IDENTIFICATION OF STUDENTS IN

LATE ELEMENTARY GRADES WITH READING DIFFICULTIES

by

CHENG-FEI LAI

A DISSERTATION

Presented to the Department of Educational Methodology, Policy, and Leadership
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2012

DISSERTATION APPROVAL PAGE

Student: Cheng-Fei Lai

Title: Identification of Students in Late Elementary Grades with Reading Difficulties

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Educational Methodology, Policy, and Leadership by:

Akihito Kamata	Chairperson
Gerald Tindal	Member
Gina Biancarosa	Member
Robert O'Brien	Outside Member

and

Kimberly Andrews Espy	Vice President for Research & Innovation/Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2012

DISSERTATION ABSTRACT

Cheng-Fei Lai

Doctor of Philosophy

Department of Educational Methodology, Policy, and Leadership

June 2012

Title: Identification of Students in Late Elementary Grades with Reading Difficulties

Piecewise latent class growth analysis (LCGA) was used to examine growth patterns in reading comprehension and passage reading fluency on easyCBM, a popular formative assessment system. Unlike conventional growth modeling, LCGA takes into account the heterogeneity of growth and may provide reliable predictions for later development. Because current methods for classifying students are still questionable, this modeling technique could be a viable alternative classification method to identifying students at risk for reading difficulty. Results from this study suggested heterogeneity in reading development. The latent classes and growth trajectories from the LCGA models were found to align closely with easyCBM's risk rating system. However, results from one school district did not fully generalize across another. The implications for future research on examining growth in reading are discussed.

CURRICULUM VITAE

NAME OF AUTHOR: Cheng-Fei Lai

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene OR
San Diego State University, San Diego CA
University of Georgia, Athens GA

DEGREES AWARDED:

Doctor of Philosophy, Educational Methodology, Policy, and Leadership, 2012, University of Oregon
Master of Science, Rehabilitation Counseling, 2005, San Diego State University
Bachelor of Science, Psychology, 2001, University of Georgia

AREAS OF SPECIAL INTEREST:

Structural Equation Modeling
Growth Modeling
Universal Design Principles on Instrument and Instruction
Test Validity and Reliability
Test Accommodation and Accessibility Issues

PROFESSIONAL EXPERIENCE:

Research Assistant, Behavioral Research and Teaching (BRT), University of Oregon, Eugene, OR, 2007 to present

Disability Specialist, Office for Students with Disabilities, University of Texas at Arlington, TX, November 2005 – July 2007

GRANTS, AWARDS, AND HONORS:

David Moursund Scholarship , University of Oregon College of Education Scholarships and Awards Committee, 2011-2012

Graduate Teaching Fellowship, Educational Methodology, Policy & Leadership, University of Oregon, Fall 2007 to Spring 2012

Travel Grant, Educational Methodology, Policy & Leadership, University of Oregon, 2010- 2012

PUBLICATIONS:

Nese, J. F. T., Biancarosa, G., Anderson, D., Lai, C. F., & Tindal, G. (2012). Within-year oral reading fluency with CBM: A comparison of models. *Reading and Writing*. doi: 10.1007/s11145-011-9304-0

Anderson, D., Lai, C., Alonzo, J., & Tindal, G. (2011). Examining a grade level math CBM designed for exceptionally low performing students. *Educational Assessment*.

ACKNOWLEDGMENTS

I wish to express sincere appreciation to my examining committee: Dr. Akihito Kamata for his guidance and scholarly advice, Dr. Gerald Tindal for his continued support and encouragement, Dr. Gina Biancarosa for her mentorship and constructive criticism, and Dr. Robert O'Brien for his insightful questions and suggestions.

Special thanks to Dr. Deni Basaraba, who edited this manuscript and consistently provided excellent feedback, Dr. Julie Alonzo, for believing in me and providing mentorship since I started the doctoral program, Dr. Leilani Sáez for sharing her depth of knowledge in reading, and Drs. Joseph Nese and Chalie Patarapichayatham, for their time navigating the complexities of the statistical analyses. My parents, husband, and close friends, provided me endless support throughout my education and graduation. For them, I am most grateful.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Response to intervention (RTI) is a promising framework for identifying students at risk for reading difficulties (Fletcher, Coulter, Reschly, & Vaughn, 2004; L. Fuchs, Fuchs, & Speece, 2002; Vaughn & Fuchs, 2003). The essential components of RTI include school-wide systematic screening, progress monitoring, and implementing instructional interventions designed to meet student needs (Johnson, Mellard, Fuchs, & McKnight, 2006). Systematic and careful implementation of RTI could reduce the number of children at-risk for failing in schools and special education referrals. Despite advances made in developing accurate screening procedures (Compton et al., 2006; Jenkins, 2003; Speece & Case, 2001), identifying students at risk for reading difficulties at the later elementary grades is still a concern (Jenkins, Hudson, & Johnson, 2007). It may not be reasonable to assume that measures commonly used to identify students in early elementary grades are equally effective for identifying students in the later elementary grades with reading difficulties due to differences in reading development between early and later elementary students (Speece, 2005).

In early elementary grades, students focus on learning to read and in later elementary grades transition to reading to learn (Chall, 1996). The differences in reading development between students in the early and the later elementary grades could lead to different reading problems as well. Some recent research also supports the idea that reading for older elementary students is different, with reading comprehension, word reading, and fluency as important factors (Speece et al., 2010) as opposed to phonological awareness, phonics, decoding, word recognition, word reading fluency, and spelling, which are important predictors of student reading performance in the early grades (National Institute for Child Health and Human Development [NICHD], 2000).

Reading fluency, particularly Oral Reading Fluency (ORF), is perhaps the most frequently researched measure in the CBM literature (Buck & Torgesen, 2003; Good, Simmons, & Kame'enui, 2001; Hintze & Silberglitt, 2005; McGlinchey & Hixson, 2004; Shapiro, Keller, Santoro, & Hintze, 2006; Shaw & Shaw, 2002; Stage & Jacobsen, 2001). Although ORF is a strong indicator of reading proficiency for students in the early elementary grade (e.g., grades 1-3) (Fuchs et al., 2001), the strength of ORF as an indicator for students in the later elementary grades (e.g., grades 4-6) is unclear (Johnson, Pool, & Carter, 2011). In fact, some studies have found the relation between ORF and comprehension is weaker in the later elementary grades than the early elementary grades (Jenkins & Jewell, 1993) and that the magnitude of the relation between ORF and comprehension decreases over time (Yovanoff, Duesbery, Alonzo, & Tindal, 2005).

Reading achievement can be relatively unstable as students progress into later elementary grades (Phillips, Norris, Osmond, & Maynard, 2002). In the later elementary grades, subgroups of students have been found to struggle in reading, including: (a) students with poor reading foundational skills (i.e., instructional casualties), (b) English language learners (ELL), (c) students requiring ongoing interventions, and/or (d) students with late-emergent reading disabilities (Johnson et al., 2011). These subgroups suggest that reading becomes increasingly heterogeneous and multidimensional at the later elementary grades (Speece et al., 2011). Their reading developmental profiles (e.g., low initial status and steep growth, high initial status and flat growth, etc.) may also appear different from those of early elementary students.

Effective RTI begins with screening mechanisms that can identify students at risk for academic failure. As part of a universal screening process, many educators use curriculum-based measurement (CBM) to screen all students in a school in the fall, winter, and spring. Although great advances have been observed in the development of universal screeners, much of the progress is limited to early elementary grades, particularly in kindergarten through second grade (Compton et al., 2006; Jenkins, 2003; Speece & Case, 2001). Screeners used in later elementary grades have also been found less accurate in identifying students who need additional instructional supports compared to early elementary grades (Jenkins, Hudson, & Johnson, 2007). Because current methods for classifying students as at-risk for academic failure are still unsatisfactory, growth modeling techniques, such as growth mixture modeling (GMM) that can identify subgroup of students with different developmental profiles, offer alternative classification methods for identifying students at risk for reading difficulties (Boscardin, Muthén, Francis, & Baker, 2008).

Although conventional growth modeling such as random coefficient modeling (e.g., Raudenbush & Byrk, 2002) and latent growth curve modeling (Meredith & Tisak, 1990) has been more widely used in education research, the assumption that students come from a single population with one average growth trajectory may not be reasonable. GMM relaxes the single population assumption and allows different classes of individuals to vary around different mean growth curves. GMM, therefore, may be a more reasonable model to use for modeling heterogeneity in reading development for students in later elementary grades. Moreover, recent research indicates there is growing interest in identifying unobservable groups of students with different growth patterns (Boscardin et al., 2008, Jordan et al., 2006; Muthén et al., 2002; Kreisman, 2003).

This study seeks to extend the current literature on curriculum-based measurement (CBM) screening assessments in reading for students in later elementary grades by examining these research questions: (a) Are there latent classes and reading growth trajectories on easyCBM, a popular CBM reading comprehension and passage reading fluency measures?; (b) Do the latent classes and growth trajectories align closely with easyCBM's risk rating system, and what are the patterns of reading skills?; and (c) Are the results from this study generalizable to another school with a different RTI model?

Conventional growth modeling and latent class growth analysis (LCGA) were used to examine these research questions. Conventional growth modeling assumes that individuals come from a single population with one growth trajectory. LCGA, a special type of GMM, on the other hand, assumes that there are different growth trajectories for each unobservable class and that all individual growth trajectories within a class are homogeneous. For both conventional growth modeling and LCGA models, the effects of covariates (e.g. student characteristics, effects of intervention, and/or initial reading skills) based on developmental profiles were examined. A series of LCGA models were conducted using several distal outcomes to explore the degree of consistency between latent class formations with the easyCBM risk rating system: (a) varying levels of risk status in Fall 2010 (F10RISK); (b) levels of risk in Winter 2011 (W11RISK); (c) levels of risk in Spring 2011 (S11RISK). Model fit were evaluated using a combination of fit indices and other criteria. The reading profiles of the latent classes derived from the final models were examined descriptively to identify students with different reading skills.

Current screening mechanisms in place at the later elementary grades have been unsatisfactory, with unacceptable rate of accuracy (Jenkins et al., 2007) and could miss students who do not display reading deficits until later (Catts, Hogan, & Adlof, 2005). Students progressing into later elementary grades may be heterogeneous in terms of their reading comprehension and fluency development. LCGA can capture such heterogeneity and offer an alternative technique to identify students at risk for reading difficulties. Greater understanding of reading development and profiles based on the latent class formation can help teachers identify students at risk for reading difficulties more easily, design targeted intervention plans, and subsequently, help build a more effective RTI model in schools.

CHAPTER II

LITERATURE REVIEW

Response to Intervention (RTI)

Until the recent reauthorization of the Individuals with Disabilities Education Act (2004), the IQ Discrepancy-based model has been the primary approach used to identify students with reading difficulties (Fletcher, 2011). Many have criticized the IQ discrepancy-based model, which relies on a battery of assessments at one point in time to make decisions about special education services eligibility, as an ineffective (Fletcher et al., 1998; Francis et al., 2005; Sternberg & Grigorenko, 2002), wait-to-fail model (Compton, Fuchs, Fuchs, & Bryant, 2006) that does not inform teachers the reasons for students' poor academic performance (MacMillan & Siperstein, 2002). As an alternative, Response to Intervention (RTI) has emerged as a promising framework for identifying of children at risk for reading difficulties (Fletcher, Coulter, Reschly, & Vaughn, 2004; L. Fuchs, Fuchs, & Speece, 2002; Vaughn & Fuchs, 2003).

Unlike the IQ discrepancy model, where the focus is on student low achievement (usually in reading) that is discrepant from the expected level of achievement, RTI enables the identification at-risk students through a prevention-oriented, multi-tiered, instructional approach that includes school-wide systematic screening, progress monitoring, and differentiated instructional intervention (Johnson, Mellard, Fuchs, & McKnight, 2006). In this approach, students who are identified as at-risk for academic failure despite strong general education instruction (Tier 1) during the screening process are provided with targeted, group-based, intensive instruction that is aligned to their needs and level of achievement (Tier 2). Students who do not respond to Tier 2 interventions (i.e., non-responders) are provided with more intensive individualized instructional interventions that aim to remediate students' deficits and their progress with acquiring critical skills is monitored closely and with greater frequency (Tier 3). With these components in place, educators can view RTI as a prevention-oriented "pre-referral treatment model" (Mather & Kaufman, 2006) whose goal is to identify students who are at risk for failing to meet grade level expectations and providing intensive instructional supports *before* they actually fail (Compton et al., 2006). For RTI to be successful, therefore, it is critical that the tools used to identify students at risk for academic failure have a high degree of accuracy in distinguishing those who are more likely to need secondary interventions (Tier 2) from those who can be successful receiving the general education program (Johnson et al., 2006).

Currently, the identification and intervening mechanisms for students in the early elementary grades with reading difficulties have been relatively successful (Compton et al., 2006; Jenkins, 2003; Speece & Case, 2001). It was not until recently that similar procedures for identifying students with similar difficulties in the later elementary grades began receiving more attention (e.g., Biancarosa & Snow, 2004). Many concerns, however, remain regarding the valid identification of these students (Jenkins, Hudson, & Johnson, 2007). One issue receiving increased attention is the applicability of measures commonly used in the early elementary, such as Oral Reading Fluency (ORF), to identify older students. Empirically, ORF has proven to be an efficient and valid tool to assess elementary school students' overall reading competence (e.g., Fuchs, Fuchs, Hosp, & Jenkins, 2001; Speece & Ritchey, 2005). However, it may not be reasonable to assume that measures commonly used to identify students in early elementary grades, such as

ORF, to be equally effective for identifying students in the later elementary grades with reading difficulties due to differences in reading development between early and later elementary students (Speece, 2005).

Reading Development of Students Over Time

Reading development between younger and older elementary students is quite different (Chall, 1996). According to Chall (1996), students in grades one and two learn the association of letters of the alphabet and their corresponding sounds (i.e., understand the spelling-sound system). In grades two to three, students consolidate what is learned in grades one and two and begin to recognize words composed of increasingly complex phonic elements and to read stories composed of increasingly complex words using their decoding knowledge (Chall, 1996). Specifically, younger elementary students are generally expected to develop skills and knowledge in phonological awareness, phonics, letter–sound relationships, word recognition, word reading fluency, and spelling (NICHD, 2000).

As Chall (1996) pointed out, what students in the early elementary grades focuses on the foundational components of speech and print – phonemes, phonological units, letter sounds – and the systematic relations between these components (i.e., *learning to read*). Beyond early elementary grades, students transition to a stage where they are *reading to learn* new information and gain knowledge. Learning word meanings (i.e., vocabulary) and building prior knowledge becomes the primary instructional goals as children in late elementary grades (four through six) begin to learn new knowledge, information, thoughts, and experiences by reading (Chall, 1996). Also beginning in grade four, the reading demands increase as students are also challenged to understand text in greater depth, to manage more challenging aspects of lower level processes such as mastering decoding and spelling complex polysyllabic words, and become familiar with the orthographic patterns associated with Anglo-Saxon, Greek, and Latin root morphemes to understand new words (Leach, Scarborough, & Rescorla, 2003).

Reading deficits of later elementary grade students. Because of the differences in reading development between students in the early and the later elementary grades, reading problems for these two groups may be different as well. Reading deficits of younger students, for example, are commonly associated with decoding, word recognition, word reading fluency, and spelling (NICHD, 2000). Older students, in contrast, may experience deficits similar to those of younger students, as well as difficulties with comprehension, vocabulary, and/or oral language (Catts, Adlof, & Weismer, 2006).

Results from a recent study (Speece et al., 2010) in which researchers examined a series of screening batteries (reading comprehension, oral-language, word recognition, word decoding, phonological processing, auditory memory, and spelling) intended to effectively identify at-risk readers in grade four students suggest that screening for reading problems at grade four requires a multivariate approach; skills such as word reading, fluency, and reading comprehension ability may be important factors that contribute to reading competence and that reading problems in fourth graders' reading skills (Speece et al., 2010). This study suggests that reading skills for older students are more complex than those targeted by most reading instruction programs and research in the early elementary grades, such as phonological awareness, phonics, decoding, word recognition, word reading fluency, and spelling (NICHD, 2000). These findings also suggest that single screening measures, such as ORF, may no longer be adequate or appropriate for identifying reading difficulties in the later elementary grades; a multivariate screening approach that incorporates multiple measures that target a number of critical reading skills may be a more reasonable alternative (Speece et al., 2010). Just as the skills required for proficient reading in the later elementary grades are more complex, so too is it possible that the range of difficulties older students experience while learning to read are more varied and complex.

Reading over time: The role of reading fluency and other dimensions. Reading fluency, particularly ORF, is perhaps the most frequently researched measure in the CBM literature, with many studies examining the predictive relation between performance on ORF and state standardized reading tests (Buck & Torgesen, 2003; Good, Simmons, & Kame'enui, 2001; Hintze & Silberglitt, 2005; McGlinchey & Hixson, 2004; Shapiro, Keller, Santoro, & Hintze, 2006; Shaw & Shaw, 2002; Stage & Jacobsen, 2001; see Appendix A for details). Although ORF can be an important indicator of reading competence that teachers can use to help plan systematic, targeted instruction (Fuchs et al., 2001; Kame'enui, Simmons, Good, & Harn, 2001; Schilling, Carlisle, Scott, & Zheng, 2007), whether fluency continues to play an important role in reading proficiency for students in the older grades is unclear (Johnson et al., 2011). Some research (Burns, et al., 2002) suggests, for example, that it is possible that reading fluency is still important at the later elementary grades. Burns et al. (2002) examined the relation between fluency rate and comprehension for 49 grade fourth grade students by having them orally read four passages with incrementing percentages of scrambled words (0-30%) and then administered a series comprehension questions related to the passages.

On average, students who successfully answered the comprehension questions read 49.68 words correct per minute, which led the authors to recommend that students need to read a minimum of 50 words correct per minute for comprehension to occur (Burns et al., 2002).

Declining contributions of ORF in the later elementary grades. Though the Burns et al. (2002) study documented a minimum rate reading fluency necessary for comprehension, the importance of reading fluency may decrease as students progressed into later elementary grades. Furthermore, not many have examined the relationship between fluency and comprehension beyond the early elementary grades (Denton et al., 2011). In fact, the relation between ORF and comprehension appears to weaken as students progress to higher grade levels, as documented by a study conducted by Jenkins and Jewell (1993) with a sample of 335 students in grades two through six. Consistently, the authors found correlations of decreasing magnitude between ORF and the Gates-MacGinitie, with correlations of .83, .88, and .86 in grades two, three, and four, dropping to .67 at grade six. Similarly, the correlations between reading aloud and Metropolitan Achievement Tests (MAT) total reading started at .87 at grade two but declined to .60 at grade six (Jenkins & Jewell, 1993). The increasingly weaker correlation between performance on ORF and achievement tests at the higher grade levels could indicate that ORF may be less sensitive in measuring reading ability of older students (Jenkins & Jewell, 1993).

Some studies have documented that reading fluency may play a less significant role in explaining the overall reading ability of older elementary students (Yovanoff, Duesbery, Alonzo, & Tindal, 2005). For example, Yovanoff et al. (2005) examined the relations between vocabulary and ORF to reading comprehension on five cross-sectional samples of students in grades 4-8 (*n* range = 900-1200). Results from a series of nested models testing for invariance of the multiple regression models suggest that ORF and vocabulary tasks were significant predictors of reading comprehension (Yovanoff, et al., 2005). However, the relative importance of ORF decreased in the higher grades, as indicated by the best model fit where all coefficients in grade four was freely estimated, while grades 5-8 were constrained to be equal (RMSEA ranges= .015-.060; AGFI ranges = .946-.989) [RMSEA less than .05 and AGFI close to 1.0 suggest good model fit] (Yovanoff et al., 2005). This study suggests that although ORF and vocabulary contribute significantly to reading comprehension in the later elementary and middle school grades the magnitude of their contributions and, subsequently, their importance, could change over the late elementary years (e.g., grades 5-6).

Stability of reading development over time. Do students who struggle with reading remain poor readers over time? Juel (1988) examined this question empirically in a small-scale longitudinal study by following the reading and writing development of 54 children from grades 1-4 and found that students' trajectories are remarkably stable. Juel (1988) found, for example, that the probability a poor reader in grade one would continue to remain a poor reader in grade four was .88. In contrast, the probability that a student with average reading skills in the beginning grade one would be a poor reader at the end of fourth grade was .12. Likewise, the probability of students with average reading ability continuing to demonstrate average reading ability in grade four was .87. Students with poor reading ability in the beginning in grade one, however, only had a probability of .13 of being categorized as an average reader in grade four.

Juel's (1988) study supports the notion that reading achievement is relatively constant, but this may be true only for students in the lower elementary grades. A study examining reading achievement for six years (grades 1-6), for example, found reading ability to be less stable over time (Phillips, Norris, Osmond, & Maynard, 2002). Using longitudinal data of 187 students, Phillips et al. (2002) found some groups of students that did not have the stable trajectories suggested by like Juel's (1988) findings. For example, one group of students in their sample did not fulfill the expectation that above- and below-average readers in grade one tend to remain so in grade six (*n*=24; 13% of the sample); specifically, students categorized as having above-average reading ability had only a .52 probability of remaining in the same group, and a .48 probability that they would drop to average. Although Phillips et al. (2002) found that a majority of students stayed in the same category in first and sixth grades (good readers remained good, average remained average, and so on), they also found that the trajectories of approximately 30% students were altered across the six grades, leading them to conclude that reading-achievement over time may not be as stable as Juel (1988) suggested.

Students at Risk for Reading Difficulties in Later Elementary Grades

Reading problems for older elementary students are heterogeneous and multidimensional (Speece et al., 2011). Struggling readers at these grade levels are likely to experience difficulties in more than one area (Johnson et al., 2011). Subsequently, their reading developmental profiles (e.g., low initial status and steep growth, high initial status and flat growth, etc.) may appear different from the profiles of early elementary students. Furthermore, a few subgroups of students could be expected to struggle in the later elementary grades, including: (a) students with poor reading foundational skills (i.e., instructional casualties), (b) English language learners (ELLs), (c) students requiring ongoing interventions, and/or (d) students with late-emergent reading disabilities (Johnson et al., 2011). Students who lack foundational reading skills in the early elementary grades, for example, make up the instructional casualties group. Because not all schools can deliver strong reading programs, this group of students may not have received strong reading instructions and may require supports in the later grades (Foorman, Breier, & Fletcher, 2003; Vaughn et al., 2008).

In recent years, the number of English Language Learners (ELLs) in schools has increased at a rapid rate (Maxwell, 2009; Education Week, 2009). In fact, the fastest growing segment of the public school population is ELLs. It is estimated that by 2015, ELLs enrollment in U.S. schools will reach 10 million and, by 2025, nearly one out of every four public school students will be an ELL student (NCELA, 2007). Moreover, ELLs in grade four have demonstrated consistently lower performance than non-ELL students on the 2009 and 2011 National Assessment of Educational Progress (NAEP) reading tests, with the average scale scores of 188 in both years compared to 225 and 224 respectively (NCES, 2009; 2011), and 70% of ELLs are categorized as being *Below Basic* proficiency compared to 30% of non-ELLs (NCES, 2011). Therefore, ELLs are another subgroup of students that may emerge as struggling readers in the later elementary grades and higher.

The third group of struggling students includes those who cannot keep up with their peers in the general education program despite participating in intensive reading interventions in the earlier grades (Johnson et al., 2011). Students from any of these groups may require on-going intervention in later grades before they are able to successfully perform at grade-level benchmarks.

It is also possible that students struggling to read proficiently in the later elementary grades have a late-emerging reading disability (RD). Students in this group may exhibit: (a) typical performance and progress during the early elementary grade but develop reading problems in later grades and/or (b) different profiles of reading performance over time (Compton, Fuchs, Fuchs, Elleman, & Gilbert, 2008; Leach et al., 2003; Lipka, Lesaux, & Siegel, 2006; Catts, Hogan, & Adlof, 2005). These students may rely almost entirely on memorization of words, thereby appearing to be proficient readers for the first several years of school, until this strategy becomes ineffective and insufficient (Juel, 1991). Researchers (Leach et al., 2003; Lipka et al., 2006) estimate that approximately 40% of children with RD in of older struggling students have a late-emerging reading disability.

In a preliminary study that examined the reading deficits of 31 students in fourth and fifth grades identified later for RD, for example, Leach et al. (2003) observed that these group of students appear to heterogeneous with regard to their strengths and weaknesses in various components of reading; some participating students had weak comprehension and word-level processing skills, some experienced only difficulty with word-level processing skills (dyslexic), while others experienced difficulty only with comprehension. The comprehension-only deficit profile could be unique to students with late-emerging reading problems. When prior achievement at grade three was examined through school records, Leach et al. (2003) found that students identified later as having RD had higher achievement scores than those who were identified as having RD earlier, leading them to conclude that these students were in fact students with late-emerging RD, not just students whose RD had been identified late.

In another similar study, Lipka et al (2006) examined reading development in a longitudinal study of 22 fourth grade four students with RD who had been followed since kindergarten. The researchers (Lipka et al., 2006) conducted repeated-measures ANOVAs and a series of pairwise comparisons to investigate the patterns of emergence in RD, reading ability, and risk status across grades K-4 and found two distinct groups of students: students with RD and typical readers. Results indicated that students with RD and typical readers had different reading development, and that a heterogeneous pattern of emergence of reading difficulties can be observed among students with RD (Lipka et al., 2006). Within the group of students with RD, Lipka et al. (2006) found three subgroups: poor readers, or students' who consistently performed below the 25th percentile across grades K - 4; borderline readers, or students' whose performance fluctuated across grades K-3, and students with late-emerging RD, or students performed above the 35th percentile across grades K-3. The follow-up pairwise comparisons revealed that the starting point of the late-emerging RD group was different from that of the borderline readers and poor readers. Throughout grades K-4, students with late-emerging RD exhibited positive rapid growth and performed higher on a reading achievement test (WRAT-3) than students in the borderline and poor-readers groups. In grade two, however, students with late-emerging RD displayed a sharp negative growth trend and by grade three, there was a significant difference in students' WRAT-3 scores that favored the late-emerging RD group and typical readers. Throughout grades K-4, the trajectories of the borderline and poor-readers groups were similar, with borderline students outperforming poor readers through all grades. Beginning in grade two, however, the differences between the two groups were no longer statistically significant. In grade four, there were no longer significant differences between the borderline, poor-readers, and late-emerging RD groups. In general, the late-emerging RD group had word reading skills in the average range of reading achievement and did not fall below the 25th percentile until grade four. Though the studies conducted by Leach et al. (2003) and Lipka et al. (2006) provide compelling results regarding the possibility of late-emerging RD characteristics in later elementary grades, the relatively small samples limit the generalizability of the results of these preliminary studies.

Universal Screening Procedures for Identifying At-risk Students in Elementary Grades

Successful implementation of RTI begins with accurate screening procedures. Educators are recommended to benchmark using screening measures, such as curriculum-based measurement (CBM), that are administered to all students in a school in the fall, winter, and spring as part of a universal (also known as school-wide) screening process (Johnson et al., 2006). Universal screening is the first step in identifying students at risk for academic failure (Hughes & Dexter, 2011).

Important qualities of screening measures: Specificity and sensitivity. Although correct classification of all readers (i.e., specificity) is ideal, however, it should be noted that the screening process will incorrectly identify some students as being at risk for academic failure (i.e., false positives) and some as not at risk when they should be (i.e., false negatives) (Johnson et al., 2011). There is always a trade-off between false positive and false negative classifications, so increasing either sensitivity or specificity inevitably results in a reduced value for the other. Good screeners should ideally minimize such occurrences and have desirable degrees of specificity and sensitivity (Jenkins, 2003). Sensitivity refers to the degree to which a screener identifies students who are "at-risk" who in fact perform unsatisfactorily on a future criterion measure (i.e., true positives). Specificity refers to the accuracy with which a screener is able to identify students who are "not-at-risk" and who later perform satisfactorily on a future criterion measure (i.e., true negatives). No screening measure can achieve 100% true-positive rates (identifying all students who are at risk) because errors will always occur (Compton, et al., 2006; Jenkins, 2003). Therefore, it is more reasonable to aim for a screening mechanism that yields the highest possible true-positives rates (highest specificity) and an acceptable number of false positives of 5-10%, or sensitivity of .90-.95 (Jenkins, 2003). In RTI, it is critical to avoid false negative errors so that students most in need of assistance can receive instructional interventions in an effort to prevent them from developing reading problems in later grades (Johnson et al., 2011).

Examining the adequacy of ORF as a screening measure. To date, ORF CBM is one of the most commonly used measures for universal screening purposes (Deno, 2002; Shapiro, Keller, Lutz, Santoro, & Hintze, 2006; Silberglitt, Yeo, & Cormier, 2010; Speece, Case, & Malloy, 2003; Wayman, Wallace, Wiley, Ticha, & Espin, 2007). As stated previously, the strong empirical and theoretical link between ORF and general reading proficiency (Fuchs, Fuchs, Hosp, & Jenkins, 2001) has encouraged much of the research to focus on ORF CBM. Although great advances have been observed in the development of universal screeners, much of the progress is limited to early elementary grades, particularly in kindergarten through grade two (Compton et al., 2006; Jenkins, 2003; Speece & Case, 2001).

ORF as a screener in the later elementary grades. Despite the wide acceptance of using ORF as a valid and effective screening tool for younger students , ORF screeners may not be as valid and effective for older students in later elementary grades (Jenkins, Hudson, & Johnson, 2007; Yovanoff et al., 2005), especially those with late-emerging reading difficulties (Compton et al., 2008). A review by Jenkins, Hudson, and Johnson (2007), for example, revealed only a handful of studies have examined the effectiveness of screeners for accurately classifying students at risk or not at risk for poor reading outcomes in grades three through four (Buck & Torgesen, 2003; Good, Simmons, & Kame'enui, 2001; McGlinchey & Hixson, 2004; Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008; Stage & Jacobsen, 2001). Jenkins et al. (2007), for example, found 15 classification studies for grades K-2, compared to only three studies examining the accuracy of classifications of ORF in grade three, and two in grade four.

At third grade, the accuracy of ORF at screening at-risk students is similar to the early elementary grades. Good, Simmons, & Kame'enui (2001) examined the relation between ORF administered in spring and the Oregon state assessment on a sample of 364 third graders. In their study, Good and colleagues (2001) reported a sensitivity rate of .83 and a specificity rate of .94, meaning that 83% of the students with *reading problems* were **correctly** identified as *at-risk* and 94% of students identified as having *adequate reading skills* were **correctly** identified as being *not at-risk*. Buck and Torgesen (2003) obtained similar findings in their examination of the relation between third grade ORF scores in spring term and reading comprehension subtest scores of the Florida

Comprehensive Assessment Test (FCAT) for a sample of 1,102 students in 13 schools in one Florida school district. Buck and Torgesen (2003) concluded that ORF measures can still accurately predict whether or not a given student will attain an acceptable score on the FCAT reading test for third graders, with sensitivity of .77 and specificity of .92. It is important to note, however, that only spring ORF scores were examined in the Good et al. (2001) and Buck and Torgesen (2003) studies.

These sensitivity values are not consistent, however, when ORF is administered at a different time, as demonstrated by Roehrig et al. (2008). In their study, Roehrig et al. (2008) examined the usefulness of ORF (DIBELS) as a predictor of future reading comprehension achievement (FCAT) on a sample of 16,539 third graders. They found that the sensitivity is at .91 when ORF and FCAT were administered concurrently in spring, but was reduced to .74 when predicting FCAT scores from ORF that was administered in the fall.

Moreover, the sensitivity and specificity rates reported in the Buck and Torgesen (2003) and Good et al. (2001) studies may not be entirely representative, as pointed out by Jenkins et al. (2007), because both groups computed sensitivity and specificity using only students who were categorized as the least and the most at-risk for later reading difficulties, excluding at least 40% of the sample. When ORF and state test scores of the entire sample of fourth graders were included, both sensitivity and specificity were in the moderate range of .60s and .70s (McGlinchey & Hixson, 2004; Stage & Jacobsen, 2001). With only moderate overall correct classification rate, reaching specificity rates only as high as .70s, this means that at least 25% of students were misclassified and likely did not receive the additional instructional supports (i.e., Tier 2 interventions) they needed to be successful.

In general, specificity and sensitivity of ORF screeners used in grades three and four are not an acceptable level. Jenkins (2003), for example, recommends the highest specificity value possible and sensitivity rates of .90-.95, compared to the earlier grades, where some screeners have reached acceptable accuracy rates (e.g. sensitivity as high as 1.00 and specificity of .98; see Jenkins et al., 2007, for full report). Such differences, then, suggest a much higher rate of false-positive and false-negative errors when screening students at-risk in later elementary grades. Moreover, little is known regarding the classification accuracy that can be achieved using ORF cut-points (Jenkins, 2003). This point is evident in Stage and Jacobson's study (2001), where sensitivity and specificity (.66 and .76) were lower when the ORF cut-point was set at 100 correct words per minute, versus sensitivity and specificity of .31 and .96 when the cut-point was 50 correct words per minute. It is clear, in other words, that more research is needed to examine classification accuracy of ORF and other CBM screeners currently used by educators in the later elementary grades.

Growth Mixture Modeling

Because current methods for classifying students are still questionable, growth modeling techniques that can identify subgroup of students with different developmental profiles offer alternative classification methods to identifying students at risk for reading difficulties (Boscardin, Muthén, Francis, & Baker, 2008). Growth mixture modeling (GMM) is emerging in educational reading and mathematics research as a method for identifying subgroups of individuals from a larger heterogeneous population with distinct developmental trajectories (Boscardin et al., 2008; Jordan, Kaplan, Olah, & Locuniak, 2006; Muthén, Khoo, Francis, & Boscardin, 2002). More widely used conventional growth modeling, such as random coefficient modeling (e.g., Raudenbush & Byrk, 2002) and latent growth curve modeling (Meredith & Tisak, 1990), assumes that the parameters of interest (i.e., growth factors like initial status and slope) originate from a single population with one average growth trajectory. GMM relaxes the single population assumption, allowing for differences across unobserved subgroups to be estimated using latent trajectory classes (i.e., categorical latent variables). Rather than considering individual variation around a single mean growth curve, GMM allows different classes of individuals to vary around different mean growth curves.

Both random coefficient modeling and latent growth curve modeling allow for the modeling of unconditional and conditional models using time-invariant (e.g. student characteristics) and time-varying (e.g. intervention program) covariates, respectively, based on continuous distribution functions. Conventional growth modeling (and other methods like regression, factor analysis, and structural equation modeling) can also be viewed as a variable-centered approach (Muthén & Muthén , 2000), where the goal is to identify variables that can significantly predict outcomes of interest and to describe the relationship between the dependent and independent variables. When the focus is the relationship among individuals and the purpose is to classify individuals into heterogeneous groups that are composed of individuals with similar characteristics, the approach is person-centered (Muthén & Muthén , 2000) and includes methods such as cluster analysis, latent class analysis, and finite mixture modeling.

Examining the appropriateness of conventional growth modeling. It may not always be reasonable to use conventional growth modeling because of the assumption that all individuals are drawn from a single population with common parameters. Using the results from Burns et al. (2002) study on determining the minimum reading fluency rate necessary for comprehension to illustrate this assumption, the authors reported the following statistics: a mean of 49.68 words per minute, with maximum and minimum rate as high as 90 and 13 words per minute. These statistics suggest a wide range of fluency rates (Burns et al., 2002) and a broad spectrum of readers. It may not be reasonable to assume, however, that students with 13 words per minute (poor readers) have the same rate of growth as the average reader. Similarly, higher performing readers with a fluency rate of 90 correct words per minute (good readers) may not have the same trajectory as the average reader. Using the same logic, the growth profiles of poor, average and good readers may be different: poor readers with low initial ORF scores, may have steep growth due to effective intervention and more room to grow, whereas good readers may start with high initial ORF scores, but display small or flat growth because there is not much room for growth (i.e., ceiling effects).

GMM offers researchers a more flexible approach in modeling growth, including the possibility to: (a) incorporate a categorical latent class variable to account for unobserved heterogeneity in the larger population, (b) model patterns of growth separately for each latent class by allowing model parameters (e.g. initial status and growth trajectories) to vary across latent classes, (c) examine the relationship between parameters of latent classes and other variables (i.e., class-invariant or class-specific covariate effects), and/or (d) predict a distal outcome based on background characteristics and class membership. Additionally, GMM can be extended to a general GMM model (Muthén et al., 2002), which allows each class to have different forms of growth (e.g. linear and curvilinear).

In current education research that utilizes GMM, many researchers are interested in examining differences in individuals and categorizing them into distinct subgroups based on characteristics such as education programs (Kreisman, 2003), intervention programs (Muthén et al., 2002), and socio-economic status (Jordan, et al., 2006). Therefore, instead of using conventional growth modeling to investigate these research questions, it may be more reasonable to use a growth mixture modeling approach to identify multiple unobserved subgroups (i.e., latent classes) with distinct growth trajectories, to describe longitudinal change within each unobserved subgroups, and to examine differences in change between and within unobserved subgroups (Muthén, 2004). GMM may also be an appealing approach to capture information about the latent classes and reading growth profiles on repeated measures within a year (i.e., fall, winter, and spring benchmark CBM).

Growth mixture modeling (GMM) in educational research. Even though the popularity of using mixture modeling like GMM is still in its infancy in education research (Bilir, Binici, & Kamata, 2008), there is growing interest in identifying unobservable groups of students with different growth patterns (Boscardin et al., 2008, Jordan et al., 2006; Muthén et al., 2002; Kreisman, 2003). This framework of modeling growth has been chosen over more conventional growth single-class approach because mixture modeling allows researchers the flexibility in specifying model specifications and assumptions (Muthén , 2004).

Furthermore, conventional growth modeling may not always be reasonable when modeling growth of reading development because of the underlying assumption that manifest growth trajectories are a sample from a single population of individuals characterized by a mean intercept and a mean slope. Mixture modeling techniques relax this assumption and allow for individual differences in growth rates. Identification of groups of students with different growth rates may be particularly important for the intervention referral decision process. With the limited resources available in many schools today, school teachers and administrators need to make informed decisions about how to allocate more instructional resources (e.g., personnel, instructional time, opportunity to learn in small groups, etc.) to support students in greater need of more intense intervention and less support to those who are likely to make progress with minimal support.

Application of GMM in early literacy research. There is, for example, a growing body of evidence supporting that students with reading difficulties consist of subgroup of students with different reading profiles and trajectories (Boscardin et al., 2008; Muthén, Khoo, Francis, & Boscardin, 2002). Muthén, Khoo et al. (2002), for example, applied GMM techniques to examine the relation between students' phonemic awareness skills in kindergarten and their first-grade word recognition development (n=409). Four distinct groups of students were found. 21% of the sample consisted of students who did not show growth at the end of kindergarten and continued to do poorly in word recognition development during first grade. 7% of the sample displayed rapid growth but remained at low reading levels at the end kindergarten. Two other classes of students were also identified (49% and 23% respectively) who were at average and above average levels in reading upon exiting kindergarten, and continued to perform well in first grade (Muthén , Khoo et al., 2002).

There is also evidence that multiple groups of students with different are present among students in kindergarten and early elementary grades from a 3-year longitudinal study (*n*=411) conducted by Boscardin et al. (2008). Specifically, Boscardin et al., (2008) examined the relationship between the development of precursor skills (i.e., phonological awareness and rapid naming) and the development of word recognition skill in later grades and found up 10 subgroups of students; students who were the most at risk for reading difficulties had a distinct developmental pattern. Results indicated that students who were identified as having difficulties acquiring phonological awareness skills in kindergarten stayed in the same developmental trajectory throughout the three years of the study. Similar to many studies examining students at risk for reading difficulties (Jenkins et al., 2007), these pioneer studies in reading research that applied latent growth modeling in have primarily focused on the development of students in early elementary grades. Boscardin and colleagues (2008) suggested that the growth mixture models are a potential mechanism for identifying and classifying students with reading difficulties that can “minimize anomalies and unfairness that are consequences of using an arbitrary cutoff for classification purposes” (p. 203).

Implications of Current Universal Screening for Later Elementary Students

The screening procedures for students beyond the early elementary grades have been unsatisfactory, with unacceptable rates of accuracy in identifying students at risk for reading difficulty (Jenkins et al., 2007). One possible reason for this could be that “children continue to develop on the very skills we use as screens, but our methods rarely take this development into account" (Speece,2005, p. 488). Furthermore, the current RTI screening methods miss students who do not display reading deficits until later (Catts, Hogan, & Adlof, 2005).

Furthermore, inaccuracy in classifying students may deprive students who are truly in need of intervention. An accurate classification system can help school administrators allocate resources appropriately by providing students with the greatest levels of need access to intervention programs designed to meet their needs. Educators need to be aware that there may be a group of students who demonstrate average achievement in the elementary grades, but who experience a decline in age-appropriate word reading ability beyond early elementary grades (Compton, et al., 2008; Leach et al., 2003; Lipka, et al., 2006). Additionally, accurate screening mechanisms prevent educators and schools from operating at the "wait-to-fail" model by identifying at-risk students and providing intensive instructional supports *before* they actually fail (Compton et al., 2006). Because schools and educators face many challenges, from facing today's emphasis on accountability to the tight economy, efficient resource allocation is increasingly important.

As proposed by Boscardin and colleagues (2008), "the use of growth mixture models to identify and classify students with reading difficulties minimizes anomalies and unfairness that are consequences of using an arbitrary cutoff for classification purposes" (p. 203). Mixture models offer an alternative to avoid the some problems associated with arbitrary decisions made to classify students as having a reading disability. In addition to being an alternative technique for identifying students with reading difficulties, mixture modeling also provides information like reading developmental patterns, growth rate, and direction. Information on these dimensions of reading can help teachers create homogenous instructional groups that are more effective for students by targeting students' weaknesses (Wesson, 1992), make informed decisions about instructional planning, and guide decisions for subsequent progress monitoring. Research on student growth could also be extended to program evaluation of different educational services on the basis of changes in growth rates or growth patterns, such as special education and reading intervention programs. Moreover, the application of these techniques could also help administrators at the school, district, and state levels determine when resources are most likely to have the greatest impact on student performance.

CHAPTER III

METHOD

Participants

Data for this study were collected from a cohort sample of 2117 students during the 2009-2010 (grade four) and 2010-2011 (grade five) school years enrolled in two school districts in the Pacific Northwest. District 1 has 1,299 students and District 2 (for cross validation) has 818 students. The districts were similar in terms of student characteristics. District 1 had almost equal proportions of male and female students. District 2 had slightly more male students (57%) than female students (43%). In both districts, the majority of the students are White, followed by Hispanic. Both districts have small proportion of students who received special education (17-20%), and an even smaller proportion (less than 5%) of students who were categorized as English language learners (see Table B1 in Appendix B for full demographics).

Procedures

Data for this study are comprised of fall, winter and spring easyCBM Multiple Choice Reading Comprehension (MCRC) measure and Passage Reading Fluency (PRF) scores for 2009-10 and 2010-11. Scores in 2009-10 were scores of students when they were in grade four, and scores in 2010-11 were scores for the same group of students when they were in grade five. The grade four measure scores were used in the conventional growth analysis and LCGA models. The grade five measure MCRC and PRF scores were used together to derive one risk rating (*low*, *medium*, or *high risk*), which in turn serve as distal outcomes in the LCGA models.

Measures

easyCBM, originally developed in 2006, currently has more than 60,000 users across 50 states. The fluency measures (e.g., Passage Reading Fluency) are administered via paper- pencil and the comprehension and vocabulary measures are administered online. easyCBM is administered at three time points during the school year (fall, winter, and spring) to identify student at risk for academic difficulties.

Multiple choice reading comprehension (MCRC). The MCRC measures are administered online. This measure requires that students first read an original, narrative passage (approximately 1,500 words) before answering 20 multiple-choice questions based on the story. Students receive credit for each question correctly answered. Each question has one question stem and three possible answers: the correct answer and two incorrect, but plausible, distractors. Alternate-form reliability coefficients during field testing of MCRC was about .59 (Alonzo & Tindal, 2009). The reported correlation between MCRC for with fourth grade state reading test was between .55-.55 and for grade 5 ranged from .53-.56 (Saez, et al., 2010).

Passage reading fluency (PRF). This measure requires that students read aloud a short, original narrative passage on a single side of a sheet of paper within 60-seconds. Examiners follow along on their own test protocol, marking as errors any words skipped or read incorrectly. If a student pauses more than three seconds on a word, the examiner supplies the word and marks it as incorrect; self- corrections are counted as correct. The passages were written to be at middle of the year reading level for each grade. The score, total words read correctly, is calculated by subtracting the number of errors from the total words read (maximum total possible = approximately 250 words).

The PRF measures were developed to be of equivalent difficulty for each grade level (see see Alonzo & Tindal (2007b) for details). Alternate-form reliability coefficients during field testing of PRF passages ranged between .87 and .97 (Alonzo, Mariano, Nese, & Tindal, 2010; Alonzo & Tindal, 2009), and test-retest reliability coefficients ranged between .91 and .97 (Alonzo & Tindal, 2009). The reported correlation between PRF with the grade four Oregon statewide reading assessment (OAKS) ranged from .64-.67 and for grade five ranged from.65-.67 (Saez, et al., 2010).

Data Analytic Plan

Structure of data. Data for this study has a hierarchical structure, in which individual measurements of the easyCBM PRF and MCRC responses are nested within students, students are nested within in classrooms, classrooms nested within in schools, and schools nested within districts. Despite the nested structure, only data at the responses and student level were analyzed. Students in District 1 on average had higher scores on both the easyCBM MCRC and PRF measures than students in District 2. In grade four, students in District 1 performed at 115.41 words correct per minute (WCPM) in fall, 136.18 WCPM in winter and 145.32 WCPM in winter on the PRF measures. On the MCRC measures, students in District 1 had average scores of 12.64 in fall, 14.36 in winter, and 14.26 in spring. Students in District 2 started slightly lower on both measures, with 100.71 WCPM in fall, 125.70 WCPM in winter, and 133.49 in spring on the PRF measures. On the MCRC measures, students in District 2 had average scores of 11.80 in fall, 13.41 in winter, and 13.49 in spring. Similar trend was observed when students entered grade five for the two districts (see Table B2 in Appendix B for full descriptives).

Data analysis approach. Piecewise growth analysis and LCGA models were used to examine growth over time. All analyses were conducted using Mplus version 5 (Muthén & Muthén, 1998-2007). Conventional growth modeling and latent class growth analysis were used to examine change within one academic year, followed by model comparisons.

Maximum Likelihood (ML) was employed as an estimation method for this study. ML minimized the discrepancies between the sample variance/covariance matrix and the model-implied estimate of the population variance/covariance matrix. Tests of normality were used to examine distribution properties, kurtosis, and multivariate normality assumption for all variables.

Purpose of this Study

This study seeks to extend the current literature on CBM screening assessments in reading for later elementary grades by exploring latent classes and reading growth profiles on the easyCBM reading comprehension and passage reading fluency measures (only three time points within a year), and examining the patterns of reading skills based on the latent class formation. The research questions proposed in this study were examined using a series of piecewise unconditional growth modeling and Latent Class Growth Analysis (LCGA) to evaluate and compare the characteristics of reading development (homogenous vs. heterogeneous growth patterns).

Piecewise growth modeling. A piecewise approach to modeling growth was used in this study because of the non-linearity of growth patterns the observed means across 2009-10. This approach to modeling non-linear growth was also used by some researchers (Christ, Silberglitt, Yeo, & Cormier, 2010; Nese, Biancarosa, Anderson, Lai,

& Tindal (2012) on data with three time points. Similar to Christ et al. (2010) and Nese et al. (2012), three time points were divided into two segments: fall-winter and winter-spring. Two time variables were coded to reflect change from fall to winter (0, 1, 1), and to reflect change from winter to spring (0, 0, 1). Some constraints were used in the analysis, including specifying zero correlations between slopes for fall-winter and winter-spring and equal error variances for the two time segments.

In the unconditional piecewise growth model (see Figure 1), the mean and the variance of intercepts (initial reading comprehension/fluency) and the slopes for the two time segments (growth trajectories) across individual students are estimated. These growth parameter estimates are continuous latent variables that capture the unobservable heterogeneity across individual differences in intercepts and slopes. The model is defined by two parts, the measurement model and the structural model. In the measurement model,

$$y_{ti} = \eta_{0i} + \eta_{1i}a_{ti} + \eta_{2i}a_{ti} + \varepsilon_{ti}\,, \tag{1}$$

y_{ti} is the observed outcome measure (e.g. easyCBM reading comprehension/fluency) of student i at time point t; η_{0i} is the intercept, η_{1i} is the slope from fall to winter for student i, η_{2i} is the slope from fall to winter; ε_{ti} is the error score for student i, $\varepsilon_{ti} \sim N(0,\sigma_{\varepsilon}^{2})$; and α_{ti} is the associated time score. When α_{0i} is the initial time point, the intercept can be interpreted as the mean score. In the structural model,

$$\eta_{0i} = \alpha_{00} + \zeta_{0i}, \tag{2}$$

$$\eta_{1i} = \alpha_{10} + \zeta_{1i}, \text{ and} \tag{3}$$

$$\eta_{2i} = \alpha_{20} + \zeta_{2i}, \tag{4}$$

where α_{00}, α_{10} and α_{20} are the mean intercept and mean slopes for fall-winter and winter-spring, and where ζ_{0i}, ζ_{1i}, and ζ_{2i} are deviations from the mean intercept and mean slopes for individual slopes and intercepts, $\begin{bmatrix}\zeta_{0i}\\ \zeta_{1i}\\ \zeta_{2i}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}\sigma_{00}^2 & & \\ \sigma_{10} & \sigma_{11}^2 & \\ \sigma_{20} & \sigma_{21} & \sigma_{22}^2\end{bmatrix}\right)$.

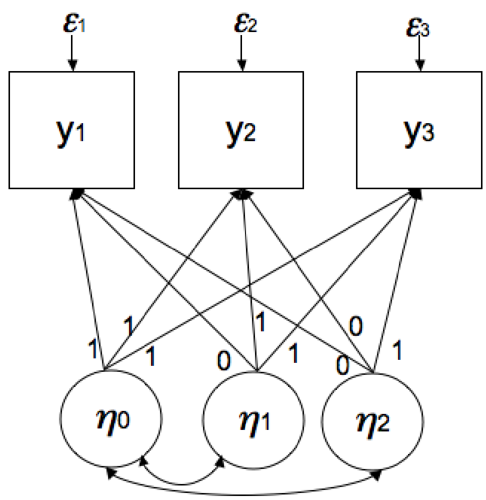


Figure 1. Unconditional piecewise growth model.

The Goodness-of-fit index χ^2, a test of exact model fit were used to evaluate the model fit. A small and insignificant χ^2 suggests good model fit because the null hypothesis assumes that the model being tested does fit the data. However, the χ^2 test is not a good indicator of model fit when the sample size is large. With large samples, small differences can yield a significant χ^2, indicating the rejection of a perfectly fitting model or a possibly good model. Considering the relatively large sample size of this study, the results of the χ^2 test were not considered to be critical in evaluating model fit for this analysis. Instead, a combination of absolute and incremental fit indices were used to evaluate the model fit, including the root-mean-square error of approximation (RMSEA), the standardized root-mean square residual, the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI). RMSEA values in the vicinity of 0.05 indicate close fit, values near 0.08 suggest fair fit, and values above 0.10 indicate poor model fit (Hu & Bentler, 1999, & Kline, 2005). Hu and Bentler (1999) recommended TLI and CFI values of .95 or better as indicators of a model fit. SRMR values of .06 through .08 (Hu & Bentler, 1999) or less than .10 are generally considered favorable (Kline, 2005).

The conditional piecewise growth model allows for the inclusion of covariates (*SPED* in this study) in an effort to explain the variation among students in intercepts and slopes (see Figure 2). Similar to the growth model without covariates, it is also defined by two parts, the measurement model and the structural model. The measurement model remained the same, but the structural model, however, is different. When *SPED* is included as a covariate, the structural model is,

$$\eta_{0i} = \alpha_{00} + \gamma_{01}(SPED_i) + \zeta_{0i}, \tag{5}$$

$$\eta_{1i} = \alpha_{10} + \gamma_{11}(SPED_i) + \zeta_{1i}, \text{ and} \tag{6}$$

$$\eta_{2i} = \alpha_{20} + \gamma_{21}(SPED_i) + \zeta_{2i}, \tag{7}$$

where γ_{01} is the coefficient of the covariate *SPED* effects (*SPED*=1, non*SPED*=0) on the intercept and γ_{11} and γ_{21} the coefficient of the effects of the covariate *SPED* on the slopes. α_{00} is the mean intercept, and α_{10} and α_{20} are the mean of fall-winter and winter-spring slopes. ζ_{0i}, ζ_{1i}, and ζ_{2i} are deviations from the mean intercept and mean slope for individual slopes and intercepts, and they have the same distribution properties of the structural model as the conditional piecewise growth model as presented in equations 2-4.

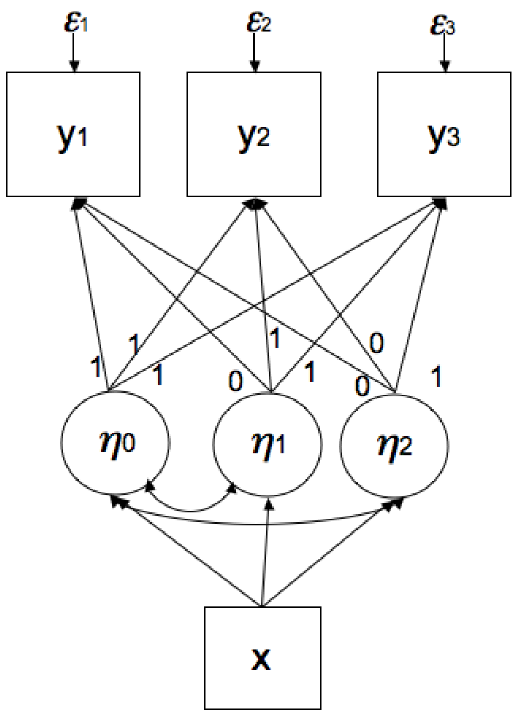
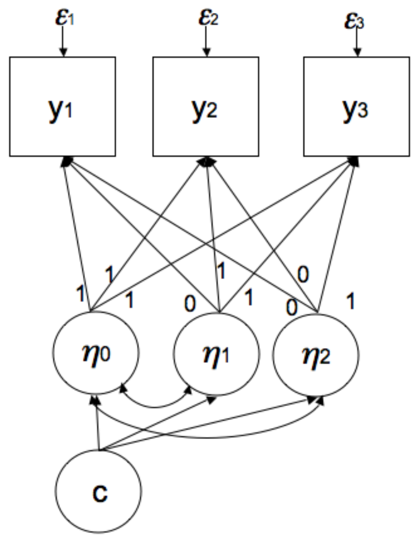


Figure 2. Conditional piecewise growth model.

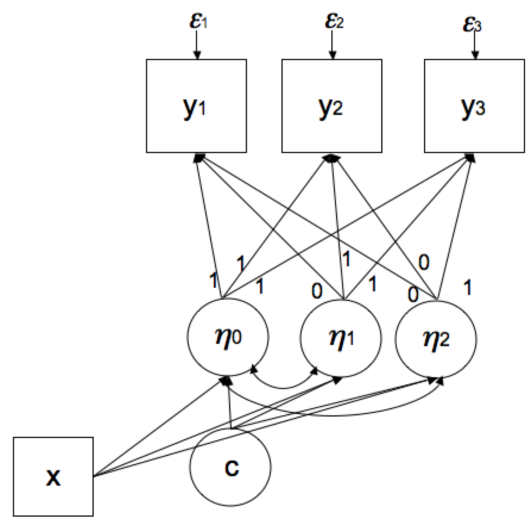
Latent class growth analysis (LCGA). LCGA was used to examine heterogeneous growth patterns in the data. LCGA is a growth modeling technique developed by Nagin (1999) and is founded on the assumptions that: (a) there are different growth trajectories for each unobservable class, and (b) that all individual growth trajectories *within* a class are homogeneous. Unlike conventional growth modeling techniques that assume individual differences in both the slope and intercept, which are estimated using random coefficients, LCGA fixes the same growth parameters to be equal within each class (i.e., all individual growth trajectories within a class are homogeneous). Such an approach is acceptable given that individual differences are captured by the multiple trajectories included in the model.

LCGA is essentially a special type of GMM, whereby the variance and covariance estimates for the growth factors within each class are invariant (i.e., all individual growth trajectories within a class are homogeneous). Some features of LCGA are appealing, including quick convergence due to the zero variance restriction and the potential to be highly practical when the model fits the data (Muthén, 2000). In this study, three models were examined for each measure: (1) unconditional piecewise LCGA model (no covariates), (2) conditional piecewise LCGA models with a covariate (*SPED*), and (3) extended LCGA models with distal outcomes (see Figures 3a-c). Distal outcomes (2 risk categories: *low or some risk* vs *high risk*) or (3 risk categories; *low*, *some* or *high risk*) were predicted by the latent classes in the extended LCGA models.

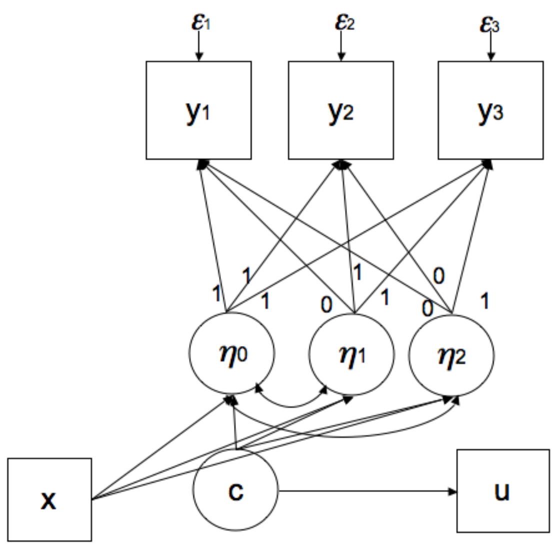
Starting the LCGA analyses with the unconditional model allows for the examination of the individual growth trajectories and class distribution (Jung & Wickrama, 2008). Because of the non-linearity growth based on the observed means, a piecewise growth approach were incorporated into the LCGA models. In the conditional LCGA model, the effects of covariates on the growth factors (intercept and slope) were examined for significance. In this study, piecewise LCGA was used to model change in grade four students who took the easyCBM PRF and MCRC measures separately.



(a)



(b)



(c)

Figure 3. (a) Unconditional, (b) conditional (covariate=*SPED*), and (c) conditional piecewise LCGA with distal outcome (F10RISK, W11RISK or S11RISK).

LCGA with distal outcome (F10RISK, W11RISK or S11RISK).

The unconditional piecewise LCGA model is also composed of a measurement model and structural model. The measurement part is,

$$y_{ti} \,|(C_i = c) = \eta_{0i} + \eta_{1i} a_{ti} + \eta_{2i} a_{ti} + \varepsilon_{ti}\,, \qquad (8),$$

where y_{ti} is the observed outcome of student i at time point t. y_{ti}, however, is conditional to the class member C_i. η_{0i} is the intercept, η_{1i} is the fall-winter slope for student i, η_{2i} is the winter-spring slope for student i, a_i is the time score, and ε_i is the error score for student i in class c, $\varepsilon_{ti} \sim N(0, \sigma_\varepsilon^2)$. The structural model is defined as

$$\eta_{0i} = \alpha_{c0}, \quad (9)$$

$$\eta_{1i} = \alpha_{c1}, \text{ and} \quad (10)$$

$$\eta_{2i} = \alpha_{c2}, \quad (11)$$

where α_{c0}, α_{c1} , and α_{c2} are the mean intercept, the means of slope for fall-winter and winter-spring within class c. There are no residual terms in equations 9 through 11 because the intercept and slopes are fixed effects within the latent classes. Thus, the structural part of the unconditional LCGA model differs from the structural model of the unconditional traditional growth because now all variances and covariances within a class are fixed at zero.

The conditional LCGA model also allows the inclusion of covariates of interest to explain the variation among students in intercepts and slopes. The conditional piecewise LCGA has the same measurement model as the unconditional LCGA (see Equations 8). When the covariate *SPED* is included to the model, the structural model is now

$$\eta_{0i} = \alpha_{c0i} + \gamma_{01}(SPED), \quad (12)$$

$$\eta_{1i} = \alpha_{c1i} + \gamma_{11}(SPED), \text{ and} \quad (13)$$

$$\eta_{2i} = \alpha_{c2i} + \gamma_{21}(SPED), \quad (14)$$

where γ_{01}, γ_{11}, and γ_{21} are the coefficient of the *SPED* effects on the intercept, the fall-winter and winter-spring slopes specific to class *c*. α_{c0i} is the mean intercept, α_{c1i} and α_{c2i} are the mean fall-winter and winter-spring slopes specific to each class *c*.

Finally, the relationship between latent class membership and easyCBM risk categories were examined by extending the unconditional/conditional LCGA models with these distal outcomes: students' risk status (l*ow*, *some*, or *high*) on easyCBM assessments in grade 5 (i.e. fall 2010, winter 2011, and spring 2011). In this model, the latent classes in the model predict the distal outcomes, which can be expressed as logistic regression with class variable *c* and a student characteristic *x* covariates can be expressed by

$$P(y_i = 1 | c_1 = k, x_i) = \frac{1}{1 + e^{-(\tau_k + \kappa_k x_i)}} \tag{15}$$

where the main effect of *c* is captured by the class-varying thresholds τ_k and κ_k is a class-varying slope for covariate *x*. As an example, students' who were at *low* and *some risk* are coded as 0, those who were at *high risk* are coded as 1.

In all of the models, the decision to select the best model fit were made using the Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), Adjusted BIC (ABIC), Lo-Mendell-Rubin Likelihood ratio test (LMR), and the Adjusted LMR likelihood ratio test. Models with lower AIC, BIC, and ABIC suggest better model fit. The LMR and Adjusted LRM likelihood ratio test compares the *k* model to a *k*-1 class model (Nylund, Asparaouhov & Muthén, 2007), with significant values indicating better fit. Entropy values were used to judge the precision of the classifications, with entropy values close to 1 as evidence of good classification (Muthén, 2004).

Although these criteria are used to determine the optimal number of latent classes, other decisive factors will also be taken into consideration, including: (a) the final stage loglikelihood values at local maxima, (b) the overall interpretability of the model based on class proportion of students in each class (no less than 1% of the total sample, (c) the estimated posterior probabilities and most likely latent class membership, and (d) the visual plots of latent classes of estimated means and observed individual trajectories will also be used to select the model with the optimal class structure.

Cross Validation Study

Once the final models for District 1 are established, conventional growth modeling and LCGA analyses were repeated using District 2 data. The stability of the latent classes and growth trajectories were examined in order to determine the generalizability of the results.

Potential Threats to Validity

Some issues could weaken the validity of this study, including (a) statistical conclusion validity, (b) internal validity, (c) construct validity, and (d) external validity (Shadish, Cook, & Campbell, 2002).

Threats to statistical conclusion validity. One potential threat to statistical conclusion validity for this study includes unreliability of measures and restriction of range. With only three benchmark measure scores (fall, winter, and spring) for only two measures (PRF and MCRC) examined in the models, it is possible that reliability could be compromised due to limited data points.

Threats to internal validity. Some potential threats to internal validity include selection bias, maturation, instrumentation, and practice effects.

Selection. Because only students who took the three easyCBM benchmark measures were selected for this study, this sample could be systematically different from a sample of students without complete data. Additionally, a considerably large number of students were excluded from the study because of missing data. Finally, the sample is composed of two districts (one for cross validation purposes) that had adopted a RTI model. This sample could be systematically different from a sample of students from another district without RTI model or drastically different model.

Maturation. One inevitable potential threat is maturation. Students experience many natural changes such as getting older, becoming more experienced, and acquiring more skills and knowledge as a result of instruction. However, the threat to maturation may be reduced because the cohort of students is from the same district, is approximately the same age, and could have similar maturational status (Murray, 1998). Furthermore, the underlying goal of this study is to examine student performance across time and recognizes that student exposure to hours of instructions is a natural part of longitudinal studies.

Instrumentation. Both the PRF and MCRC measures were designed to contain appropriate grade-level content, be sensitive to showing growth in a discrete skill area over short periods of time (minimum of 1-2 weeks of instruction), and be comparable in difficulty. The three benchmark measures used in this study, however, were not equated and thus may not be truly equivalent in difficulty.

Practice effects. The last potential threat to internal validity is testing or practice effects. Students could have the opportunity to practice or gain familiarity with measures similar to the benchmark measures, especially those who received secondary interventions (Tier 2 or 3) and were monitored more frequently with alternate forms of the measures. However, this threat could be minimal due to the fact that all the easyCBM benchmark measures are distinct for each term, and there are ten alternate progress monitoring forms available for teachers.

Threats to construct validity. The constructs of text-based (passage) reading fluency and reading comprehension are operationalized using only one fluency- and one comprehension-based measure because the data used in this study contain only responses on easyCBM benchmark measures. Consequently, this could lower the construct validity of the study because these measures only represent one method of assessing reading fluency and comprehension respectively.

Threats to external validity. This study was conducted using the responses from students in two districts that could have adopted different or similar RTI models. The generalizability of the results to other schools with different RTI models could be limited. The cross validation study using data from another school district proposed here could reduce this threat.

CHAPTER IV

RESULTS

Main Sample (District 1): MCRC

Before examining for latent classes and growth trajectories in the sample, a linear and non-linear piecewise approach to modeling growth was used because non-linear growth patterns were evident in the observed means for the 2009-2010 school year. The AIC, BIC, and ABIC fit indices from the piecewise growth showed model improvement compared to linear growth model. Additionally, CFI, TLI, RMSEA, and SRMR fit indices also supported the conclusion that the piecewise models were a better fit for the MCRC measures. When *SPED* was added as a covariate, *SPED* explained variation among students in the intercept, but not the slopes. Furthermore, all fit statistics did not indicate model improvement by including *SPED* as a covariate (see Table C1 in Appendix C for statistics).

Several models were tested to examine the latent classes and reading growth trajectories on the easyCBM reading comprehension (MCRC) measures, which relate to the first research question of this study. These models include (i) unconditional LCGA, (ii) conditional LCGA, (iii) LCGA (both conditional and unconditional) models with distal outcomes. Note that class names such as Classes 1, 2, 3, and 4 in a LCGA model reported in the following section do not always translate to Class 1 being the highest performing group, and correspondingly, Class 4 being the lowest performing group. Also, the classes in each model were introduced in the order of lowest to highest performing group. Finally, classes and growth trajectories were presented along with the 2009-10 easyCBM norm percentiles as reference (see Table B3 in Appendix B for 2009-10 norm; Tables C2 and C3 for of all final piecewise LCGA models results in Appendix C).).

Unconditional LCGA. A 4-class solution from an unconditional piecewise LCGA model was identified as the best solution based on the criteria described in Chapter III (see Figure 4 below). The entropy value for this model was not high. Class 4 (*n*=162, 38%) was the highest performing group, but showed a slight decrease in performance from fall to winter (η_1=-.57, $p < .05$) and minimal growth from winter to spring (η_2=.50, $p < .05$). Class 3 (*n*=444, 38%), the majority of students, started at slightly below the 50th percentile and displayed small and non-significant growth overall across the school year (η_1=.88, $p < .05$; η_2=-.17, $p > .05$). Class 2 (*n*=333, 29%) was the second lowest group and showed a trajectory of above the 20th percentile. Class 2 displayed the steepest growth from fall to winter in this 4-class unconditional piecewise LCGA model (η_1= 3.42, $p < .05$), but almost no growth from winter to spring (η_2=.05, $p < .05$). Class 1 (*n*=219, 19%) was the lowest performing group, with their trajectory slightly above the 10th percentile throughout the year. Class 1 displayed steeper growth from fall to winter (η_1= 2.32, $p < .05$) and negative, non-significant growth from winter to spring (η_2=-.36, $p > .05$).

Conditional LCGA. In addition to examining unconditional LCGA models, the inclusion of special education status (*SPED*; *n*=226, 17%) to the model was examined to determine the effect of the covariate on the growth trajectories and latent class formation. Unlike the unconditional model, a 2-class solution was found to be the best model (see Figure 5 below). Based on the results of this model, there were two groups of students.

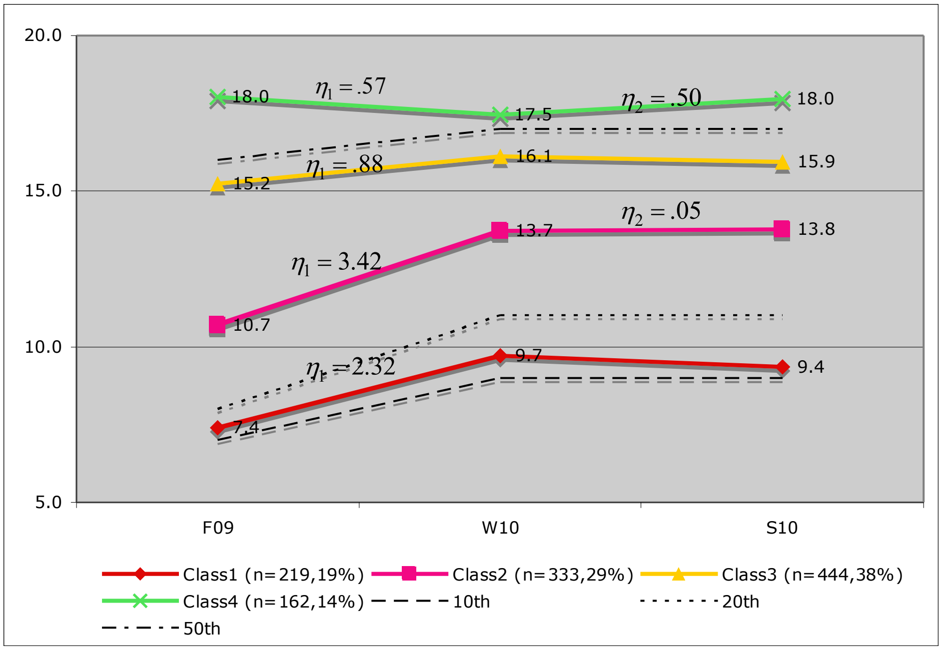


Figure 4. Unconditional piecewise LCGA model based on MCRC growth for District 1.

Both classes displayed non-significant growth from winter to spring. Class 2 (n=613, 53%) displayed a trajectory below the 50th percentile, with minimal growth fall-winter and winter-spring terms (η_1=.55, $p < .05$; η_2=-.09, $p > .05$). The other group of students (Class 1; n=545, 47%) displayed a trajectory similar to the 20th percentile, with steeper growth from fall to winter (η_1=3.10, $p < .05$) than Class 1, and non-significant negative growth from winter to spring (η_2=-.35, $p > .05$). Although two latent class trajectories were found in the conditional piecewise LCGA model, the inclusion of *SPED* as a covariate was only a significant predictor of the intercepts on both groups, indicating that students' initial fall MCRC scores can be explained by whether a student receives special education instruction (i.e. *SPED* effect). However, *SPED* it did not significantly predict the slopes for the Class 2 (γ_{11}= -.48, $p > .05$; γ_{21}=.64, $p > .05$) and Class 1 (γ_{11}=-.45, $p >$ .05; γ_{21}=.62, $p > .05$), suggesting that the fall-winter and winter-spring growth rates between students who received special education instruction and students who received general education instruction were similar.

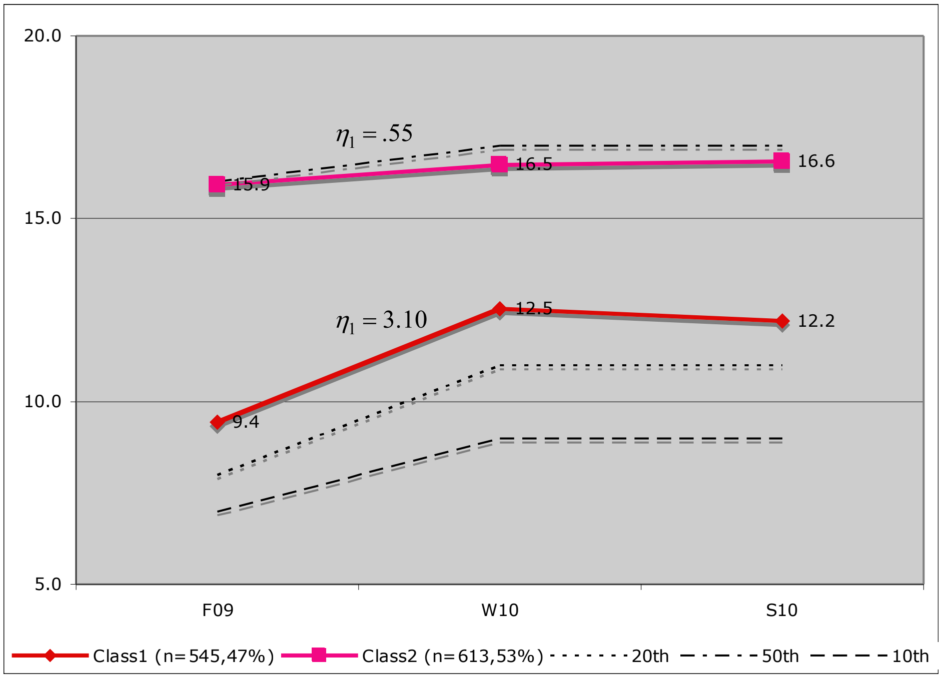


Figure 5. Conditional piecewise LCGA model based on MCRC growth for District 1.

LCGA with distal outcomes. To determine whether class membership was related to the risk status of students in grade five, four extended piecewise LCGA models with distal outcomes were conducted using easyCBM risk levels. In this study, the risk status were computed based on the two-tests risk ratings in this study (see Figure 6). Students with MCRC and PRF scores in the 0^{th}-10^{th} and 11^{th}-20^{th} percentile were categorized as *high risk* students. The distal outcomes of interest include risk status in fall 2010 (F0RISK), winter 2011 (W11RISK), and spring 2011 (S11RISK). Distal outcomes consisted of three ordinal categories: *low*, *some*, and *high risk* and were coded 0, 1, and 2 accordingly.

In a conditional model with F10RISK as the distal outcome, I examined whether the *SPED* covariate would still contribute to the prediction of the growth parameters, the latent classes, and the risk level in the fall of grade five. Then, three different unconditional models with distal outcomes F10RISK, W11RISK, and S11RISK (see Figure 7) were examined to determine if the unconditional piecewise LCGA models were sufficient in predicting students' risk level in grade five.

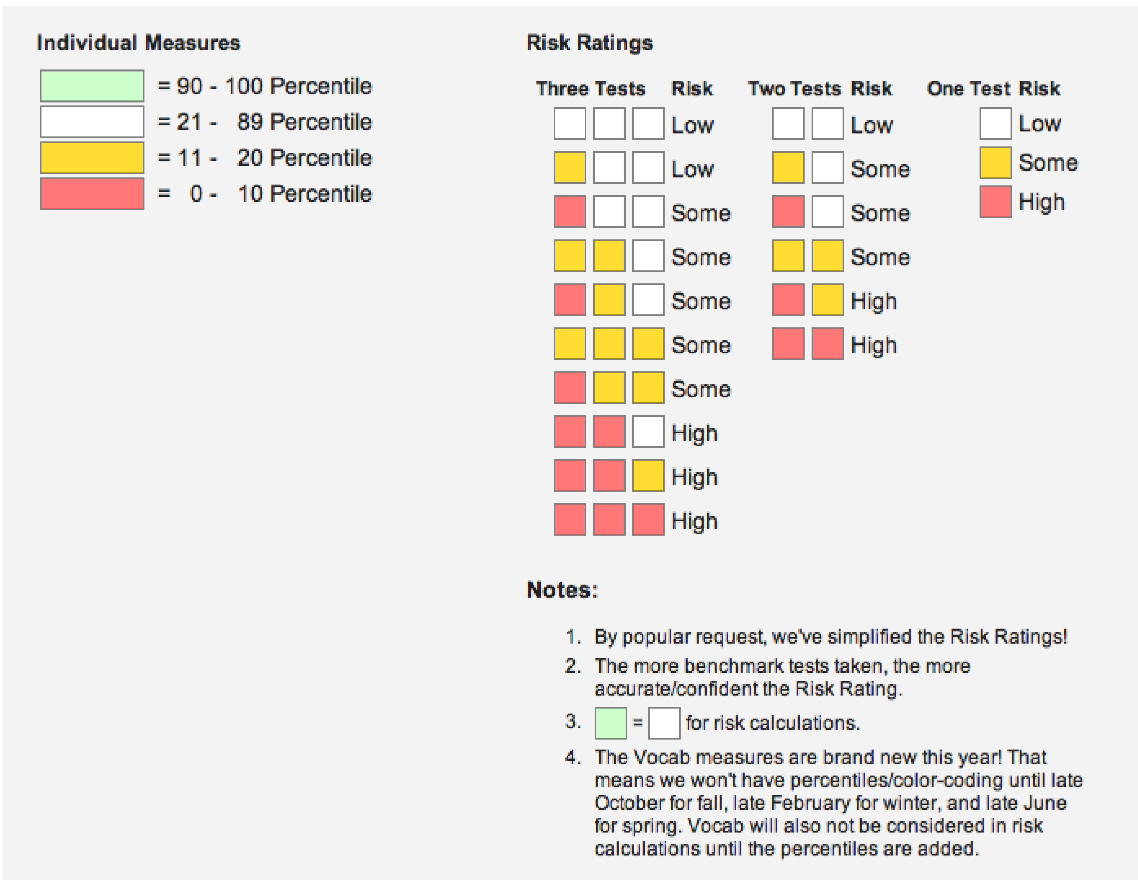


Figure 6. easyCBM risk rating system.

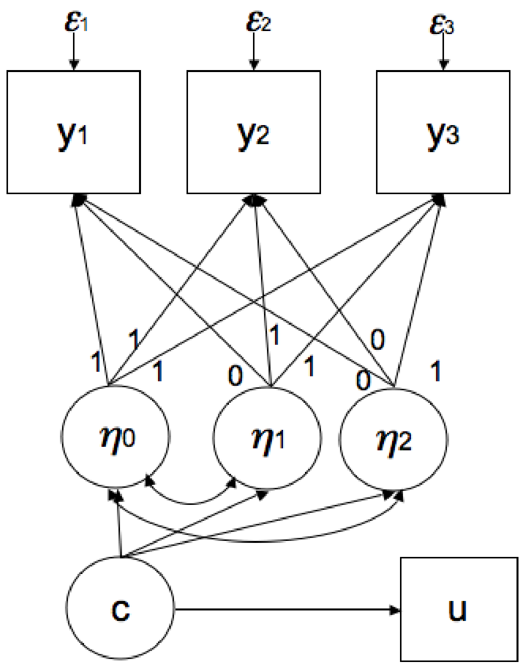


Figure 7. Unconditional piecewise LCGA with distal outcome (F10RISK, W11RISK, or S11RISK).

In the first extended conditional piecewise LCGA model, *SPED* was added as the covariate and risk status in the fall of grade five (F10RISK) was used as the distal outcome. A 3-class solution was found to have the best model fit (see Figure 8). Similar to the conditional model without a distal outcome, all classes displayed non-significant growth from winter to spring. Class 3 (n=519, 45%) displayed minimal growth over the year, and displayed a trajectory that was similar to the 50th percentile (η_1=.32, $p < .05$; η_2=.13, $p > .05$). In Class 2 (n=411, 35%), students had a trajectory between the 20th and 50th percentile, and demonstrated positive growth from fall to winter (η_1=1.93, $p < .05$) but then negative growth from winter to spring (η_2=-.16, $p > .05$). Students in Class 1 started with the lowest fall scores (n=237, 20%) and displayed negative growth trajectories (η_1=2.68, $p < .05$; η_2=-.59, $p > .05$). Similar to the 2-class conditional piecewise model, the effects of *SPED* on the observed differences in growth rates from fall to winter or winter to spring were not statistically significant: Class 3 (γ_{11}= .19, $p >$ .05; γ_{21}=-.77, $p > .05$), Class 2 (γ_{11}=-.16, $p > .05$; γ_{21}=.47, $p > .05$), and Class 1 (γ_{11}=-.28, $p > .05$; γ_{21}=.75, $p > .05$). The effects of *SPED* on the intercepts for all three groups were also not statistically significant.

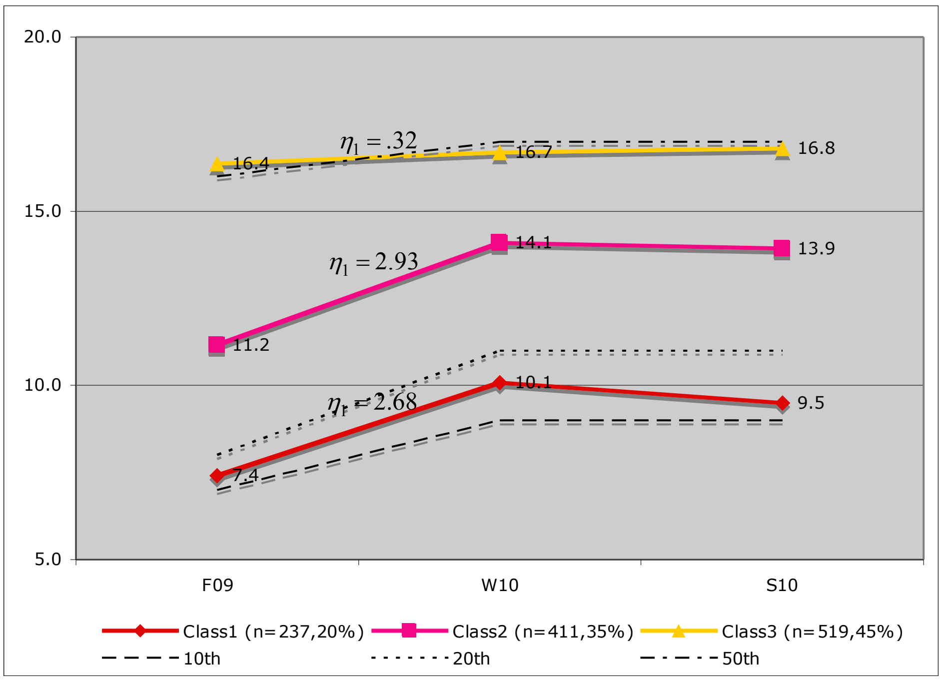


Figure 8. Conditional piecewise LCGA model Distal Outcome (F10RISK) based on MCRC growth for District 1.

Three unconditional piecewise LCGA models with F10RISK, W11RISK and S11RISK as distal outcomes were examined to determine if the model without covariates could sufficiently predict students' growth trajectories and their risk status in fall, winter, and spring of grade five. Similar to the unconditional piecewise LCGA results, the 4-class solution was the best model fit for the extended model with F10RISK as the distal model (see Figure 9). The proportion of students in each class, along with the trajectories, were similar to the unconditional piecewise LCGA results as well. Class 4 (n=161, 15%)

consisted of students who performed above the 50th percentile across the year and demonstrated negative growth from the fall to the winter (η_1=-.56, p < .05) and non-significant negative growth from the winter to the spring (η_2=-.48, p > .05). In Class 3, students performed close to the 50th percentile range (n=439, 38%), with growth from fall to winter (η_1=.90, p < .05), and non-significant negative growth from winter to spring (η_2=-.12, p > .05). Class 3 performed between the 20th and 50th percentile (n=352, 30%) showed the steepest growth in this 4-class model during fall to winter (η_1=3.35, p < .05) and then negative growth (η_2=-.06, p < .05) from winter to spring. Class 1 (n=215, 18%) was the lowest performing group with trajectories close to the 10th percentile, showed steep, positive growth from fall to winter (η_1=2.31, p < .05), followed by non-significant negative growth from winter to spring (η_2=-.29, p > .05).

For the unconditional piecewise LCGA models with W11RISK and S11RISK as the distal outcomes (see Figures 10 and 11), the 4-class solution was also the best fit. However, the entropy of the model was not high. The four latent classes and their trajectories, as well as the proportions of students in each class, were similar with W11 RISK and S11RISK as the distal outcomes to those obtained from the unconditional piecewise LCGA model with F10RISK as the distal outcome. This suggests that the four latent classes and their trajectories are stable across the year. In general, all classes displayed positive growth from fall to winter, except the highest performers. The magnitude of the positive growth ranged from .80 to 3.52, with the two lowest

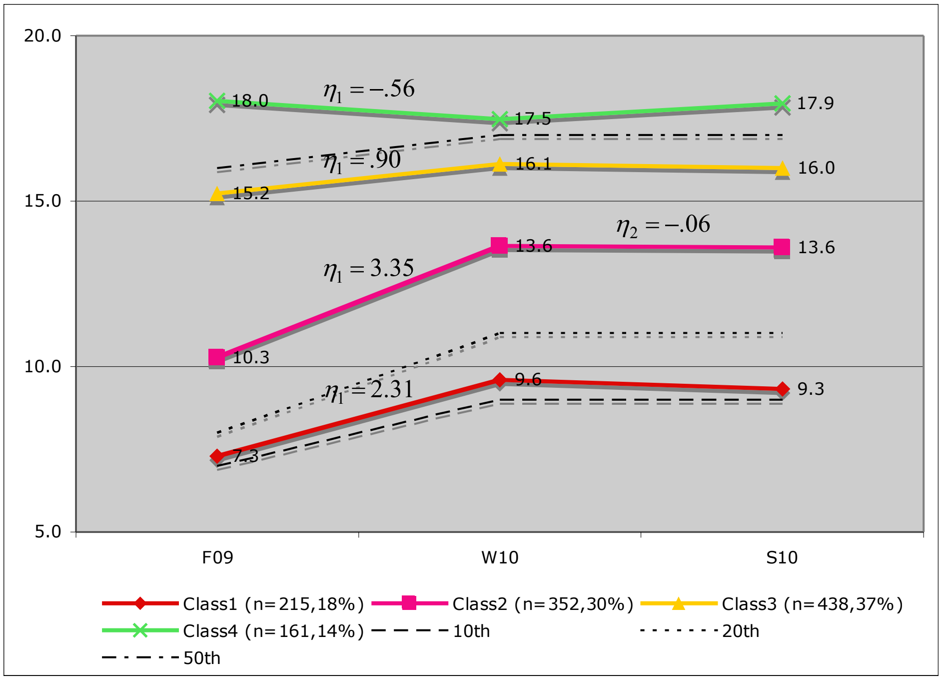


Figure 9. Unconditional Piecewise LCGA with distal outcome (F10RISK) based on MCRC growth for District 1.

performing groups displaying the most growth. Minimal growth from winter to spring was observed for all classes, with the majority of the classes demonstrating non-significant growth.

Alignment of latent classes with easyCBM risk ratings. To examine whether the four classes from the unconditional piecewise LCGA models with distal outcomes aligned closely with the easyCBM risk rating system, I examined the frequencies of students within each class identified as having *high*, *some*, or *low risk* (as defined by easyCBM two-tests Risk Ratings). In all of these models, the proportion of *high*, *some*,

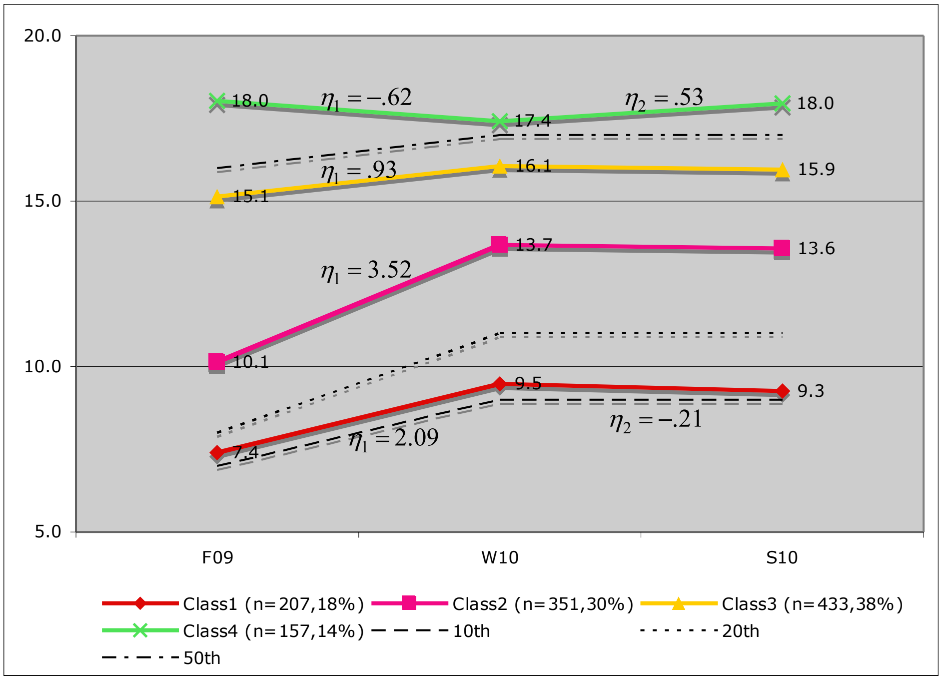


Figure 10. Unconditional Piecewise LCGA with distal outcome (W11RISK) based on MCRC growth for District 1.

and *low* risk students in each four classes were quite similar (see Figures 12-14). With F10RISK as the distal outcome, two classes (Class 1 and Class 4) had the majority of students with *low risk*. Class 3 comprised 69.89% of *low risk* students, 29.26% of *some risk*, and .85% of *high risk* students. Class 2 had the majority of the *high risk* students (49.30%), a moderate number of students who were categorized as being at *some risk* (41.86%), and few students (8.84%) considered to be at *low risk*. As described earlier, students in Class 2 were those with the lowest growth trajectory. Of the 111 *high risk* students in the sample, 106 (95.5%) of the students categorized as being at *high risk* were captured in Class 2. Similar results were obtained when W11RISK and S11RISK

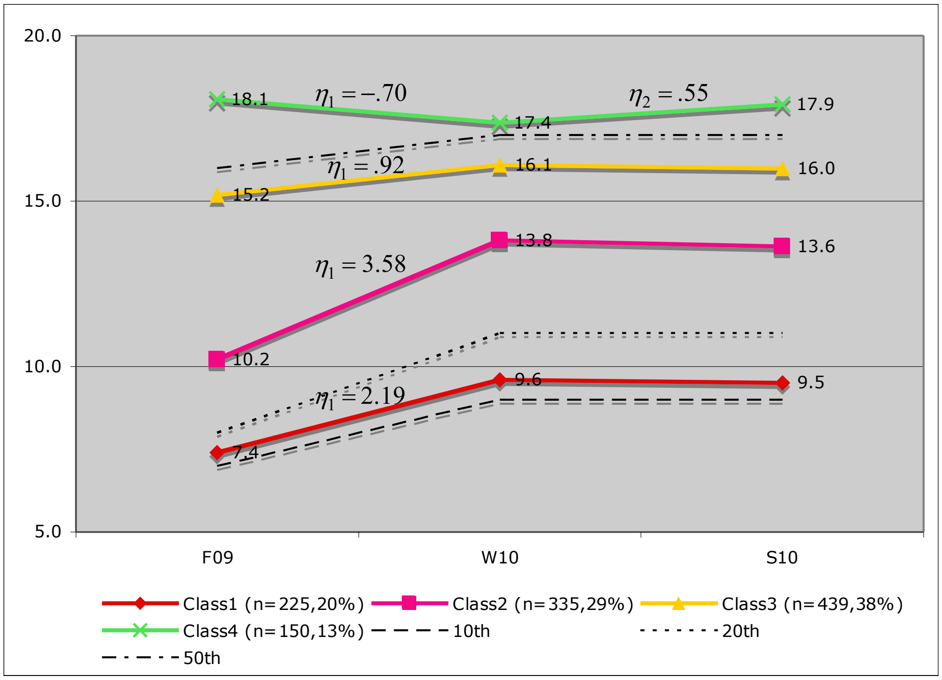


Figure 11. Unconditional Piecewise LCGA with distal outcome (S11RISK) based on MCRC growth for District 1.

were the distal outcomes, where the class with the lowest growth trajectory predicted the majority (82.35% and 80.83%, respectively) of the *high risk* students in the sample. These results seemed to suggest that the latent classes and MCRC growth trajectories from the unconditional piecewise LCGA with F10RISK, W11RISK and S11RISK as distal outcomes aligned closely with the easyCBM two-test risk ratings and could provide useful information to teachers and school administrators, particularly with the majority of the *high risk* students who were successfully predicted in the following year.

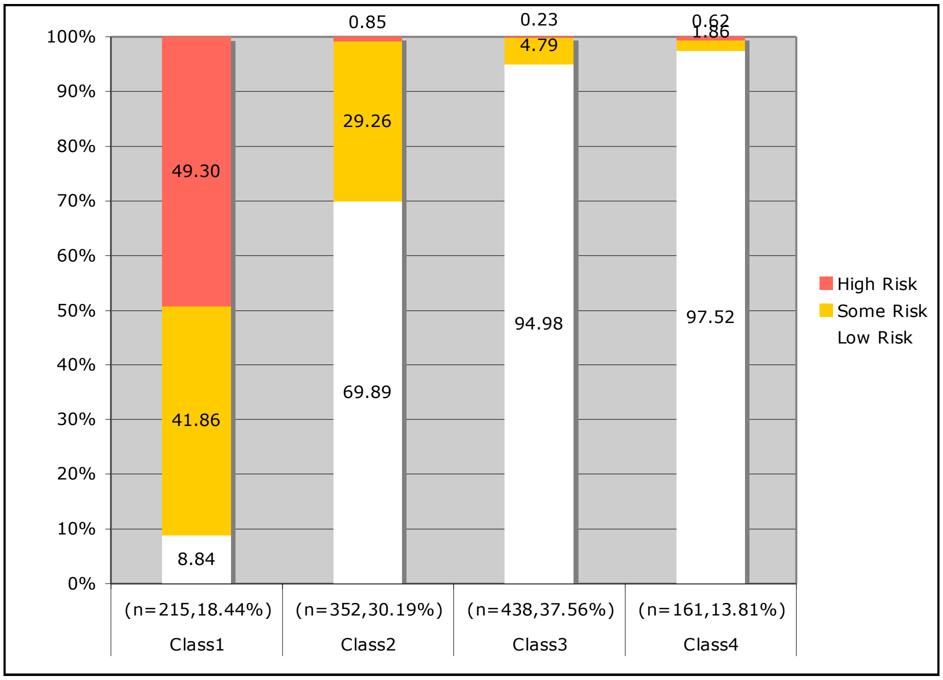


Figure 12. Percentage of students in each of the four latent classes identified based on MCRC unconditional piecewise LCGA with F10RISK as distal outcome results for District 1.

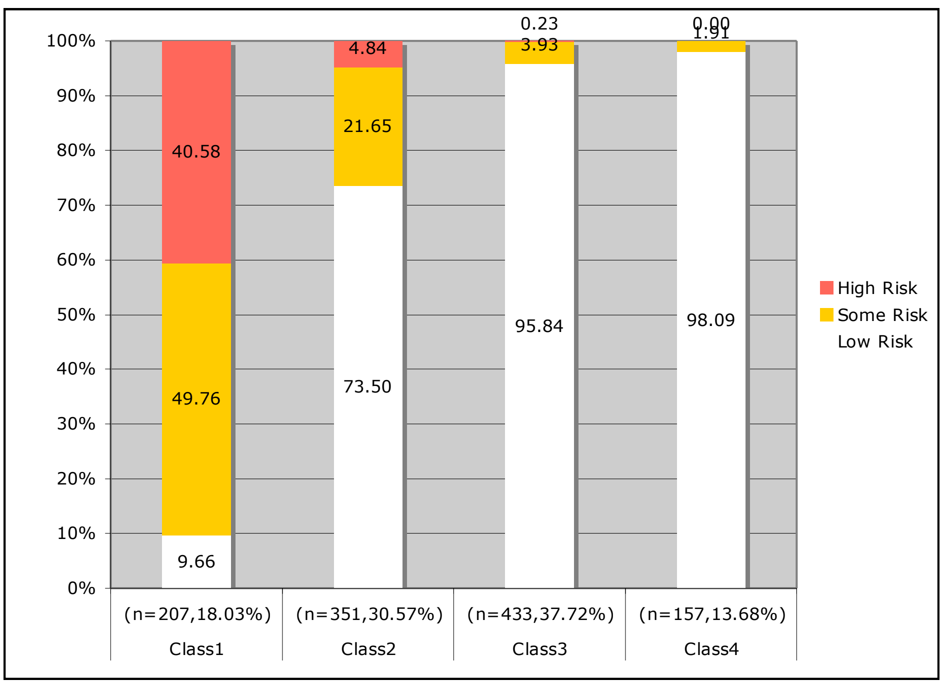


Figure 13. Percentage of students in each of the four latent classes identified based on MCRC unconditional piecewise LCGA with W11RISK as distal outcome results for District 1.

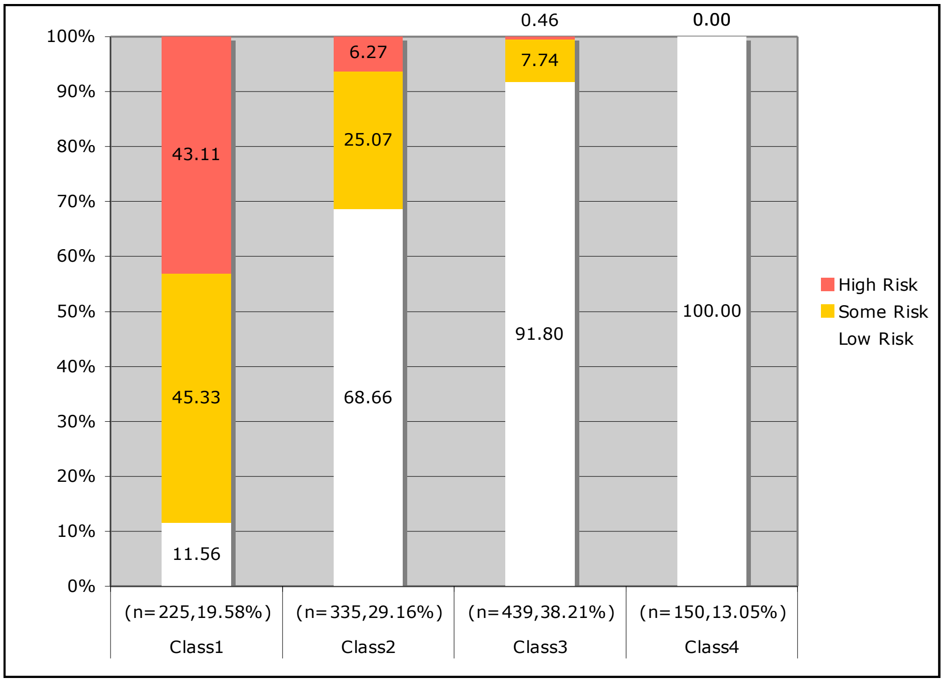


Figure 14. Percentage of students in each of the four latent classes identified based on MCRC unconditional piecewise LCGA with S11RISK as distal outcome results for District 1.

Main Sample (District 1): PRF

The same series of models used to examine the latent classes and reading growth trajectories on the easyCBM MCRC measures were used with the passage reading fluency (PRF) measures as well. Again, non-linear growth patterns were examined. Piecewise growth models were supported by fit indices as having better model fit than the non-linear growth models. Again, several LCGA models were examined, including unconditional, conditional, and models with distal outcomes (see Tables C4 and C5 for of all final piecewise LCGA models results in Appendix C).

Unconditional LCGA. A 7-class solution was the best solution for the unconditional piecewise LCGA results modeling growth in PRF (see Figure 15). All classes demonstrated significant positive growth from the fall to winter. All but Class 2 displayed significant, but smaller growth from winter to spring. Classes 3, 4, and 5 displayed the steepest growth, ranging from 22-25 words correct per minute (WCPM) while other classes (the highest two performing classes and the two lowest performing classes) displayed smaller positive growth (14.81-17.80 WCPM). Much smaller growth rates in WCPM were observed from winter to spring, however, particularly for the bottom four classes (Classes 4, 3, 2, and 1); growth rates of 3-9 WCPM were observed in the bottom four classes compared to growth rates of 14–17 WCPM for the highest three classes. Class 1 had trajectories lower than the 10th percentile, and Classes 2 and 3 had trajectories between the 10th and 50th percentile.

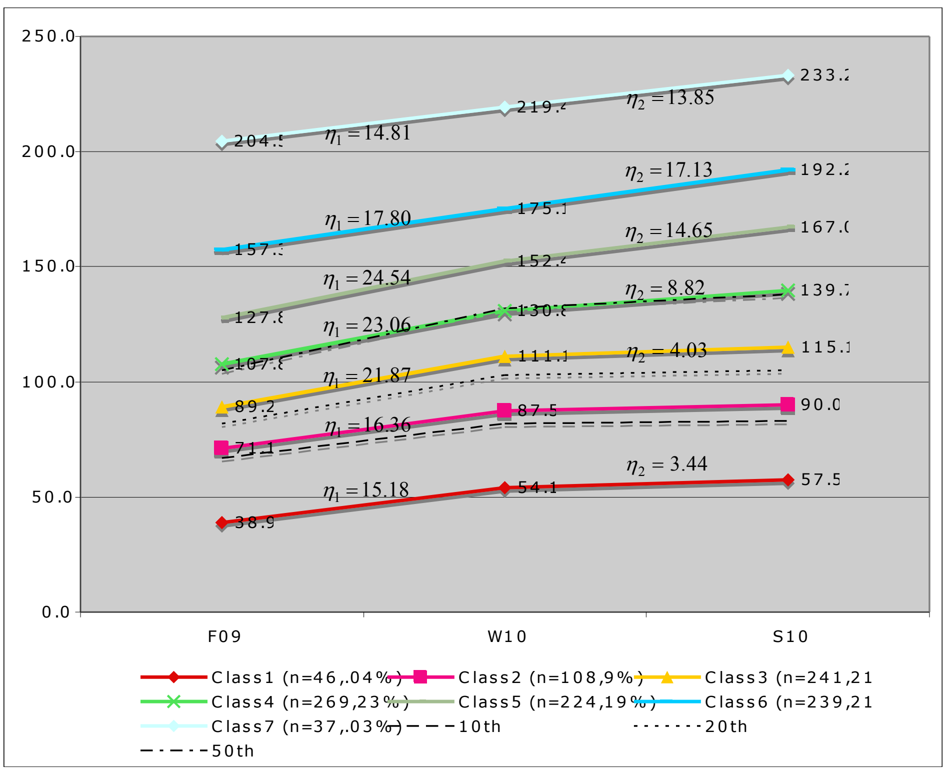


Figure 15. Unconditional piecewise LCGA model based on PRF growth for District 1.

Conditional LCGA. When *SPED* was added to the model, the 3-class solution was found to have the best model fit (see Figure 16). Three groups of students were observed in this model: high (*n*=456, 39%), medium (*n*=367, 32%), and low (*n*=341, 29%) performers. These groups displayed greater growth from the fall to winter than from the winter to spring in general. The fall to winter growth for the three classes ranged from 18-24 WCPM, and winter to spring growth ranged from 4-16 WCPM. The low performers displayed the least growth from winter to spring (η_2=3.57, $p < .05$) and the high performers displayed the highest growth (η_2=16.32, $p < .05$). *SPED* however was only a significant and consistent predictor of the intercepts for all three classes. No consistent pattern of statistically significant differences as observed in student's growth rates from fall to winter or from winter to spring across the three classes.

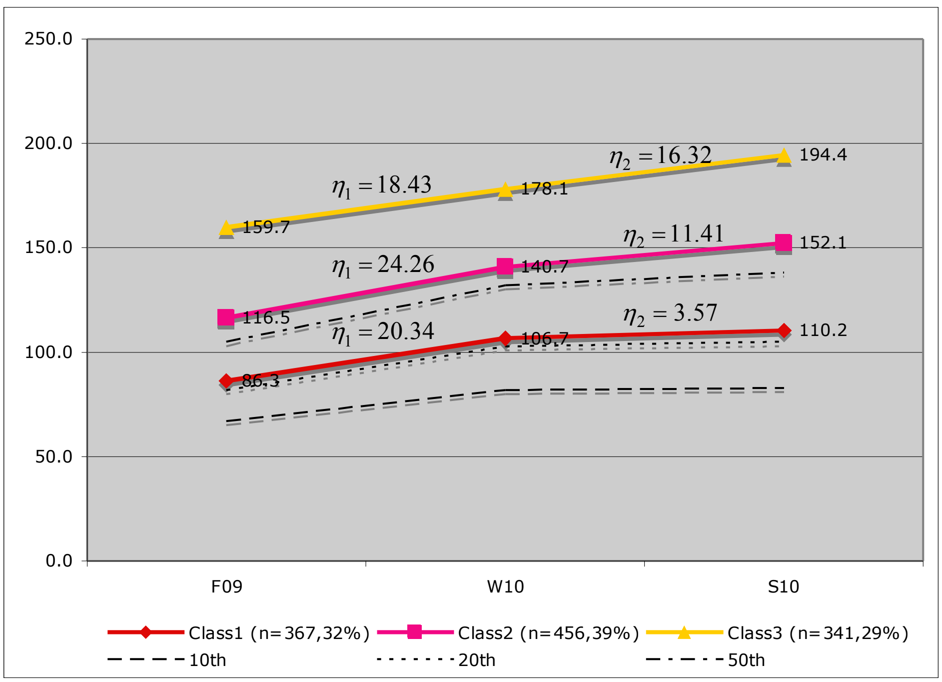


Figure 16. Conditional piecewise LCGA model based on PRF growth for District 1.

LCGA with distal outcomes. With *SPED* as a covariate and F10RISK as the distal outcome, the 3-class solution remained the best model fit (see Figure 17). Again, three groups of students were observed: high (*n*=232, 18%), medium (*n*=525, 40%), and low (*n*=542, 42%) performers. Similar to the 3-class conditional piecewise model, *SPED* did not significantly explain the growth factors consistently from the fall to winter or winter to spring for all three classes.

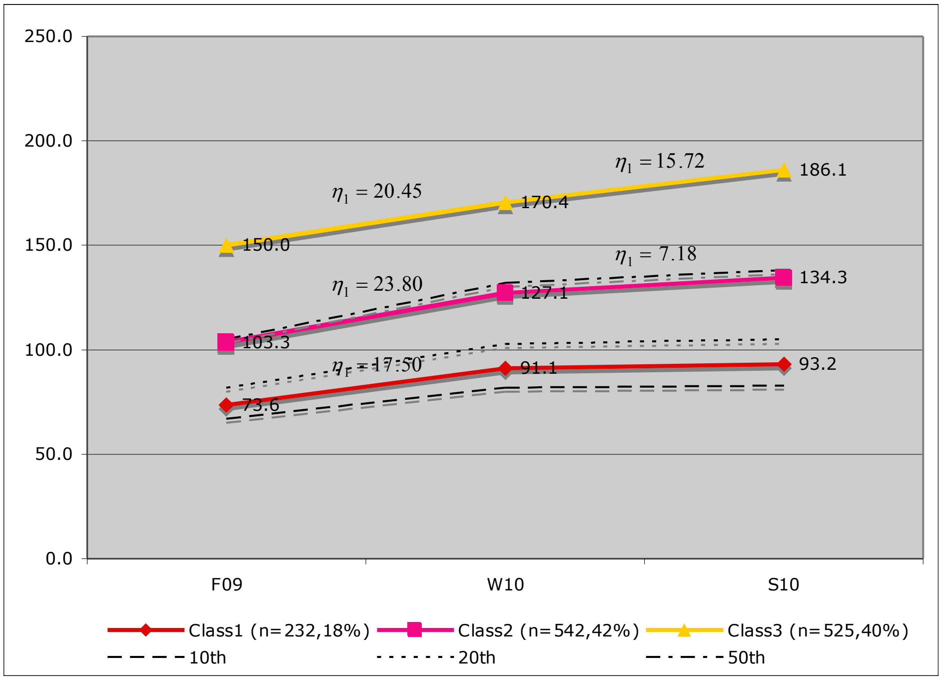


Figure 17. Conditional piecewise LCGA model with distal outcome (F10RISK) based on PRF growth for District 1.

A 4-class solution had the best model fit for the extended unconditional model with F10RISK as the distal model (see Figure 18). The highest performing students (Class 4; *n*=275, 24%) displayed relatively consistent rates of growth across the year (η_1=17.33, $p < .05$; η_2=17. 12, $p < .05$). The second highest performing group (Class 3, *n*=342, 29%) displayed trajectories above the 50th percentile. Class 3 displayed the steepest fall-winter growth in this 4-class model, (η_1=24.58, $p < .05$), and then almost half the growth from winter to spring (η_2=12.56, $p < .05$). Class 2 (*n*=414, 36%) were the second lowest performing group that displayed trajectories above the 20th percentile. This group of students displayed greater growth from fall to winter and less growth from winter to spring (η_1=22.23, $p < .05$; η_2=5.46, $p < .05$). Class 1 (*n*=135, 12%) was the lowest performing group and displayed a trajectory below the 10th percentile. The growth rates from fall to winter and winter to spring for this class were the smallest compared to the other three classes (η_1=16.22, $p < .05$; η_2=3.3, $p < .05$). All four classes demonstrated positive growth from the fall to winter and winter to spring. Fall-winter growth rates ranged from 16 to 24 WCPM and winter-spring growth rates ranged from 3 to 17 WCPM. Only the highest performing students had similar rates of growth from fall to winter and winter to spring. The two lower-performing groups, in contrast, displayed much lower growth rates of only 3-5 WCPM from the winter to spring.

With W11RISK and S11RISK as the distal outcomes, the 4-class solution also had the best model fit (see Figures 19 and 20), with each class displaying similar trajectories. The only exception was the winter to spring slope for the second highest performing group in the model with W11RISK as the distal outcome; a much smaller growth rate of only 5-6 WCPM was observed in this model compared to a growth rate of 12-13 WCPM observed in other two models with F10RISK and S11RISK as distal outcomes. Overall, the highest and lowest performing groups across the three models displayed less growth (15-18 WCPM) from fall to winter compared to the second highest and lowest performing groups (21–25 WCPM). In general, smaller rates of growth were observed from winter to spring compared to fall to winter growth, with the lowest performing group displaying minimal growth of only 2–3 WCPM.

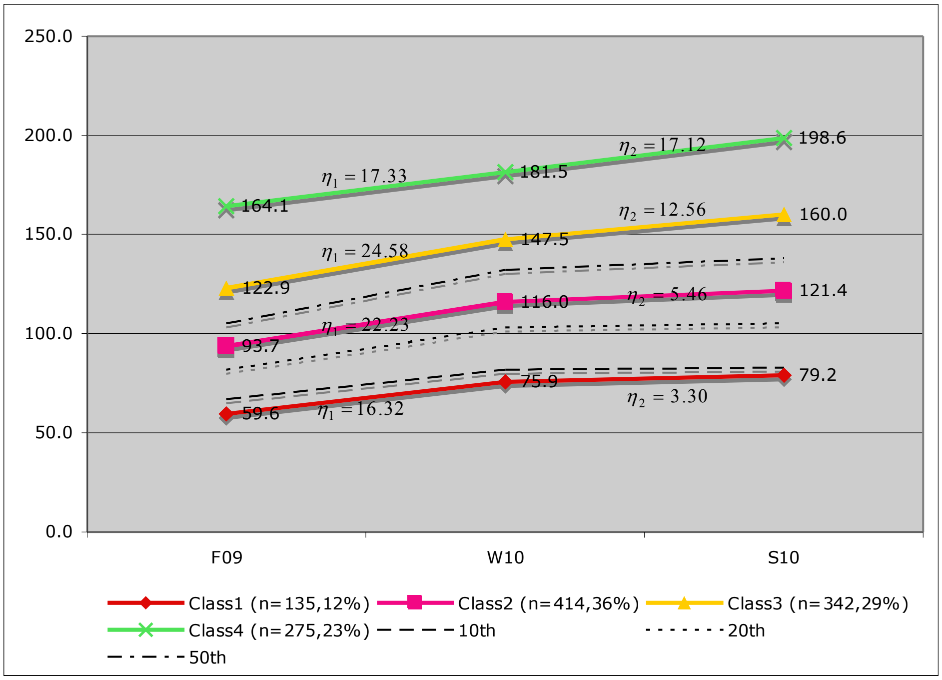


Figure 18. Unconditional Piecewise LCGA with distal outcome (F10RISK) based on PRF growth for District 1.

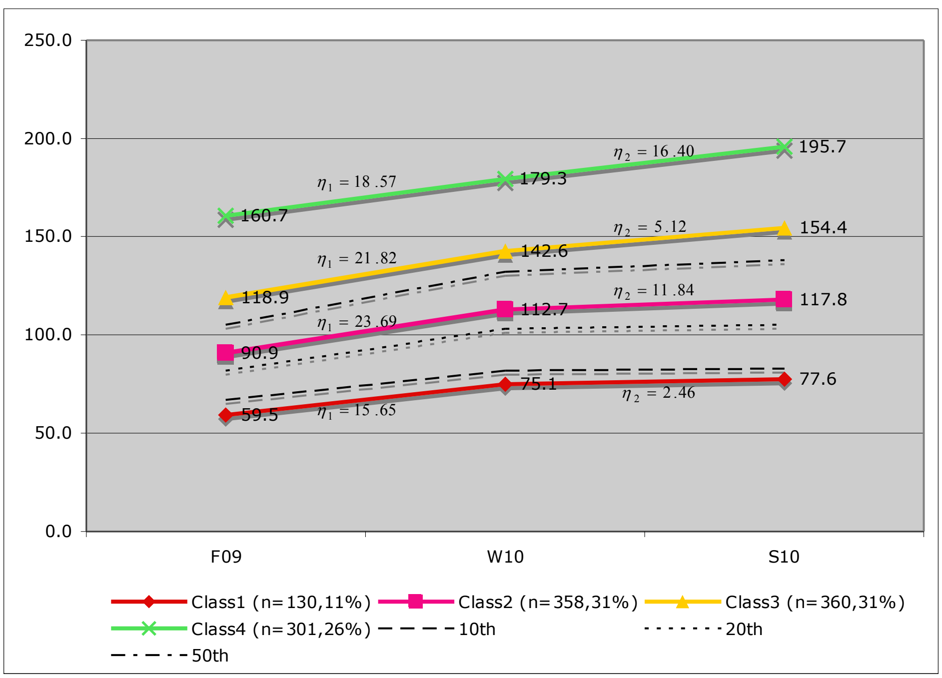


Figure 19. Unconditional Piecewise LCGA with distal outcome (W11RISK) based on PRF growth for District 1.

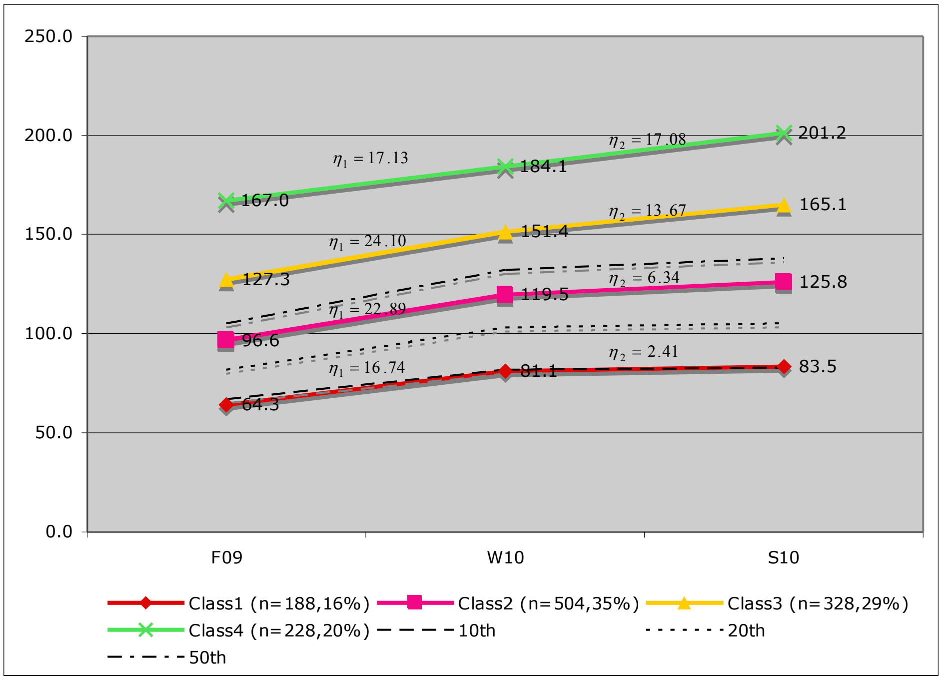


Figure 20. Unconditional Piecewise LCGA with distal outcome (S11RISK) based on PRF growth for District 1.

Alignment of latent classes with easyCBM risk ratings. The frequencies of students within each class identified as having *high*, *some*, or *low risk* were examined to determine the alignment between the latent classes and the risk ratings from the easyCBM system. In all three models with the three distal outcomes, the class with the lowest trajectory consistently contained the majority of the *high-risk* students (see Figures 21-23). The proportion of the *high risk* students in each class, however, gradually decreased from fall to spring (70%, 56%, and 20%), indicating that, over the year, these fifth grade students became more fluent readers of connected text passages. Across all three models, the class with the lowest trajectory captured majority of the *high-risk* students in the fall (n=110, 95%), winter (n=74, 73%), and spring (n=103, 86%). The results of these models suggested that the four latent classes and PRF growth trajectories had close alignment to the easyCBM risk ratings. Similar to the unconditional models with the three distal outcomes based on MCRC measures, two classes consistently predicted the *low risk* students in the models with distal outcomes using PRF measures (F10RISK distal outcome: Classes 2 and 3; W11RISK distal outcome: Classes 1 and 3; W11RISK distal outcome: Classes 1 and 4).

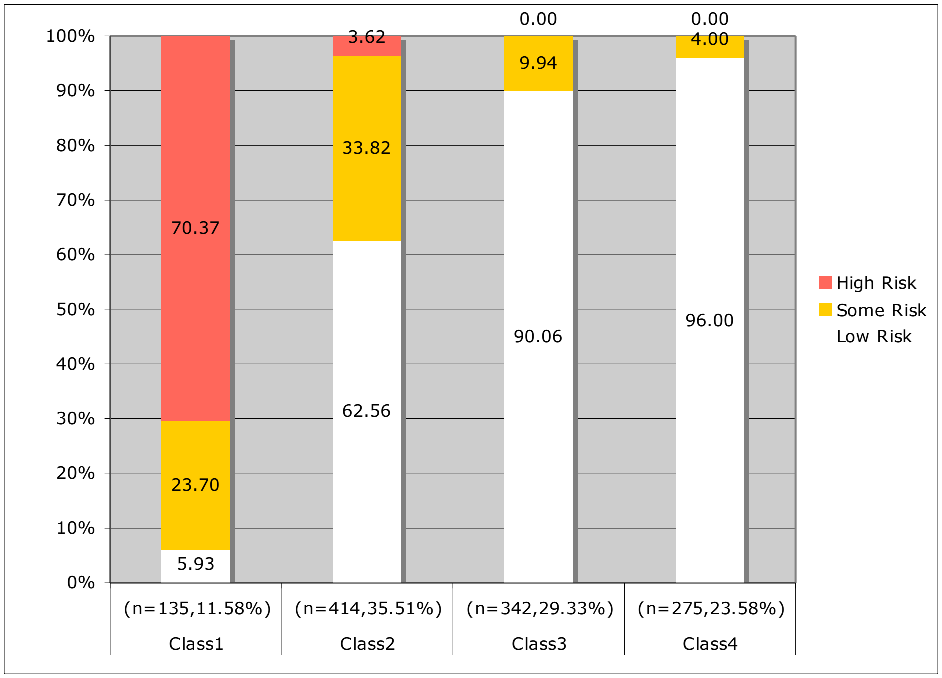


Figure 21. Percentage of students in each of the latent classes identified based on PRF unconditional piecewise LCGA with F10RISK as distal outcome results for District 1.

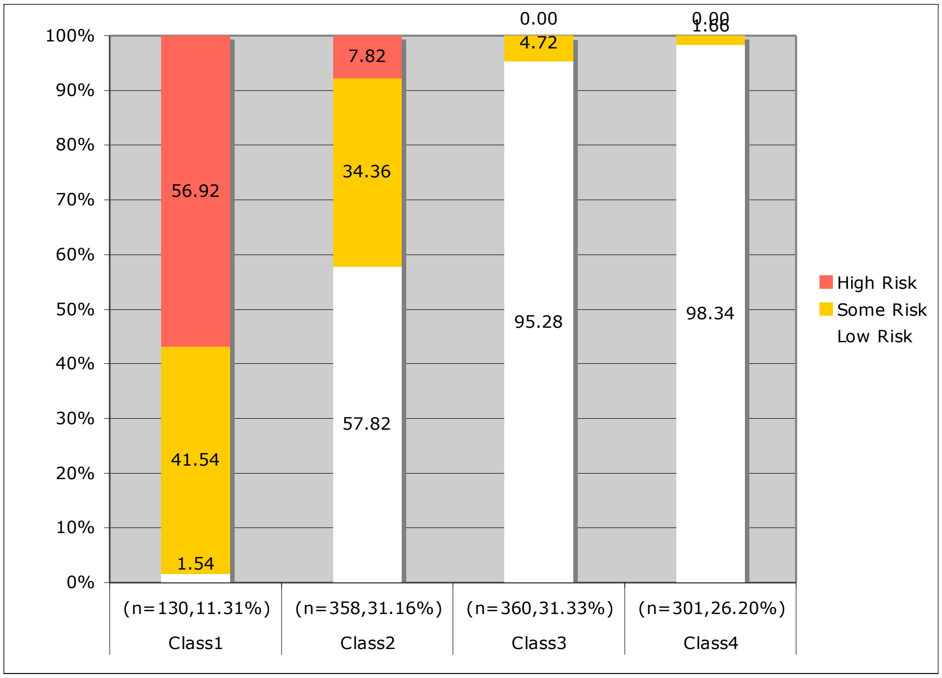


Figure 22. Percentage of students in each of the latent classes identified based on PRF unconditional piecewise LCGA with W11RISK as distal outcome results for District 1.

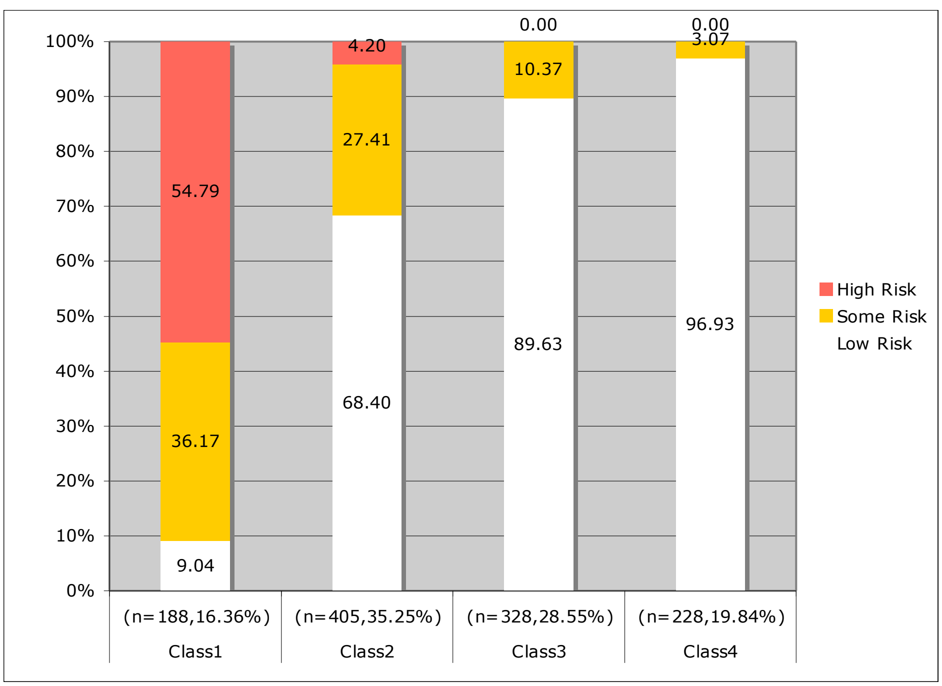


Figure 23. Percentage of students in each of the latent classes identified based on PRF unconditional piecewise LCGA with S11RISK as distal outcome results for District 1.

Cross Validation Sample (District 2): MCRC

Unconditional, conditional, and LCGA models with distal outcomes (F10RISK, W11RISK, and S11RISK) were examined (see Tables D1 and D2 in Appendix D for full model results).

Unconditional LCGA. Similar to District 1, a 4-class solution was identified as the best solution based on the unconditional piecewise LCGA results (see Figure 24). Class 4 (n=84, 12%) was the highest performing group but showed non-significant negative growth in over the year (η_1=.01, p > .05; η_2=-.09, p > .05). Class 3 (n=265, 37%) was the second highest performing group and contained the majority of students, but displayed only small growth from fall to winter (η_1=.99, p < .05) and even less growth from winter to spring (η_2=.65, p < .05). Class 2 (n=176, 26%) was the second lowest performing group and demonstrated steepest growth (compared to all other classes) from fall to winter in this 4-class model (η_1=3.24, p < .05), and non-significant minimal growth from winter to spring (η_2=.22, p > .05). Class 1 (n=189, 25%) was the lowest performing group and displayed steeper growth from fall to winter (η_1=1.45, p < .05) than from winter to spring (η_2=-.76, p > .05).

Conditional LCGA. Results from the conditional piecewise LCGA results for District 2 were very similar to the results obtained from data for District 1 (see Figure 25). Class 2 (n=387, 54%) displayed more growth from fall to winter (η_1=1.01, p < .05) and less growth from winter to spring (η_2=.42, p < .05). Class 1 (n=327, 46%) had steeper growth from fall to winter compared to Class 2 (η_1=2.44, p < .05), but non-significant negative growth from winter to spring (η_2=-.51, p > .05). Similar to District 1 results, *SPED* was a significant predictor of the intercept for Class 1, but it did

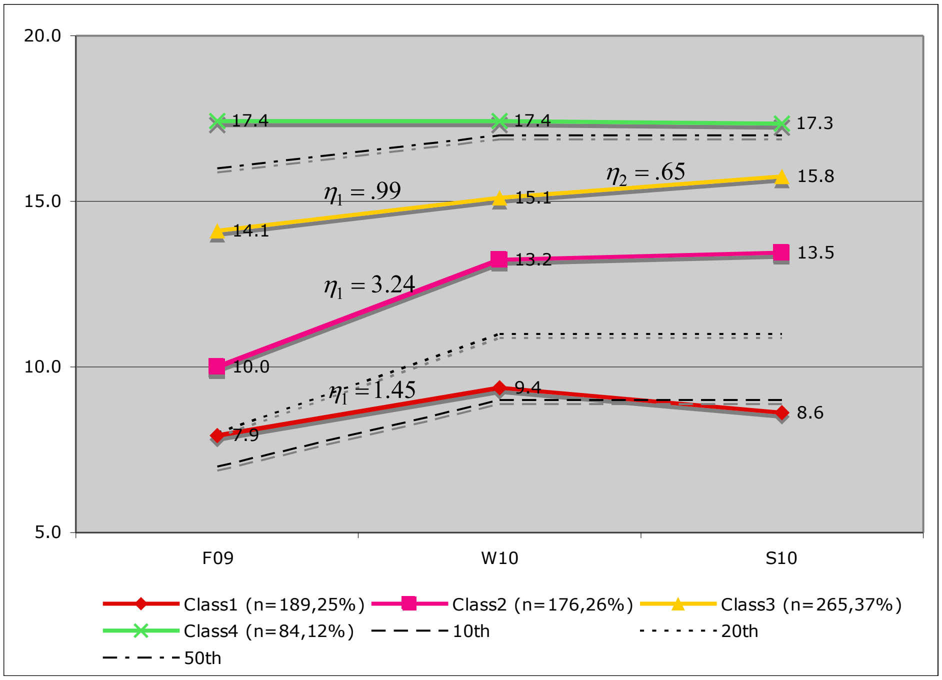


Figure 24. Unconditional piecewise LCGA based on MCRC growth for District 2.

not significantly predict the intercept, the slopes from fall to winter or from winter to spring of Class 2 (γ_{11}= -.19, p > .05; γ_{21}=.47, p > .05) and Class 1 (γ_{11}=-.50, p > .05; γ_{21}=.53, p > .05).

LCGA with distal outcomes. With *SPED* as a covariate in this extended model with F1RISK as a distal outcome, the 2-class solution had the best model fit (see Figure 26). Class 2 (n=407, 57%) displayed steeper growth from fall to winter and non-significant smaller growth from winter to spring (η_1=1.12, p < .05; η_2=.37, p > .05). Class 1 (n=308, 43%) demonstrated more positive growth from fall to winter than Class 2 (η_1=2.44, p < .05), but non-significant negative growth from winter to spring (η_2=-.56, p > .05). *SPED* did not predict any of the growth parameters in this model.

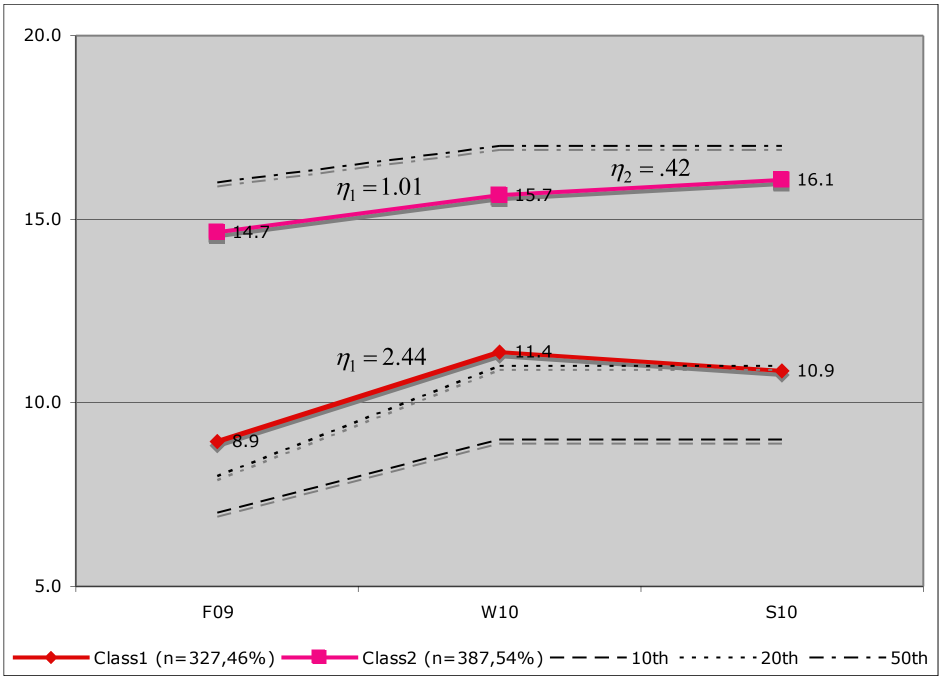


Figure 25. Conditional piecewise LCGA based on MCRC growth for District 2.

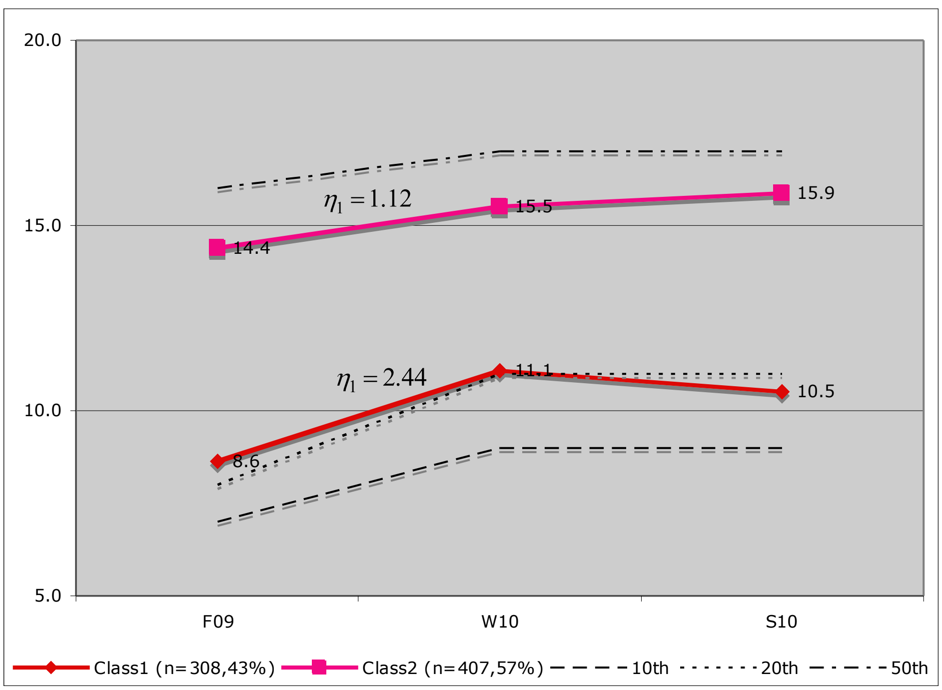


Figure 26. Conditional piecewise LCGA with distal outcome (F10RISK) based on MCRC growth for District 2.

With *SPED* removed from the model, three unconditional piecewise LCGA models with F10RISK, W11RISK and S11RISK as distal outcomes were conducted to examine if the unconditional models could predict the growth parameters and students' risk levels in grade five. Unlike the results from District 1, a 3-class solution was identified as having the best model fit when F10RISK was the distal outcome (see Figure 27). Class 3 (n=150, 21%) was the highest performing group and displayed non-significant small growth over the year (η_1=.07, $p > .05$; η_2=.19, $p > .05$). Class 2 (n=327, 46%), the medium performing group, contained the majority of students and displayed steeper growth from fall to winter (η_1=2.02, $p < .05$), than from winter to spring (η_2=.44, $p < .05$). Class 1 (n=237, 33%) displayed growth similar to Class 2 from fall to winter (η_1=2.00, $p < .05$), but non-significant negative growth from winter to spring (η_2=-.48, $p > .05$).

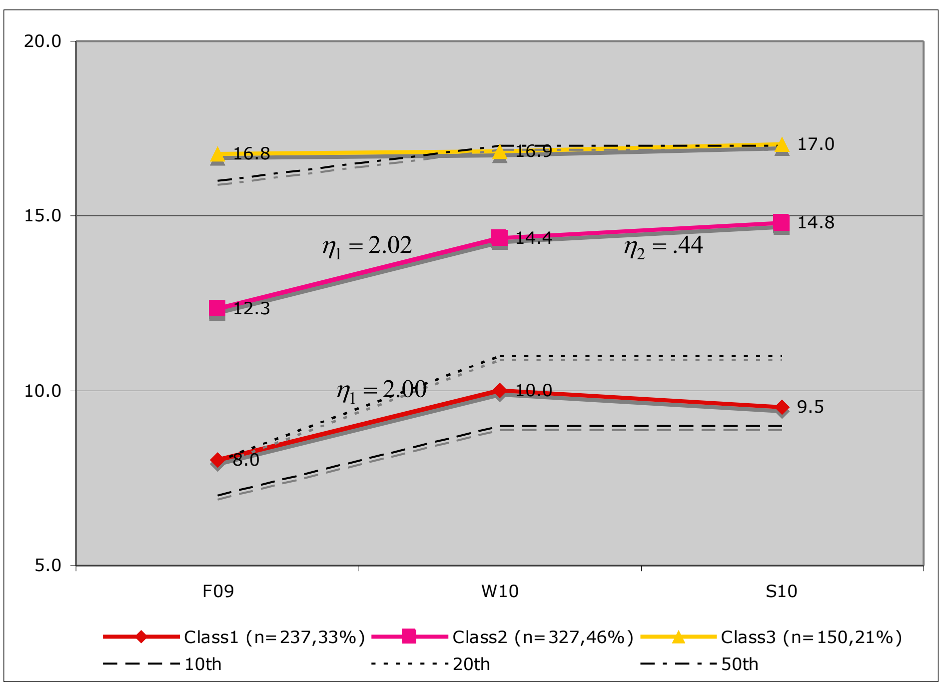


Figure 27. Unconditional piecewise LCGA with distal outcome (F10RISK) based on MCRC growth for District 2.

The 3-class solution was also found to have the best model fit when W11RISK was included in the model as the distal outcome (see Figure 28). This 3-class solution differed from the 4-class solution that was found to have the best model fit for the data from District 1. The three classes were also similar in terms of trajectories and proportions of students to the 3-class solution with F11RISK. Class 3 displayed non-significant small growth over the year (η_1=.13, p > .05; η_2=.18, p > .05), Class 3 displayed steeper growth from fall to winter and less from winter to spring (η_1=1.14, p < .05; η_2=.42, p < .05), and the lowest performing group (Class 1) displayed positive growth from fall to winter, but non-significant negative growth from winter to spring (η_1=1.9, p < .05; η_2=-.42, p > .05).

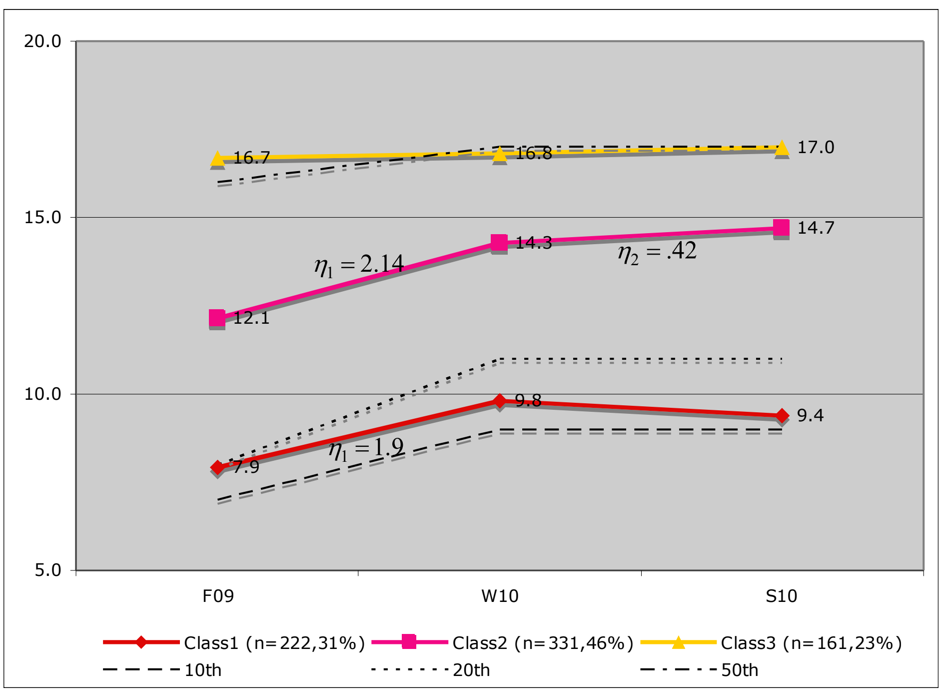


Figure 28. Unconditional piecewise LCGA with distal outcome (W11RISK) based on MCRC growth for District 2.

When S11RISK was the distal outcome, a 4-class solution emerged as having the best model fit (see Figure 29). This solution was similar to the results from District 1.

Class 4 (n=108, 15%) was the highest performing group. Students in Class 4 displayed non-significant growth over the year (η_1=-.03, p > .05; η_2=.02, p > .05). Class 3 (n=121, 37%) was the second highest performing group and students displayed more growth from fall to winter than winter to spring (η_1=3.0, p < .05; η_2=.73, p > .05). Class 2 (n=230, 32%) was composed of students with the second lowest growth trajectory who displayed similar rates of growth as Class 3 from fall to winter (η_1=3.05, p < .05), then negative growth from winter to spring (η_2=-.59, p > .05). The lowest performing students comprised Class 1 (n=121, 17%) and displayed positive growth from fall to winter (η_1=1.06, p < .05), then non-significant negative growth from winter to spring (η_2=-.12, p > .05).

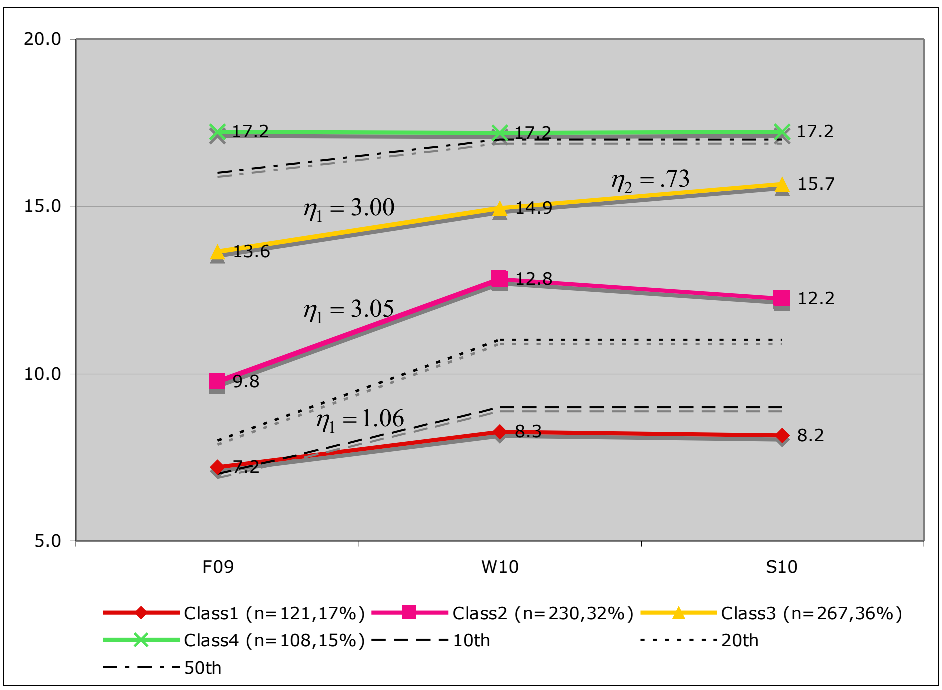


Figure 29. Unconditional piecewise LCGA with distal outcome (S11RISK) based on MCRC growth for District 2.

Alignment of latent classes with easyCBM risk ratings. Similar to the descriptive analysis of data from District 1, the frequencies of students within each class were examined to determine the alignment between easyCBM's risk levels and the latent classes derived from the models with distal outcome. In the model with F10RISK as the distal outcome, the class with the lowest trajectory (Class 1) had the majority of the *high risk* students (43.88%). When W11RISK was the distal outcome, the class with the lowest trajectory (Class 1) has 36.94% of *high-risk* students. Finally, with S11RISK as the distal outcome in the model, the class with the lowest trajectory (Class 3) had 60.33% of *high risk* students (see Figures 30-32). The class with the lowest growth trajectory in all of the models with F10RISK, W11RISK and S11RISK distal outcomes was able to predict the majority of the *high-risk* students (92.86%, 82%, and 73% respectively) in the sample. The percentages of *high-risk* students predicted by these models were similar to the results from District 1, especially for the models with F10RISK and W11RISK as outcomes.

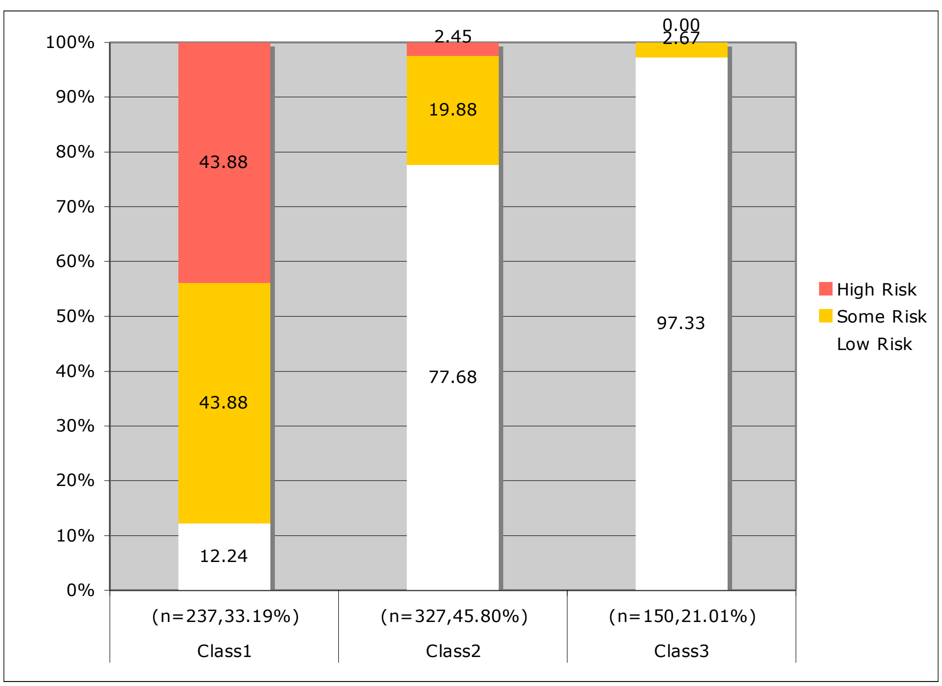


Figure 30. Percentage of students in each of the latent classes identified based on MCRC unconditional piecewise LCGA with F10RISK as distal outcome results for District 2.

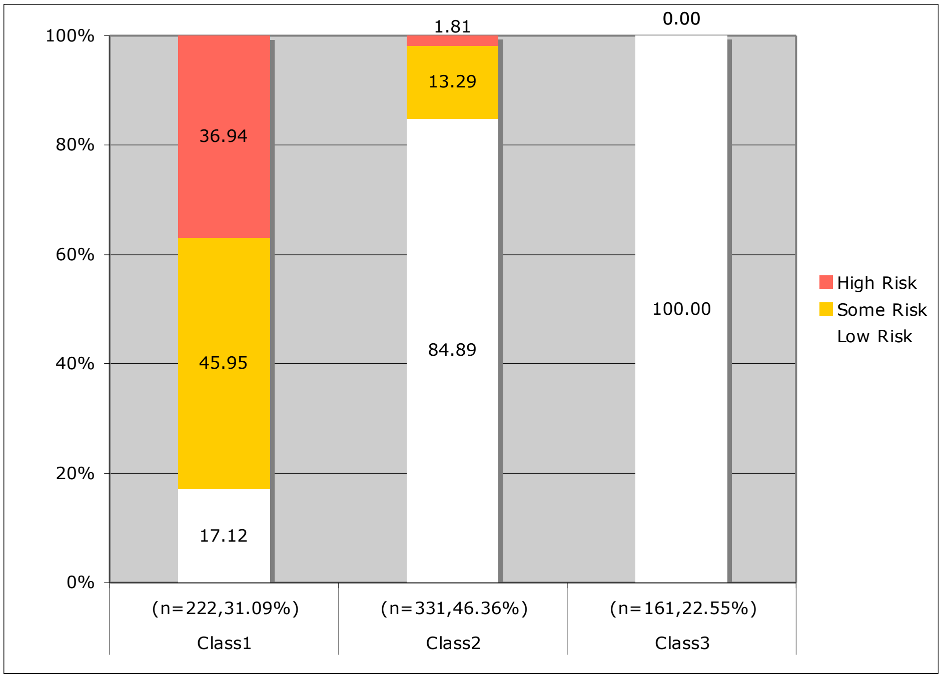


Figure 31. Percentage of students in each of the latent classes identified based on MCRC unconditional piecewise LCGA with W11RISK as distal outcome results for District 2.

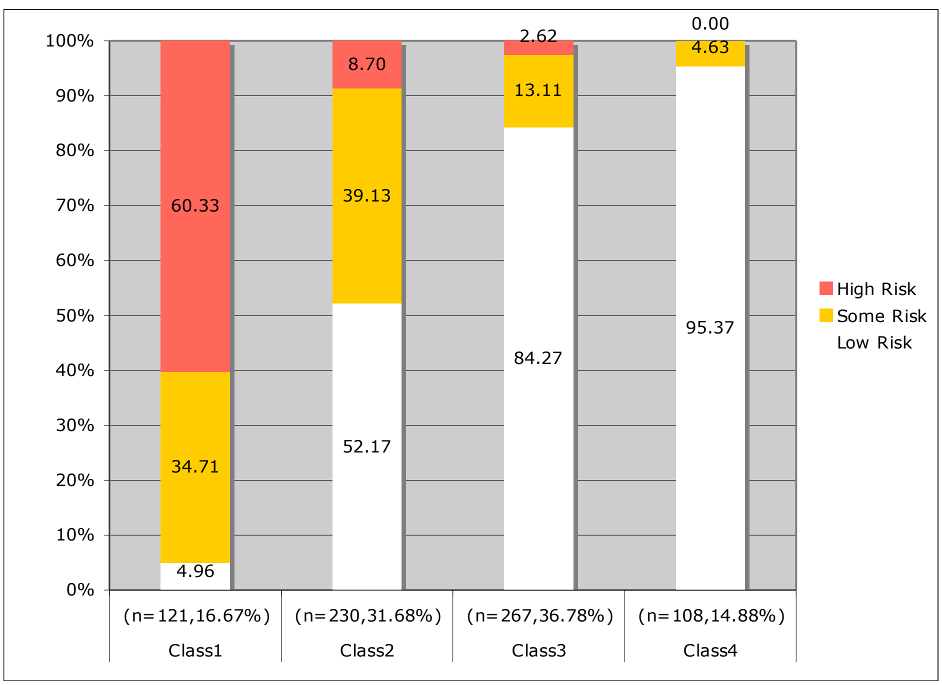


Figure 32. Percentage of students in each of the latent classes identified based on MCRC unconditional piecewise LCGA with S11RISK as distal outcome results for District 2.

Cross Validation Sample (District 2): PRF

Unconditional, conditional, and LCGA models with distal outcomes (F10RISK, W11RISK, and S11RISK) were examined (see Tables D3 and D4 in Appendix D for full model results).

Unconditional LCGA. After examining the latent classes and reading growth trajectories on the MCRC measures for students from District 2, the same series of models were conducted to examine the same questions on the PRF measures. Unlike the 7-class model obtained with data from District 1, a 6-class solution was the best solution for the unconditional piecewise LCGA modeling growth of PRF in District 2 (see Figure 33). All classes had significant positive growth of approximately 15-29 WCPM from fall to winter. All classes also displayed significant smaller growth from winter to spring, except Classes 4 and 1. The highest and lowest performing groups (Classes 6 and 1) displayed smaller growth from fall to winter at 15-16 WCPM, compared to the other classes at 23-29 WCPM. From winter to spring, a general trend of smaller growth rates was observed for all classes. Only Classes 3 and 2 displayed the highest growth (11 and 18 WCPM respectively) and students in other classes experienced smaller rates of growth (less than 5 WCPM).

Conditional LCGA. With *SPED* added to the model as a covariate, a 2-class solution was found to have the best model fit (see Figure 34). This class formation was different from the 3-Class solution obtained with data from District 1. There were two groups of students: Class 2 (*n*=342, 48%) and Class 1 (*n*=372, 52%). Both groups displayed steeper growth from fall to winter than from winter to spring. The fall to winter growth was 29 WCPM for Class 2 and 24 WCPM for Class 1. The winter to spring

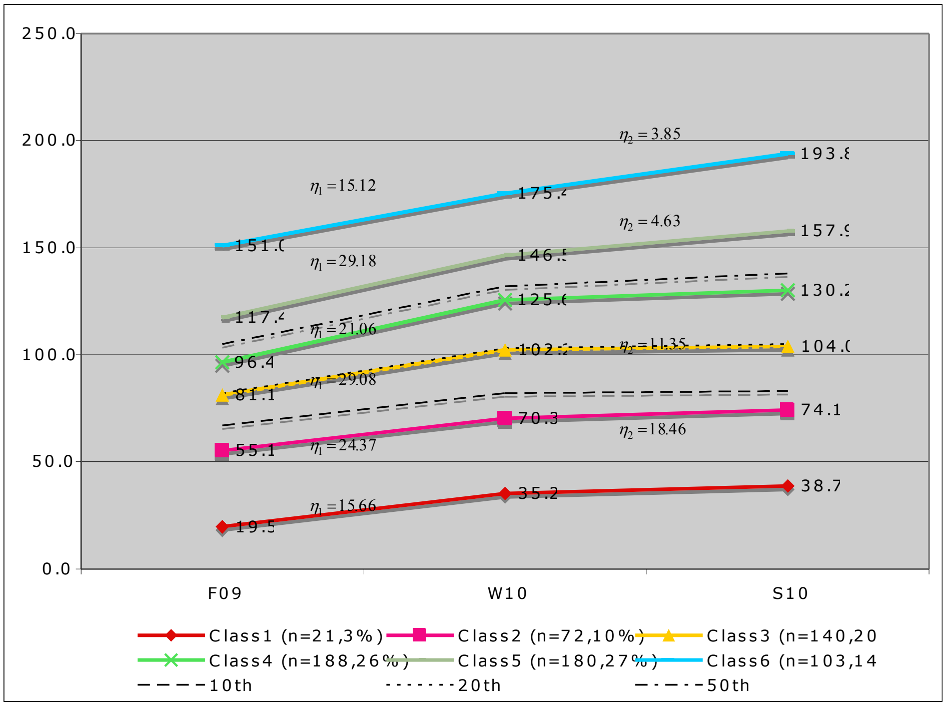


Figure 33. Unconditional piecewise LCGA based on PRF growth for District 2.

growth for Class 2 was higher (12 WCPM) than Class 1 (3 WCPM). Similar to previous conditional models with F11RISK distal outcome, *SPED* significantly did not predict all the growth parameters, except for the intercept and the winter-spring slope.

LCGA with distal outcomes. In a conditional piecewise LCGA model with *SPED* as a covariate and F10RISK as the distal outcome, a 4-class solution had the best model fit (see Figure 35), which differed from the 3-class solution for District 1. In this model, Class 4 (*n*=141, 20%) was high performing group. Class 3 (*n*=284, 40%) was the majority and the second highest performing group. Class 2 (*n*=187, 26%) was the second lowest performing group. Class 1 (*n*=103, 14%) was the lowest performing students. The *SPED* as a covariate in this model had non-significant effect on the growth parameters, except on the intercept for highest and lowest group of students (Classes 4 and 1) and on the fall-winter slope for Classes 2 and 3.

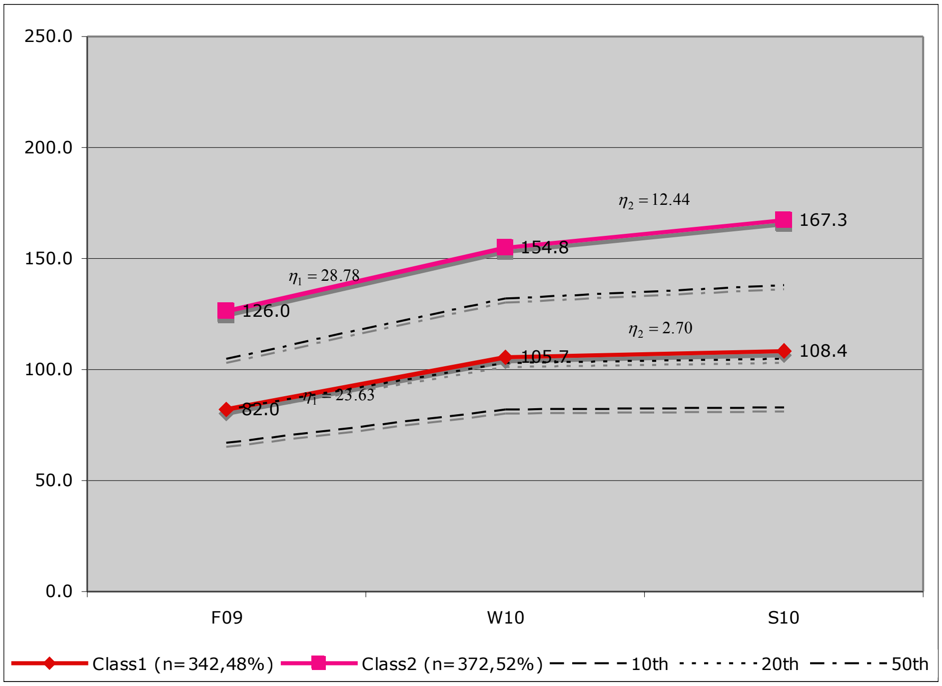


Figure 34. Conditional piecewise LCGA based on PRF growth for District 2.

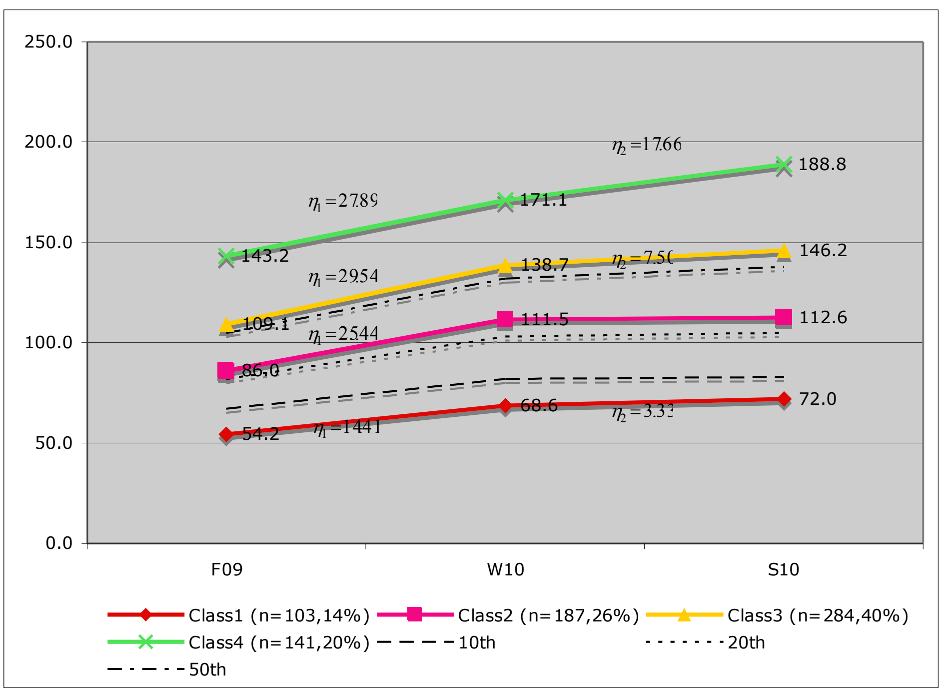


Figure 35. Conditional piecewise LCGA with distal outcome (F10RISK) based on PRF growth for District 2.

SPED was subsequently dropped in the following three unconditional piecewise LCGA models with F10RISK, W11RISK and S11RISK as distal outcomes. With F10RISK as the distal outcome, a 6-class solution was identified as having the best model fit (see Figure 36), as opposed to a 4-class solution that was found to have the best model fit for data from District 1. Classes 6 (*n*=105, 15%) and 5 (*n*=191, 27%) were the two highest-performing groups. Classes 4 (*n*=193, 27%) and 3 (*n*=132, 18%) were students who performed at and above the 20th percentile range over the year. The last two classes, Classes 2 (*n*=72, 10%) and 1 (*n*=21, 3%), were composed of students with the lowest performance trajectories; both groups performed below the 10th percentile over the year.

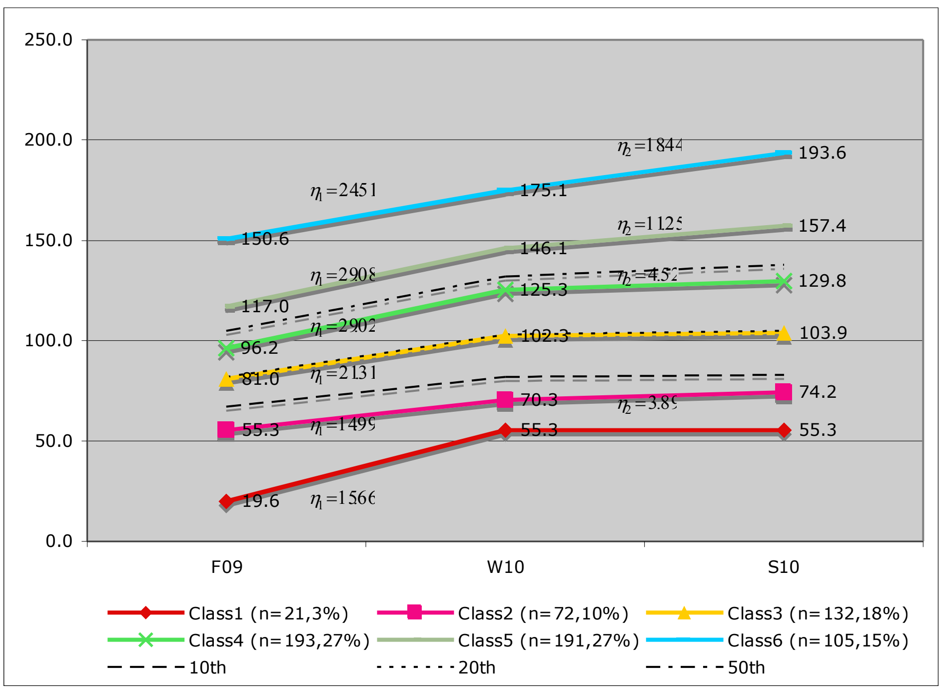


Figure 36. Unconditional piecewise LCGA with distal outcome (F10RISK) based on PRF growth for District 2.

In general, students in all classes displayed more growth from fall to winter than from winter to spring. The lowest two groups (Classes 1 and 2) displayed less growth from fall to winter (between 15 and 16 WCPM) compared to the other four higher performing groups (in the range 21-29 WCPM). As for growth from winter to spring, only the highest two performing groups (Classes 6 and 5) had growth in the range of 10-19 WCPM, but other groups displayed much smaller growth (between 2 and 5 WCPM).

With W11RISK and S11RISK as the distal outcomes, the 4-class solution had the best model fit (see Figures 37 and 38), with each class displaying similar trajectories. The four groups consisted of the high performers, second highest performing group, second lowest group, and the lowest group (Class 3; *n*=104, 15%). The trajectories and proportion of students in each Class in these two models (W11RISK and S11RISK) were similar, where in general, steeper growth was observed from fall to winter compared to winter to spring. The highest three higher groups displayed growth between 20 and 30 WCPM while the lowest group had a growth rate of only 15 WCPM. From winter to spring, only the highest group displayed growth of 18 WCPM, but all other groups only showed growth in the range of 3-8 WCPM. Although the 4-class solution was also found to be the best model fit for data from District 1, the trajectories and proportions of students in each class were different.

Alignment of latent classes with easyCBM risk ratings. The class formations derived from the three unconditional piecewise LCGA models with distal outcomes were examined to determine the alignment between easyCBM risk rating system and the proposed latent classes. In the model with F10RISK as the distal outcome, the lowest three performing groups of students (Classes 1, 5 and 6) included 25%, 72%, and 100% of *high-risk* students, respectively. These classes also captured the majority of students who were categorized as being at *high-risk*, predicted 29%, 46%, and 19% the *high-risk* students in the sample. The three highest-performing classes (Classes 2, 3, and 4)

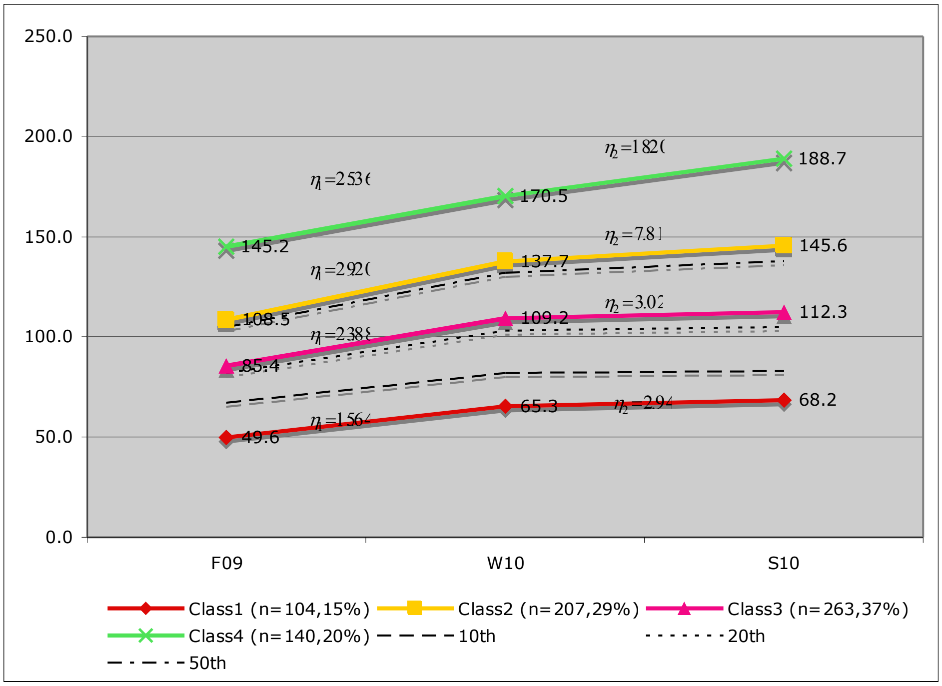


Figure 37. Unconditional Piecewise LCGA with distal outcome (W11RISK) based on PRF growth for District 2.

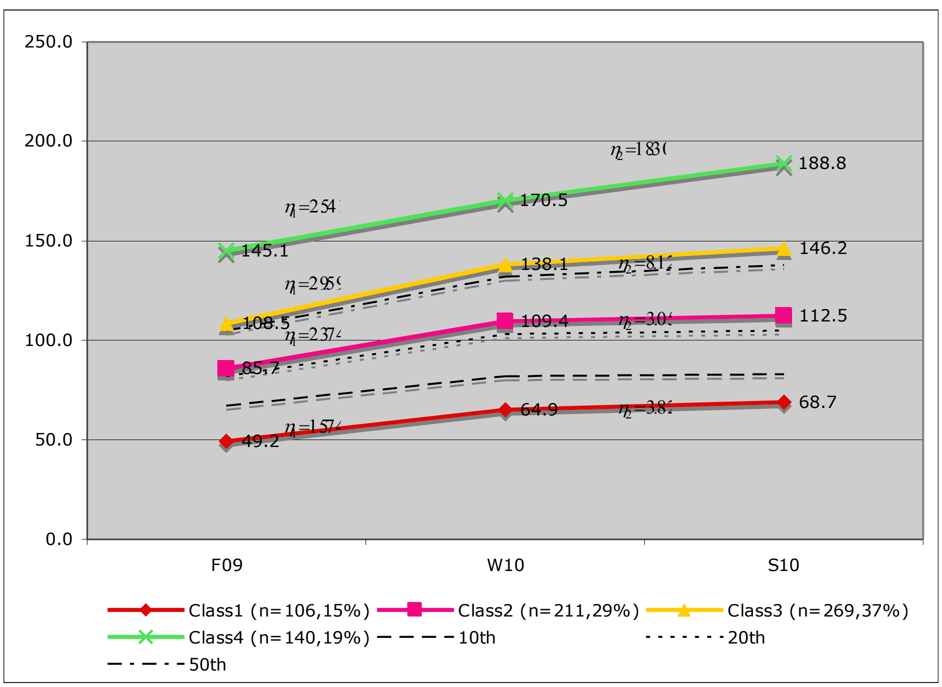


Figure 38. Unconditional piecewise LCGA with distal outcome (S11RISK) based on PRF growth for District 2.

consisted of primarily students with *low risk* status (see Figure 39). The models with W11RISK and S11RISK as distal outcomes suggested that the class with the lowest trajectory was able to capture between 58% and 68% of the *high risk* students in the sample (see Figures 40 and 41).

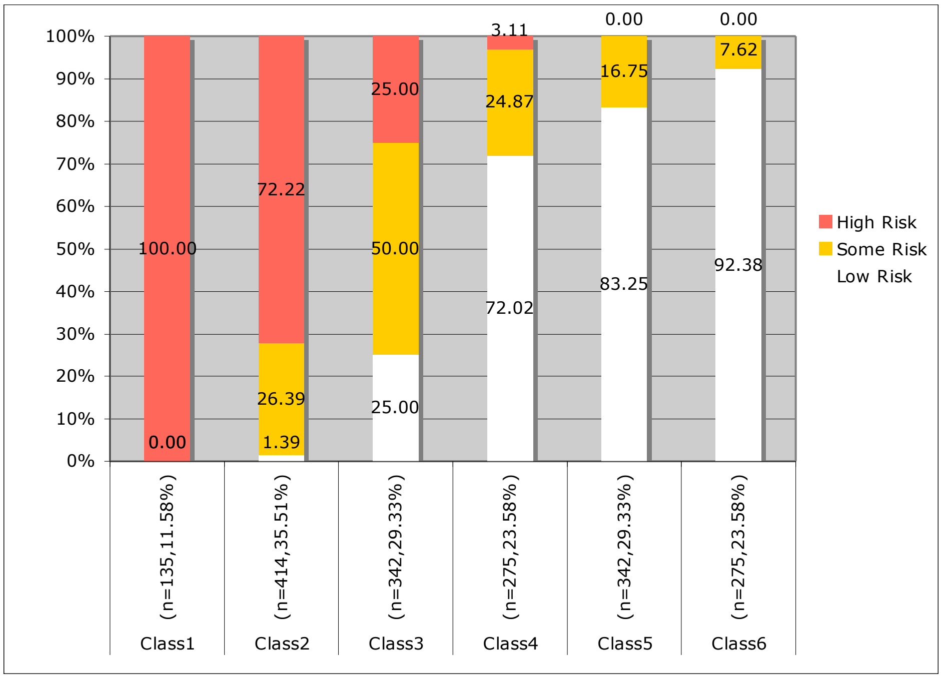


Figure 39. Percentage of students in each of the latent classes identified based on PRF unconditional piecewise LCGA with F10RISK as distal outcome results for District 2.

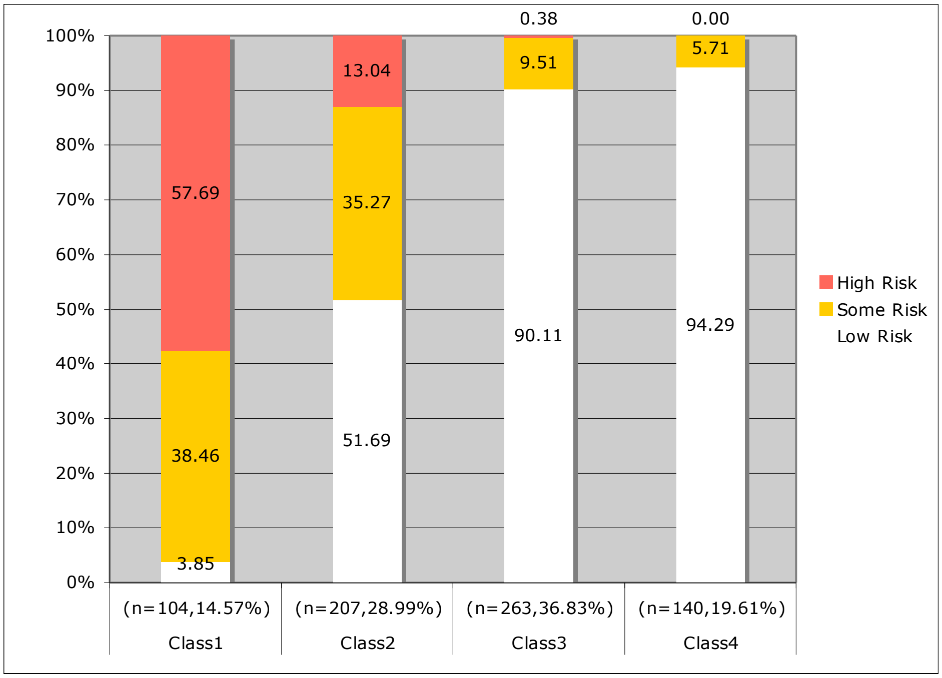


Figure 40. Percentage of students in each of the latent classes identified based on PRF unconditional piecewise LCGA with W11RISK as distal outcome results for District 2.

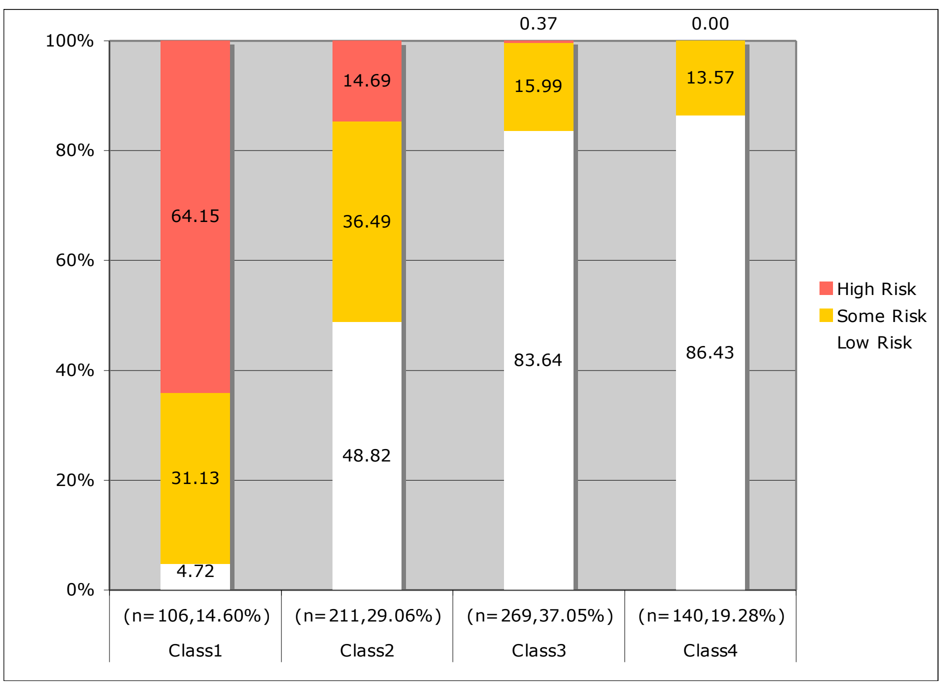


Figure 41. Percentage of students in each of the latent classes identified based on PRF unconditional piecewise LCGA with S11RISK as distal outcome results for District 2.

CHAPTER V

DISCUSSION

Summary

Results from this study suggested that there was heterogeneity in reading development. The latent classes from the piecewise LCGA models align with the easyCBM risk rating systems. Although the LCGA models yielded multiple latent classes, the number of latent classes varied by measures, models and districts.

Multiple latent classes and growth trajectories. A series of LCGA models were fit to examine the latent classes and reading growth trajectories of reading development using the easyCBM reading comprehension (MCRC) and passage reading fluency (PRF) measures. Results from two districts support that there was heterogeneity in reading development as evidenced by the multiple latent classes in all of the models. The heterogeneity of reading development as evidenced in this study is more closely aligned with the RTI framework, which assumes the existence of heterogeneity in students and places them into 3-tier instruction programs. As stated earlier, conventional growth models that assume individuals are from one single population with one average trajectory may not be reasonable when there is heterogeneity of growth in the population. In general, the LCGA models that used PRF measures typically suggested more latent classes (up to six and seven) than the models that used MCRC measures (up to four), which could be due to scale differences between the PRF and MCRC measures. In the PRF measures, the scale ranged from zero to 250, whereas the scale for the MCRC measure was more restrictive, ranging from 0 to 20, resulting in less opportunity for variability in student performance.

Alignment of latent classes and growth trajectories with easyCBM risk ratings. Results from all the unconditional LCGA models that used the latent classes to predict distal outcomes (i.e., risk levels in the fall, winter and spring of grade five) provided some evidence of the alignment between the latent classes and growth trajectories and easyCBM's risk rating system. For example, the trajectories of the latent classes aligned the easyCBM risk ratings of *low*, *some*, and *high risk* across the districts. In general, the class with the lowest initial fall scores (i.e., the lowest class) in all the models across MCRC and PRF measures for both districts was composed primarily of *high risk* students. There was generally one class that consisted of students performing at the average level, which corresponded to students considered as being categorized as *some risk* in the easyCBM system. Finally, there were also classes composed primarily of high performing students, which corresponded to the students easyCBM categorized as *low risk*.

The class formations from these models that used the latent classes to predict distal outcomes (risk levels in the fall, winter and spring of grade five) provided preliminary convergent evidence supporting the predictive validity of the risk ratings of the easyCBM system. This aspect of the study signifies an early effort to examine the validity of the easyCBM risk rating system as no validity studies have been conducted to investigate this feature. The relatively high percentage of *high risk* students captured by the class with the lowest initial fall MCRC and PRF scores across districts supports the intent of these measures as screeners, which is to identify students who are at risk for reading difficulties (Alonzo & Tindal, 2007). It is important to examine the validity of the risk rating system of easyCBM because it is a popular formative assessment system with over 60,000 users in the country. Furthermore, teachers are using the risk rating system to help create instructional groups and modify instructions as part of the RTI program.

Generalizability of the models. Although all the LCGA models yielded multiple latent classes, the number of latent classes varied by measures, models and districts. The unconditional models, conditional models and unconditional models with S11RISK as the distal outcome that used the MCRC measures were the only models that generalize across the districts. The unconditional models with W11 and S11RISK as distal outcomes that used the PRF measures were the only models that generalize across the districts.

In all of the conditional models, *SPED* as a covariate evidently did not have significant effects on the growth factors, especially the growth rates from fall to winter and winter to spring. In general, the model with F10RISK as the distal outcome across measures and districts seemed to be most predictive of *high risk* students, predicting 86-96% of the *high risk* students in the district. Across the districts and measures, the models with F10RISK as distal outcome seemed to predict a higher percentage of *high risk* students: 93%-96% for the model using MCRC measures and 86%-95% for the model using PRF measures. Overall, all of the models with distal outcomes (F10RISK, W11RISK and S11RISK) that used the MCRC measures predicted a higher percentage of *high risk* students (73%-96%) compared to PRF measures (68%-95%; see Appendix E).

Limitations

It is should be noted that there are several limitations in this study. First, the results from this study can only be generalized to students with complete grade four and grade five benchmark data that share similar characteristics with District 1 and 2. Students without complete benchmark data from dissimilar districts across both years could have different demographic and performance characteristics. Second, data about the intervention programs and curricula used in both districts were not included in the analysis. Over an academic year, students' reading development is likely to be heavily influenced by many factors such as the core, supplemental, and/or intervention programs being implemented in the building. These data could help explain the growth trends of greater rates of growth observed from fall to winter growth and smaller rates of growth observed from winter to spring. Other potential variables to investigate include student-group assignments, program dosage, and instructional time (Kame'enui, Simmons, & Coyne, 2000). Third, the choice to include special education status as the only covariate used in all the models is another potential limitation of the study. Students with special education status are only one small proportion of the population that is likely to experience difficulty learning to read and receiving additional instructional services and supports. Not only is it possible that students have been incorrectly identified as needing special education services but research has also indicated that students from specific demographics subgroups, such as males and/or children from particular minority subgroups are often over-identified (Artiles & Trent, 1999; Hibel, Farkas, & Morgan, 2010). Other student characteristics such as English Language proficiency, and/or gender could also be examined for significance.

Finally, results from the series of piecewise LCGA in District 1 did not fully generalize to District 2. This aspect of the study could be improved by replicating the study using two comparable districts with similar curricula, RTI program, and student population (e.g., percent of students eligible for free and reduced lunch). It might be worthwhile to plan a future study that examines the stability of latent classes and trajectories of several cohorts of late elementary school students across years.

Implications for Future Research

In both districts, the class with the lowest initial fall status (i.e., "lowest class") in all of the unconditional LCGA models (with and without distal outcomes) was composed of scores on the MCRC measures that fell between the 10th and 20th percentile. On the same set of models that used PRF measures, the "lowest class" for both districts generally included students whose fall score fell below the 10th percentile. In all unconditional models with distal outcomes, the majority of the students in all of the "lowest classes" were *high risk* students, based on the analysis that examined the alignment between the latent classes and the easyCBM risk rating system. Results from all unconditional models (regardless of the distal outcomes) suggest that students in the "lowest class" were the target students for Tiers 2 and 3 interventions, which would be informative to school administrators and district in identifying students at risk for reading difficulties.

Non-linear within-year growth. A non-linear growth trend on the reading fluency measures was evident, where steeper growth was observed from fall to winter compared to the growth observed from winter to spring. The findings were similar to previous studies (Christ et al., 2010; Nese et al., 2012) that also found evidence of non-linear growth on measures of reading fluency. The steeper fall to winter growth pattern was also true for the models that used MCRC measures where the class with the lowest initial fall scores showed steeper growth from fall to winter growth and minimal or non-significant growth from winter to spring. This non-linear growth trend could be further examined by investigating seasonal effects (i.e., lack of exposure formal instruction during the summer) in future studies so that reasons for the steeper growth observed from fall to winter can be more fully understood.

Stability of reading development over time. Despite the overall steeper fall to winter growth trend from most of the classes, especially the classes with lower initial fall MCRC and PRF scores, the relative positions of the classes remained constant over the year. The higher performing groups continued to have scores in the higher range across the year, and the lower performing groups continued to have lower scores throughout the year. Although there is preliminary evidence of relatively stable growth trends of the latent classes within a year, it may still be unwarranted to conclude that reading achievement is constant as Juel (1988) proposed in her study. Because only within-year growth was examined in this study, the results here only reveal a small fraction of reading development. To examine the stability of reading development over time, a multi-year cohort study similar to those conducted by others (Phillips et al., 2002; Leach et al., 2003; Lipka et al., 2006) is warranted. However, because the LCGA models suggested the existence of multiple latent classes, future research should consider the use of latent transition analysis (LTA) that allows the examination of movement between subgroups over time.

The relation between reading fluency and comprehension. The overall higher percentage of *high risk* students that were captured by the "lowest class" in the models compared to the MCRC measures compared to the PRF measures may be due to the greater emphasis on reading comprehension and less on reading fluency in the later elementary grades (Chall, 1996). It is also possible that the role of reading fluency decreases in the later elementary grades, as suggested by Yovanoff et al. (2005).

However, the approach used in this study whereby growth in reading comprehension and reading fluency was modeled separately did not directly test for the declining role of reading fluency (i.e. PRF). One possible way to investigate this aspect of reading in the future is modeling growth using both MCRC and PRF measures concurrently using parallel process growth mixture model to examine the association between PRF and MCRC over time. Applying this model to examine within- and between-year reading growth would provide greater insights into the role of reading fluency and comprehension on subgroups of students, with special interest on the lower performing groups, across the early and later elementary grade levels.

Sensitivity and specificity. Despite the overall high percentages of *high risk* students that were captured by the "lowest class" in the LCGA models, a small proportion of students with *some* or *low* risk were also captured by the "lowest class". The rate of true-positives of the "lowest class" from these models could be explored in future studies by examining the specificity-sensitivity rate analysis. As stated earlier in this paper, it is important for screening measures like the easyCBM MCRC and PRF to have high specificity and sensitivity of .90-.95 (Jenkins, 2003). A comparison of the optimum specificity and sensitivity rates for MCRC and PRF could shed light on the similarities or differences at identifying students at risk for reading difficulties.

***SPED* as a covariate.** Another interesting finding of this study relates to the inclusion of *SPED* as a covariate to the models. When *SPED* was added as a covariate to the model the number of latent classes in the unconditional models reduced to fewer classes in the conditional models. In the models that used MCRC measures, for example, four classes were reduced to two. In the models that used PRF measures, models with six or seven classes were reduced to models with two or three classes. The fewer classes in the conditional models could be because special education status sufficiently explained the variance between students, especially on their initial fall MCRC and PRF scores. Additionally, the non-significant effects of *SPED* on the growth parameters (intercept, fall-winter slope, and winter-spring slope) in most of the conditional models may indicate that students in *SPED* programs and students in general education programs did not differ on their rates of growth from fall to winter and/or winter to spring. Consequently, the two- or three-class solutions in the conditional models can be considered as adequately capturing the heterogeneity of students' initial fall scores and growth trajectories.

Conclusion

There are many aspects of reading development that can be examined in the future, including the examination of other covariates, the association between non-linear growth and seasonal effects, the stability of reading development over time as influenced by reading fluency and reading comprehension across early and later elementary grades, and the specificity-sensitivity rates of the MCRC and PRF measures. Investigating these aspects of the reading measures of a popular formative system like the easyCBM can have significant implications. District and school administrators can make informed decisions about the allocation of adequate resources to serve students in the "lowest class" who have demonstrated the greatest need for additional instructional support. Teachers and RTI teams can also use this information to plan appropriate reading intervention programs and create strategic homogeneous instructional groups by targeting students' weaknesses identified as *high risk*.

APPENDIX A

STUDIES THAT EXAMINED CORRELATION BETWEEN ORF AND STATE TESTS

Authors	Grade	Sample Size	State Test	Correlation
Good, Simmons, & Kame'enui (2001)	K-3	364	Oregon	.60s
Stage & Jacobsen (2001)	4	173	Washington	.40s
Shaw & Shaw (2002)	3	52	Colorado	.70-.80s
Buck & Torgesen (2003)	3	1102	Florida	.70s
McGlinchey & Hixson (2004)	4	1362	Michigan	.40s-.80s
Hintze & Silberglitt (2005)	1-3	1766 (5 cohorts)	Minnesota	.40s-.60s
Shapiro, Keller, Santoro, & Hintze (2006)	3-5	n_1=617 n_2=782	Pennsylvania	60s-.70s

APPENDIX B

SAMPLE INFORMATION

Table B1

Demographic Information

		District 1 (n = 1299)		District 2 (n = 818)	
		n	%	n	%
Gender	Male	586	50.2	407	56.9
	Female	581	49.8	308	43.1
Ethnicity	American Indian/Alaskan Native	23	2.0	14	2.0
	Asian/Pacific Islander	40	3.4	15	2.1
	Black	33	2.8	9	1.3
	Hispanic	139	11.9	126	17.6
	White	833	71.4	486	68.0
	Multi-Ethnic	56	4.8	32	4.5
	Decline/Missing	16	1.4	14	2.0
	Missing	27	2.3	19	2.7
ELL	No	1122	96.1	710	99.3
	Yes	45	3.9	5	0.7
SPED	No	969	83.0	570	79.7
	Yes	198	17.0	145	20.3

Table B2

Descriptive Statistics

Measure Scores	Main Sample (District 1)					Cross Validation Sample (District 2)				
	n	Min	Max	*M*	*SD*	*n*	Min	Max	*M*	*SD*
Grade 4										
Fall09PRF	1113	6	263	115.41	37.99	686	1	205	100.71	33.54
Fall09MCRC	1109	0	20	12.64	4.29	687	1	20	11.80	4.00
Wint10PRF	1126	5	269	136.18	37.58	704	4	245	125.70	36.65
Wint10MCRC	1118	2	20	14.36	3.43	704	0	20	13.41	3.57
Spr10PRF	1138	9	340	145.32	42.60	714	5	256	133.49	41.26
Spr10MCRC	1129	1	20	14.26	3.58	712	0	20	13.49	3.92
Grade 5										
Fall10PRF	1167	3	356	154.16	42.23	715	1	251	139.26	38.63
Fall10MCRC	1167	0	20	14.10	3.35	715	0	20	13.31	3.44
Wint11PRF	1152	6	332	160.09	42.37	710	7	340	147.22	39.62
Wint11MCRC	1137	1	20	16.01	3.14	708	0	20	15.39	3.38
Spr11PRF	1135	4	353	173.08	40.42	711	0	282	163.77	39.33
Spr11MCRC	1139	0	20	14.66	2.85	712	0	20	14.10	3.06

Note: PRF is passage reading fluency; MCRC is multiple choice reading comprehension.

Table B3

2009-10 Norm Percentiles

	Fall09		Winter10		Spring10	
Percentile	PRF	MCRC	PRF	MCRC	PRF	MCRC
10th	67	7	82	9	83	9
20th	82	8	103	11	105	11
50th	105	16	132	17	138	17
75th	129	16	155	17	168	17
90th	155	18	176	18	194	19

APPENDIX C

STATISTICS

Table C1

Growth Model Results for District 1

	MCRC			PRF		
	Linear	Piecewise	Conditional Piecewise	Linear	Piecewise	Conditional Piecewise
Fit Statistics						
AIC	17369.522	17228.822	19243.113	30808.1	30653.108	35039.979
BIC	17399.849	17274.312	19278.549	30838.458	30698.645	35075.414
ABIC	17380.791	17245.725	19256.314	30819.4	30670.058	35053.18
CFI	0.878	1	0.141	0.958	1	0.098
TLI	0.878	1	-0.03	0.958	1	-0.083
RMSEA	0.203	0	0.435	0.213	0	0.775
SRMR	0.222	0	0.302	0.105	0	0.371
Fixed Effect						
Intercept, η_0	12.906	12.609	13.164	116.831	115.065	120.873
FW slope, η_1	0.835	1.729	1.628	15.392	20.752	21.243
WS slope, η_2	--	**-0.054**	**-0.126**	--	10.044	9.786
Covariate Effect						
Intercept, η_0	--	--	-3.078	--	--	-31.984
FW slope, η_1	--	--	**0.518**	--	--	**-2.815**
WS slope, η_2	--	--	**0.188**	--	--	**-3.637**

Random Effect						
Variance (Intercept, η_0)	11.412	13.7	13.164	1213.157	1312.437	--
Variance (FW slope, η_1)	0.23	2.165	1.628	17.257	65.898	--
Variance (WS slope, η_2)	--	0.653	-0.126	--	93.405	--
Covariance (Intercept, η_0 & FW slope, η_1)	--	-4.439	--	74.58	-46.6	--
		(r=-0.815)		(r=0.515)	(r=-0.158)	
Covariance (Intercept, η_0 & WS slope, η_2)	--	**0.196**	--	--	153.009	--
		(r=0.066)			(r=0.437)	
Error Variance	5.766	4.797	13.284	171.359	126.024	1381.087

Note. Non statistically significant values in boldface. FW = Fall-Winter. WS = Winter-Spring.

Table C2

Fit Statistics for all Final Piecewise LCGA Models Using MCRC Measures for District 1

Statistic	1 Class	2 Class	3 Class	4 Class
Unconditional				
AIC	18459.907	17028.998	16819.716	16721.002
BIC	18480.125	17074.488	16890.479	16817.037
ABIC	18467.42	17045.901	16846.01	16756.686
Entropy	-	0.820	0.728	0.712
LMR *(p value)*	-	0	0.5216	0.0001
Adjusted LMR *(p value)*	-	0	0.5273	0.0002
Conditional				
AIC	18218.458	16911.669		
BIC	18253.839	16992.54		
ABIC	18231.605	16941.719		
Entropy	-	0.812		
LMR *(p value)*	-	0		
Adjusted LMR *(p value)*	-	0		
Conditional-Distal (F10RISK)				
AIC	20028.572	18373.321	18049.982	
BIC	20074.132	18474.565	18206.91	
ABIC	20045.545	18411.038	18108.444	
Entropy	-	0.838	0.782	
LMR *(p value)*	-	0	0	
Adjusted LMR *(p value)*	-	0	0	

Unconditional-Distal (F10RISK)				
AIC	20270.022	18484.25	18147.322	18042.648
BIC	20300.395	18550.058	18248.566	18179.327
ABIC	20281.337	18508.766	18185.039	18093.566
Entropy	-	0.843	0.773	0.743
LMR *(p value)*	-	0	0	0
Adjusted LMR *(p value)*	-	0	0	0
Unconditional-Distal (W11RISK)				
AIC	20001.87	17863.002	17568.894	17458.041
BIC	20032.15	17918.515	17654.687	17574.114
ABIC	20013.092	17883.575	17600.69	17501.059
Entropy		0.844	0.761	0.738
LMR *(p value)*		0	0	0
Adjusted LMR *(p value)*		0	0	0
Unconditional-Distal (S11RISK)				
AIC	19670.271	17933.469	17653.401	17546.436
BIC	19695.508	17988.991	17739.209	17662.529
ABIC	19679.627	17954.052	17685.212	17589.474
Entropy		0.839	0.750	0.734
LMR *(p value)*		0	0.0001	0
Adjusted LMR *(p value)*		0	0.0001	0

Note. LMR = Lo-Mendel-Rubin test.

Table C3

Final Piecewise LCGA Models Using MCRC Measures Results for District 1

			Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Best Model fit		4-class	2-class	3-class	4-class	4-class	4-class
	n		1158	1158	1167	1167	1149	1150
Class 1	Fixed Effect							
		Intercept, η_0	7.40	9.44	7.40	7.30	7.40	7.41
		FW slope, η_1	2.32	3.10	2.68	2.31	2.09	2.19
		WS slope, η_2	**-0.36**	**-0.35**	**-0.59**	**-0.29**	**-0.21**	**-0.10**
	Covariate Effect							
		Intercept, η_0	-	-1.40	**-0.11**	-	-	-
		FW slope, η_1	-	**-0.48**	**-0.28**	-	-	-
		WS slope, η_2	-	**0.64**	**0.75**	-	-	-
	Random Effect							
		Error Variance	8.08	10.12	8.36	8.05	7.82	8.18
	Class Proportion		0.19	0.47	0.20	0.18	0.18	0.20
	PP		0.87	0.95	0.92	0.89	0.90	0.89
	Odds Ratio	Class1 on SPED	-	3.31	-	-	-	-
	(Last Class as Reference)							
	Probability	Low Risk	-	-	0.14	0.12	0.14	0.14
		Some Risk	-	-	0.43	0.41	-	-
		High Risk	-	-	0.44	0.47	-	-
		Some & High Risk	-	-	-	-	0.86	0.86
Class 2	Fixed Effect							
		Intercept, η_0	10.30	15.93	11.16	10.29	10.14	10.23
		FW slope, η_1	3.42	0.55	2.93	3.35	3.52	3.58
		WS slope, η_2	**0.05**	**0.09**	**-0.16**	-0.06	**-0.11**	**-0.18**

	Covariate Effect							
		Intercept, η_0	-	-0.70	**-0.04**	-	-	-
		FW slope, η_1	-	**0.45**	**-0.16**	-	-	-
		WS slope, η_2	-	**-0.62**	**0.47**	-	-	-
	Random Effect							
		Error Variance	6.78	3.92	7.01	7.10	7.16	6.82
	Class Proportion		0.29	0.53	0.35	0.30	0.31	0.29
	PP		0.82	0.94	0.88	0.86	0.83	0.82
	Probability	Low Risk	-	-	0.76	0.70	0.74	0.69
		Some Risk	-	-	0.23	0.28	-	-
		High Risk	-	-	**0.01**	**0.02**	-	-
		Some & High Risk	-	-	-	-	0.27	0.31
Class 3	Fixed Effect							
		Intercept, η_0	15.24		16.36	15.22	15.13	15.17
		FW slope, η_1	0.88		0.32	0.90	0.93	0.92
		WS slope, η_2	**-0.17**		**0.13**	**-0.12**	**-0.11**	**-0.11**
	Covariate Effect							
		Intercept, η_0	-		**-0.01**	-	-	-
		FW slope, η_1	-		**0.19**	-	-	-
		WS slope, η_2	-		**-0.77**	-	-	-
	Random Effect							
		Error Variance	3.63		3.23	3.60	3.71	3.69
	Class Proportion		0.38		0.44	0.38	0.38	0.38
	PP		0.84		0.91	0.85	0.85	0.86
	Probability	Low Risk	-		0.97	0.95	0.96	0.91
		Some Risk	-		0.03	0.05	-	-
		High Risk	-		**0.00**	**0.00**	-	-
		Some & High Risk	-		-	-	0.04	0.09

Class 4	Fixed Effect							
		Intercept, η_0	18.03			18.03	18.04	18.07
		FW slope, η_1	-0.57			-0.56	-0.62	-0.70
		WS slope, η_2	0.50			**0.48**	0.53	0.55
	Covariate Effect							
		Intercept, η_0	-			-	-	-
		FW slope, η_1	-			-	-	-
		WS slope, η_2	-			-	-	-
	Random Effect							
		Error Variance	1.37			1.38	1.39	1.38
	Class Proportion		0.14			0.14	0.14	0.13
	PP		0.79			0.80	0.80	0.81
	Probability	Low Risk	-			0.98	0.98	**1.00**
		Some Risk	-			**0.02**	-	-
		High Risk	-			**0.01**	-	-
		Some & High Risk	-			-	**0.02**	**0.00**

Note. Model 1 = Unconditional, Model 2 = Conditional, Model 3 = Conditional-Distal, Model 4 = Unconditional-Distal (F10RISK), Model 5 = Unconditional-Distal (W11RISK), Model 6 = Unconditional-Distal (S11RISK), PP = Average Latent Class Probabilities for Most Likely Latent Class Membership. Non-statistically significant values are in boldface. FW = Fall-Winter. WS = Winter-Spring.

Table C4

Fit Statistics for all Final Piecewise LCGA Models Using PRF Measures for District 1

Statistic	1 Class	2 Class	3 Class	4 Class	5 Class	6 Class	7 Class
<u>Unconditional</u>							
AIC	34413.4	32783.4	31892.7	31342.7	31001.9	30764.9	30764.9
BIC	34433.6	32828.9	31963.5	31438.8	31123.3	30911.6	30911.6
ABIC	34420.9	32800.3	31919.0	31378.4	31047.0	30819.5	30819.5
Entropy		0.835	0.864	0.857	0.870	0.871	0.871
LMR *(p value)*		0	0.0013	0.1333	0.0322	0.0167	0.0167
Adjusted LMR *(p value)*		0	0.0016	0.1383	0.0342	0.0181	0.0181
<u>Conditional</u>							
AIC	34015.3	32369.0	31582.8				
BIC	34050.7	32449.9	31709.3				
ABIC	34028.5	32399.1	31629.9				
Entropy		0.846	0.836				
LMR *(p value)*		0	0.0801				
Adjusted LMR *(p value)*		0	0.0826				
<u>Conditional-Distal (F10RISK)</u>							
AIC	38209.9	36302.9	35267.4				
BIC	38251.3	36396.0	35412.2				
ABIC	38225.9	36338.8	35323.2				
Entropy		0.832	0.865				
LMR *(p value)*		0	0.0034				
Adjusted LMR *(p value)*		0	0.0036				

Unconditional-Distal (F10RISK)				
AIC	38634.6	36712.3	35514.4	34888.9
BIC	38660.4	36769.2	35602.3	35007.8
ABIC	38644.5	36734.2	35548.3	34934.7
Entropy		0.819	0.869	0.857
LMR *(p value)*		0	0.0003	0.0567
Adjusted LMR *(p value)*		0	0.0004	0.06
Unconditional-Distal (W11RISK)				
AIC	35369.0	33357.1	32318.1	31774.1
BIC	35394.2	33412.6	32403.9	31890.1
ABIC	35378.3	33377.7	32349.9	31817.1
Entropy		0.837	0.875	0.865
LMR *(p value)*		0	0.0006	0.0918
Adjusted LMR *(p value)*		0	0.0007	0.0961
Unconditional-Distal (S11RISK)				
AIC	35918.4	33572.6	32526.3	31976.5
BIC	35948.6	33628.1	32612.1	32092.6
ABIC	35929.6	33593.2	32558.1	32019.5
Entropy		0.841	0.881	0.861
LMR *(p value)*		0	0	0.0794
Adjusted LMR *(p value)*		0	0	0.0836

Note. LMR = Lo-Mendel-Rubin test.

Table C5

Final Piecewise LCGA Models Using PRF Measures Results for District 1

			Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Best Model fit		7-class	3-class	3-class	4-class	4-class	4-class
	n		1164	1164	1299	1299	1149	1150
Class 1	Fixed Effect							
		Intercept, η_0	38.89	86.31	73.60	59.57	59.47	64.34
		FW slope, η_1	15.18	20.34	17.50	16.32	15.65	16.74
		WS slope, η_2	3.44	3.57	**2.05**	3.30	**2.46**	2.41
	Covariate Effect							
		Intercept, η_0	-	-34.14	-18.50	-	-	-
		FW slope, η_1	-	-5.29	**-0.743**	-	-	-
		WS slope, η_2	-	**0.50**	**3.65**	-	-	-
	Random Effect							
		Error Variance	237.67	368.06		416.87	355.83	395.19
	Class Proportion		0.04	0.32	0.18	0.12	0.11	0.16
	PP		0.93	0.94	0.94	0.96	0.94	0.93
	Probability	Low Risk	-	-	0.56	0.43	**0.02**	0.09
		Some & High Risk	-	-	0.44	0.57	0.98	0.91
		Some Risk	-	-	-	-	-	-
		High Risk	-	-	-	-	-	-
Class 2	Fixed Effect							
		Intercept, η_0	71.13	159.67	103.34	93.73	90.87	96.60
		FW slope, η_1	16.36	18.43	23.80	22.23	21.82	22.89
		WS slope, η_2	**2.53**	16.32	7.18	5.46	5.12	6.34
	Covariate Effect							
		Intercept, η_0	-	-15.62	-14.09	-	-	-
		FW slope, η_1	-	**-1.169**	**-2.452**	-	-	-
		WS slope, η_2	-	**-3.014**	**-1.932**	-	-	-

	Random Effect							
		Error Variance	132.22	624.64		202.94	191.06	200.97
	Class Proportion		0.09	0.29	0.42	0.36	0.31	0.35
	PP		0.90	0.96	0.92	0.91	0.93	0.94
	Probability	Low Risk	-	-	0.99	0.97	0.59	0.68
		Some & High Risk	-	-	**0.01**	**0.03**	0.41	0.32
		Some Risk	-	-	-	-	-	-
		High Risk	-	-	-	-	-	-
Class 3	Fixed Effect							
		Intercept, η_0	89.20	159.67	149.97	122.90	118.87	127.31
		FW slope, η_1	21.87	18.43	20.45	24.58	23.69	24.10
		WS slope, η_2	4.03	16.32	15.72	12.56	11.84	13.67
	Covariate Effect							
		Intercept, η_0	-	-15.62	**-3.807**	-	-	-
		FW slope, η_1	-	**-1.169**	**-2.02**	-	-	-
		WS slope, η_2	-	**-3.014**	**-1.069**	-	-	-
	Random Effect							
		Error Variance	135.89	624.64		206.31	212.52	214.51
	Class Proportion		0.21	0.29	0.40	0.29	0.31	0.29
	PP		0.88	0.96	0.96	0.89	0.89	0.89
	Probability	Low Risk	-	-	**1.00**	**1.00**	0.95	0.89
		Some & High Risk	-	-	**0.00**	**0.00**	0.05	0.11
		Some Risk	-	-	-	-	-	-
		High Risk	-	-	-	-	-	-
Class 4	Fixed Effect							
		Intercept, η_0	107.78			164.14	160.69	167.01
		FW slope, η_1	23.06			17.33	18.57	17.13
		WS slope, η_2	8.82			17.12	16.40	17.08

	Covariate Effect						
		Intercept, η_0	-		-	-	-
		FW slope, η_1	-		-	-	-
		WS slope, η_2	-		-	-	-
	Random Effect						
		Error Variance	146.35		586.49	590.60	598.09
	Class Proportion		0.23		0.23	0.26	0.20
	PP		0.87		0.94	0.96	0.93
	Probability	Low Risk	-		**1.00**	0.98	0.97
		Some & High Risk	-		**0.00**	0.02	**0.03**
		Some Risk	-		-	-	-
		High Risk	-		-	-	-
Class 5	Fixed Effect						
		Intercept, η_0	127.84				
		FW slope, η_1	24.54				
		WS slope, η_2	14.65				
	Covariate Effect						
		Intercept, η_0	-				
		FW slope, η_1	-				
		WS slope, η_2	-				
	Random Effect						
		Error Variance	154.00				
	Class Proportion		0.19				
	PP		0.85				
	Probability	Low Risk	-				
		Some & High Risk	-				
		Some Risk	-				
		High Risk	-				

Class 6	Fixed Effect		
		Intercept, η_0	157.26
		FW slope, η_1	17.80
		WS slope, η_2	17.13
	Covariate Effect		
		Intercept, η_0	-
		FW slope, η_1	-
		WS slope, η_2	-
	Random Effect		
		Error Variance	296.73
	Class Proportion		0.21
	PP		0.92
	Probability	Low Risk	-
		Some & High Risk	-
		Some Risk	-
		High Risk	-
Class 7	Fixed Effect		
		Intercept, η_0	204.55
		FW slope, η_1	14.81
		WS slope, η_2	13.85
	Covariate Effect		
		Intercept, η_0	-
		FW slope, η_1	-
		WS slope, η_2	-
	Random Effect		
		Error Variance	593.10
	Class Proportion		0.03
	PP		0.95
	Probability	Low Risk	-
		Some & High Risk	-

Some Risk	-
High Risk	-

Note. Model 1 = Unconditional, Model 2 = Conditional, Model 3 = Conditional-Distal, Model 4 = Unconditional-Distal (F10RISK), Model 5 = Unconditional-Distal (W11RISK), Model 6 = Unconditional-Distal (S11RISK), PP = Average Latent Class Probabilities for Most Likely Latent Class Membership. Non-statistically significant values are in boldface. FW = Fall-Winter. WS = Winter-Spring.

APPENDIX D

CROSS VALIDATION SAMPLE STATISTICS

Table D1

Fit Statistics for all Piecewise LCGA Models Using MCRC Measures for District 2

Statistic	1 Class	2 Class	3 Class	4 Class
Unconditional				
AIC	11624.17	10866.32	10696.73	10659.09
BIC	11642.45	10907.46	10760.73	10745.94
ABIC	11629.75	10878.88	10716.27	10685.61
Entropy		0.80	0.76	0.69
LMR *(p value)*		0.00	0.01	0.05
Adjusted LMR *(p value)*		0.00	0.01	0.05
Conditional				
AIC	11492.94	10810.47		
BIC	11524.94	10883.60		
ABIC	11502.71	10832.80		
Entropy		10832.80		
LMR *(p value)*		0.00		
Adjusted LMR *(p value)*		0.00		
Conditional-Distal (F10RISK)				
AIC	12457.35	11522.81		
BIC	12493.92	11605.12		
ABIC	12468.52	11547.96		
Entropy		0.83		
LMR *(p value)*		0.00		
Adjusted LMR *(p value)*		0.00		

<u>Unconditional-Distal (F10RISK)</u>				
AIC	12588.58	11580.73	11373.13	
BIC	12611.44	11631.02	11450.86	
ABIC	12595.56	11596.09	11396.88	
Entropy		0.83	0.79	
LMR *(p value)*		0.00	0.00	
Adjusted LMR *(p value)*		0.00	0.00	
<u>Unconditional-Distal (W11RISK)</u>				
AIC	12509.92	11503.18	11280.87	
BIC	12532.78	11553.47	11358.60	
ABIC	12516.91	11518.54	11304.62	
Entropy		0.83	0.80	
LMR *(p value)*		0.00	0.00	
Adjusted LMR *(p value)*		0.00	0.00	
<u>Unconditional-Distal (S11RISK)</u>				
AIC	13139.49	12152.47	11926.60	11863.79
BIC	13167.02	12212.13	12018.38	11987.70
ABIC	13147.97	12212.13	11954.88	11901.96
Entropy		0.82	0.78	0.73
LMR *(p value)*		0.00	0.00	0.03
Adjusted LMR *(p value)*		0.00	0.00	0.03

Note. LMR = Lo-Mendel-Rubin test.

Table D2

Final Piecewise LCGA Models Using PRF Growth Measures for District 2

			Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Best Model fit		4-class	2-class	2-class	3-class	3-class	4-class
	n		714	714	715	715	715	727
Class 1	Fixed Effect							
		Intercept, η_0	7.93	8.94	8.64	8.02	7.92	7.21
		FW slope, η_1	1.45	2.44	2.44	2.00	1.90	1.06
		WS slope, η_2	**-0.76**	**-0.51**	**-0.56**	**-0.48**	**-0.42**	**-0.12**
	Covariate Effect							
		Intercept, η_0	-	-0.83	**-0.54**	-	-	-
		FW slope, η_1	-	**-0.50**	**-0.50**	-	-	-
		WS slope, η_2	-	**0.53**	**0.63**	-	-	-
	Random Effect							
		Error Variance	8.99	10.20	5.60	9.45	9.14	7.00
	Class Proportion		0.25	0.46	0.43	0.33	0.31	0.17
	PP		0.90	0.95	0.95	0.94	0.94	0.89
	Probability	Low Risk	-	-	0.24	0.15	0.19	0.08
		Some Risk	-	-	-	-	-	0.36
		High Risk	-	-	-	-	-	0.57
		Some & High Risk			0.76	0.85	0.81	-
Class 2	Fixed Effect							
		Intercept, η_0	9.99	14.65	14.39	12.34	12.14	9.77
		FW slope, η_1	3.24	1.01	1.12	2.02	2.14	3.05
		WS slope, η_2	**0.22**	0.42	**0.37**	0.44	**0.42**	**-0.59**
	Covariate Effect							
		Intercept, η_0	-	**-0.45**	**-0.25**	-	-	-
		FW slope, η_1	-	**-0.19**	**-0.20**	-	-	-
		WS slope, η_2	-	**0.47**	**0.45**	-	-	-

	Random Effect							
		Error Variance	6.25	4.93	9.97	5.60	5.76	8.05
	Class Proportion		0.26	0.54	0.57	0.46	0.46	0.32
	PP		0.76	0.93	0.95	0.90	0.91	0.84
	Odds Ratio	Class1 on SPED	-	0.27	-	-	-	-
	(Last Class as Reference)							
	Probability	Low Risk	-		0.88	0.77	0.84	0.54
		Some Risk	-		-	-	-	0.37
		High Risk	-		-	-	-	0.09
		Some & High Risk			0.12	0.23	0.16	-
Class 3	Fixed Effect							
		Intercept, η_0	14.12			16.78	16.68	13.64
		FW slope, η_1	0.99			**0.07**	**0.13**	1.30
		WS slope, η_2	0.65			**0.19**	**0.18**	0.73
	Covariate Effect							
		Intercept, η_0	-			-	-	-
		FW slope, η_1	-			-	-	-
		WS slope, η_2	-			-	-	-
	Random Effect							
		Error Variance	4.01			2.24	2.33	3.99
	Class Proportion		0.37			0.21	0.23	0.37
	PP		0.80			0.88	0.87	0.81
	Probability	Low Risk	-			0.97	**1.00**	0.85
		Some Risk	-			-	-	0.13
		High Risk	-			-	-	**0.02**
		Some & High Risk				0.03	**0.00**	-

Class 4	Fixed Effect			
		Intercept, η_0	17.43	17.23
		FW slope, η_1	**0.01**	**-0.03**
		WS slope, η_2	**-0.09**	**0.02**
	Covariate Effect			
		Intercept, η_0	-	-
		FW slope, η_1	-	-
		WS slope, η_2	-	-
	Effect			
		Error Variance	1.52	1.77
	Class Proportion		0.12	0.15
	PP		0.86	0.86
	Probability	Low Risk	-	0.95
		Some Risk	-	0.05
		High Risk	-	0.00
		Some & High Risk		-

Note. Model 1 = Unconditional, Model 2 = Conditional, Model 3 = Conditional-Distal, Model 4 = Unconditional-Distal (F10RISK), Model 5 = Unconditional-Distal (W11RISK), Model 6 = Unconditional-Distal (S11RISK), PP = Average Latent Class Probabilities for Most Likely Latent Class Membership. Non-statistically significant values are in boldface. FW = Fall-Winter. WS = Winter-Spring.

Table D3

Fit Statistics for all Piecewise LCGA Models Using PRF Measures for District 2

Statistic	1 Class	2 Class	3 Class	4 Class	5 Class	6 Class
Unconditional						
AIC	21208.58	20302.63	19513.06	19097.81	18886.12	18738.46
BIC	21226.86	20343.77	19577.05	19184.65	18995.82	18871.01
ABIC	21214.16	20315.19	19532.60	19124.32	18919.61	18778.93
Entropy		0.79	0.88	0.88	0.87	0.88
LMR *(p value)*		0.00	0.00	0.08	0.01	0.01
Adjusted LMR *(p value)*		0.00	0.00	0.08	0.01	0.01
Conditional						
AIC	21053.92	20186.06				
BIC	21085.91	20259.20				
ABIC	21063.69	20208.39				
Entropy		0.78				
LMR *(p value)*		0.00				
Adjusted LMR *(p value)*		0.00				
Conditional-Distal (F10RISK)						
AIC	22018.33	20910.551	20123.369	19686.503		
BIC	22054.90	20992.852	20251.393	19860.25		
ABIC	22029.50	20935.697	20162.486	19739.59		
Entropy		0.842	0.875	0.884		
LMR *(p value)*		0	0.0048	0.0075		
Adjusted LMR *(p value)*		0	0.0052	0.0079		

<u>Unconditional-Distal (F10RISK)</u>						
AIC	22172.98	20984.12	20203.27	19755.26	19538.12	19394.64
BIC	22195.85	21034.41	20281.00	19860.42	19670.72	19554.67
ABIC	22179.97	20999.49	20227.02	19787.39	19578.63	19443.54
Entropy		0.85	0.87	0.88	0.88	0.89
LMR *(p value)*		0.00	0.09	0.02	0.00	0.04
Adjusted LMR *(p value)*		0.00	0.09	0.02	0.00	0.05
<u>Unconditional-Distal (W11RISK)</u>						
AIC	22114.33	20939.56	20120.91	19695.16		
BIC	22137.20	20989.85	20198.64	19800.33		
ABIC	22121.32	20954.92	20144.66	19727.30		
Entropy		0.85	0.87	0.88		
LMR *(p value)*		0.00	0.02	0.04		
Adjusted LMR *(p value)*		0.00	0.02	0.04		
<u>Unconditional-Distal (S11RISK)</u>						
AIC	22515.66	21385.09	20558.41	20131.57		
BIC	22538.60	21435.57	20636.42	20237.12		
ABIC	22522.73	21400.64	20582.44	20164.09		
Entropy		0.83	0.87	0.88		
LMR *(p value)*		0.00	0.00	0.04		
Adjusted LMR *(p value)*		0.00	0.00	0.04		

Note. LMR = Lo-Mendel-Rubin test.

Table D4

Final Piecewise LCGA Models Using PRF Growth Measures for District 2

			Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Best Model fit		6-class	2-class	4-class	6-class	4-class	4-class
	n		714	714	715	715	715	727
Class 1	Fixed Effect							
		Intercept, η_0	19.53	82.05	54.24	19.61	49.63	49.16
		FW slope, η_1	15.66	23.63	14.41	15.66	15.64	15.74
		WS slope, η_2	**3.54**	2.70	3.33	**3.53**	2.94	3.82
	Covariate Effect							
		Intercept, η_0	-	22.65	-12.92	-	-	-
		FW slope, η_1	-	**-2.03**	**3.22**	-	-	-
		WS slope, η_2	-	8.86	**-0.15**	-	-	-
	Random Effect							
		Error Variance	189.41	659.19	376.62	190.22	404.44	399.73
	Class Proportion		0.03	0.52	0.14	0.03	0.15	0.15
	PP		0.99	0.93	0.97	0.99	0.96	0.97
	Probability	Low Risk	-	-	**0.02**	**0.00**	**0.03**	0.05
		Some & High Risk	-	-	0.98	**1.00**	0.97	0.95
		Some Risk	-	-	-	-	-	-
		High Risk	-	-	-	-	-	-
Class 2	Fixed Effect							
		Intercept, η_0	55.14	126.05	86.02	55.29	85.37	85.67
		FW slope, η_1	15.12	28.78	25.44	14.99	23.88	23.74
		WS slope, η_2	3.85	12.44	**1.14**	3.89	3.02	3.05
	Covariate Effect							
		Intercept, η_0	-	-76.34	**-4.45**	-	-	-
		FW slope, η_1	-	-11.19	-7.04	-	-	-
		WS slope, η_2	-	-9.33	**4.32**	-	-	-

	Random Effect							
		Error Variance	138.60	523.72	139.62	138.91	151.98	151.51
	Class Proportion		0.10	0.48	0.26	0.10	0.29	0.29
	PP		0.96	0.94	0.93	0.95	0.92	0.91
	Odds Ratio		-	0.57	-	-	-	-
	Probability	Low Risk	-		0.38	**0.02**	0.53	0.49
		Some & High Risk	-		0.62	0.98	0.48	0.51
		Some Risk	-		-	-	-	-
		High Risk	-		-	-	-	-
Class 3	Fixed Effect							
		Intercept, η_0	81.13		109.15	80.97	108.55	108.53
		FW slope, η_1	21.06		29.54	21.31	29.20	29.59
		WS slope, η_2	**1.80**		7.50	**1.65**	7.81	8.12
	Covariate Effect							
		Intercept, η_0	-		**-8.37**	-	-	-
		FW slope, η_1	-		**-1.10**	-	-	-
		WS slope, η_2	-		**2.84**	-	-	-
	Random Effect							
		Error Variance	118.03		185.17	121.20	186.62	186.32
	Class Proportion		0.20		0.40	0.18	0.37	0.37
	PP		0.90		0.91	0.92	0.92	0.91
	Probability	Low Risk	-		0.81	0.26	0.89	0.83
		Some & High Risk	-		0.19	0.74	0.11	0.17
		Some Risk	-		-	-	-	-
		High Risk	-		-	-	-	-

Class 4	Fixed Effect							
		Intercept	96.43		143.20	96.24	145.19	145.05
		Slope1	29.18		27.89	29.02	25.36	25.41
		Slope2	4.63		17.66	4.52	18.20	18.30
	Covariate Effect							
		Intercept, η_0	-		19.79	-	-	-
		FW slope, η_1	-		-26.56	-	-	-
		WS slope, η_2	-		**5.43**	-	-	-
	Covariate Effect							
		Intercept, η_0	119.69		401.49	117.08	411.76	412.37
		FW slope, η_1	0.26		0.20	0.27	0.20	0.19
		WS slope, η_2	0.90		0.95	0.88	0.95	0.95
	Probability	Low Risk	-		0.92	0.73	0.95	0.87
		Some & High Risk	-		0.08	0.27	0.05	0.13
		Some Risk	-		-	-	-	-
		High Risk	-		-	-	-	-
Class 5	Fixed Effect							
		Intercept, η_0	117.42			117.02		
		FW slope, η_1	29.08			29.08		
		WS slope, η_2	11.35			11.25		
	Covariate Effect							
		Intercept, η_0	-			-		
		FW slope, η_1	-			-		
		WS slope, η_2	-			-		
	Random Effect							
		Error Variance	157.91			158.94		
	Class Proportion		0.27			0.27		
	PP		0.89			0.89		
	Probability	Low Risk	-			0.82		
		Some & High Risk	-			0.18		

		Some Risk	-	-
		High Risk	-	-
Class 6	Fixed Effect			
		Intercept, η_0	150.99	150.61
		FW slope, η_1	24.37	24.51
		WS slope, η_2	18.46	18.44
	Covariate Effect			
		Intercept, η_0	-	-
		FW slope, η_1	-	-
		WS slope, η_2	-	-
	Random Effect			
		Error Variance	393.63	394.55
	Class Proportion		0.14	0.15
	PP		0.96	0.96
	Probability	Low Risk	-	0.92
		Some & High Risk	-	0.08
		Some Risk	-	-
		High Risk	-	-

Note. Model 1 = Unconditional, Model 2 = Conditional, Model 3 = Conditional-Distal, Model 4 = Unconditional-Distal (F10RISK), Model 5 = Unconditional-Distal (W11RISK), Model 6 = Unconditional-Distal (S11RISK), PP = Average Latent Class Probabilities for Most Likely Latent Class Membership. Non-statistically significant values are in boldface. FW = Fall-Winter. WS = Winter-Spring.

APPENDIX E

PERCENTAGE OF HIGH RISK STUDENTS CAPTURED BY THE LOWEST CLASS IN UNCONDITIONAL LCGA MODELS WITH DISTAL OUTCOMES

	MCRC				PRF			
	District 1		District 2		District 1		District 2	
Distal Outcome	Class Solution	% High Risk Students captured by Lowest class	Class Solution	% High Risk Students captured by Lowest class	Class Solution	% High Risk Students captured by Lowest class	Class Solution	% High Risk Students captured by 3 Lowest classes
F10RISK	4-class	95.50	3-class	92.86	4-class	86.36	6-class	94.64
W11RISK	4-class	82.35	3-class	93.18	4-class	72.55	4-class	68.18
S11RISK	4-class	80.83	4-class	73.00	4-class	85.83	4-class	68.00

REFERENCES CITED

Alonzo, J., Mariano, G. J., Nese, J. F., & Tindal, G. (February 2010). *Reliability of the easyCBM® reading assessments*. Paper presented at the Pacific Coast Research Conference, San Diego, CA.

Alonzo, J., & Tindal, G. (2007). *The development of word and passage reading fluency measures for use in a progress monitoring assessment system* (Technical Report No. 40). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2009). *Alternate form and test-retest reliability of easyCBM reading measures* (Technical Report No. 0906). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Artiles, A.J., & Trent, S.C. (1999). The overrepresentation of minority students in special education: A continuing debate. *The Journal of Special Education*, *27*, 134-142.

Biancarosa, C., & Snow, C. E. (2006). *Reading next—A vision for action and research in middle and high school literacy:* A report to Carnegie Corporation of New York (2nd ed.).Washington, DC: Alliance for Excellent Education.

Bilir, M. K., Binici, S., & Kamata, A. (2008). Growth Mixture Modeling: Application to reading achievement data from a large-scale assessment. *Measurement and Evaluation in Counseling and Development*, *41*, 104-119.

Boscardin, C. K., Muthén, B., Francis, D. J., & Baker, E. L. (2008). Early identification of reading difficulties using heterogeneous developmental trajectories. *Journal of Educational Psychology*, *100*, 192–208.

Buck, J., & Torgesen, J. (2003). *The relationship between performance on a measure of oral reading fluency and performance on the Florida Comprehensive Assessment Test.* Technical Report No.1, Florida Center for Reading Research.

Burns, M. K., Tucker, J. A., Hauser, A., Thelen, R. L., Holmes, K. J., & White, K. (2002). Minimum reading fluency rate necessary for comprehension: A potential criterion for curriculum-based assessments. *Assessment for Effective Intervention, 28*, 1-7.

Catts, H.W., Adolf, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49*, 278–293.

Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In H. Catts & A. Kamhi (Eds.), *Connections between language and reading disabilities* (pp. 23-36). Mahwah, NJ: Erlbaum.

Chall, J.S. (1996). *Stages of reading development*, New York, NY: McGraw-Hill.

Christ, T. J., Silberglitt, B., Yeo, S., & Cormier, D. (2010). Curriculum-based measurement of oral reading: An evaluation of growth rates and seasonal effects among students served in general and special education. *School Psychology Review, 39*, 447-462.

Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology, 98*, 394–409.

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

Deno, S. L. (2002). Problem solving as "best practices." In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (1), (pp. 37-56). Washington, DC: National Association of School Psychologists.

Denton, C. A., Barth, A. E., Fletcher, J. M., Wexler, J., Vaughn, S., Cirino, P. T., Romain. M., & Francis, D. J. (2011). The relations among oral and silent reading fluency and comprehension in middle school: Implications for identification and instruction of students with reading difficulties. *Scientific Studies of Reading, 15*, 109-135.

Education Week (2009) Portrait of a population: How English-language learners are putting schools to the test. Quality Counts. *Education Week, 28*, 53.

Fletcher, J. M. (2011). Identifying Learning Disabilities in the Context of Response to Intervention: A Hybrid Model. Retrieved October 28, 2011 from http://www.rtinetwork.org/learn/ld/identifyingld

Fletcher, J. M., Coulter, W. A., Reschly, D. J., & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. *Annals of Dyslexia, 54*, 304-331.

Fletcher, J. M., Francis, D. J., Shaywitz, S. E., Lyon, G., Foorman, B. R., Stuebing, K. K., Shaywitz, B. A. (1998). Intelligent testing and the discrepancy model for children with learning disabilities. *Learning Disabilities Research and Practice, 13,* 186–203.

Foorman, B. R., Breier, J. I., & Fletcher, J. M. (2003). Interventions aimed at improving reading success: An evidence-based approach, *Developmental Neuropsychology, 24*, 613–639

Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A., & Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities, 38,* 98–108.

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*, 239-256.

Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly, 25*, 33-45.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance of decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third grade high-stakes outcomes. *Scientific Studies of Reading, 5*, 257-288.

Hibel, J., Farkas, G., & Morgan, P. (2010). Who is placed into special education? *Sociology of Education, 83,* 312.332.

Hintze, J. M., & Silberglitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review, 34*, 372-386.

Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Structural Equation Modeling, 6*, 1-55.

Hughes, C. & Dexter, D. D. (2011). Universal Screening Within a Response-to-Intervention Model. Retrieved October 31, 2011 from http://www.rtinetwork.org/learn/research/universal-screening-within-a-rti-model

Jenkins, J. R. (2003, December). *Candidate measures for screening at-risk students*. Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention symposium, Kansas City, MO. Retrieved October 24, 2011, from http://www.nrcld.org/symposium2003/jenkins/index.html

Jenkins, J. R., Hudson, R. F., Johnson, E. S. (2007). Screening for at-risk readers in a Response-to-Intervention (RTI) framework. S*chool Psychology Review. 36*, 582-600.

Jenkins, J. R., & Jewell, M. (1993). Examining the validity of two measures for formative teaching: Reading aloud and maze. *Exceptional Children, 59*, 421-432.

Johnson, E., Mellard, D.F., Fuchs, D., & McKnight, M.A. (2006). *Responsiveness to intervention (RTI): How to do it.* Lawrence, KS: National Research Center on Learning Disabilities.

Johnson, E. S., Pool, J. & Carter, D. R. (2011). Screening for Reading Problems in Grades 4 Through 12. Retrieved from http://www.rtinetwork.org/essential/assessment/screening/screening-for-reading-problems-in-grades-4-through-12

Jordan, N. C., Kaplan, D., Olah, L N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development, 77*, 153-175.

Juel, C. (1991). Beginning Reading. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 759-788). Mahwah, NJ: Erlbaum.

Juel, C. (1998). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*, 437-447.

Jung, T. & Wickrama, K. A. S. (2008). An introduction to latent class growth: Analysis and growth mixture modeling. *Social and Personality Psychology Compass*, *2*, 302-317.

Kame'enui, E. J., Simmons, D. C., & Coyne, M. D. (2000). Schools as host environments: Toward a schoolwide reading improvement model. *Annals of Dyslexia, 50,* 31-51.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Kame'enui, E. J., Simmons, D. C., Good, R. H., & Harn, B. A. (2001). The use of fluency-based measures in early identification and evaluation of intervention efficacy in schools. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 307–331). Timonium, MD: York.

Kreisman, M. B. (2003). Evaluating academic outcomes of Head Start: An application of general growth mixture modeling. *Early Chidlhood Research Quarterly, 18*, 238-254.

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

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

MacMillan, D. L. & Siperstein, G. N. (2002). Learning disabilities as operationally defined by schools. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds). *Identification of Learning Disabilities: Research to Practice* (pp. XX-XX). Mahwah, NJ: Laurence Erlbaum.

Mather, N. & Kaufman, N. (2006). Introduction to the special issues, part one: It's about the what, the how well, and the why. *Psychology in the Schools, 43*, 747-753.

McGlinchey, M. T. & Hixson, M. D. (2004). Using curriculum-based measurement to predict performance on state assessments of reading. *School Psychology Review, 33*, 193-203.

Maxwell, L. (2009) Shifting landscape: immigration transforms communities, *Education Week, 28*(17). Retrieved October 23, 2011 from www.edweek.org/go/qc09

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*, 107–122.

Muthén, L. K., & Muthén, B. O. (1998-2007). *Mplus User's Guide. Fifth Edition.* Los Angeles, CA: Muthén & Muthén.

Murray, C. (1998). *Design and analysis of group-randomized trials*. New York: Oxford University Press.

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345-368). Newbury Park, CA: Sage.

Muthén , B., Brown, C. H., Booil Jo, K. M., Khoo, S., Yang, C., Wang, C., Kellam, S. G., Carlin, J. B., & Liao, J. (2002). General growth mixture modeling for randomized preventive interventions. *Biostatistics, 3*, 459-475.

Muthén, B., Khoo, S. T., Francis, D., & Boscardin, C. (2002). Analysis of reading skills development from kindergarten through first grade: An application of growth mixture modeling to sequential processes. In S. R. Reise & N. Duan (Eds.), *Multilevel modeling: Methodological advances, issues, and applications* (pp. 71–89). Mahwah, NJ: Erlbaum.

Muthén, B., & Muthén, L. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, *24*, 882–891.

Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods, 4*, 139-157.

Nese, J. F. T., Biancarosa, G., Anderson, D., Lai, C. F., & Tindal, G. (2012). Within-year oral reading fluency with CBM: A comparison of models. *Reading and Writing.* doi: 10.1007/s11145-011-9304-0

National Center for Education Statistics (2009). *The Nation's Report Card: Reading 2009* (NCES 2010–458). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

National Center for Education Statistics (2011). *The Nation's Report Card: Reading 2011* (NCES 2012–457). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs (NCELA), (2007). English Language Learners Face Unique Challenges. Retrieved from www.ncela.gwu.edu/expert/fastfaq/4.html.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of scientific research literature on reading and its implications for instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office. Retrieved from http://www.nichd.nih.gov/publications/nrp/report.htm

Nese, J. F. T., Biancarosa, G., Anderson, D., Lai, C. F., Alonzo, J., & Tindal, G. (2012). Within-year oral reading fluency with CBM: A comparison of models. *Reading and Writing, 25*, 887-915. doi: 10.1007/s11145-011-9304-0

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*(4), 535-569.

Phillips, L. M., Norris, S. P., Osmond, W. C., Maynard, A. M. (2002). Relative reading achievement: A longitudinal study of 187 children from first through sixth grades, *Journal of Educational Psychology, 94*, 3–13.

Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Roehrig, A. D., Petscher, Y., Nettles, S. M., Hudson, R. F, & Torgesen, J. K. (2008). Accuracy of DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology, 46,* 343–366.

Shadish, W. R., Cook, Thomas, D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Schilling, S. G., Carlisle, J. F., Scott, S. E., & Zheng, J. (2007) Are fluency measures accurate predictors of reading achievement? *The Elementary School Journal, 107*, 429-448.

Shapiro, E., Keller, M., Lutz, J., Santoro, L., & Hintze, J. (2006). Curriculum-based measures and performance on state assessment and standardized tests. *Journal of Psychoeducational Assessment, 24*, 19-35.

Shaw, R. & Shaw, D. (2002) *DIBELS Oral Reading Fluency-Based Indicators of Third Grade Reading Skills for Colorado State Assessment Program* (CSAP). (Technical Report) Eugene, OR: University of Oregon.

Speece, D. L. (2005). Hitting the moving target known as reading development: Some thoughts on screening children for secondary interventions. *Journal of Learning Disabilities, 38*, 487-493.

Speece D. L., Case L. P., (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology*, *93*, 735-749.

Speece, D. L., Case, L. P., & Molloy, D. E. (2003). Responsiveness to general education instruction as the first gate to learning disabilities identification. *Learning Disabilities Research & Practice, 18*, 147-156.

Speece, D. L. & Ritchey, K. D. (2005) A longitudinal study of the development of oral reading fluency in young children at risk for reading failure. *Journal of Learning Disabilities, 38*, 387–399.

Speece, D. L, Ritchey, K. D. Silverman, R., Schatschneider, C., Walker, C. Y., & Andrusik, K. N. (2010). Identifying children in middle childhood who are at risk for reading problems. *School Psychology Review, 39*, 258-276.

Stage, S. A., & Jacobsen, M. D. (2001). Predicting student success on a state-mandated performance-based assessment using oral reading fluency. *School Psychology Review, 30*, 407-419.

Sternberg, R. J., & Grigorenko, E. L. (2002). Difference scores in the identification of children with learning disabilities: It's time to use a different method. *Journal of School Psychology, 40,* 65–83.

Vaughn, S., Fletcher, J. M., Francis, D. J., Denton, C. A., Wanzek, J., Wexler, J., Cirino, P. T., Barth, A. E., & Romain, M. A. (2008). Response to intervention with older students with reading difficulties. *Learning and Individual Differences, 18*, 338–345.

Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice, 18*, 137-146.

Wayman, M. M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41,* 85-120.

Wesson, C. L. (1992). Using curriculum-based measurement to create instructional groups. *Preventing School Failure, 36*, 17-20.

Yovanoff, P., Duesbery, L., Alonzo, J. & Tindal, G. (2005). Grade-level invariance of a theoretical causal structure predicting reading comprehension with vocabulary and oral reading fluency. *Educational Measurement: Issues and Practice, 24*, 4-12.