. The .17 levels difference on the KRA between these classes is not statistically significant (as indicated by overlapping confidence intervals). The GORT reading comprehension scores differentiated between adjacent classes consistently across the four classes. The degree to which each of the three measures differentiated particular levels of readers differently does not reduce the trustworthiness of the classification of readers into these four levels. Rather, it indicates that none of the individual measures of reading comprehension could provide the same precision and reliability of classification across the range of achievement as a composite classification utilizing information from all three measures. The four latent classes described provide one such composite classification.

5.2. Profiles of below-average comprehenders

To address our research question, we examined the component skill profiles of students within the low average and struggling comprehender classes and identified five distinct skill profiles. In so doing, we fitted sets of LCA models with increasing numbers of latent classes based on students' performance on the nine measures of component reading skills (GORT rate and accuracy; TOWRE sight word efficiency and phonemic decoding efficiency; PPVT vocabulary; and WLPB-R listening comprehension, letter-word identification, word attack, and reading vocabulary). We fitted models to the combined subsample of *Struggling* and *Low Average Comprehenders*, the combination of which we refer to as *Below-average Comprehenders*, and to each of these two subsamples separately. We report here the results for the *Below-average Comprehenders,* because similar skill profiles were found whether we conducted the analysis with the combined sample or combined findings based on separate analyses with the two subsamples (i.e., interpreted the profiles that emerged from both of the two separate analyses). Results indicated substantial heterogeneity in component skills within the population of below-average comprehenders. Table 4 provides the goodness-of-fit statistics for the three-class, four-class, five-class, and six-class models, as well as the estimated within-class correlations fitted for selected measures. The five-class solution was chosen, based on goodness-of-fit statistics as well as substantive concerns. In particular, we found that the addition of a fourth and fifth class each yielded a statistically significant improvement in the model fit, as indicated by the bLRT (Δ-$2LL = 53.8$; $\Delta df = 10$; $p < .0001$ and Δ-$2LL = 68.1$; $\Delta df = 10$; $p < .0001$, respectively) and supported by declines in BIC and AIC. The addition of a sixth class also led to a statistically significant improvement in the model fit according to the bLRT (Δ-$2LL = 59.2$; $\Delta df = 10$; $p < .0001$), a result supported by declines in BIC and AIC. However, the addition of a sixth class led to the split of a meaningfully interpretable and moderately sized class (*Weak Reading Comprehenders,* described below) into two small classes (of only 10 and 7 students) which were challenging to interpret as theoretically or practically distinct from each other. Specifically, the two classes both demonstrated means on all of the component skill measures that were in the average range or above, while one of the two was distinguished by above-

[3] As noted in the Methods section above, the stratified sampling plan prevents us from making generalizations about the prevalence of these four classes in the broader population of adolescent readers. Thus, these proportions should not be interpreted as the proportion of adolescent readers in the population performing at each level of reading comprehension, but solely as the proportion of our sample classified into each latent class. Making broader generalizations about prevalence would require a simple random sample of adolescent readers or analytic techniques (e.g., weighting) that are beyond the scope of the current research questions and analyses.

Table 3
Estimated means on reading comprehension measures and corresponding 95% confidence intervals in brackets for each latent class, based on four-class model describing levels of reading comprehension achievement ($n = 319$).

	Struggling comprehenders	Low average comprehenders	Average comprehenders	Advanced comprehenders
GORT (Standard scores)	80.61 [78.47, 82.75]	87.91 [84.58, 91.24]	98.27 [94.31, 102.23]	109.55 [104.82, 114.28]
WLPB-R (Standard scores)	86.34 [84.35, 88.33]	97.21 [94.55, 99.87]	99.55 [94.94, 104.16]	128.52 [123.13, 133.91]
KRA (Achievement levels)	1.53 [1.41, 1.65]	2.95 [2.85, 3.05]	4.43 [4.26, 4.60]	4.60 [4.38, 4.82]
Proportion of sample classified	.38	.23	.23	.16
n Classified	121	74	72	51

average means on the WLPB-R reading measures (i.e., letter-word identification, word attack, and reading vocabulary). Given our interest in sources of reading difficulty, we concluded that distinguishing between readers with strong decoding and those with very strong decoding was not an important distinction and could be due to idiosyncratic factors. In each model, the assumption of conditional independence of indicators was relaxed for the eight pairs of measures listed in Table 4. Including each of these within-class correlations had both a theoretical justification and improved the goodness-of-fit of the model, as evidenced by a series of Likelihood Ratio Tests (all p's<.05).

Table 5 presents the fitted means and corresponding 95% confidence intervals on the nine component reading skill measures for each of the five latent classes of below-average comprehenders, and Fig. 1 displays these fitted means and confidence intervals in graphic form. Although the five classes were found to be somewhat ordinal in the severity and multiplicity of their weaknesses, their skill profiles were not parallel; that is, when each class was compared to the next most similar class, it was found to differ significantly on some indicators but not on others. This indicates that each class shares some areas of strength and weakness with other classes, but also that each presents a unique profile. Below, we describe the skill profile of each class and interpret it relative to those of the other classes and to meaningful benchmarks on national norms, drawing on the confidence intervals reported in Table 5 and displayed in Fig. 1 to determine which apparent differences are statistically significant.

The class that performed lowest on all measures was designated as *Readers with Severe Global Weaknesses* (n = 28 and 14% of sample) and is indicated by the red line in Fig. 1. This class demonstrated skills that were more than one standard deviation below national norms on all measures. As shown by the confidence intervals in Fig. 1, the means of this class were significantly different from a standard score of 85 on each measure. This class showed particularly low performance on word reading accuracy measures, scoring on average more than two standard deviations below national norms on GORT passage accuracy, phonemic decoding efficiency, and word attack. This class performed significantly below all other classes on all of the reading accuracy and reading fluency measures (as indicated by the non-overlapping confidence intervals for these indicators).

The next class was designated *Readers with Moderate Global Weaknesses* (n = 71 and 36% of sample) and is indicated by the green line in Fig. 1. Like the class of *Readers with Severe Global Weaknesses*, this class demonstrated below-average performance on all measures, though to a lesser degree, typically performing a standard deviation below national norms. Although the *Readers with Severe Global Weaknesses* had significantly lower performance than the *Readers with Moderate Global Weaknesses* on the reading accuracy and fluency measures, the two classes had equally low performance on the vocabulary and listening comprehension measures. (As shown by the overlapping confidence intervals in Fig. 1, the apparent differences between the two classes' estimated means on these measures were not statistically significant). It is also worth noting that both classes demonstrated relative weaknesses on tasks involving decoding of pseudo-words (word attack and phonemic decoding efficiency) compared to their performance on tasks involving sight words (letter-word identification and sight word efficiency), suggesting that they may over-rely on knowledge of known words rather than automatized decoding skill while reading. Together, these two classes present difficulties in multiple component skills with similar levels of difficulty with language comprehension but differing degrees of difficulty with reading accuracy and reading fluency skills.

The third class, designated *Dysfluent Readers* (n = 57 and 29% of sample) and indicated by the orange line in Fig. 1, demonstrated language comprehension and word reading accuracy skills in the average range but fluency skills (at both the word and passage levels) that were below-average. Specifically, their mean on WLPB Reading Vocabulary and WLPB Listening Comprehension was not significantly different from the national average (as shown by the confidence interval overlapping a standard score of 100). However, their performance on the four timed measures was significantly lower than the higher-performing classes, not significantly different from the *Readers with Moderate Global Weaknesses*, and substantially below their own performance on all of the other measures. Moreover, *Dysfluent Readers* performed significantly better than *Readers with Severe Global Weaknesses* on all assessed component skills, but were distinguishable from *Readers with Moderate Global Weaknesses* only in their language comprehension skills and on the WLPB letter-word identification task.

Alternately, the class designated *Weak Language Comprehenders* (n = 21 and 11% of sample) and indicated by the blue line in Fig. 1 consistently demonstrated skills in the average range, except for a relative weakness in listening comprehension. As shown by the confidence intervals in Fig. 1, the only measure on which this class performed significantly below the national mean of 100 was the WLPB Listening Comprehension measure. This class performs significantly better than the two lowest classes on all assessed

Table 4
Goodness-of-fit statistics and results of bootstrapped log-likelihood ratio test for Latent Class Analysis describing skill profiles of below-average comprehenders ($n = 195$).

	3-class solution	4-class solution	5-class solution	6-class solution
Goodness-of-fit statistics				
−2LL	12,796.534	12,742.778	12,674.654	12,615.500
AIC	12,890.534	12,856.778	12,808.655	12,769.500
BIC	13,044.365	13,043.339	13,027.946	13,021.521
Δ-2LL (Δ df = 10)	121.082***	53.756***	68.124***	59.154***
Within-class correlations				
GORT rate with GORT accuracy	.40	.37	.37	.39
GORT rate with SWE	.41	.44	.55	.51
PDE with WA	.51	.51	.51	.54
SWE with LWID	.22	.22	.25	.29
SWE with PDE	.46	.45	.43	.44
SWE with WA	.21	.22	.22	.24
PPVT with RV	.47	.38	.38	.30
LWID with PDE	.39	.37	.39	.43
WA with LWI	.51	.49	.50	.48

*** p<.0001.

Table 5
Estimated means and corresponding 95% confidence intervals on component reading skills, based on four-class latent class model describing below-average comprehenders, with areas of relative weakness (n = 195).

	Severe global weaknesses	Moderate global weaknesses	Dysfluent readers	Weak language comprehenders	Weak reading comprehenders
GORT					
Rate	71.61 [65.39, 77.82]	81.78 [79.53, 84.03]	84.98 [82.77, 87.19]	99.64 [97.09, 102.19]	112.43 [107.87, 117.00]
Accuracy	65.27 [58.63, 71.92]	80.36 [76.40, 84.32]	84.96 [82.37, 87.54]	105.38 [100.87, 109.88]	106.68 [100.27, 113.09]
TOWRE					
Sight word efficiency	77.87 [70.93, 84.81]	88.58 [86.01, 91.14]	89.21 [89.86, 94.57]	100.24 [97.10, 103.37]	98.02 [91.96, 104.07]
Phonemic decoding efficiency	66.00 [60.57, 71.43]	83.63 [77.36, 89.90]	86.58 [83.47, 89.70]	99.29 [94.73, 103.86]	100.79 [95.95, 105.63]
PPVT	78.74 [74.51, 82.97]	80.06 [75.26, 84.86]	93.76 [90.36, 97.15]	99.71 [96.50, 102.93]	98.82 [92.08, 105.57]
WLPB					
Listening comprehension	77.80 [73.02, 82.58]	78.15 [73.06, 83.25]	97.30 [91.42, 103.18]	92.55 [85.50, 99.61]	104.47 [95.51, 112.43]
Letter-word identification	76.97 [70.09, 83.85]	88.73 [85.46, 92.01]	96.07 [92.99, 99.14]	106.38 [102.89, 109.87]	110.45 [104.53, 116.37]
Word attack	67.10 [57.52, 76.69]	85.66 [78.86, 92.46]	94.00 [89.05, 98.96]	114.29 [108.21, 120.36]	112.90 [106.61, 119.19]
Reading vocabulary	79.44 [75.86, 83.03]	84.48 [80.88, 88.09]	95.59 [91.39, 99.78]	100.27 [95.70, 104.84]	100.11 [94.80, 105.42]
Fitted probability of classification	.15	.35	.30	.11	.10
Proportion of below-average sample classified	.14	.36	.29	.11	.10
N of sample classified	28	71	57	21	18

skills, but is distinguishable from *Dysfluent Readers* in their reading accuracy and reading fluency. Furthermore, they show a relative strength in reading accuracy as is witnessed by the overlap of the confidence interval for WLPB word attack with a standard score of 115 (1 *SD* above the mean).

The final class, designated as *Weak Reading Comprehenders* (n = 18 and 9% of sample) and indicated by the purple line in Fig. 1, demonstrated component reading skills that were all in or above the average range, suggesting that their specific weakness lie primarily in the task of reading comprehension itself. As shown in Fig. 1, the confidence intervals for this class do not cover or fall below the national mean of 100 on any measure. Although *Weak Reading Comprehenders* significantly outperformed the lowest three classes on all assessed skills, they are distinguishable from *Weak Language Comprehenders* primarily in their above-average passage reading rates, which were more than two-thirds higher than national norms. The extremely high reading rate of *Weak Reading Comprehenders* suggests an additional hypothesis: they may be reading through text at a speed that is not conducive to strategic comprehension. These competing hypotheses are considered further in the discussion.

To explore the extent to which these five skill profiles might be specific to *Struggling Comprehenders* or *Low Average Comprehenders*,

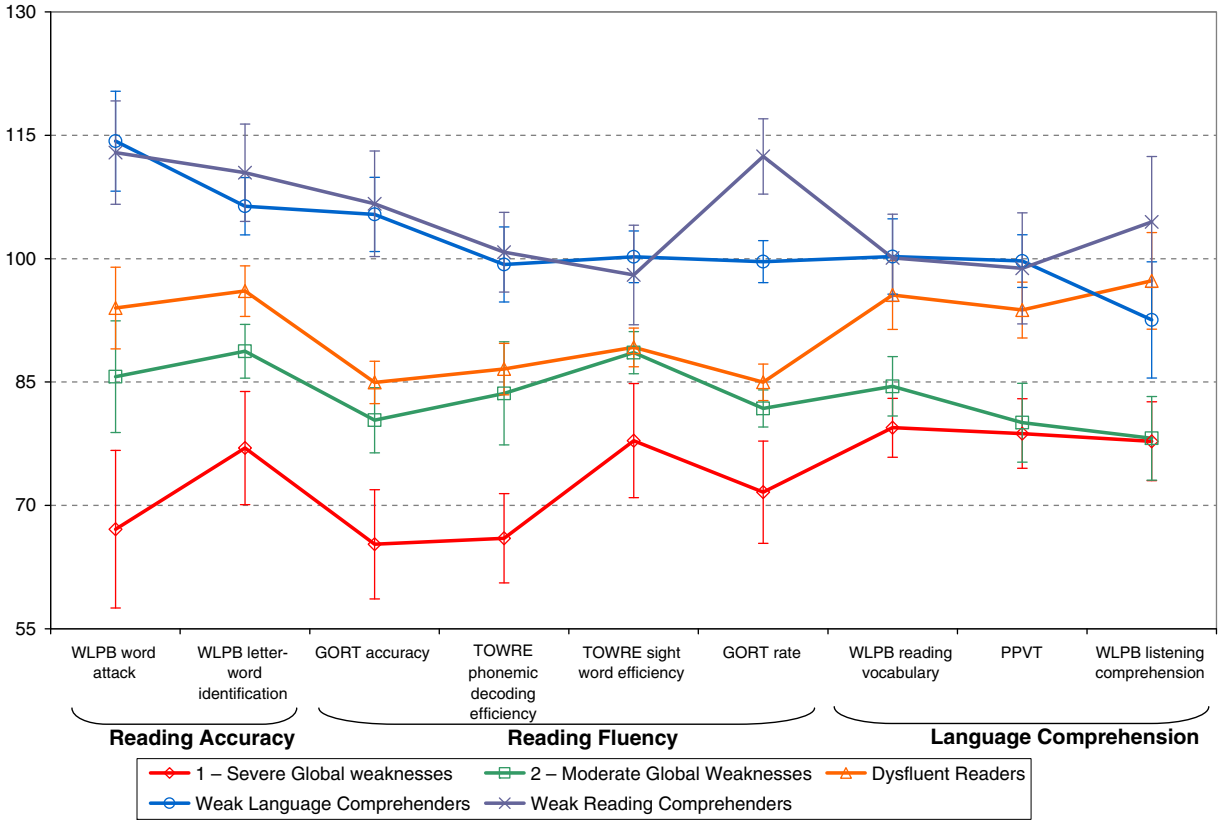

Fig. 1. Line plot with mean standard score on each component reading skill and associated 95% confidence intervals for the five profiles of below-average comprehenders, based on the five-class solution (n = 195).

Table 6
Percentage of below-average comprehenders by comprehender class (low average vs. struggling) and component skill profile, with number of students in parentheses (n = 195).

	Comprehender class	
Component skill profile	Low average	Struggling
Severe global weaknesses	4.1% (3)	20.7% (25)
Moderate global weaknesses	14.9% (11)	49.6% (60)
Dysfluent readers	40.5% (30)	22.3% (27)
Weak language comprehenders	18.9% (14)	5.8% (7)
Weak reading comprehenders	21.6% (16)	1.7% (2)

we investigated whether students' classification into one of the five skill profiles differed as a function of their comprehender class (struggling vs. low), as determined by the first set of LCA models described above. An additional five-class LCA model was fitted that included students' comprehender class as a dichotomous predictor of skill profile class and compared this to a model in which comprehender class was unrelated to skill profile class. A log-likelihood ratio test confirmed that comprehender class did indeed significantly predict membership in skill profile class (Δ-2LL = 70.562; Δdf = 4; $p < .0001$), indicating that the differences in the distributions of *Low Average* and *Struggling Comprehenders* assigned to the five skill profiles was statistically significant.

To provide insight into the relationship between comprehension level and component skill profile, Table 6 displays the number and percentage of *Low Average* and *Struggling Comprehenders* in the sample who fell into each of the skill profile classes. As the table demonstrates, the most prevalent profile among Struggling Comprehenders was the *Readers with Moderate Global Weaknesses* profile: about half (49.6%) fell into this profile. Despite nearly equal numbers of students from the two classes of comprehenders, the Dysfluent Readers profile was the most prevalent profile among Low Average Comprehenders (40.5%) but was only a distant second in prevalence among Struggling Comprehenders (22.3%).[4] Interestingly, a single class of comprehenders predominated in each of two of the skill profiles. Specifically, *Readers with Severe Global Weaknesses* were comprised almost exclusively of *Struggling Comprehenders*, whereas *Weak Reading Comprehenders* were comprised almost exclusively as *Low Average Comprehenders*.

Additionally, to investigate the extent to which the KRA achievement levels relate to component skill profiles, we cross-tabulated the prevalence of each component skill class by KRA level. As shown in Table 7, students scoring Proficient or below on the KRA can appear in any of the *Below-average Comprehender* classes, with the only exception being that no students scoring Unsatisfactory on the KRA could be classified as *Weak Reading Comprehenders*. Despite the pervasiveness of all the classes within each achievement level, some classes were more common than specific levels. For example, students scoring at KRA level 1 (unsatisfactory) were much more likely to fit the *Severe* or *Moderate Global Weaknesses* classes than the others (34% and 53% respectively), while students scoring at KRA level 2 (basic) were more likely to fit the *Moderate Global Weaknesses* or *Dysfluent Readers* classes (47% and 31% respectively). KRA level 3 students, deemed "proficient" by Kansas standards, most often fit the *Dysfluent Readers* class (41%), but just as many level 3 students were split between the *Weak Language* and *Weak Reading Comprehender* classes (19% and 22% respectively). However, it should be noted that more than 10% of the KRA level 2 students also fit these two profiles, demonstrating relative strength in most component skills. These results suggest that a great deal of heterogeneity in strengths and weaknesses of component reading skills exists at all three of these KRA achievement levels.

[4] Despite our inability to make generalizations about prevalence in the overall population of the readers as noted in the Methods section, we are able to make generalizations about the prevalence of skill profiles within each comprehender class. Because the variable used to stratify the sample, the KRA scores, was included as an indicator of comprehender class, the estimates of prevalence of component skill profiles within comprehender class provided here take into account the stratification of the sample. One way of understanding this is to think of each of the comprehender classes as representative of a specific population to which we can generalize our findings concerning the prevalence of component skill profile.

6. Discussion

The challenges facing the field of adolescent literacy are considerable and have been well documented (e.g., Adelman, 2006; Allensworth & Easton, 2005; Biancarosa et al., 2007; Lemke et al., 2005; McCombs, Kirby, Barney, Darilek, & Magee, 2005; National Center for Educational Statistics, 2005; PEW Center on the States, 2008; Vanneman et al., 2009). To date, very few interventions or programs have been found to be effective with ASRs, especially with those attending urban schools, and as a result, the achievement gap has not significantly narrowed, particularly for certain subgroups of learners (Carnegie Council on Advancing Adolescent Literacy, 2010; Denton & Vaughn, 2008; Edmonds et al., 2009; Hess, 2008; Slavin et al., 2008). One of the reasons for the lack of progress in closing the achievement gap may be due to the fact that interventions are not closely aligned with the comprehensive and diverse needs of ASRs. The overarching goal of this study was to identify and bring clarity to the reading skill profile and instructional needs of unique clusters of ASRs and closely examine the profile each cluster presented. With this body of knowledge, we hypothesize that literacy interventions might be crafted that are better aligned with and responsive to the needs of ASRs.

In order to answer our research question, "Do adolescents with below-average comprehension exhibit differentiated profiles of component reading skills including word reading accuracy, word-level and passage-level fluency, and oral language?" and to expand upon the general comments above, we highlight the unique clusters of ASRs found in the Latent Class Analysis.

We found that there were, indeed, unique subgroups of ASRs in our sample of urban students with specific instructional needs, including students with specific weakness in listening comprehension, those with specific weaknesses in reading comprehension, others who struggled with fluency, and students with moderate and severe levels of global weaknesses. This finding supports the notion that it is unlikely there is a single underlying source of poor comprehension (e.g., Cain & Oakhill, 2006; Perfetti & Hart, 2001). Just how individual component skill profiles relate to comprehension is an area in need of clarity, as not all poor comprehenders experience deficits in all reading component skills (Cornoldi, De Benie, & Pazzaglia, 1996). In fact, in a study by Cornoldi et al. (1996), the researchers speculated that there may be subtypes of poor comprehenders. In essence, our works brings some measure of clarity to this question.

By using data from a diverse sample drawn from urban schools and statistical techniques that allow for inferences to the larger population, we provided more robust evidence for claims about the heterogeneity in ASRs that can inform both policy and practice decisions.

The five profiles of component skills found among below-average comprehenders demonstrate considerable heterogeneity. Although the profiles were somewhat ordinal in the severity and multiplicity of their weaknesses, the classes were not strictly parallel indicating that adolescent readers are not simply better or worse on all skills. Rather, they are distinguished also by their specific strengths and weaknesses. Only two of the classes were particularly similar with respect to relative "peaks" and "valleys" though dissimilar with respect to severity of all difficulties: those with severe global weaknesses and those with moderate global

Table 7
Proportions of profiles of below-average comprehenders by Kansas reading achievement levels ($n = 195$).

	Below-average comprehender profile					
KRA level	Severe global weaknesses	Moderate global weaknesses	Dysfluent readers	Weak language comprehenders	Weak reading comprehenders	Total
1	18	28	6	1	0	53
2	7	32	21	6	2	68
3	3	11	30	14	16	74
Total	28	71	57	21	18	195

weaknesses. Other classes were distinguished from these two classes and from each other by particular skills. *Dysfluent Readers* showed weaknesses only on speeded measures (TOWRE and GORT); as shown in Fig. 1, they are the most disordinal relative to the other classes in that their performance on the speeded measures is essentially the same as the two lowest classes whereas their performance on the language and unspeeded word reading measures is much more similar to the two higher-performing classes. *Weak Language Comprehenders* were distinguished by average to above-average performance on all component skills except listening comprehension, which was a half of a standard deviation below norms. *Weak Reading Comprehenders* demonstrated pervasive strengths, performing at or above average on all components skills. It is probable that students in this class demonstrate weaknesses in skills not assessed in this battery of component skills, including potential difficulties with strategic processing of extended text, limited experience with particular genres of texts, or limitations in background knowledge necessary for comprehension of the passages on the reading comprehension measures. It is also possible that these students have all the requisite skills and knowledge, but are disengaged from academic reading tasks to an extent sufficient to suppress their reading comprehension performance.

The two lowest below-average comprehenders classes (*Readers with Severe Global Weaknesses* and *Readers with Moderate Global Weaknesses)* demonstrated relative weaknesses on tasks involving decoding of pseudo-words (word attack and phonemic decoding efficiency) compared to their performance on tasks involving sight words (letter-word identification and sight word efficiency), suggesting that they may over-rely on knowledge of known words rather than automated decoding skill while reading. It is notable that our analyses distinguished two groups of such readers, whereas only the study by Buly and Valencia (2003) has done so previously.

The two top classes of below-average comprehenders were relatively similar, but the *Weak Reading Comprehenders* read much faster than the *Weak Language Comprehenders* (or any of the other classes). The extremely high reading rate of *Weak Reading Comprehenders* (more than two-thirds higher than national norms) suggests that what appears to be a strength may in fact be their real weakness; that is, they may be reading through text at a such a fast speed that it is not conducive to strategic comprehension. Previous studies have not found groups of such readers, who in essence show no weakness on the components identified by the Simple View of Reading, but who nevertheless struggle with reading comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990). The current study confirms theories that such readers do indeed exist (Biancarosa & Snow, 2004; Biancarosa et al., 2007).

One other take on the profiles is to look at how their vocabulary and decoding skills are associated or disassociated. Unlike prior studies (e.g., Buly & Valencia, 2003) there was little evidence for "automatic word callers," i.e., readers with accurate and fluent word reading skills but low vocabulary and comprehension. However, that is not to say that vocabulary levels were always commensurate with word reading skills. If you look at the right-most two indicators in Fig. 1, you will see that the relative levels of reading vocabulary and word attack are strikingly different for the different classes—the *Readers with Several Global Weaknesses* have vocabulary skills that are much better than their word attack skills (by about 2/3 *SD*), the *Readers with Moderate Global Weaknesses* and *Dysfluent Readers* have vocabulary and word attack skills that are more or less commensurate with each other, and the top two profiles have decoding skills that are much higher (almost 1 *SD*) than their vocabulary skills.

The group of *Dysfleunt Readers* is consistent with previous findings of the only other study to assess oral reading fluency as a dimension separate from word reading accuracy (Buly & Valencia, 2003; Kieffer, Biancarosa, Christodoulou, Mancilla-Martinez, & Snow, 2007; Lesaux & Kieffer, 2008). Together, these findings suggest that the fluency assessments that are so often used as curriculum-based measures in earlier grades can also provide information about grade-level reading achievement at higher-grade levels (Fuchs, Fuchs, Hosp, & Jenkins, 2001).

Discussions of the role and impact of reading fluency must be tempered by issues of definition and measurement. For example, is reading fluency accuracy in word-level reading skills, proficiency in word-level reading skills and prosody, or proficiency in word-level reading skills, prosody, and comprehension (e.g., see Kuhn, Schwanenflugel, & Meisinger, 2010; Rasinski, Reutzel, Chard, & Linan-Thompson, 2011 for a recent discussion of this issue).

As discussed above relative to the proportions of types of comprehenders found, the proportion of students demonstrating each below-average comprehender component skill profile (e.g., *Dysfluent Readers and Readers with Severe Global Weaknesses*) might have been different given a different sampling plan or operationalization of below-average comprehension; however, it is nonetheless accurate to conclude that these five distinct skill profiles exist, though at rates that may differ within the larger population. Given our sampling plan and analyses, our findings speak directly to the proportion of students demonstrating specific profiles relative to levels of performance on the stratifying measure (see Table 7). When we look at the five profiles of component skills found among below-average comprehenders, results suggest that a great deal of heterogeneity in strengths and weaknesses of component reading skills exists at each of the three lowest KRA achievement levels. Among below-average comprehenders who achieved a score of 3 on the KRA, nearly 20 percent of these students demonstrated either severe or moderate global weaknesses, another 40 percent struggled with fluency, while the rest appeared to struggle with comprehension alone. Among students scoring a 2 on the KRA, more than half showed either moderate or severe global weaknesses, and almost a third struggled on fluency measures; very few struggled with comprehension alone. Only among those students scoring a 1 on the KRA was there anything approaching homogeneity in terms of their component reading skills; over 85% showed moderate or global weaknesses. Given the high number of ASRs in urban school populations, there is likely a great deal of heterogeneity in the component reading skills with which these students struggle, but greater proportions of students scoring at the lowest levels of state tests like the KRA imply greater proportions of students with profiles similar to those students in the most needy subgroups found (i.e., *Dysfluent Readers, Readers with Severe or Moderate Global Weaknesses*).

As described earlier, the theoretical foundation for our study was the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough,

1990). While acknowledging the complexity of the reading process, this theory holds that reading comprehension is the result of skilled and efficient word-level reading skills and language comprehension. Generally, we found that the Simple View theory holds for the ASRs in our study. That is, the poor comprehenders in our study presented reading skill profiles that included readers with poor word-level skills (WL) but adequate language comprehension, those with adequate word-level skills but poor language comprehension (LC), and those with both poor word-level skills and language comprehension (WL/LC). Thus, we extend the theoretical utility of the Simple View of Reading for ASRs by showing that ASRs do indeed reflect readers with poor skills in word-level, language comprehension, or both areas. This is a significant finding with implications for policy and intervention.

An additional finding was that, as expected, there are multiple levels of adolescent comprehenders in urban schools and those levels are aligned with the five levels designated by the Kansas Reading Assessment. However, we found four levels: Struggling Comprehenders, Low Average Comprehenders, Average Comprehenders, and Advanced Comprehenders. Additionally, all three measures differentiated between each pair of comprehender classes. This finding indicates that none of the individual measures of reading comprehension could provide the same precision and reliability of classification as a composite score from all three measures. Thus, in part, comprehension measures themselves may account for classification differences.

7. Conclusion

While research studies during the past decade have added to our understanding of the reading profiles of struggling adolescent learners, collectively they have fallen short of providing us with sufficient information that enables researchers or practitioners to differentiate various subgroups of learners from among this larger group of underachievers. In the absence of reliably defined subgroups, intervention work is adversely affected. The size of the achievement gap (particularly for students attending urban schools) and the shortness of instructional time available makes it imperative that the instruction provided to these at-risk learners be exceedingly well designed and delivered. However, a prerequisite to designing effective interventions is a clear understanding of the defining characteristics of the learners who will be the target of the interventions. Hence, an overriding purpose of this study has been to differentiate various subgroups of struggling adolescent learners so that responsive interventions can be developed that help narrow the achievement gap. The analysis described in this article seems to support the development of comprehensive reading programs that present reading instruction in a balanced fashion. That is, both the word-level and comprehension needs of ASRs must be addressed.

7.1. Future research

Questions remain concerning effective interventions for struggling adolescent readers, especially those attending urban schools. The results of several recent large-scale random assignment experiments shows just how difficult it is to move the needle on elementary and adolescent struggling reader achievement; few studies have been able to show results with the moderate to large effects needed to close the achievement gap (e.g., Corrin, Somers, Kemple, Nelson, & Sepanik, 2008; Gamse, Jacob, Horst, Boulay, & Unlu, 2008; James-Burdumy et al., 2010). Thus, additional research is needed to answer the question, *Do interventions that are closely aligned with the reading component skill profile and needs of subgroups of ASRs identified in the study result in statistically and socially significant reading proficiency gains?* In addition, research is needed that explores the question, *Why do students with adequate reading skills in all the identified key components of reading perform poorly on standardized measures of reading comprehension?* Student characteristics like motivation for reading, engagement in school, strategy knowledge, and information processing skills might help inform our knowledge of this subtype of ASR. Finally, research is needed that answers the often-stated question, *What works for whom under what conditions?* This question is largely overlooked in large-scale studies that focus on mean score gains of the targeted sample under study.

The current study focused on ASRs in urban schools in one state in the Midwest; this focus is certainly a limiting factor. However, findings from the study suggest that ASRs do indeed present clear and heterogeneous profiles of reading strengths and weaknesses. Given the large numbers of ASRs in urban schools, policy and practice need to address the varied strengths and weaknesses of this population through better identification and diagnostic systems and multiple and comprehensive intervention approaches.

Acknowledgement

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education by grant R305G04011 to *The University of Kansas Center for Research on Learning*. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.
Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college.* Washington, DC: U.S. Department of Education.
Allensworth, E. M., & Easton, J. Q. (2005). *The on-track Indicator as a predictor of high school graduation.* Chicago: Consortium on Chicago School Research.
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.
Berman, I., & Biancarosa, G. (2005). *Reading to achieve: A governor's guide to adolescent literacy.* Washington, DC: NGA Center for Best Practices.
Biancarosa, G., Palincsar, A. S., Deshler, D. D., & Nair, M. (2007). Adolescent literacy: Myths and realities. In D. D. Deshler, A. S. Palincsar, G. Biancarosa, & M. Nair (Eds.), *Informed choices for struggling adolescent readers: A research-based guide to instructional programs and practices* (pp. 11–36). Newark, DE: International Reading Association.
Biancarosa, G., & Snow, C. E. (2004). *Reading next—A vision for action and research in middle and high school literacy: A report from Carnegie Corporation of New York.* Washington, DC: Alliance for Excellent Education.
Buly, M. R., & Valencia, S. (2003). *Meeting the needs of failing readers: Cautions and considerations for state policy.* Seattle: University of Washington, Center for the Study of Teaching and Policy.
Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *The British Journal of Educational Psychology, 76*, 683–696.
Carnegie Council on Advancing Adolescent Literacy (2010). *Time to act: An agenda for advancing adolescent literacy for college and career success.* New York, NY: Carnegie Corporation of New York.
Cassidy, J., & Cassidy, D. (2009). What's hot for 2009?*International Reading Association, 26*(4), 1–9 February/March.
Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research, 45*, 1142–1157.
Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In H. W. Catts & A. G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 25–40). Mahwah, NJ: Lawrence Erlbaum Associates.
Chall, J. S. (1983). *Stages of reading development.* New York, NY: McGraw-Hill.
Christenbury, L., Bomer, R., & Smagorinsky, P. (Eds.). (2009). *Handbook of adolescent literacy research.* NY: Guilford.
Connor, C. M., Morrison, F. J., Fishman, B. J., Ponitz, C. C., Glasney, P. S., Underwood, P., et al. (2007). The ISI classroom observation system: Examining the literacy instruction provided to individual students. *Educational Researcher, 38*(2), 85–99.
Connor, C. M., Morrison, F. J., Fishman, B. J., Schatschneider, C., & Underwood, P. (2007). Algorithm-guided individualized reading instruction. *Science, 315*(5811), 464–465.
Cornoldi, C., De Benie, R., & Pazzaglia, F. (1996). Profiles of reading comprehension difficulities: An analysis of single cases. In C. Cornoldi & J. Oakland (Eds.), *Reading comprehension difficulties: Processes and interventions* (pp. 113–136). Mahwah, NJ: Erlbaum.
Corrin, W., Somers, M. -A., Kemple, J., Nelson, E., & Sepanik, S. (2008). *The enhanced reading opportunities study: Findings from the second year of implementation (NCEE*

2009–4036). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
Council of Great City Schools (2001). *A decade of outcomes in the nation's urban schools 1990–1999: A report on urban student achievement and course taking.* Washington DC: Council of the Great City Schools.
Curtis, M. B. (2002). *Adolescent literacy: A synthesis of research.* Boston: Lesley College, The Center for Special Education.
Denton, C. A., & Vaughn, S. (2008). Reading and writing intervention for older students with disabilities: Possibilities and challenges. *Learning Disabilities Research and Practice*, *23*(2), 61–62.
Deshler, D. D. (2006). Building a strong literacy foundation for high school excellence. In D. Clark (Ed.), *The challenge of high school reform* (pp. 20–27). Washington, DC: The Aspen Institute.
Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test-third edition.* Circle Pines, MN: American Guidance Service.
Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Klinger Tackett, K., et al. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*, *79*(1), 262–300.
Ehri, L. C. (1997). Sight word learning in normal readers and dyslexics. In B. A. Blachman (Ed.), *Foundations of reading acquisition and dyslexia* (pp. 163–189). Mahwah, NJ: Erlbaum.
Ehri, L. C. (1998). Grapeheme-phoneme knowledge is essential for learning to read words in English. In J. Metsala, & L. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 3–40). Mahwah, NJ: Lawrence Erlbaum.
Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading disabilities in at-risk children. *Journal of Educational Psychology*, *90*, 37–55.
Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239–256.
Gamse, B. C., Jacob, R. T., Horst, M., Boulay, B., & Unlu, F. (2008). *Reading first impact study final report (NCEE 2009–4038)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10.
Hess, F. M. (2008). *Still at risk: What students don't know, even now.*Common Core Washington, DC.
Hock, M. F., Brasseur, I. F., Deshler, D. D., Catts, H. W., Marques, J., Mark, C. A., et al. (2009). What is the reading component skill profile of adolescent struggling readers in urban schools? *Learning Disability Quarterly*, *32*(1), 21–38.
Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, *2*, 127–160.
James-Burdumy, S., Deke, J., Lugo-Gil, J., Carey, N., Hershey, A., Gersten, R., et al. (2010). Effectiveness of selected supplemental reading comprehension interventions: Findings from two student cohorts (NCEE 2010–4015). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
Kamil, M. L. (2003). *Adolescents and literacy: Reading for the 21st century.* Washington, DC: Alliance for Excellent Education.
Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008). Improving adolescent literacy: Effective classroom and intervention practices: A Practice Guide (NCEE #2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education Retrieved from. http://ies.ed.gov/ncee/wwc.
Kansas State Department of Education (2005). Report card 2004–2005 Retrieved February 12, 2007, from. http://online.ksde.org/rcard/.
Kieffer, M. J., Biancarosa, G., Christodoulou, J. A., Mancilla-Martinez, J., & Snow, C. E. (2007). *Shades of struggle: Heterogeneity among urban adolescent struggling readers.* Chicago, IL: American Educational Research Association April.
Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge, United Kingdom: Cambridge University Press.
Kintsch, E. (2005). Comprehension theory as a guide for the design of thoughtful questions. *Topics in Language Disorders*, *25*, 51–64.
Konold, T. R., Juel, C., McKinnon, M., & Deffes, R. (2003). A multivariate model of early reading acquisition. *Applied Psycholinguistics*, *24*, 89–112.
Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly*, *45*(2), 230–251 Dx.doi.org/10.1598/RRQ.45.2.4.
LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293–323.
Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis.* Boston: Houghton Mifflin.
Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology*, *95*(2), 211–224.
Lee, J., Grigg, W. S., & Donahue, P. L. (2007). The nation's report card: Reading 2007: National assessment of educational progress at grades 4 and 8 (NCES No. 2007–496). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
Lemke, M., Sen, A., Johnston, J. S., Pahlke, E., Williams, T., Kastberg, D., et al. (2005). Characteristics of U.S. 15-year-old low achievers in an international context: Findings from PISA 2000 (NCES 2006–010). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
Lesaux, N. K., & Kieffer, M. J. (2008). *Qualitative differences or a question of degree? The prevalence and sources of reading comprehension difficulties among language minority learners and their classmates in urban middle schools.*Paper presented at the Annual Meeting of the Society for the Scientific Study of Reading, Asheville, NC July.
Lippman, L., Burns, S., & McArthur, E. (1996). Urban schools: The challenge of location and poverty. National Center for Education Statistics.
Lo, Y., Mendell, N., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika*, *88*, 767–778.
Lubke, G. H., & Muthén, B. O. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, *10*, 21–39.
Magidson, J., & Vermunt, J. K. (2002). Latent class models for clustering: A comparison with K-means. *Canadian Journal of Marketing*, *20*, 37–44.
McCardle, P., & Chhabra, V. (2004). *The voice of evidence in reading research.* New York: Brookes.
McCardle, P., Scarborough, H. S., & Catts, H. W. (2001). Predicting, explaining, and preventing reading difficulties. *Learning Disabilities Research and Practice*, *16*, 230–239.
McCombs, J. S., Kirby, S. N., Barney, H., Darilek, H., & Magee, S. (2005). *Achieving state and national literacy goals, a long uphill road: A report to Carnegie Corporation of New York.* Santa Monica, CA: RAND.
Mellard, D. F., Fall, E., & Mark, C. (2008). Reading profiles for adults with low-literacy: Cluster analysis with power and speeded measures. *Reading and Writing*, *22*, 975–992.
National Center for Educational Statistics (NCES) (2004). *The nation's report card: Reading 2002.* Washington, DC: U.S. Department of Education.
National Center for Educational Statistics (NCES) (2005). The nation's report card: Reading 2005. Washington, D.C: U.S. Department of Education http://nces.ed.gov/nationsreportcard/nrc/reading_math_2005/s0002.asp?printver.
National Center for Educational Statistics (NCES) (2009). The nation's report card: Reading 2009 trial urban district assessment. Washington, D.C.: U.S. Department of Education http://nces.ed.gov/nationsreportcard/pdf/dst2009/2010459.pdf.
National Institute of Child Health and Human Development (2000). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00–4769). Washington, DC: Government Printing Office.
Nylund, K. L., Asparouhov, T., & Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*, 535–569.
Perfetti, C. A., & Hart, L. (2001). The lexical bases of comprehension skill. In D. Gorflen (Ed.), *On the consequences of meaning selection* (pp. 67–86). Washington DC: American Psychological Association.
Perfetti, C. A., Marron, M. A., & Foltz, P. W. (1996). Sources of comprehension failure: Theoretical perspectives and case studies. In C. Cornoldi, & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 137–165). Mahwah, NJ: Erlbaum.
PEW Center on the States (2008). *One in 100: Behind bars in America 2008.* Washington, DC: The PEW Charitable Trusts.
Phelps, S. (2005). *Ten years of research on adolescent literacy, 1999–2004: A review.* Chicago: Learning Point Associates.
Rashotte, C. A., MacPhee, K., & Torgesen, J. K. (2001). The effectiveness of a group reading instruction program with poor readers in multiple grades. *Learning Disability Quarterly*, *24*, 119–134.
Rasinski, T. V., Reutzel, D. R., Chard, D., & Linan-Thompson, S. (2011). Reading fluency. In M. Kamil, P. D. Pearson, E. B. Moje, & P. F. Afflerbach (Eds.), *Handbook of reading research*, *vol. IV*, New York: Routledge.
Sabatini, J. P. (2002). Efficiency in word reading of adults: Ability group comparisons. *Scientific Studies of Reading*, *6*(3), 267–298.
Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., et al. (2007). *Interventions for adolescent struggling readers: A meta-analysis with implications for practice.* Portsmouth, NH: RMC Research Corporation, Center on Instruction.
Schultz, G. F. (1993). Socioeconomic advantage and achievement motivation: Important mediators of academic performance in minority children in urban schools. *The Urban Review*, *25*(3), 221–232.
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 distribution of reading ability. *The New England Journal of Medicine*, *326*, 145–150.
Slavin, R. E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. *Reading Research Quarterly*, *43*(3), 290–322.
Snow, C. (2002). *Reading for understanding: Toward an R & D program in reading comprehension.* Santa Monica, CA: RAND.
Tofighi, D., & Enders, C. K. (2006). Identifying the correct number of classes in a growth mixture models. In G. R. Hancock (Ed.), *Mixture models in latent variable research*. Greenwich, CT: Information Age.
Torgesen, J. K. (1999). Assessment and instruction for phonemic awareness and word recognition skills. In H. W. Catts, & A. G. Kamhi (Eds.), *Language and reading disabilities* (pp. 128–149). Needham Heights, MA: Allyn and Bacon.
Torgesen, J. K., Rashotte, C. A., & Alexander, A. (2001). Principles of fluency instruction in reading: Relationships with established empirical outcomes. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 334–355). Parkton, MD: York Press.
Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency.* Lake Oswego, OR: Northwest Evaluation Association.
United States Census Bureau (2002). *2000 Census Bureau of population and housing, summary population and housing characteristics.* Washington, DC: Author.
Vanneman, A., Hamilton, L., Baldwin-Anderson, J., & Rahman, T. (2009). Achievement gaps: How Black and white students in public schools perform in mathematics and reading on the National Assessment of Educational progress, (NCES 2009–455). Washington, DC: National Center for Educational Statistics, Institute of Education Sciences, U.S. Department of Education.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: multivariate evidence for a convergent skills model of reading development. *SSSR Journal*, *11*(1), 3−32.
Wechsler, D. (1991). *Wechsler intelligence scale for children* (3rd ed.). San Antonio, TX: Psychological Corporation.
Wiederholt, J. L., & Bryant, B. R. (2001). *GORT4: Gray Oral Reading Test.* Austin, TX: PRO-ED.
Woodcock, R. (1991). *Woodcock language proficiency battery-revised.* Allen, TX: DLM Teaching Resources.
Woodcock, R. W. (1997). Woodcock diagnostic reading battery: Examiner's manual. In R. W. Woodcock (Ed.), *Woodcock diagnostic reading battery*. Ithaca, IL: Riverside Publishing.
Yang, C. (2006). Evaluating latent class analyses in qualitative phenotype identification. *Computational Statistics and Data Analysis*, *50*, 1090−1104.