2012

Repeated Reading: Testing Alternative Models for Efficient Implementation

Gregory Paul Lewis
Utah State University

REPEATED READING: TESTING ALTERNATIVE MODELS FOR

EFFICIENT IMPLEMENTATION

by

Greg Lewis

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Education

Approved:

D. Ray Reutzel, Ph.D. Major Professor	Jamison Fargo, Ph.D. Committee Member
Sylvia Read, Ph.D. Committee Member	Cindy Jones, Ph.D. Committee Member
Sarah Clark, Ph.D. Committee Member	Mark R. McLellan, Ph.D. Vice President for Research and Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2012

ABSTRACT

Repeated Reading: Testing Alternative Models for Efficient Implementation

by

Greg Lewis, Doctor of Philosophy

Utah State University, 2012

Major Professor: D. Ray Reutzel, Ph.D.
Department: Teacher Education and Leadership

Repeated reading has been used for over 30 years. In the publication of the National Reading Panel Report, repeated reading was listed as an effective strategy for developing fluency. Yet, repeated reading's efficacy has been recently questioned. Understanding the "how-to" of efficiently using evidence-based practices would allow teachers to deliver successful, time-sensitive instruction and intervention to students. This study was based on two research questions. First was a gain score (increase between a student's first read and their final repeated reading), a better model and therefore a better criterion than the currently popular criterion of reaching a set words-read-correctly-per-minute (WRCM) hot read, such as Samuels' criterion of 95 WRCM. The study's second question was exploring which demographic variables, such as age, ethnicity, gender, current reading ability, and socioeconomic status (SES), played a significant role in predicting the effectiveness of using weekly repeated reading scores as a predictor of benchmark reading measures at midyear and end-of-year outcome measures. The study used a unique theoretical multilevel path model to explore repeated reading.

A complex model was developed to study (a) the growth of a student's ability to read words with speed and accuracy and (b) how student demographic features affect growth rates. It was found that a hot read advancement criterion provided a better model fit than the hypothesized advancement criterion of a student's increase or gain between cold and hot reads. Student growth during repeated reading was found to be constant once a minimum WRCM criterion was reached. While repeated reading was shown to be a strategy that worked equally well for all students, the strategy was shown to be highly-effective for English-language learners and showed promise in helping to closing the achievement gap. Limitations were discussed and recommendations provided.

(114 pages)

PUBLIC ABSTRACT

Repeated Reading: Testing Alternative Models for Efficient Implementation

by

Greg Lewis, Doctor of Philosophy

Utah State University, 2012

An investigation was conducted to determine the best criterion for advancement to a new reading passage during the commonly used classroom strategy of repeated reading. Knowing when to move students to a new passage during the repeated reading process was considered of value to teachers in efficiently using student learning time. The study also explored the effect of demographic variables (age, ethnicity, gender, SES, and beginning reading ability) on predicting outcome measures based on repeated reading scores.

A complex multi-level structural equation model was developed to study (a) the growth of student's ability to read words with speed and accuracy and (b) how student demographic features affect growth rates. This multilevel structural equation model used included two phases: multilevel growth modeling and path analysis. It was found that using a hot (or final) read-advancement criterion provided a better model fit than the hypothesized advancement criterion of a student's increase or gain between cold (beginning) and hot (final) reads. Also, student growth during repeated reading was found to be constant once a minimum words-read-correctly-per-minute (WRCM) criterion of 25 was reached. Repeated reading was shown to be a strategy that worked equally well for all students with little variance explained by the slope of student growth. Furthermore, the strategy was shown to be effective for non-White learners and showed promise in helping to close the achievement gap.

DEDICATION

This work is dedicated to my wonderful wife, Debbie, who is my best friend and supporter. She has a pure love of all children. It is hoped this work will help more children as they learn to read.

ACKNOWLEDGMENTS

I would like to thank Ogden City Schools for allowing access to the data set for the research. The dedication and willingness of the district teachers to track student data with the hope of improved student achievement need to be recognized and are greatly appreciated.

My doctoral committee has been very encouraging and supportive throughout the process. I want to publically thank my chair, Dr. D. Ray Reutzel, for his guidance not just with the dissertation but as a mentor in research and instruction.

Last, I wish to thank my family who gave willingly so that I might complete this huge project. My wife and children have been patient and supportive; they have sacrificed much time and attention so I might complete this project.

Greg Lewis

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Carol laboriously read the words from the small reader, *"/s/………/i/………/d/………; /s/…/i/…/d/…; /s/…/id/; Sid."* Then more quickly but with a very flat *tone "is…" "not…" "on…" "the…"* Then again more slowly*, "/m/........., /a/…..…, /t/………; /m/…, /a/…, /t/…, /m/…/at/, /m/…/at/, mat.* The teacher, Mrs. Gonzales, replied, *"Good, now read the whole sentence."* Carol hesitantly reads, *"S…id, is, not, on, the m…at."*

This teacher-student interaction is common and indeed expected at the end of the year in kindergarten or the beginning of the first-grade year in classrooms across the country. As a kindergartener, Carol would be learning how to recognize letters and then say the sounds in sentences from left to right blending them together to pronounce words. As Carol follows this process, she would be progressing through normal stages of learning to read (Chall, 1976; Ehri, 1985).

Unfortunately, Carol is not in kindergarten or first grade. Carol is a third-grade student. Because of high transiency, she has never consistently attended one school long enough to have experienced sustained instruction in reading. Since her attendance during the first 3 years of schooling has been marked by frequent relocation, Carol does not qualify for any special services under current federal guidelines. Carol's only chance will be the interventions the teacher will provide.

As an experienced teacher in a high-risk school where students face many challenges caused by poverty, Mrs. Gonzales has taught many students like Carol. She knows there is a narrow window of opportunity to help Carol before the family is once again uprooted and Carol starts at another new school (Champagne, 2008). Mrs. Gonzales knows she needs to apply efficient, evidence-based strategies to help Carol become a fluent reader. But as a teacher, she has many questions: What is the best intervention for Carol? How should the intervention be most effectively administered? How will Mrs. Gonzales know it is working?

Many students face what Kameenui (1993) referred to as the "tyranny of time." There is a window of opportunity as a teacher works with struggling students. For some students, such as those who are highly transient, the time available for intervention and instruction is very short. The consequence of using the wrong instruction/intervention or even ineffectually using proper instruction/interventions can have long-lasting negative results in the life of a child. Teachers too often are simply given a list of possible interventions. They know the strategies have been shown to be effective but the teacher often lacks basic knowledge such as: when should the strategy be used, what is the appropriate duration of strategy application, and what is the most appropriate materials (passages) to use. As a result, it is not surprising that strategies are often misapplied and outcomes are very different from classroom to classroom and from those desired.

Without this defining guidance, the teacher is left with only his or her own judgment as guide. Mrs. Gonzales sighs and selects an intervention she believes will help Carol—a well-used, scientifically based reading strategy known as "repeated reading" that she has seen used and tried herself with other challenged students. Did she make the proper choice? And, how can she most efficiently use the strategy? Available research has limited answers to her questions.

Oral Repeated Reading

Oral repeated readings have been used in various ways over time to develop student's reading ability (Rasinski, 2006). Recently repeated reading has been associated more closely with developing reading "fluency." In the National Reading Panel (NRP) report (2000), repeated reading was listed as a strategy for the effective student practice of guided, repeated oral reading with teacher feedback. Multiple studies have shown the effectiveness of repeated reading with high risk groups (Freeland, Skinner, Jackson, McDaniel, & Smith, 2000; Mastropieri, Leinart, & Scruggs, 1999). The publication of the NRP report led to a "rediscovery" of fluency as an important part of the curriculum. This led to the wide-spread implementation of oral repeated reading in a variety of classroom settings. However, the practice was implemented with varied fidelity and techniques. Repeated reading is currently a commonly used strategy in schools (Gamse, Bloom, Kemple, Jacob, & Institute of Education Sciences, 2008). Yet research has been limited to only certain aspects of the strategy.

Current Reviews of Repeated Oral Reading

Recently, two meta-analysis studies have reviewed the research on repeated reading (Kuhn & Stahl, 2000; Therrien, 2004). While positing that oral repeated reading resulted in significant growth, both studies identified weaknesses in the research including:

- A majority of the studies had baseline or multiple baseline designs. The design of these studies limits the research to a few students. Also, Kuhn and

Stahl posit that reliance on one design is "problematic" when extrapolating research to the entire population.

- Relatively few studies had control groups. Lack of controls limits the ability to make nontreatment comparisons.
- The greater part of the studies involved only children with learning disabilities. Therefore, ability to extrapolate all findings to the general population is limited.
- While the learning-disabled children studied showed significant growth using repeated reading, these students still fell further behind their peers in overall performance. The widening of this achievement gap calls into question the effectiveness of current practice in repeated reading.

In order to determine if any new findings have been published since these meta-analyses or if significant studies were not included, two different databases were searched for reported research on repeated reading: a search of digital dissertations and a wider Educational Research Information Center (ERIC) electronic database search of all articles on repeated reading. Exclusionary criteria for the search were articles that were only descriptive of repeated reading and other meta-analysis. The inclusionary criterion was simply an article with a recognizable type of common research design (i.e., quantitative, qualitative, or single subject). A vote counting procedure was used to organize the data for analysis. In the review of the articles reporting original research, no attempt to apply criteria to determine the quality of the research was made. If an experimental quality requirement was used, few articles would be included as indicated when Chard (2009) previously reviewed articles and found that no research meeting rigorous experimental standards was currently available. This analysis was simply to determine the focus of the studies and student demographics included in the sample.

In the electronic bibliographic search of Dissertation Abstracts Internationals (DAI) conducted in March of 2009, 154 dissertations were found which focused on oral reading fluency. Of those 154 dissertations, 85 met the inclusionary criteria. The research questions addressed in these 85 dissertations on the benefits of repeated readings were grouped into four categories.

1. Forty-one studies (48%) focused on the effects of using the repeated reading strategy on different types of disabled and/or high risk groups.
2. Nineteen studies (22%) focused on the effectiveness in combining repeated reading with other reading fluency strategies.
3. Eight dissertations (9%) explored the effects of different types and levels of reading materials used when repeatedly reading.
4. Seventeen studies (20%) provide comparisons of repeated reading with other reading interventions.

None of the reviewed dissertations focused on the components or steps of repeated reading. The dissertations only looked at the types of disabled students with which it was effective, other strategies with which it could effectively be combined, which types of reading materials were the most effective, and whether it was as effective as other strategies.

In September of 2009, a wider search including the phrase “repeated reading” was conducted using the ERIC database. There were 150 abstracts found. Articles based on the dissertations search previously tallied were not excluded from this search. Sixty of the articles were either simply descriptive or a meta-analysis of previous research. The remaining 90 articles were tallied into categories.

1. Fifty-nine of the studies (66%) focused only on at-risk students. Most of these designs used a single-subject research model and measured the effect of repeated reading with students with different disabilities.
2. Thirty-eight studies (42%) studied the effectiveness of combining repeated reading with another strategy.
3. Thirteen of the studies (14%) compared the use of repeated reading against the effectiveness of another strategy.
4. Thirteen of the studies (14%) researched the effectiveness of procedures used by teachers in using repeated reading.
5. Twenty-six of the articles (29%) studied variables such as the audience and the type of text used.

With the exception of the 13 studies on effective procedures, the research questions listed in the ERIC search paralleled the same research questions found in the review of dissertations. All of the reported research described the same basic repeated reading process as originally outlined by Samuels (1979). In describing the 13 articles relating research on specific components of the procedures for repeated reading that is the focus of this study, the articles posited the following findings/recommendations.

1. Accuracy should be the focus of the first reading. Speed should be the focus

starting on the second reading (Bell, 1990).

2. Distributed practice produces a higher return than massed (one sitting) practice. Repeated reading is best spread over days rather than all practice on one day (Durgunoglu, 1993; Krug, 1990).
3. One-on-one practice with an adult is more effective than individual practice (Folley, 1987; Hindon & Paratore, 2007; Nichols, Rupley, & Rasinski, 2009). Paired practice is more effective than individual practice (Fuchs et al., 2001).
4. Repeated reading of passages in context is more efficient than practice with lists of words (Greene, 1988; Therrien, 2004).
5. Therrien (2004) found that using a criterion of three successive improvements was more efficient than the criterion of 90 WRCM in determining students' readiness to move to a new passage for second-grade students.

From this review of the current research, many studies supported the NRP conclusion that repeated reading helps students read faster and with more accuracy. However, the research on the most effective techniques for implementation of the strategy was inconclusive. While repeated reading has produced results, lingering questions remain to guide teachers as to how this strategy can efficiently be used and the appropriate time for its use. Information about the relative effectiveness of each component of repeated reading could prove invaluable to a teacher, such as our Mrs. Gonzales, in a busy classroom.

In the original procedure proposed by Samuels (1979), students reread a passage until the passage was read with 90% accuracy and prosody (rhythm, pitch, and intonation of speech) sounded like someone speaking. No measurable criterion for prosody was used. The teacher had to make a judgment on whether the student was ready to move on to a new passage. Later, Samuels increased the accuracy criteria to 95%. Because of the need for direct interaction between student and teacher on every repeated reading, this criterion is seldom used in current classroom practice. In this study, three different practices to set criteria for moving to a new passage were observed by the researcher being used within the sampled second through fifth-grade classrooms: (a) students repeatedly read a passage until the accuracy of WRCM in the passage was higher than 90-95%; (b) students reread passages a set number of times; and (c) students reread until they reach a target goal—reaching a set increase criteria.

In applying the second finding listed in the ERIC search, the research reports that if the "set number of times" criteria is used a student should reread a passage more than three times but less than seven (Kuhn & Stahl, 2000). An example of the current use of this criterion can be found in a currently popular fluency program. In "The Six-minute Solution" (Sopris West Educational Services, 2005), students reread a passage once per day for 5 days and then move to a new passage.

If the third or "target goal" criterion is used, students reread until their WRCM reaches a number predetermined by the teacher. This final WRCM is commonly referred to as the student's "hot" read. In the "Read Naturally" program, teachers are encouraged to set the goal of a thirty word increase between cold and hot read below fifth grade. Above fifth grade a 40-word increase between reads is recommended (Read Naturally, 2010).

Current research on which criterion is most effective in determining when a student should move to a new passage during repeated reading has been very limited. While a set number of repeated readings of a passage has been shown to be more effective than reading for a percent accuracy (Therrien, 2004) as a criterion for moving on to a new passage, no research is available to show if a predetermined gain score or increase between cold (initial) and hot (final) repeated readings might be a better criterion for moving to a new passage. No evidence has been published for the effectiveness of a target goal such as Read Naturally's (2010) goal for increases between cold and hot reads. Also, the only demographic information currently used in practice by one of the fluency programs currently available to help teachers set student goals is whether the student is above or below fifth grade.

If the teacher could more accurately know how much growth is required between the first "cold" read and the last "hot" read to reach a predetermined improvement target over time, the teacher could be much more prescriptive and therefore more efficient in the use of time and material. For Carol, our example student who might only be in one school for a short period of time, a time-effective strategy is critical. In common practice observed in the target district, the strategy of repeated reading has been administered in a uniform dose to all students regardless of their need or individual progress. What if teachers could more effectively prescribe the method and duration of this strategy? Also, is this criterion equally effective with all students regardless of their demographics such as gender, socioeconomic status (SES), and English-language development?

A Proposed Model

Using the theory that students get better at reading through practice originally proposed by Samuel's Automaticity Theory, a model can be made of expected growth during repeated reading (see Figure 1).

As seen in Figure 1, a student, repeatedly reading a passage, increases the number of words read correctly per minute (WRCM) each time they read a passage as indicated by the line marked C, which represents the "Gain Score." When a new passage is introduced, the WRCM drops as the students are unfamiliar with the new words and sentence structure—indicated by line D. However, each time a new passage is introduced, the students should retain some new knowledge and therefore will begin the

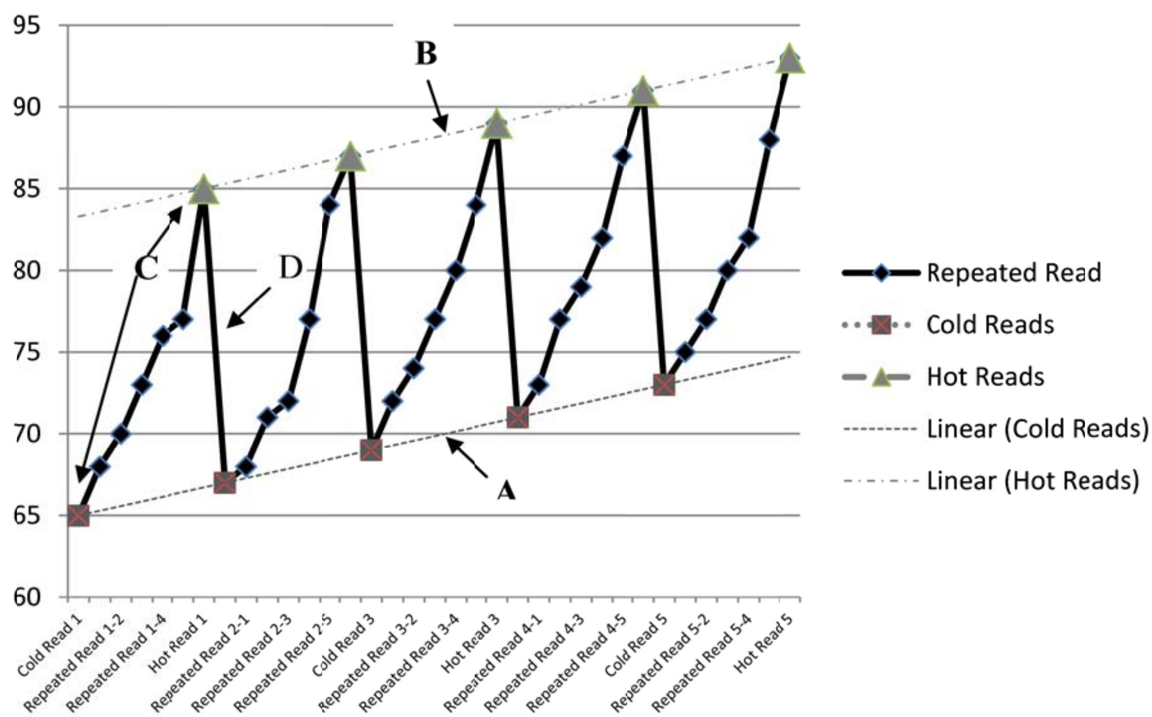

Figure 1. Expected model of WRCM using repeated reading.

new passage at a slightly higher WRCM that on the previously practiced passage. This growth between initial or “cold read” is represented by line A. Similarly by the same reasoning, there should be a gain between the final or “hot read” for each passage read as represented by line B.

In a “cold” read, students have never seen the text passage before. The student reads the text for 1 minute and the WRCM is calculated. The student then rereads the passage several times calculating the WRCM and recording the score for each read. Based on Samuel’s automaticity theory, each time a student rereads the same passage, more of the mental process moves to the “automatic” stage and less cognitive capacity is required. Therefore, the student is able to read more words correctly within the one minute time period. The final repeated reading of the passage, referred to as a “hot” read, should have the most words read correctly of the series because students should have become more automatic due to less cognitive capacity being used to blend words as the difficult words are becoming sight words.

This premise posits a pattern should be created where the WRCM increases from the cold read each time the student repeatedly reads the passage. The WRCM should drop again when a new passage is introduced due to the cognitive capacity again required to decode and blend unfamiliar words. However, the learning from the previous passage should affect the new cold read positively as students have become more skilled in decoding and blending and more automatic in their recognition of words.

In this model, there are three different measures which could possibly be used to predict student growth: the cold read score or growth between cold reads as represented by line A in Figure 1; the hot read score or growth between hot reads as represented by line B in Figure 1; or the gain score or average growth between each cold and hot read as represented by line C of Figure 1. If the premise of this model is correct and provides a good model fit for the data, it should be possible to posit which indicator best predicts or models student growth by comparing path models.

Purpose of the Study

The purpose of this study was twofold. First, it would be valuable for teachers to know which indicator or measure during repeated reading more accurately predicts student reading growth. Thus the most accurate advancement criteria of when it would be best for students to move to a new repeated reading passage could be identified. Second, an understanding of which student variables impact growth during repeated reading is of value to teachers providing targeted and time-sensitive interventions for students. As an operating hypothesis for this study, it is predicted that actual student growth or the gain scores between cold and hot reads should provide the best model fit in modeling learning during repeated reading as designated by Line C of the model. Thus, the gain scores between cold and hot reads would be a more effective criterion for moving to a new practice passage than the criterion of reaching a specific WRCM on the last or “hot” repeated read or tracking the gain between initial cold reads. By determining the best fit criterion, it will help teachers determine the optimal “dosage” for repeated reading to achieve maximum growth possible in a shorter amount of time.

To illustrate the benefit of more efficiently determining optimal dosage for fluency practice one can think of a doctor prescribing a drug, what level must be reached in order to gain maximum benefit? And, at what dosage level does further use of the drug become useless or even dangerous? Likewise, using the best fit criterion will save time and increase effectiveness of the repeated reading strategy.

This study used a mixed model analysis. Multilevel structural equation modeling was used in the analysis of variables that affect the model. Student growth, or the measure of words read correctly per minute (WRCM) over time, was the first level of the model. The student demographic variables were the second level of the model with variables clustered by teacher. The results were then placed within a path model to determine goodness of model fit in predicting two outcome measures: the midyear DIBELS Oral Reading Fluency (DORF) and the high-stakes state Adequate Yearly Progress (AYP) reading assessment. The DORF is a benchmark assessment that has been normed with a large sample size and is used in many school districts across the United States as well as the sample district to identify students needing reading interventions (Cummings, Kennedy, Otterstedt, Baker, & Kame'enui, 2011). The high stakes AYP reading assessment is a criterion referenced test (CRT) that is state normed and intended to determine if students have learned the reading objectives outlined in the state core curriculum. The high-stakes assessment is equated to make comparison of student growth between grade levels and from year to year possible.

If an indicator such as the gain score measurement provides the best model fit, teachers may use this information to prescribe optimal repeated reading practice. For example, a student improves x-number of words per minute between the first and last read of a passage. This becomes their "gain score." Over time, the gain score should predict an increase y-number of words in their reading ability of unpracticed passages (cold read). Teachers would be able to mathematically prescribe the duration of the intervention necessary to obtain a set WRCM growth.

Also, teachers may use the effects of different demographic variables in their prescription. Being able to make a more precise prescription of the amount of repeated reading practice needed to reach specific within text fluency will allow teachers to assign just enough practice to achieve the needed benefit without wasting student time in over practice that will not transfer to unfamiliar grade level text reading (end-of-year benchmarks). If the null hypothesis from questions one and two prove to be untrue (see Table 1), recommendation for effective use of the repeated reading strategy can be made for classroom use. These recommendations should be of great value to classroom teachers in using repeated reading to improve reading rate and accuracy for all students

Table 1

Research Questions and Hypotheses

Question	Hypothesis/null hypothesis
1. Is the gain score between cold (initial) reads and hot (final) repeated reading of a passage a better model and therefore a better criterion than the currently popular criterion of reaching a set WRCM hot read or a WRCM cold read?	*Hypothesis:* Since the gain score between cold and hot reads should be closely correlated to individual student growth, it should correlate more closely to benchmark measures than the currently practiced criterion of reading a passage to a predetermined criteria. *Null hypothesis:* There will be no difference between path models for the criterion of the gain score between hot and cold reads and the criterion of cold read performance or the criterion of hot read performance in predicting outcome reading scores.
2. What factor do the demographics such as age, ethnicity, gender, current reading level and SES play in predicting the effectiveness of using the increase between cold and hot reads as a predictor of benchmark reading measures?	*Hypothesis:* Individual characteristic of students directly influence the rate at which students' progress in learning to read. Many student demographics are closely inter-related. The characteristics that should best predict oral reading growth are SES and current reading level. *Null hypothesis:* There will be no difference between the independent demographic variables in the regression equation predicting oral reading growth during repeated reading.

CHAPTER II

REVIEW OF THE LITERATURE

In this review of the literature on repeated reading, two major topics have been studied. First, in order to fully understand the repeated reading strategy and its relationship with developing fluency, it is necessary to have a working definition of fluency. The currently accepted definitions have developed over time and within historical settings. With the intention of personally defining fluency and its relationship with repeated reading, the theoretical underpinnings need to be understood. As repeated reading has developed as a strategy over time, benefits and limitations of implementation can be outlined. Second, a discussion on the research describing the procedure of repeated reading is presented. The purpose of this second discussion is to outline the current knowledge about repeated reading procedures and situate this new research into that body of knowledge.

Developing a Theory for Repeated Reading

History of Fluency

Oral reading fluency has been part of education in the United States since schools began in the original colonies and settlements. In early society, books were scarce and the ability to read to others was highly treasured. In this setting, elocution was the goal of reading instruction in early schools (Rasinski, 2006). Being able to read fluently was viewed as being able to say the words clearly in a way that could inform and entertain others.

As books became more available, being able to read silently for learning and enjoyment rose in importance (Rasinski, 2006). Previous to this shift, students had been trained to only recite well—without necessarily understanding what they were reading. In 1891, Horace Mann claimed that more than eleven twelfths of all the children in reading classes did not comprehend the meaning of the words as they read (Rasinski, 2006).

Soon the quantity of text available shifted. More books were accessible and the purpose of oral reading changed. Rather than learning to recite text to entertain others, students were encouraged to read silently with a focus on comprehension and personal enjoyment. Instead of being the end result of instruction, oral reading's primary use was as a method of checking students' word recognition and speed after silent reading" (Eldredge, Reutzel, & Hollingsworth, 1996). At this point, oral reading for instruction or entertainment was no longer seen as the purpose of reading. Consequently, silent reading took its place as the most important goal (Allington, 1983).

Automaticity Recognized

Even with the shift to silent reading, references to fluency, still defined as good elocution, continued to appear in the literature. At the turn of the century, fluency instruction began to be defined as the ability to read words with ease—or with accuracy and speed. Huey (1908, p. 104) outlined the idea that was to develop into automaticity theory when he stated:

> Perceiving being an act, it is, like all other things that we do, performed more easily with each repetition of the act. To perceive an entirely new word or other combination of strokes requires considerable time, close attention, and is likely to be imperfectly done, just as when we attempt some new combination of movements, some new trick in the gymnasium or new serve in tennis. In either

> case, repetition progressively frees the mind from attention to details, makes facile the total act, shortens the time, and reduces the extent to which consciousness must concern itself with the process.

In short, Huey suggested repetitious readings were necessary to develop a readers' ability to decode text with "automaticity," meaning without excessive cognitive effort. He suggested that just as an athlete or a musician practices to improve performance, readers should practice to perfect their reading performance. There is little evidence that his theory about the value of repeated readings enjoyed wide-spread acceptance during his lifetime. These instructional ideas were almost forgotten as behaviorism shifted the focus of instruction away from the mental processes happening within the brain. Fluency instruction remained buried until the cognitive revolution occurred mid-century.

In 1974, LeBerge and Samuels published their landmark study that formalized the procedure of repeated reading as a way to practice fluency. In order to help struggling special education students, they developed a procedure in which students repeatedly read an assigned passage until the student reached a predetermined goal. Since their articles were published, repeated reading has been used with a variety of struggling students (Chard, Vaughn, & Tyler, 2002). The procedures they developed are still used widely in classroom in much the same format they proposed (Kuhn & Stahl, 2000).

Rediscovering Fluency Instruction

Still, fluency remained a low priority in daily classroom instruction. In 1983, Allington referred to fluency as "the neglected goal of reading instruction." Silent reading maintained its dominance and typically the only oral reading taking place during a school day was round robin reading of basal stories. However, fluency regained prominence in reading instruction when it was included as one of the five pillars of reading instruction listed in the NRP's (2000) report. While the report was never intended to create a comprehensive list of all components of an effective reading program, the panel attempted to compile a list of effective strategies that had been tested and documented in educational literature. Fluency again became a hot topic and is once more included in teacher professional development and in daily reading instruction.

Repeated Reading Effectiveness Research

The publication of the NRP (2000) report led to a "rediscovery" of fluency and subsequently the widespread implementation of oral repeated reading in a variety of classroom settings. When the report listed repeated reading as an effective strategy for building fluency by using guided, repeated oral reading with feedback, multiple studies showing the effectiveness of repeated reading were provided. Positive effects of repeated reading found in current literature are listed in Table 2.

A Developing Theory for Repeated Reading

While much documentation exists to show the effectiveness of repeated reading, understanding why it is an effective strategy is necessary to determine how it can best be used in the classroom. The supporting theory has developed over time.

Automaticity as a Theory

When Samuels (LaBerge & Samuels, 1974; Samuels, 2006a, 2006b) proposed the repeated reading strategy, he posited automaticity theory as an explanation for how good

Table 2

Positive Impact of Oral Repeated Reading

Finding	Researchers
Effective with high-risk groups	Freeland et al. (2000); Mastropieri et al. (1999)
Improvement in student's reading rate and accuracy	Chomsky (1976); Dahl (1974); Samuels (1979); Dowhower (1987)
Improvement in students' reading comprehension	Dowhower (1987); O'Shea, Sindelar, & O'Shea (1985)
Increase in students' reading vocabulary	Elley (1989); Koskinen & Blum (1984)
Increase in students' confidence about their ability to read	Koskinen & Blum (1984); Topping (1987); Trachtenberg & Ferruggia (1989)
Similar gains not only in below-level readers, but also among instructional- and mastery-level readers	O'Connor, White, & Lee Swanson (2007); Sindelar, Monda, & O'Shea (1990); Therrien (2004)
Effectively used in a variety of formats:	
Teacher-led	Dowhower (1987); O'Shea, Sindelar, & O'Shea (1987)
Paraprofessional-led	Mercer, Campbell, Miller, Mercer, & Lane (2000)
Teacher-led	Dowhower (1987); O'Shea et al. (1987)
Peer-led	Rasinski, Padek, Linek, & Sturtevant (1994)
Whole-class	Sindelar et al. (1990)
Small-group	Kuhn (2005)
Pull-out	Simmons, Fuchs, Fuchs, Mathe, & Hodge (1995)
Rate increases	Rasinski (1989)
Miscues decrease	Rasinski (1989)
Increase in ability to read meaningful phrases	Rasinski (1989)
Effective regardless of student level of expertise	NRP (2000)

readers became good readers. As Samuels was out jogging, he began thinking about the struggle his special education students were experiencing while reading. In thinking about the cause, he knew these students were lacking in both reading accuracy and speed. He thought about who the best trained and skilled people were in society and concluded it was athletes and musicians. These athletes and musicians had developed the ability to perform with a high degree of accuracy and speed. The next question was, how did they become so skilled? The answer to the question seemed obvious and was a restatement of Huey's previous observation. The athletes and musicians spent many hours repeating the task correctly until it became "automatic." He posited that struggling readers could improve their reading by using the same techniques. This automaticity theory neatly fit into the information-processing theory being espoused as part of the cognitive revolution taking place concurrently. Kuhn and Stahl (2000) stated that automaticity theory accounts for two main components of fluent oral reading: accurate decoding and sufficient rate.

Automaticity Theory as Part of Cognitive Theory

According to Samuels (2006a, 2006b), automaticity theory as part of the information processing theory makes five assumptions. First, the human mind has only limited capacity to perform difficult tasks. The information processing theory posits that humans can only process five to nine items of information at one time in the brain's working memory. Second, in order to perform difficult tasks, mental attention must be expended. Simply put, some tasks require more attention because of their difficulty or novelty. Third, the effort caused by difficulty or novelty consumes most if not all of the capacity of working memory. Fourth, with continued practice, the cognitive effort, or conscious attention, required to perform these tasks becomes less, thus opening up some capacity in working memory to do more than just to say words. Last, when the amount of effort used to perform a task is reduced through practice, the person is able to process a parallel task in addition to saying words, such as understanding what the word means.

A good example of this process can be seen by observing kindergarten students tying their shoes for the first time. The mental effort required is obvious. It is easy to see the concentration needed to complete the mental process the children are trying to complete. One can almost hear them saying, “First make a bunny ear, and then loop around….” At first, the process requires the child’s full concentration—any distraction will cause them to start over. In contrast, an adult tying a shoe lace does so quickly with little or no conscious attention. In fact, the adult will rarely remember even tying their shoes because they were thinking of something else at the time. The process has moved to “automaticity” and requires little cognitive thought.

In the opening example of chapter one, the struggles of a beginning-reading, third-grade student were illustrated. Like a student learning to tie shoes, the beginning reader must use most of the available cognitive space to decode the graphemes and verbalize the words. Indeed, little space is left over for meaning. If the beginning reader is asked to explain what they just read, the student will usually respond, “I don’t know.” Automaticity theory’s attempts to explain how to move processes to the automatic level (little or no attention demands) so cognitive capacity is freed up for the parallel task of comprehension.

As a cognitive theory, information processing theory is quite simple. However, reading is more complex than can be explained by a simple linear model. Samuels’ automaticity theory and later Topping’s deep fluency model attempt to provide a deeper explanation of the cognitive processes involved in learning to read. All of these theories and models discuss the need for learners to move mental processes to an automatic level freeing cognitive space for more difficult tasks.

Student failure to automatize mental activities leads to difficulties with higher level tasks. Stanovich (1980) posited in his interactive compensatory explanation that students compensate for weak automatic word-decoding processes by relying upon context-bound strategies that also require significant amounts of cognitive resources.

An example of student's ineffectively compensating for demanding decoding processes can be found in a common, misguided fix-it strategy for reading. During the height of whole language instruction, students were frequently told to look for context clues to help them with decoding words and creating meaning, or "look at the picture" to help figure out the word or understand the meaning. Recently, it was found that only poor readers "look at the picture" for help with words and meaning. Pictures can be misleading but more importantly, looking at the pictures to help with decoding words requires more attention capacity than simply decoding the words. Given the nature of the English language, using context cues is effective and efficient only 25% of the time (Topping, 2006b). Good readers use more efficient processes in decoding and comprehending.

Alternative Explanatory Theories

Other theories have been proposed to explain the results attributed to repeated reading. Moskal (2006) posited that the use of a partner when repeated reading moves the process into the realm of collaborative learning. Moskal also purports that as students chart their improvement, they become self-managed learners which supports what Dewey saw as optimal for increasing work ethic and motivation. Thus, the process and the resulting gains can be attributed to more of a constructionist bent.

Schreiber (1980) posited that instead of developing a pattern for thinking, as proposed by the Automaticity Theory, an alternate explanation would be that practice gave more awareness to the "prosodic" features of text.

Finally, from a constructivists view, repeated reading is viewed as a process that is artificial because in "the real world" learners do not repeatedly read passages. Therefore, time should be focused on less artificial activities.

Since theories guide research and implementation, it is important to keep these alternative explanations as part of the discussion. Reading is very complex. It is important to remember that fluency is only one piece of the puzzle. As Topping (2006a) stated, "Fluency has little value in itself. Its value lies in what it enables" (p. 124). One theory simply cannot account for the complexity involved in reading.

Creating a Model for Repeated Reading

Since LeBerge and Samuel's (1974) proposed the automaticity model, others have developed and deepened the idea expanding new models to explain how fluency develops. For example, Perfetti (1985) proposed that lower level processes (like word identification) must reach a minimum performance level before higher level processes (such as questioning the author) can be performed simultaneously during reading (verbal efficiency theory).

Another example of a deeper explanatory model was formed when Logan (1997) posited his instance theory of automatization—a model for memory retrieval and how it functions in fluency. First, a student's focus is on storing details of the letter or word in memory (obligatory encoding). Second, the student must focus enough attention on the details to retrieve previous exposures to the stimulus from memory (obligatory retrieval).

Next, the student codes and stores a new memory “trace” for each exposure. Memory traces are laid down each time the task is executed. As the number of trials on a task increase, the strength of the number of memory traces also increases (instance representation). Finally, the student’s reaction time decreases as a result of the practice and repetition. The reaction time is dependent on: the amount of practice, the level of consistency in the task environment, and the number of relevant instances of the task recorded in memory (power law). As the number of memory traces increases, performance becomes reliant on memory retrieval rather than problem solving.

A synthesis of ideas on how fluency develops can be found in Topping’s deep fluency model (Topping, 2006a). Topping proposed three levels of fluency development: surface, strategic, and deep. In the surface fluency stage, students develop automaticity. They are able to read with appropriate speed and accuracy. As readers move to the strategic fluency stage, they achieve basic comprehension and they are able to read with appropriate prosody (expression). In the final or deep fluency stage, readers develop what Reutzel (2006) called metafluency. Metafluency describes the process readers undergo as they monitor their accuracy and comprehension and adjust the rate and prosody to compensate.

Over time, a sophisticated model has developed to explain how a strategy like repeated reading helps students develop fluency. This model is deeply rooted in cognitive theory as it attempts to illuminate the mental processes a student uses as they read fluently.

The Strategy of Repeated Reading

The Process

Samuels (2006b) described how his and LeBerge's automaticity theory moved to an instructional procedure. He reminisced, "Our article on automaticity was only a theory, with no practical suggestions in it, and I had always thought that a good theory should have some practical aspects" (p. 25). Thus he connected repeated reading as an application process for automaticity theory.

Samuels (1979) reported his original methodology for repeated reading as he worked with special education students. The student repeatedly read short passages (about 250 words) until they reached a speed of ninety-five words per minute. Once they had achieved the minimum rate criteria, they moved on to a new passage. Each time they read, they recorded the number of words read per minute on a chart. Thus, they were able to track their improvement. Samuels reported an increase in accuracy, speed, and expression; a decrease in the number of readings required to reach criterion; and the initial or cold reads were better—an indication of transfer.

Since Samuels' initial trial, multiple experiments have reported features found to be important in the repeated reading process (Blachowicz, 2000; Chard et al., 2002; Kuhn, 2000; NRP, 2000):

1. Brief daily practice
2. Repeated oral reading of passages
3. Overlap of shared words across passages
4. Consistency in text context

5. Controlled text difficulty
6. Provision of corrective feedback
7. Teacher-modeled text reading
8. Audio-taped modeled reading
9. Repeated reading with a partner
10. Cross-age tutoring with a partner
11. Specified performance criterion levels of fluency

Reason to Hope

Once the foundational principles are understood, principles emerge that guide the implementation of repeated reading. Yet, these principles are still incomplete and might easily be improved upon.

The specified performance criterion outlined as an essential feature of repeated reading was the basis for the questions found in this project. While Samuels' procedure specified a 95 WRCM criterion, recent research by Rashotte and Torgesen (1985) has suggested a criterion of a specific number of repetitions (e.g., three to five) is more effective. The focus for this projects research question was if the improvement between cold and hot reads, or gain score, is a better predictor than reaching another specific criterion such as a final goal score. The implication of a more effective criterion for moving to a new passage during repeated reading is a more flexible ability for teachers to prescribe criterion for students. This will allow for more efficient use of time and materials. This would have great promise for students who face the "tyranny of time" (Kameenui, 1993) in closing their individual gap in reading ability. I believe the answers to the questions posed by this dissertation will help close the knowledge gap and make repeated reading a more efficient tool for teachers and ultimately students.

CHAPTER III

METHODOLOGY

The purpose of this study was twofold. First, it would be valuable for teachers to know which indicator or advancement criteria measure of repeated reading more accurately predicts student reading growth. Second, an understanding of which student demographic variables impact growth during repeated reading is of value to teachers providing targeted and time-sensitive interventions for students.

In this section, the methodology used in the study is presented. The characteristics of the participants and the selection process were discussed. Next the complex model design is reviewed and explained. The study used a multilevel structural equation model that included two phases: multilevel growth modeling and a path analysis. Lastly, variables including all explanatory measures used are defined and justified.

Participants

The study was conducted in an urban school district in the state of Utah. Because of its large ethnic population, high poverty rate, and the inner-city location, the district was classified as urban rather than suburban or rural. The demographics at the time of the study for the school district were (2008-09 school year) are listed in Table 3. The district was and is still considered one of the most diverse and economically challenged in the state. As part of the district's participation in the federal Reading First project and participation in the state-sponsored K-3 reading initiative, teachers had received training in scientifically-based reading research strategies. This training included best practices

Table 3

Sample District Demographics by State Classification

Group	Number or % of students
Enrollment	12,884
Female	47.6%
Low SES (free and reduced lunch)	72.0%
Ethnic minority	53.7%
English language learners	24.3%
Homeless	5.4%

instruction in fluency with modeling of repeated reading as a strategy. As teachers implemented the repeated reading strategy, they had access to two published repeated reading programs: Read Naturally and The Six-minute Solution. Teachers self-selected if they would use these programs or select their own materials for the repeated reading process. All teachers who participated used one of the three methods as part of their instruction.

Selection Process

Since the objective of the study was to determine specific student growth over time in relation to the increase in words read correctly during repeated reading, existing student data already collected for the district's Reading First project was analyzed. Grade 2-5 teachers were asked to record student growth during repeated reading currently occurring in their classrooms. The invitation to participate was extended to all 168 second- through fifth-grade teachers within the district. Kindergarten and first-grade teachers were not invited to participate as repeated reading is not a common practice until the end of the year in first grade. Teachers were not asked to change practice but simply report the existing repeated reading data. While some students completed more than one repeated reading cycle (cold read, practice reads, and a hot read) in a week, only one repeated reading cycle was reported for each student per week. This data were used along with DIBELS benchmark assessment data and the high-stakes AYP, state criterion reference tests (end-of-level assessments) available through state and local testing to perform the analysis. The only requirement for participation was teacher willingness to complete the data collection charts using their existing data and the use of an identifiable form of repeated reading occurring within the classroom. Because of the self-selection process, no control groups were used and thus the study had a quasi-experimental design. Teachers were informed that all identifiers for teacher and student would be masked before any data analysis was performed. While an unknown number of teachers started collecting data, 18 teachers from eight different schools completed and submitted the data collection request.

Model Design

In order to model the sophisticated interactions happening during student growth, it was necessary to combine a multilevel structural equation multi-level model with path modeling. The resulting multilevel structural equation model was represented in Figure 2. Few previous repeated reading studies have used a growth model and therefore have been limited to classical statistical analysis. This studies focus on growth required a sophisticated model that allowed analysis of repeated measures of student outcomes

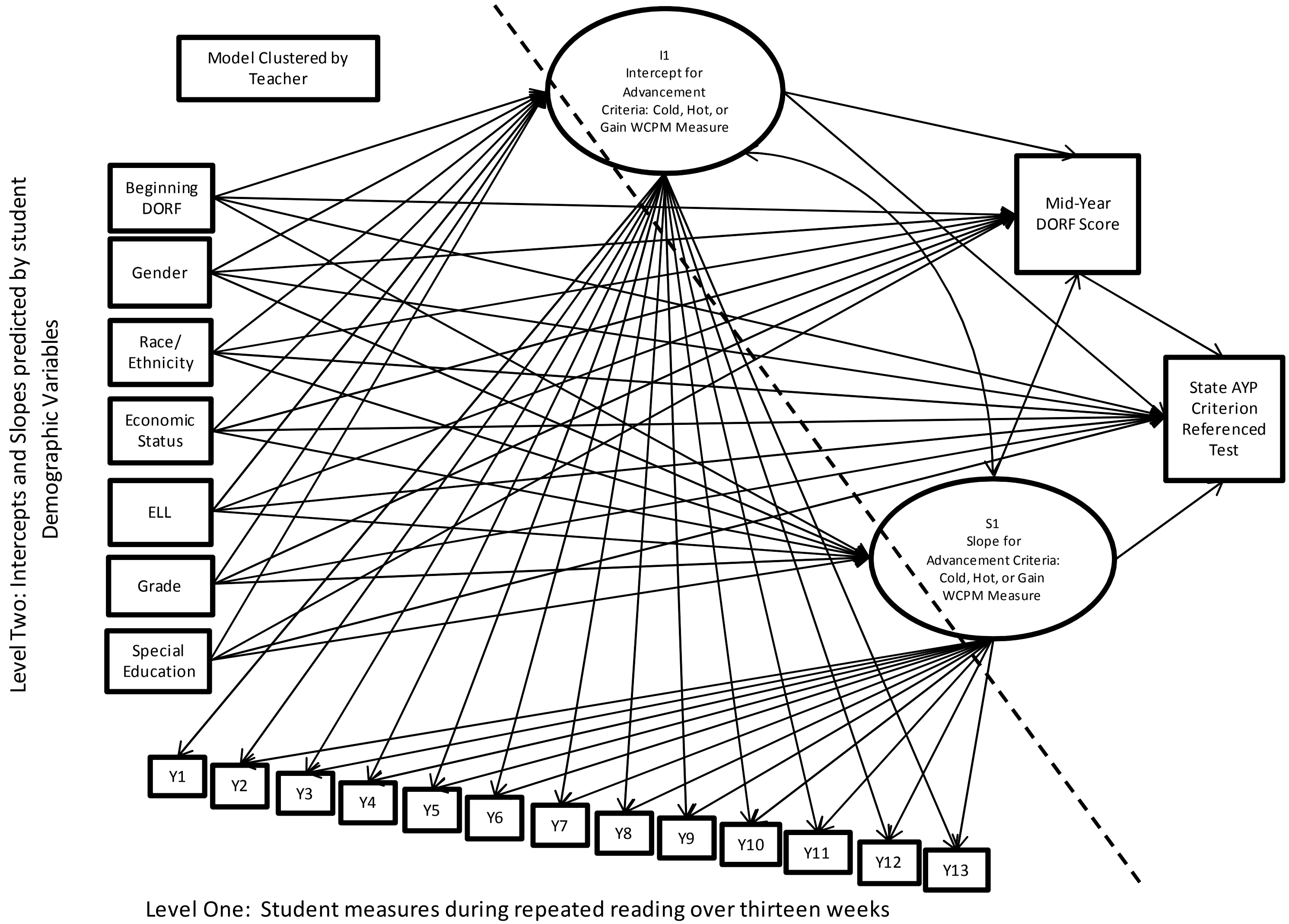

Figure 2. Multilevel structural equation model for repeated reading.

linked over time with growth trajectories identified. Therefore, a multilevel model was used for the first part of analysis to examine student growth.

The questions for this study also required a complex model in that three different advancement criteria were being compared and the student demographic variables affecting the model were being studied. A path model was selected for the second phase of the analysis to compare the goodness of fit between the advancement criteria.

Thus, there were two phases to the analysis—growth and path modeling. Using the Mplus® version 5 (Muthén & Muthén, 1998-2010) software package, student variables were first analyzed using a multilevel mixed model clustered by teacher. This first phase of the analysis created a growth factor measurement for each student and was run three times. Each of the three runs was based on one of the three advancement criteria: cold read advancement criteria model (CRACM), hot read advancement criteria model (HRACM), or GAIN SCORE ADVANCEMENT CRITERIA MODEL (GSACM). For each advancement criterion studied, a model was created. The individual student measurement for each advancement criteria, or the criterion used to determine when a student is ready to move on to a new passage for repeated reading practice (CRACM, HRACM, GSACM), was used to calculate the latent variable of intercept and slope in each model. A latent variable is a non-observed variable that is inferred through a mathematical model from other observable variables. The student's growth scores were calculated from 13 weeks of student WRCM data.

Raw scores were used to determine correlation coefficients for all nominal variables, the intercept, and the slope. Standardized scores were used to determine correlation coefficients for continuous variables.

This information then fed into the second analysis or path model. The path model allowed examination of goodness of fit for the three advancement criteria on two outcome measures: end DORF scores and high stakes AYP assessments. The path model was based on the student's growth factor for the advancement criteria predicting the student's WRCM as measured by the midyear DORF benchmark assessment; and then in turn the midyear DORF assessment predicting the student's high stakes reading assessment score. Regression analysis of student demographic characteristics also allowed study of the effects of student demographic characteristics on the outcome measures.

Since there were three different advancement criteria being compared, the analysis was completed three times. The first examination used growth between each of the initial WRCM or cold read growth trajectory to predict model fit. In this study, the growth between the cold reads will be referred to as the CRACM. The second round of analysis uses the final WRCM or hot read growth trajectory to predict model fit. The growth between hot reads will be referred to as the HRACM. Finally, gain scores between hot and cold repeated reading of a passage was used as the growth trajectory to predict model fit. This last analysis was called the GSACM. For each analysis, student repeated reading scores representing one of the achievement criteria gathered over a 13-week period represented level one in the model. To analyze the within student variance in each set, student variables were used as level two predictors clustered by teacher.

Multilevel Modeling

In the first part of the study model, multilevel modeling was used. Multilevel modeling has been used for measuring changes in student achievement over time because the assumption of independence of observations that provides the base of classical statistical analysis was unusable (Raudenbush & Bryk, 2002). Indeed, many student factors were so interrelated that a more sophisticated model than simple linear regression was required. The typical statistical methods could not account for the enormous variables within subjects such as initial starting WRCM and age. The inadequacy of typical statistical methods was compounded at another level because typical statistical analysis was not sophisticated enough to also account for the second level of variables found within the classroom that were the explanatory measures (data collected for analysis as covariates). Explanatory measures used as variables are listed in Table 4.

These explanatory measures are the demographic variables for each student. Explanatory measures were defined as follows.

1. Student gender was designated as either male or female.
2. Race/ethnicity designations changed in federally-mandated reporting during the time of this study. Originally, Hispanics were a separate count. Under current federal definition, Hispanics are now classified as White for race. Because of the high proportion of Hispanics in this study, they were not included in the White classification. The traditional race/ethnicity explanatory measure of White or non-White was used for this study with the non-White group including Hispanic, Asian, Black, Indian, and Pacific Islander.

Table 4

Design Matrix of Explanatory Measures

Demographic designation	Explanation	Code
Gender: Designated as male or female	Evidence of a gender gap has been reported in the literature (Nation's Report Card, 2008)	Nominal
Race/ethnicity: Designated as White or non-White	Achievement gap between racial groups in reading achievement is well documented (Nation's Report Card, 2008)	Nominal
Initial literacy level: Measured by the beginning DORF score (WRCM)	Disparity of initial level of literacy highly correlates with outcome measures (Walpole, Chow, & Justice, 2004)	Benchmark and continuous
SES: Designated as full pay or free/ reduced lunch	SES serves as a negative predictor of student achievement (Clements, Reynolds, & Hickey, 2004)	Nominal
English language learner: Designated as fluent English speaker or as English learner	Direct correlation in evident between language acquisition and student achievement (Kato, Albus, Liu, Guven, & Thurlow, 2004)	Nominal
Age: Measured by current grade level	Reading abilities develop through stages as children gain experience (Chall, 1976)	Continuous

3. Each student's initial literacy level was measured by the beginning of year WRCM as measured on the DIBELS benchmark.
4. SES was measured by the students paying full price for school lunch versus student who received free or reduced price lunches.
5. Students were identified as English language learners (ELLs) based on the district screening of all students. During school registration, all students were screened for factors indicating they are not native English speakers. An investigation was made for each student who might not be a native English

speaker including a language level assessment. A team considered data collected and made a determination if the student was currently an "English Language Learner."

6. Age was determined by designating students by their current grade level.
7. Special education status is determined by students being referred to special education, testing to determine eligibility, and then a team including educators and the parents determining special education standing.

The multilevel analysis constraint required complex scrutiny. In a classroom-level-only analysis, Raudenbush and Bryk (2002) posited that group averages obscure individual results and do not represent any single group or individual. The results can be clouded by aggregation bias, misestimated precision, unit of analysis problems, and impoverished conceptualization. Aggregating data at the classroom level only can easily deflect the student's growth curve. The interpretation of growth, focused solely on classroom effects, is unwarranted without taking into account differences in individual students.

Given the complexity involved in analyzing variables simultaneously at within-student data and between-student measures, multilevel structural equation modeling was deemed appropriate for the first stage of the analysis. The advantages of multilevel structural equation linear modeling include conservation of information, statistical accuracy, more power and precision, and the ability to give weight to composites of the levels of interaction (Raudenbush & Bryk, 2002). Multilevel structural equation modeling allowed analysis to be completed concurrently at more than one level. For this study's model, only two levels were needed. Level one analyzed within-student data that included the variable of growth in WRCM over time (repeated individual student measure). Level two accounted for explanatory demographic measures such as: gender, race/ethnicity, initial literacy level, SES, and grade/age. Data was clustered by teacher, a variable which manifests itself in differences such as type of instruction, amount of practice, etc. Because of the similarity in the demographics between schools used in the study, an additional level of analysis was not deemed necessary.

This investigation was considered quasi-experimental in that complete confidence cannot be reached that all relevant background variables have been identified and controlled. In using existing data, students were not randomly assigned nor was a control group created. Each student needed to serve as their own control of fixed effects thus also allowing for control of between class/teacher effects. Other threats to validity considered included (Raudenbush & Bryk, 2002):

1. Bias created by differences in student background. The more dramatic the differences, the more sensitive the analysis will be to different methods of adjustment—creating less credible inferences.
2. Heterogeneity of regression. The regression line may indicate dramatic effects for low SES students and little difference for high SES students.
3. Shrinking as a self-fulfilling prophecy.
4. Inadequate conceptualization. The model needed to account for individual growth and classroom variables. A model must reflect the most parsimonious fit that adequately describes the data.

5. Measurement limitations. For a multilevel model, the rule of thumb is that the regression analysis needs at least 10 observations for each predictor Transformations were used to compensate for positively skewed data. To avoid bias caused by random, missing data, incomplete cases were not dropped.

The complexity of the model created, including both a multilevel structural analysis and a path model, attempted to minimize these threats to validity.

Path Modeling

Path modeling is used to test theoretical relationships (Schumacker & Lomax, 2004). While not designed to discover causes, path modeling can test relationships and establishes causal relationships or the relationship between variables. Goodness-of-fit indices can be used to determine how well the model describes the relationship between the variables. Only one previous study was found (Chard et al., 2008) that used a path model to study growth in the analysis of oral reading fluency. Chard and his colleagues' study modeled the relationship between variables in predicting growth on third grade oral reading fluency and on third grade standardized tests of reading. Similarly, this study was designed to model the relationship between variables in repeated reading and the variables ability to predict student outcome measures. In this way, a path model was used to compare the goodness of fit for the three advancement criteria being studied.

Instrumentation

Measurement During Repeated Reading

A quantitative approach to the data collection was used for predictor and outcome variables. For daily progress measurement, the parametric measure of WRCM was used. This measure has commonly been used as a standard to measure student’s reading rate attenuated for accuracy (Kaminski, Cummings, Powell-Smith, & Good, 2008). The student’s cold read score, meaning an unpracticed reading of a passage, was calculated as the WRCM the first time a student read a passage. The student’s hot read score, meaning the score after multiple repetitious readings, was calculated as the WRCM the last time a student repeatedly read a passage. The number of times a student repeatedly read a passage was also recorded. Next, the gain score was calculated as the difference between the cold and hot reads. The grade levels of passages to be practiced were assigned by the teacher. Typically, passages practiced were at the student’s grade level and not at their instructional or independent reading levels. Teachers, who chose grade-level reading passages, expressed time constraints and difficulty in finding and managing quality passages as reasons why they chose not to provide practice passages at the students independent reading level. These teachers also reasoned that all of their students needed exposure to grade-level vocabulary and text. Additionally, they expressed that with pre-teaching strategies, students could be supported in using grade-level readings.

Teachers monitored the repeated reading procedure. However, students recorded and charted the results daily to monitor their own progress. The chance for inaccuracy due to student measurement and recording was tolerated to allow for the benefits of selfmonitoring and to reflect actual normal classroom implementation of repeated reading.

Benchmark Assessment

The DIBELS benchmark assessments was used to compare progress over time to determine each student's initial reading ability and their reading ability at the conclusion of the measured instruction. A high correlation has been shown between oral reading fluency (ORF) scores and comprehension scores measured on various states reading achievement test scores (Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008). The DIBELS passages have been norm-referenced and therefore provide a valuable benchmark of the individual-growth indicator for reading fluency development. The DIBELS benchmark assessment also provides cut scores that indicate a student's probability of success in subsequent schooling. The DIBELS tasks were selected based on its ability to predict future reading achievement. The normed cut scores allowed students to be identified and grouped for analysis based on their risk for failure in subsequent years.

The DIBELS benchmark assessments, used as both pre- and postassessments, were administrated by a team of outside proctors trained at the district level. Each proctor received training using training materials from the DIBELS web site. Each proctor was observed giving the assessment and rated for reliability using an administration checklist. Each proctor was also monitored daily by a team leader to maintain reliability during the assessment window. The proctors moved as an assessment team from school to school administering the DIBELS benchmark assessments.

High-stakes AYP, Criterion-Referenced Test

Each student in the state of Utah has been assessed for mastery of grade-level material using the Utah Criterion Referenced Test (CRT). The Utah CRT is the high-stakes AYP assessment used by the state. The CRT provides a reading score that was also used as an outcome measure for students reading achievement. This measure served as a second outcome measure in addition to the end DIBELS Benchmark Assessment. The CRT was administered in April and was proctored by the classroom teacher. Each teacher received training in testing protocol and used a script for test administration. The results were equated at the state level and cut scores generated to determine student's individual reading level. Equating allows comparative analysis of students from year to year and from group to group.

Time Frame for Assessments

The time frame for individual student repeated reading data collection spanned approximately 15 consecutive weeks. The initial benchmark was determined using the beginning DIBELS benchmark assessment. Teachers reported cold and hot read times for 15 weeks with the number of times each passage was read. Then the ending DIBELS benchmark assessment was collected. Criterion reference results were collected at the end of the year. The design matrix for measures is detailed in Table 5.

Analysis

The statistical analysis of results (see Appendix A) was completed using the software package Mplus® version 5 (Muthén & Muthén, 1998-2010). The syntax used for the multilevel mixed model outlined at the beginning of this section is found in Appendix B.

Table 5

Design Matrix of Measures

Assessment	Purpose	Frequency of administration
Cold read (c1-15)	WRCM weekly measured on the first time each passage was reading. (Independent variable)	Weekly
Hot read (h1-15)	WRCM measured on the last time each passage was read. (Independent variable)	Weekly
Gain score (y1-15)	Measures the gain in WRCM between cold and hot reads measured weekly. (Independent variable)	Weekly
Beginning DORF benchmark assessment (DORFB)	Standardized measure of oral reading fluency using words read correctly per minute on previously unseen passages. (Dependent variable)	3 times per year
Midyear DORF benchmark assessment (DORFE)	Standardized measure of oral reading fluency using WRCM on previously unseen passages. (Dependent variable)	3 times per year
High-stakes AYP state assessment (CRT)	Reading score from high stakes AYP CRT given to determine mastery of grade-level scope and sequence. (Dependent variable)	Spring

CHAPTER IV

RESULTS

The purpose of this study was to explore which of three fluency advancement criteria models, (a) cold read score, (b) hot read score, or (c) gain or growth scores, was the more accurate predictor of student readiness to move to a new passage for repeated reading practice. In this chapter, the reporting of results was organized around the study's two guiding research questions. First, which of three fluency advancement criteria models (cold read scores, hot read scores, or gain or growth scores) is best to determine when to move a student's fluency practice to a new passage? Second, which demographic factors (age, ethnicity, gender, current reading level, SES, etc.) best predict the growth during repeated reading fluency practice?

As an operating hypothesis for the study, it was predicted that the student growth or GSACM would provide the best model fit because it would reflect more accurately when students had achieved sufficient growth in reading rate and accuracy as compared with static cold or hot read scores alone to determine when to move a student a new reading passage for practice. The goal of this study was to determine which factors influenced students' repeated reading gains within a single text and/or across texts over time to provide teachers guidance about the number of repetitions needed for repeated reading to achieve maximum student growth possible in an optimal amount of time.

A multilevel structural equation model was used to analyze the data for this study. There were two distinct phases to the data analytic process: (Stage 1) multi-level growth modeling, and (Stage 2) path analysis. Using the *Mplus*® v. 5 software package, student scores were analyzed using a multilevel growth model clustered by teacher to estimate a latent growth factor for each student. The data set was analyzed three separate times using one of three reading fluency advancement criteria - cold read, hot read, or gain score. The latent variables, students' weekly repeated reading scores and student demographic data, were then fed into the second phase of the analysis which used path analysis. The path model allowed an examination of the goodness of fit for the three advancement criteria, cold read, hot read, or gain scores, in predicting two outcome measures: the midyear DIBELS ORF score, required by Utah state law to be reported to parents, and the state CRT high-stakes AYP assessment. Specifically, the path model fit was based on the student's latent growth factor predicting the student's WRCM as measured by the midyear DORF benchmark assessment; and then in turn predicting the student's end-of-year high stakes reading assessment score—the state end-of-level CRT reading score. Regression analysis of student demographic characteristics also allowed an examination of the effects of student demographic characteristics on the latent variables, midyear DORF measurement and end-of-level CRT measure.

The report of results was organized as follows:

- Data screening and cleaning process.
- Descriptive statistics including skewedness, means, standard deviations and correlations.
- Analysis of model fit for each reading fluency advancement criteria—cold, hot, and gain scores
- Analysis of the cold read model.

- Analysis of the hot read model.
- Analysis of the gain score model.
- Description of the correlation between the demographics and outcome variables.
- Summary of results.

Data Cleaning and Screening Processes

Visual inspection of the collected 15 weeks of student repeated reading score trend data revealed that the initial week's repeated reading scores did not follow the typical overall pattern for individual students as reflected in subsequent weeks' scores. The inconsistent trend data obtained during the first 2 weeks of the study was attributed to students' learning of a new fluency practice routine. To correct for potential novelty effects of the initial routine learning process, the first week's repeated reading scores were dropped from the final analysis.

It was also noted in the visual inspection of the data that there was significant attrition in student scores reported during the 15th week. This attrition was attributed to teachers shifting instructional time away from the fluency practice routine using repeated reading toward administration of the midyear DIBELS benchmark assessment. Because less than 25% of the students had a repeated reading score recorded during the 15th week of data collection, week 15's data were also dropped from the final analysis. Based on the findings from the visual inspection of data, the final data used to determine the student's repeated reading growth intercept and slope were determined using 13 of the 15 collected week's scores of the repeated reading procedure (Week 2 through Week 14).

The visual inspection of the raw week-to-week oral reading fluency scores revealed a trend toward a gradual elevation of the minimum threshold of WRCM on each subsequent cold read resulting in a positive growth trajectory. However, if a second through fifth-grade student read less than 25 words correct per minute, these students typically exhibited flat or little growth and repeated reading had no effect on student WRCM performance. Since having a few students performing significantly lower than other students is common in most classrooms, data for these low-performing students were retained in the data pool because they would be considered important in determining the overall fit of the growth model and relationships between the demographic variables and the repeated reading process. Therefore all individual data, including those visually identified as exhibiting no or flat growth, were retained in the data analysis.

Descriptive Statistics

The study sample included 451 nonrandomly selected second- through fifth-grade students. Due to student attrition, 51 (11%) of the sampled students were eliminated from the study as no outcome measures for these students were available. This student attrition was evenly spread among teachers and school. The remaining 400 students in the sample represented 6% of one entire school district's second- through fifth-grade population. Eighteen teachers volunteered to record and submit their students' repeated reading data for 15 weeks. They represented 12% of all second- through fifth-grade teachers in the district. These teachers taught in eight different schools of the district's 16 elementary schools.

The final sample student population closely mirrored the total district student population.

- The percentage of females in the sample was 47% compared to the district's 48%.
- The economic status of the sample group was 79% of students receiving free and reduced lunch compared to the district's slightly lower 72% calculated poverty rate.
- The sample's special education students made up 15% of the group while the district's total was slightly higher at 16%.
- The percentage of ethnic minority learners in the sample was slightly higher than the district percentage: 56% compared to 54%.
- The student attrition rate for the sample was 9.9% over the 13 weeks of the study, which was excellent compared to the larger district sample's 34.4% mobility rate (number of students enrolled less than 160 out of the 180 possible school days).

Overall, there were no unexpected differences between the sample's demographics and the demographics of the larger district population. Thus, the results of this study's sample group can be reasonably assumed to reflect the performance of the larger district student population. Also of note is the fact that this district's student population has the highest free and reduced lunch rate in the state (77%). With 52% of all district students classified as a minority for race/ethnicity, this school district has one of the most diverse populations in the region. While the study district had a minority majority, it is important to note that 47% of the population was Hispanic. Thus other race/ethnicities than White and Hispanic comprised a small group. During the analysis, it was important to remember that the outcomes from this study were obtained from students in a higher-than-normal risk group.

Descriptive statistics for the outcome and predictor variables used in the analyses are shown in Table 6. From an examination of Table 6, one can see that the mean score on the DIBELS Oral Reading Fluency (ORF) benchmark assessment increased 22.65 WRCM over the 13 weeks of the study. As practice took place daily, the reported increase from repeated reading was close to the recommended DIBELS benchmark increase for student growth over time. However, the number of students scoring proficient on the DIBELS ORF at benchmarking periods did not increase dramatically. This finding indicates that although students read more words correctly with each repeated reading in the sample population, the growth mirrored the average, expected

Table 6

Sample Measured Descriptive Statistics

Variable	Range	Min	Max.	Mean
DORFE: The midyear benchmark score on the DIBELS Oral Reading Fluency (DORF) in WRCM	175	7	182	85.58
DORFB: The beginning benchmark score on the DORF in WRCM	194	0	194	62.93
CRT: State AYP high-stakes criterion referenced test standardized for mastery at 160	60	130	190	162.61

Note. N = 400.

growth for all students in the DIBELS norming sample and thus did not result in an overall rapid closing of reading fluency gaps.

For the other dependent measure in the analysis, the high-stakes AYP state assessment mean score was 162 with over half of the sample, or 62% of students, meeting the state's proficiency standard indicating mastery of grade-level core standards. This percentage is only slightly lower than the district sample population score of 62.4%.

Goodness-of-Fit Tests for Three Models: Cold Read, Hot Read, and Gain Scores

Goodness-of-fit statistics were run multiple times, each time adding and deleting variables in an attempt to improve fit. These attempts to delete or add student demographic variables did not yield better fit results than did the initially proposed structural equation model on the comparative fit index (CFI), Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), or the standardized root mean square residual (SRMR). Thus after trying to obtain better fit indices by adding and deleting variables, the originally proposed model of independent and dependent variables (Figure 3) yielded the best goodness-of-fit results for all three advancement criteria models studied—cold read, hot read, or gain scores.

In posing the first research question, it was theorized that using growth or gain scores obtained from cold to hot reads as a level one, student growth measure would provide the best model fit. This hypothesis was based on the premise that a measure of actual student growth such as the gain score between cold and hot reads would provide a

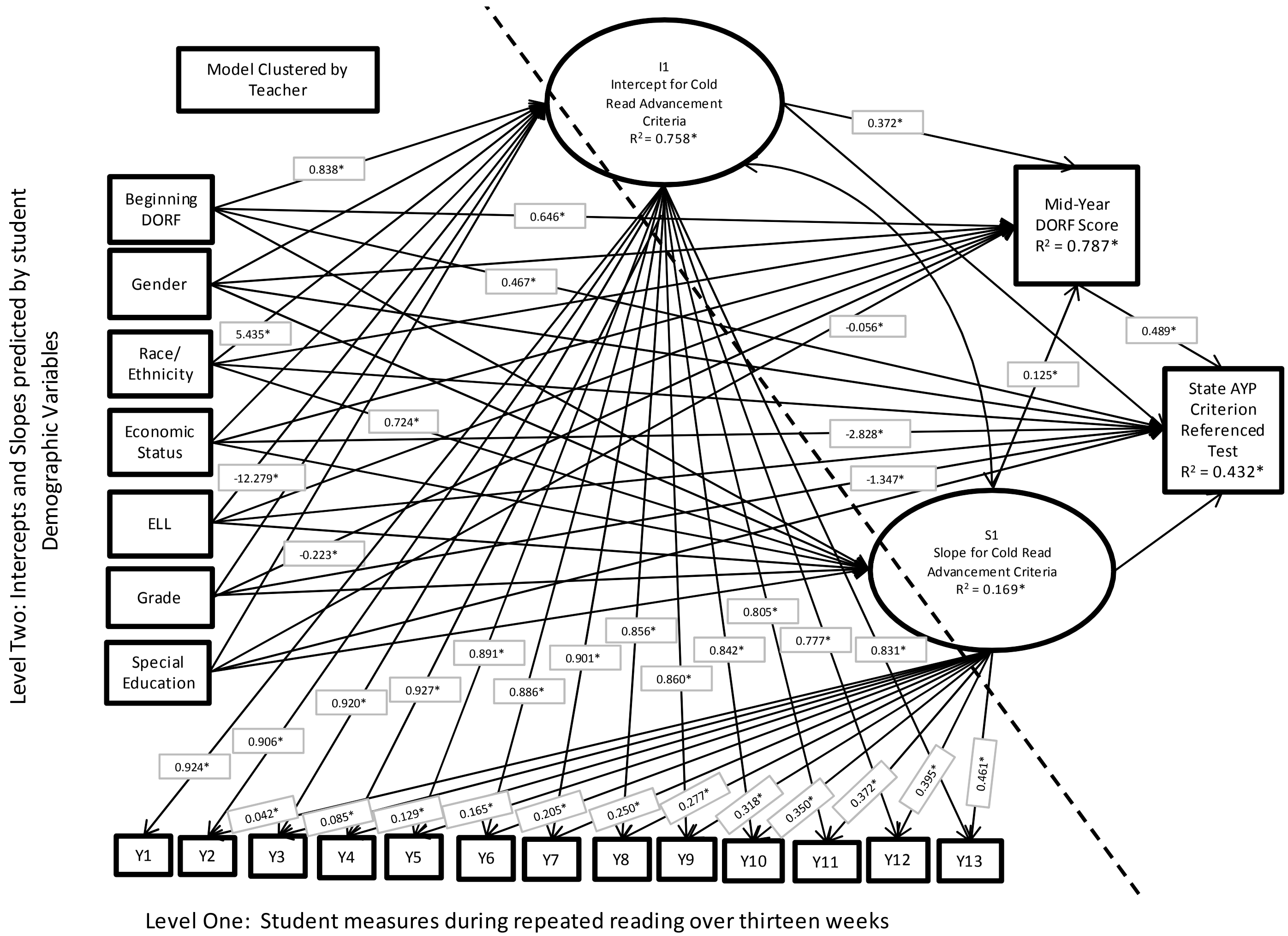

Figure 3. Cold read criteria multilevel structural equation model with correlation coefficients.

* Two-tailed *p* values less than 0.05. Standardized scores were used for continuous variable.

more accurate advancement criterion than the unpracticed cold read or the practiced hot read score. Thus, it was hypothesized that the gain score model would provide the most efficient indication of when a student was ready to move to a new passage during repeated reading. As evidenced by the data, the cold to hot read gain score model did not fit the model as well as the other advancement criteria (0.800 on the CFI with 0.95 considered a good model fit). As shown in Table 7, the hot read model provided the best fit with the CFI = .949 (very close to the desirable 0.95 standard). The cold read model providing only a slightly less desirable fit at CFI = 0.936.

The next section reports the correlation coefficients and R^2 results for each of the three advancement criteria models—gain, hot, and cold read scores.

Cold Read Advancement Criteria Model

The correlation coefficients and R^2 results calculated using the students' cold read scores as the level-one predictor variable are found in Figure 3. Because this was a multilevel model, the level one predictor variables were students' individual fluency scores and the level two predictor variables were the demographic variables for each

Table 7

Measures of Repeated Reading Model Fit

	Chi-square statistics			Goodness-of-fit indices			
Variable	Chi square	*Df*	*p* value	CFI	TLI	RMSEA	SRMR
Model based on cold reads	417.430	186	0.000	.936	.928	.056	.060
Model based on hog reads	351.697	186	0.000	.949	.943	.047	.064
Model based on gains from hot to cold reads	427.047	186	0.000	.800	.775	.057	.112

Note. Criteria for goodness-of-fit statistics use were CFI > .95, TLI > .95, RMSEA < .05, SRMR < .06.

student. The intercept, which represents the beginning WRCM, predicted the first week reading score (WRCM) well with a correlation coefficient of 0.924 ($p < .05$) words read correctly per minute. This correlation coefficient can be understood to predict one unit or one word per minute growth in the intercept. The intercept is a latent (theorized or unmeasured) variable used to simplify the relationships between the measured variables. In this case, the intercept predicts the starting average number of words read correctly per minute for all of the students in the sample. Over the course of the study, there was a gradual decrease in this coefficient to 0.83* ($p < 0.05$) in the 13th week. The gradual decrease in coefficients was understandable, as individual variation in student growth patterns overtime would likely decrease the correlation between the weekly student scores and the initial intercept.

The correlation coefficients for each of the weekly cold reads were also regressed on the slope. The slope is calculated to represent the linear gain in words read correctly per minute for the student group over the duration of the data. This pattern of resulting coefficients was the opposite of the intercept. The slope regression coefficients started small but gradually increased over the 13 weeks of the study. The coefficients regressed on the slope between the first 2 weeks produced a very weak coefficient of 0.042 ($p < 0.05$). These coefficients gradually increased over the 13 weeks until the final coefficient for the slope was a moderate 0.461 ($p < 0.05$).

Level Two Variables: Student Demographic and Beginning DORF Variables

When the cold read scores were used to construct the intercept latent variable, which represents the initial words read correctly per minute, three level-two variables reached significance for predicting the intercept: the beginning DORF WRCM score, the student's race/ethnicity, and the student's ELL status. For every 0.838 ($p < 0.05$) more words students read on the beginning DORF assessment, the intercept increased by one WRCM. Also in the cold read model, non-White students had a coefficient of 5.435 ($p < 0.05$). ELLs WRCM coefficient was 12.279 ($p < 0.05$). While there was a significant difference for ELL students on the intercept, there was not a significant difference in their outcome measures. The other four level-two variables: gender, economic status, grade, and special education status in the analysis did not reach levels of statistical significance in predicting the intercept.

Only one demographic variable reached significance when regressed on the slope: the student's race/ethnicity status. The coefficient for non-White students was 0.724 ($p < 0.05$). Significance for the other level-two variables—beginning DORF, gender, economic status, ELL, grade level, and special education status—was not found.

Midyear DORF Regression

Three level-two variables produced significant coefficients when regressed on the midyear DORF score: the beginning of the year DORF score, the student's grade, special education status. Additionally, the latent variables for intercept and slope produced significant coefficients. The coefficient for the beginning DORF score regressed on the midyear DORF score was 0.646 ($p < .05$). Since these are the same measures differentiated only by time, the strong coefficient is to be expected. The coefficient when the student's grade was regressed on the midyear DORF score was -0.223 ($p < .05$). This negative relationship between age and DORF scores is understandable as DORF scores have a higher predictive value in younger grades (Hintze, Ryan, & Stoner, 2003), thus a negative coefficient is to be expected. Special education designation resulted in a weak -0.056 coefficient. This result seems reasonable as students receiving special education services would be expected to have a lower midyear DORF score. A moderate coefficient of 0.308 ($p < .05$) was calculated when the midyear DORF score was regressed on the latent intercept. The coefficient for the latent slope was small but significant, 0.125 ($p < .05$).

High-Stakes AYP Assessment Regression

When the high stakes state AYP measure was regressed in the cold read model, the beginning DORF and the midyear DORF scores both produced similar significant coefficients. The beginning student DORF score had a coefficient of 0.467 ($p < .05$) and the midyear DORF score had a coefficient of 0.489 ($p < .05$). The demographic variables of economic status and grade were also significant in being able to predict end-of-year AYP assessment scores. Low economic status produced a coefficient of -2.828 ($p < .05$). Thus, low economic status has a strong negative prediction on student end-of-level tests. A student's grade level also had a negative relation with the end-of-year assessment. The coefficient for the student's grade level was -1.347 ($p < .05$).

R^2 for Cold Read Advancement Criterion Model

In comparing the predictability of the three models, it is valuable to compare the R^2s to determine the proportion of the variability accounted for by each model. The R^2

measure indicates how well future outcomes are likely to be predicted by the model. For the cold read advancement criterion model the R^2 measure for the intercept was 0.758 ($p < .05$). This finding means that 75.8% of the variance can be predicted by the latent intercept, which is the average beginning WRCM for all students, in the model. The R^2 for the slope was 0.169 ($p < .05$) or only 16.9% of the model variance can be can be explained by the slope—the average rate of student growth. The R^2 for the midyear DORF score was a strong predictor at 0.787 ($p < .05$) or 78.7% of the variance is predicted by the ending WRCM in the model. The R^2 for the high-stakes AYP assessment was moderate, 0.432 ($p < .05$) or 43.2% of the variance was predicted by the state CRT assessment.

Intercept Regressed on Slope

When the intercept was regressed on slope in the CRACM, the correlation coefficient estimate was 1.828. This estimate did not meet significance requirements for a two-tailed *p* value ($p = 0.610$). Therefore, the finding was that the intercept could not be used to predict the slope.

Hot Read Advancement Criteria Model

A second model using student's final practiced or "hot" read for a passage as the level one student measure was analyzed. The correlation coefficients and R square results calculated using the student's hot reads as the level-one student measure are represented in Figure 4. Again, the level-one results (individual student fluency scores) and level two demographic predictor variables are used to calculate intercepts and slopes for each

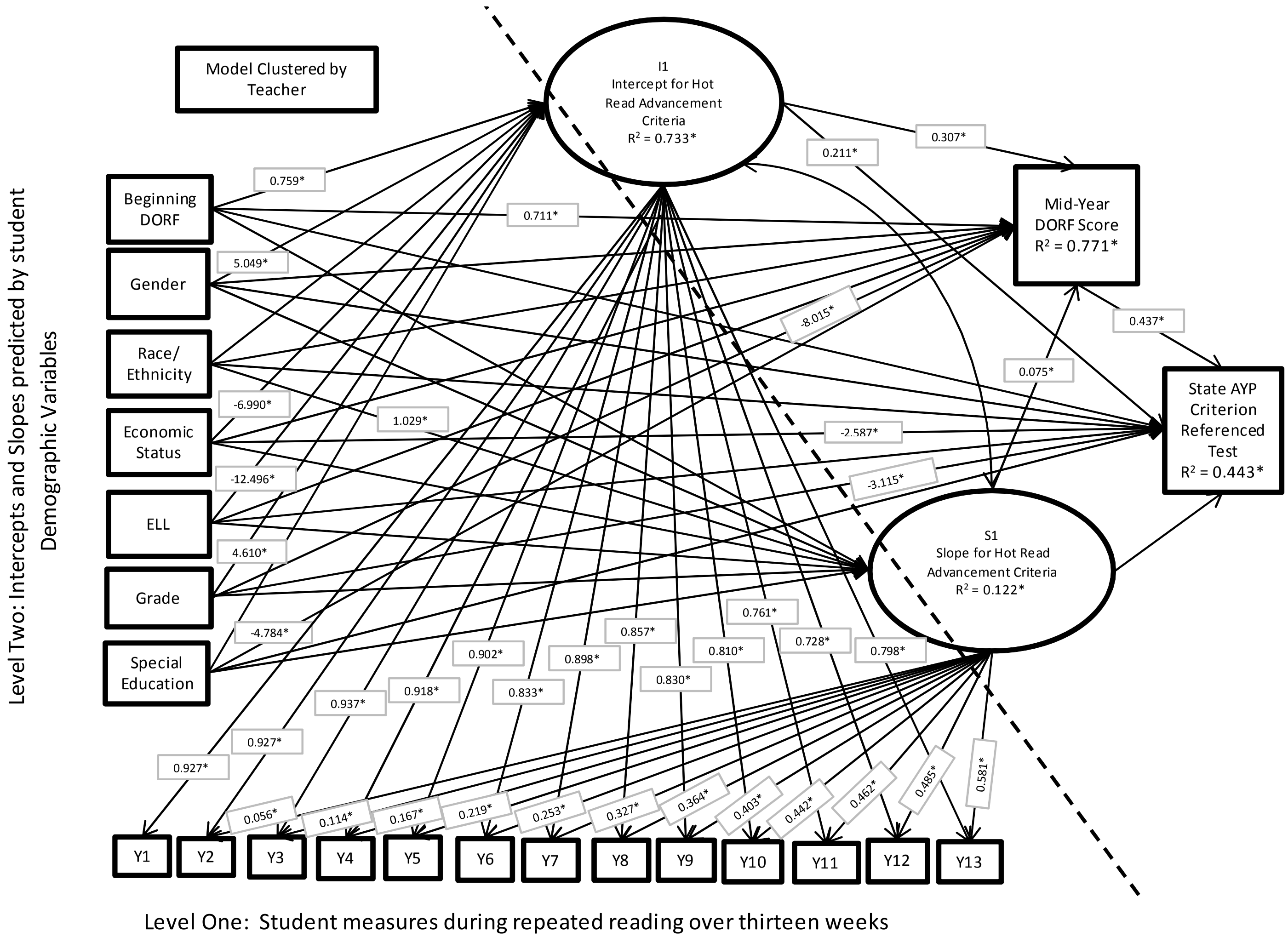

Figure 4. Hot read criteria multilevel structural equation model with correlation coefficients.

* Two-tailed *p* values less than 0.05. Standardized scores were used for continuous variable.

student in this multilevel model. Similar to the cold read model, the coefficient predicting the intercept or average beginning WRCM for the hot read model started high at 0.927 ($p < .05$). Over the 13 weeks, there was a gradual decrease—with the final coefficient of 0.798 ($p < 0.05$). This slight decrease over time reasonably reflects a weakening relationship with the calculated beginning WRCM intercept due to decreased individual variation in student growth patterns as students learn.

The correlation coefficients for each of the weekly hot reads were also regressed for the slope. The pattern was the same as the cold read model in that the coefficients from the slope regression started small and gradually increased over the 13 weeks of the study. The measure for the slope between the first 2 weeks produced a weak coefficient of 0.056 ($p < 0.05$). The coefficient gradually increased over the 13 weeks until the final moderate coefficient for the slope was 0.581 ($p < 0.05$). This gradual growth in coefficients over time was consistent with the expectation that student reads at the end of the study would more closely reflect outcome measures than beginning of study reads.

Level Two: Student Demographic Variables

When the initial hot read advancement criteria was used to construct the intercept latent variable, five of the seven demographic variables reached significance for the intercept: the beginning DORF WRCM score (0.759, $p < 0.05$), the student's gender (5.049, $p < 0.05$), the student's SES (-6.990, $p < 0.05$), the student's ELL status (-12.496, $p < 0.05$), and the students grade (4.610, $p < 0.05$). SES and ELL status both had significant negative impacts on the student's starting point (intercept).

The beginning DORF score was a strong predictor as would be expected as it is the same measure of the initial words read correctly per minute. Girls started higher than boys and students read more words correctly in older grades. Inversely, students in the low SES group and ELLs beginning WRCM were lower than their sample counterparts. Student race ethnicity and special education status did not reach significance.

Slope Regressed on Demographic Variables

In the hot read model, only one demographic variable reached significance when regressed on the slope: the student's race/ethnicity status. This is similar to the results from the cold read model. A student with non-White status' coefficient was 1.029 ($p < 0.05$). Non-White students' growth increased more over time than their White counterparts.

Midyear DORF Regression

Only two of the demographic variables and both the latent variable produced significant coefficients when regressed on the midyear DORF score: the beginning DORF score, the student's grade, and the latent variables for intercept and slope. The coefficient for the beginning DORF score regressed on the midyear DORF score was 0.711 ($p < .05$). Since these are the same measure differentiated only by time, the strong coefficient is to be expected. The coefficient when the student's grade was regressed on the midyear DORF score was -8.015 ($p < .05$). As with the similar negative relationship between age and DORF scores in the Cold Read Model, the result is understandable as DORF scores have a higher predictive value in younger grades.

Special education status impacted the midyear DORF score to a greater negative extent in the hot read model—special education designation resulted in a -4.784 coefficient which was much stronger than the -0.056 result from the cold read model. A weak coefficient of 0.307 ($p < .05$) was calculated when the midyear DORF score was regressed on the latent intercept. The coefficient for the latent slope was small but significant, 0.075 ($p < .05$).

High-Stakes AYP Assessment Regression

When the high-stakes state AYP measure was regressed in the Hot Read Model, the beginning DORF did not produce a significant coefficient. However, the midyear DORF score produced a mild, significant coefficient, 0.211 ($p < .05$).

The demographic variables of economic status and grade were also significant in being able to predict end-of-year AYP assessment scores. Low economic status produced a coefficient of -2.587 ($p < .05$). Thus, low economic status has a strong negative influence on student end-of-level tests. Grade also had a negative relation with the end-of-year assessment. The coefficient for the student's grade was -3.115 ($p < .05$). As previously stated, this negative coefficient possibly reflects a problem with the states equating of scores.

R^2 for Hot Read Model

In determining the proportion of the variability accounted for in the hot read model, the R^2 measure for the intercept was 0.733 ($p < .05$). This finding means that 73.3% of the variance can be predicted by the latent intercept or average beginning WRCM in the model. The R^2 for the slope was 0.122 ($p < .05$) or 12.2% of the model's variance can be can be explained by the slope—the average rate of student growth. The R^2 for the midyear DORF score was a strong, significant predictor of variance at 0.711 ($p < .05$). The midyear DORF predicted 71.1% of the variance. The R^2 for the high-stakes AYP assessment was moderate, 0.443 ($p < .05$). The end-of-year AYP assessment accounted for 44.3% of the variance in the model.

Intercept Regressed on Slope

When the intercept was regressed on slope in the HRACM, the correlation coefficient estimate was 5.331. This estimate did not meet significance requirements for a two-tailed *p* value ($p = 0.287$). Therefore, the finding was that the intercept could not be used to predict the slope.

Gain Score Advancement Criteria Model

A third model, gain score advancement criteria model (GSACM) using student's growth in words read correctly per minute between the initial "cold" read and the final "hot" read for a passage as the level one student measure, was analyzed. The correlation coefficients and R^2 results calculated using the student's gain score between cold and hot reads as the level one student measure are represented in Figure 5. Again, the level one results (individual student measures) become the outcomes of level two (the intercepts and slopes for each student) in this multilevel model.

The coefficient predicting the intercept or starting point for the GSACM model started at 0.576 ($p < .05$) WRCM in first repeated reading gain score predicting one unit (WRCM) change in the intercept. Over the 13 weeks, there was a gradual decrease in the

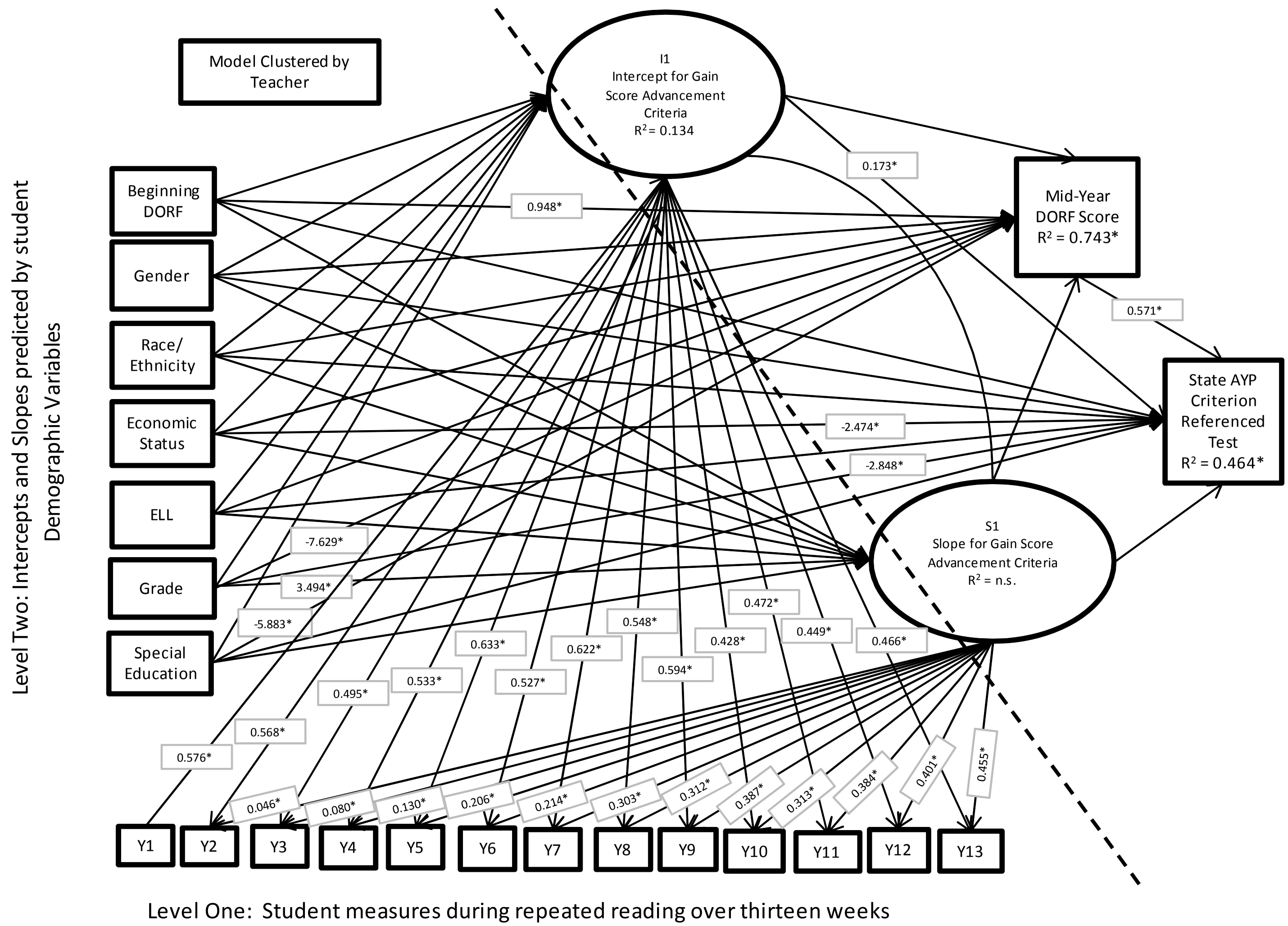

Figure 5. Gain score criteria multilevel structural equation model with correlation coefficients.

* Two-tailed p values less than 0.05. Standardized scores were used for continuous variable.

coefficient—with the final coefficient of 0.466* ($p < 0.05$). As in the other models, the gradual decrease in coefficients was understandable as individual variation in student growth patterns over time would likely decrease the correlation between the weekly student scores and the initial intercept. There was an obvious difference between the GSACM level one coefficients and the CRACM and HRACM read models' coefficients. The difference in intercept starting and ending coefficients was more than .33 WRCM greater in both of the first two studied models than the GSACM.

The correlation coefficients for each of the gain scores were also regressed for the slope. The pattern was the same as the CRACM and HRACM in that the coefficients from the slope regression started small and gradually increased over the 13 weeks of the study. The measure for the slope between the first 2 weeks produced a weak coefficient of 0.046 ($p < 0.05$). The coefficient gradually increased over the 13 weeks until the final moderate coefficient for the slope was 0.455 ($p < 0.05$). This gradual growth in coefficients over time was consistent with the expectation that final student reads would reflect closely the outcome measures. It also reflected similar growth in the CRACM and HRACM.

Level Two: Student Demographic Variables

The pattern of significance for level two predictors in the GSACM was different than the other two models. When the initial GSACM was used to construct the intercept latent variable, none of the demographic variables reached significance for the intercept.

In the GSACM, only one demographic variable reached significance when regressed on the slope: the student's grade. For every unit increase in grade, the slope increased 3.494 ($p < 0.05$) WRCM. The result indicated that grade level has a significant positive effect on the slope in the GSACM.

Midyear DORF Scores Regressed on Demographic Variables

When the GSACM was used as the advancement criteria latent variable, only three of the seven demographic variables reached significance: beginning DORF score, grade level, and special education status.

The beginning DORF variable produced a greater intercept (0.948, $p < 0.05$), or starting WRCM, than both the CRACM and HRACM. For every word students could read correctly per minute on the beginning DORF oral reading benchmark, students read 0.948 WRCM more on the initial passage in the GSACM.

For every year of school, the student's coefficient for the intercept was -7.629 units (WRCM) lower in the GSACM. Similarly, the coefficient for students receiving special education services' intercept was -5.883 ($p < 0.05$)—a significant negative correlation.

High-Stakes State Assessment Regressed on Demographic and Other Variables

Economic status and grade level both had negative coefficients when regressed on the state end-of-year assessments in the GSACM. The coefficient for economic status was -2.474 ($p < 0.05$) and the coefficient for grade level was -2.848 ($p < 0.05$). Both of these variables were also significant in the CRACM and HRACM. The high-stakes assessment also produced a weak significant coefficient when regressed on the latent intercept (0.173, $p < 0.05$). The coefficient was 0.572 ($p < 0.05$) when the high-stakes assessment was regressed on the midyear DORF score.

R^2 for Gain Score Model

For the gain score model, the R^2 measure for the intercept was 0.134 ($p < .05$). This finding means that only 13.4% of the variance can be predicted by the latent intercept in the model. The R^2 for the slope was not significant. The R^2 for the midyear DORF score was a strong predictor at 0.743 ($p < .05$) or 74.3% of the variance accounted for by the midyear DORF score. The R^2 for the high-stakes AYP assessment was moderate, 0.464 ($p < .05$) or about half of the variance in the model account for by the end-of-the-year assessment when all other variables are held constant.

Intercept Regressed on Slope

When the intercept was regressed on slope in the GSACM, the correlation coefficient estimate was -2.424. This estimate did not meet significance requirements for a two-tailed p value ($p = 0.574$). Therefore, the finding was that the intercept could not be used to predict the slope.

Correlation Matrix

In addition to the correlations for the outcome variables discussed, the correlational relationships between demographics and the outcome variables were considered of note in this study. Table 8 displays the correlation matrix among the variables.

Table 8

Correlation Matrix of Study Variables

Variable	DORFE	DORFB	CRT	Gender	Race	Economic	ELL	Grade	SpEd
CORFE	1.000	.831**	.547**	.086	-.100*	-.046	-.094	.287**	-.295**
DORFB	.831**	1.000	.377**	.100*	-.077	.026	-.046	.549**	-.222**
CRT	.547**	.377**	1.000	.095	-.219**	-.239**	-.152**	-.145**	-.281**
Gender	.086	1.00*	.09	1.000	-.010	.053	.040	.048	-.114**
Race	-.100*	-.077	-.219**	-.010	1.000	.404**	.246**	.096	-.043
Economic	-.046	.026	-.239**	.053	.404**	1.000	.142**	.187**	.028
ELL	-.094	-.046	-.152**	.040	.246**	.142**	1.000	.028	.017
Grade	.287**	.549**	-.145**	.048	.096	.187**	.028	1.000	.053
SpEd	-.295**	-.222**	-.281**	-.114*	-.043	.028	.017	.053	1.000

* Correlation is significant at the 0.05 level (2-tailed)
** Correlation is significant at the 0.01 level (2-tailed).

The significant demographic variables studied only had weak predictive value in projecting outcome measures. Since special education students had been identified using a discrepancy model and are therefore performing at lower levels than expected, identified special educations student scores understandably were weakly, negatively correlated with beginning and ending oral reading scores and the high-stakes reading state assessment.

There was also a moderate correlation ($r = .549$, $p < .01$) between beginning DORF oral reading scores and the student’s grade level. This reflects a desired pattern of growth: as the student moves through each grade level, their oral reading fluency score as measured on the DORF would be expected to increase. However, there was only a weak correlation ($r = .287$, $p < .01$) between the midyear oral reading fluency (DORF) score and the student’s grade level. This indicated that not all growth is accounted for by a student’s maturation in the model and other instructional factors might be having a greater effect at midyear.

There was a weak negative correlation (*r* = -.145, *p* < .01) between the student's grade level and the state's high-stakes AYP reading assessment. A possible explanation for this initially unexpected finding will be discussed in the next chapter.

Results for ELLs were not significant in predicting oral reading fluency as measured by the DORF, however there was a slight negative correlation between ELL status and performance on the state's high-stakes AYP reading assessment (*r* = -.145, *p* < .01). Likewise, economic status and race/ethnicity were not significant predictors of oral reading fluency as measured by the DORF but were mild negative predictors of high-stakes test performance (economic *r* = -.239, *p* < .01; race/ethnicity *r* = -.219). All of the significant correlations between demographic variables and outcome variables were anticipated in direction and strength.

Summary

This study was considered to be quasi-experimental in that students in the study sample were neither randomly selected nor assigned to treatment groups. Twelve percent of the teachers in the district volunteered to report oral repeated reading data already being collected. The study sample's demographics reflected the demographics of the study district. Demographic indicators such as gender, grade, performance on the state high-stakes AYP assessment and special education status were all within the expected range. However, the study sample and the district demographics reflect a study population that can be considered high risk as the percentage of non-White students, the percentage of ELLs, and the percentage of economically disadvantaged students was considerably higher than the state average.

Student Growth Observations

Students' DIBELS oral reading benchmark scores increased over the 15 weeks of assessment administrations. However the students' proficiency scores, the DIBELS calculation indicating the likelihood of students' need for intervention, did not change significantly. While students increased their WRCM over the 15 weeks, students' proficiency scores did not change. Thus, students WRCM scores increased at the same rate as national norms.

The first week of repeated reading measurement results did not mirror growth during the subsequent weeks. This result was attributed to novelty effects. There appeared to be a minimum threshold of WRCM for student growth using the repeated reading routine. If a second through fifth-grade student read less than 25 words correctly per minute, they exhibited flat growth during the study.

Chi-Square Statistics

As indicated by the chi-square statistics, there were significant differences in the "goodness of fit" among the three models. A target 0.95 CFI and TLI indices of fit criteria was used to evaluate the advancement criteria models. The HRACM came very close to a strong fit at 0.94 on both the CFI and TLI. The CRACM scored close to the criteria 0.93 on the CFI and 0.92 on the TLI indices. The GSACM did not provide as good of fit scoring 0.80 on the CFI and 0.77 on the TLI. A 0.06 statistic was used as the minimum criteria for the RMSEA and SRMR indices of fit. The HRACM had the lowest RMSEA score: 0.047. Yet, the HRACM had the highest SRMR score at 0.064. Both scores met the criteria for the CRACM (RMSEA = 0.056, SRMR = 0.060). For the GSACM, the RMSEA met the criteria (0.057) but the SRMR for the GSACM did not meet the standard (0.112)

The statistical syntax was adjusted multiple times trying different combinations of dependent variables attempting to increase the goodness of fit for each of the three models. None of these attempts resulted in a better goodness of fit than the original proposed models. Although many of the variables did not reach significance when regressed upon the intercept or slope, they all contributed to the overall goodness of fit. In the end, the HRACM provided the best model fit based on the Chi Square statistical analysis. The CRACM also provided an acceptable fit. The GSACM did not meet the chi square statistical requirements for model fit.

Intercepts

The significance of the intercepts, representing the average beginning WRCM, varied depending on the advancement criteria model used. For the CRACM, the student's final DORF oral reading scores were 0.838 words higher for every word read correctly on the beginning DORF assessment. In the HRACM, the student's final DORF oral reading scores were 0.759 words higher for every word read correctly on the beginning DORF assessment. For the GSACM, the student's final DORF oral reading scores were 0.034 words higher for every word read correctly on the beginning DORF assessment. Thus, the GSACM's beginning DORF intercept was significantly lower than the other models.

The beginning DORF score and the student's race/ethnicity were the only variables with significant intercepts in the CRACM. In the HRACM, the beginning DORF score, gender, SES, ELL status, and grade were all significant intercepts. In the GSACM, only the student's grade had a significant intercept. Overall, the HRACM had the highest number of significant intercepts for variables affecting the model. In the CRACM, students with a non-White race/ethnicity had an intercept higher than their White counterparts. There was not a significant difference in the other two models.

Slope

The slope, representing the increase in weekly words read correctly per minute calculated separately for each independent variable holding the other independent variables constant, varied depending on the advancement criteria model used. In the CRACM and HRACM, a student's race/ethnicity was the only variable determined to have a significant positive slope. In the GSACM, there were no significant variables affecting the slope.

The regression slope, representing the increase in WRCM on the midyear DIBELS benchmark assessment, did not vary in which variables were significant for each of the advancement criteria models used. In all three advancement criteria models, the beginning DORF oral reading assessment score, the student's grade, and the student's special education status all produced a significant slope. In all three models, the beginning DORF oral reading assessment score produced a positive slope. Additionally, the student's grade and special education status both produced a negative slope regardless of the model.

The regression slope on the increase in score on the state high-stakes AYP assessment did not vary in which variables were significant for each of the advancement criteria. For all three models, the beginning DORF oral reading assessment had a positive slope on the AYP assessment and the student's grade level had a negative slope on the AYP assessment.

Correlations

A matrix of correlations revealed significant correlations between dependent and independent variables. The beginning DORF oral reading assessment had a high correlation (0.831) with the ending DORF oral reading assessment and a weak correlation (0.377) with the high-stakes AYP assessment. The ending DORF oral reading assessment had a moderate correlation (0.547) with the high-stakes AYP assessment and a moderate effect (0.549) on the ending DORF oral reading assessment.

Race/ethnicity had a mild negative correlation with the ending DORF oral reading assessment (-0.219) and a mild negative correlation with the high-stakes AYP assessment (-0.219). There was not a significant correlation between race/ethnicity and the beginning DORF assessment.

The SES of the student did not have a significant correlation with the beginning or ending DORF oral reading assessment. SES has a mild negative correlation (-0.239) with the high-stakes AYP assessment.

Similar to the SES, the ELL status of students did not have a significant correlation with the beginning or ending DORF assessment. There was a mild negative correlation (-0.152) between ELL status and the high-stakes AYP assessment.

Special education status had a weak negative correlation with the beginning DORF assessment (-0.222), the ending DORF assessment (-0.295), and the high-stakes AYP assessment (-0.281).

Grade level had a mild positive correlation (0.287) with the ending DORF assessment. There was no significant correlation between grade level and the beginning DORF assessment or the high-stakes AYP assessment.

Gender had a weak (.100) positive correlation with the beginning DORF assessment but was not significantly correlated to the ending DORF assessment or the high-stakes assessment.

In an analysis of the correlation coefficients, the only difference between models was the latent advancement criteria variable regressed on the high-stakes AYP assessment. In the GSACM, the correlation coefficient was -0.246. In both the CRACM and the HRACM, the correlation coefficient was 0.372 when regressed on the high-stakes AYP assessment.

The noteworthy finding from the R^2 calculations was that ending DORF oral reading scores accounted for much of the variance in the models: CRACM ($R^2 = 0.743$), HRACM ($R^2 = 0.771$), and GSACM ($R^2 = 0.743$). The intercept in the CRACM ($R^2 = 0.758$) and the HRACM ($R^2 = 0.733$) also accounted from significant variance. However, the intercept for the GSACM was a low predictor ($R^2 = 0.134$) of the variance in the model.

CHAPTER V

DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

Educators have used the repeated reading format proposed by Samuels (1979) for over 30 years. In the publication of the NRP report (2000), repeated reading was listed as an effective strategy for developing fluency. Yet, repeated reading's efficacy has been recently questioned (Goodman, 2006; Marcell, 2012). As with many educational procedures, poor or partial implementation of a strategy may have led to lower student performance (Reeves, 2011). Understanding the "how-to" of efficiently using evidence-based practices allows teachers to deliver successful, time-sensitive instruction and intervention to students.

Research Question One Findings

In this investigation, there were two research questions. First, was the GSACM, reflecting the gain score between cold (initial) reads and hot (final) repeated reading of a passage, a better model and therefore a better criterion than the currently popular criterion of the HRACM, reaching a set WRCM hot read, such as Samuels' (1979) criteria of 95 WRCM? Marcell (2012) argued that a developmental aspect should be considered when examining fluency and comprehension. Since the GSACM is a measure of student growth, it was hypothesized that the GSACM would more closely predict the outcome measures than the current practice of reading a passage to a predetermined criterion. An example of a predetermined growth criterion used in the study was the Read Naturally (2010) program, a popular fluency practice program that used a gain criterion of 30

WRCM between cold and hot scores. The null hypothesis for this question was that there would be no significant difference between the three models studied, GSACM, CRACM, HRACM, in their ability to predict successful outcomes on the measures selected. If the null hypothesis were shown to be untrue, recommendations could be made to alter practices thus better informing fluency instruction.

A two-stage path model using multilevel modeling as the first stage was performed on 13 weeks of repeated reading data on 400 second through fifth-grade students to determine the best model fit for the three advancement criteria models studied. The path model was used to analyze the ability of three models to predict students' performance on two outcome measures: midyear oral reading fluency and the state high-stakes AYP assessment. Chi-square statistics were significant at the .05 critical alpha levels, $p < .000$ for each of the three models. Similar to the Chard and colleagues' (2008) path model study, the goodness of fit indices (GFI) were unique for each of the models studied. Thus the null hypothesis was rejected and confidence reached to conclude that there was a significant difference among the three advancement criterion models' abilities to predict students' performance on two outcome measures: midyear oral reading fluency and the state high-stakes AYP assessment.

It was hypothesized that the GSACM would provide the best fit and therefore the gain score would be the best indicator of when a student is ready to move to practicing a new passage with repeated reading. The rationale for the hypothesis was that the gain score, a direct measure of the increase or growth of WRCM on a currently practiced passage, would reflect the student's individual outcome performance better than the

CRACM or HRACM. This hypothesis was rejected as the GSACM did not meet the statistical requirements for a good model fit: CFI = 0.800, TLI = 0.775, RMSEA = 0.057, SRMR = 0.112; whereas, the CRACM and HRACM both had acceptable goodness-of-fit statistics with the HRACM providing a slightly better model fit. (CRACM: CFI = 0.936, TLI = 0.928, RMSEA = 0.056, SRMR = 0.060; HRACM: CFI = 0.949, TLI = 0.943, RMSEA = 0.047, SRMR = 0.064.) In explaining these results it was found that there was little difference between students in their individual gain scores over time—a restriction in range producing similar gain scores for all students except those reading initially less than 25 WRCM. Therefore the gain score did not distinguish well between students given this restriction in range and was not deemed as a sufficiently reliable criterion to use to determine when students should advance to a new passage during repeated reading.

The CRACM, while providing a good fit, was not a practical recommendation as an advancement criterion. It did not make sense to use the beginning score as an advancement criterion. Since performance on the cold read for the next passage has yet to be done, it could not be used as a criterion for when to stop repeatedly reading the previous passage and moving on to the new passage. Therefore student WRCM on the hot read remained the recommended criterion for advancing to a new passage. This study found that the HRACM provided the most logical and best fit of the models studied.

If the HRACM provided the best fit and the GSACM was not an acceptable model, the next question was how this finding could be applied to instruction. Previous research has established the desirability of distributed practice in repeated reading: Durgunoglu (1993) and Krug (1990) both found that distributed practice over a number of days was more effective than repeated reading of a passage for only one day; and Rashotte and Torgesen (1985) posited three to five repetitions as an advancement criterion; and, Therrien (2004) found that using a criterion of three successive improvements on one passage was more efficient than Samuels' original method of reaching a criterion of 90 WRCM. This study found little variation between students in WRCM growth due to repeated reading. Therefore, it was recommended from the study findings that the most efficient criterion for knowing when a student is ready to move to a new passage was the students' hot read after practiced reading. This criterion was used in the popular *Six-minute Solution* program (Sopris West Educational Services, 2005).

Research Question Two Findings

The second question in the study was about which demographic variables such as age, ethnicity, gender, beginning reading ability and SES played a significant role in predicting the effectiveness of using weekly repeated reading scores as a predictor of benchmark reading measures at midyear and end of year. It was hypothesized that student demographics would directly influence the rate at which students progressed in learning to read, accounting for a large amount of variance in each of the three models. This study provided a unique contribution to the literature as distinctive student demographic effects on repeated reading have not been reported. The null hypothesis for this question was that there would be no significance in the student demographics within the models. Again, if the null hypothesis was shown to be untrue, recommendations could be made about specific student characteristics that could inform needed alterations in fluency instruction and practice.

This study proposed a unique theoretical multilevel path model to explain the relationship between student demographics and repeated reading's effects on the outcome measures of WRCM and Utah's state high-stakes AYP assessment.

The study findings showed that several demographic variables produced significant correlation coefficients. Since some of the demographic variables were significant, the null hypothesis was rejected. However regardless of significance, all demographic variables were found to be important as it was found that if any of the demographic variables were omitted from the model, the goodness of fit statistics were reduced beneath the acceptable level. Thus, it was concluded that the combination of demographic variables contributed to the overall fit of the model.

Beginning DORF Scores

The beginning DORF score was used a baseline WRCM. As it is the same measure as the midyear DORF, a strong correlation (0.831, $p < 0.05$) was expected. Of note, the WRCM measure correlation to the high-stakes AYP assessment starts mild (0.377, $p < 0.05$) and became moderate at the midyear DORF (0.547, $p < 0.05$). These were important accountability finding as the midyear DORF scores have been used as one point of documentation as teachers reported to parents and the state legislature which first through third-grade students were reading at grade level. If the DORF proved to be a poor predictor of AYP measures, then the use of the DORF progress monitoring tools would not predict outcome measures.

Also, the finding indicated time as a factor in the correlation between WRCM measures and high-stakes assessments and may explain the difference in previous fluency to comprehension correlation research findings. Goode, Simmons, and Kame'enui (2001) reported a strong correlation between accuracy and rate fluency measures and comprehension measures of approximately 0.80. Valencia and colleagues (2010) reported a moderate correlation between the two measures of 0.40 to 0.50. This study found WRCM given at the beginning of the year was a weak predictor of high-stakes performance given at the end of the year. However, the correlation strengthened over time. It was also important to note that the beginning DORF score did not significantly predict the slope, the rate of individual student growth, in any of the models. Regardless of the beginning WRCM level, students increased their WRCM at roughly the same rate using repeated reading as a fluency practice routine.

Gender

There was a weak correlation with girls scoring higher on the beginning DORF score (0.100, $p < 0.05$), the midyear DORF score (0.086, $p < 0.05$), and the high-stakes AYP assessment (0.095, $p < 0.05$). While gender had a weak correlation with the beginning WRCM and the outcome measures, the only significant correlation coefficient for gender was on the HRACM intercept. There was no significant difference in the effectiveness of repeated reading between boys and girls in any of the models. Thus, repeated reading is equally effective for both girls and boys.

High-Risk Groups

Race/ethnic status produced an expected, weak negative correlation on the outcome measures: midyear DORF (-0.100, $p < 0.05$) and the high-stakes AYP assessment (-0.219, $p < 0.05$). This discrepancy between White and non-White students was consistent with state gap-analysis data and the National Literacy Panel's report on second-language learners (August & Shannahan, 2006). Of particular note in this study, was the positive correlation coefficient when the race/ethnicity status was regressed on the slope in two of the models (1.029, $p < 0.05$ for the HRACM and 0.724, $p < 0.05$ for the CRACM). Race/ethnicity was the only demographic variable to produce a significant coefficient with the slope. This finding indicated that non-White students actually increased approximately one word per passage more than White students between cold and hot reads. Fitzgerald, Amendum, and Gutherie (2008) found that linear fluency growth over time for ELLs, although starting lower, closely mirrors native English speakers in developing readers. This study found that ELL students' linear fluency growth was slightly higher than their L1 counterparts when using repeated reading in second through fifth grades. This was an encouraging finding in the use of repeated reading for closing the reading achievement gap between non-White and White students.

ELLs, low SES, and special education (SPED) designation all produced negative, weak correlation with the high-stakes end of year AYP assessment (ELL -0.152, $p < 0.05$; economic -0.239, $p < 0.05$; SpEd -0.281, $p < 0.05$). None of these groups produced a significant coefficient when regressed on the slope. This finding indicated that ELL students, low SES, and SPED students all progressed at the same rate as their counterparts during repeated reading practice—an indication that repeated reading practice was equally effective in raising WRCM for all of these subgroups.

Model Variance

The intercept (average starting WRCM) and midyear DORF scores accounted for most of the variance in the recommended HRACM model (73.3% and 77.1%, respectively). The model slope accounted for only 12.2% of the variance. Since the model was based on WRCM, it was expected that a high percentage of student score variance would be predicted by the intercept and midyear DORF score. The low amount of variance predicted by the slope was another indicator that students' progress at the same rate when using the repeated reading strategy.

Limitations

In all statistical analysis there are procedural limitations. This study's research design constructed and tested a path model to compare advancement criteria. Multi-level modeling was used to control for student variations. Variation caused by teacher differences was controlled by nesting students within a teacher variable. Growth curves could then be calculated predicting outcomes on the midyear DORF benchmark assessment and on the high-stakes AYP state criterion referenced test. The design was considered quasi-experimental in that existing student data was used and teacher participation was voluntary. Since there was no control group, confidence could not be reached that all relevant background variables were identified and controlled.

Sample Limitations

The sample included 400 students ranging from second to fifth grade. Twelve percent of teachers in the study district volunteered to submit existing repeated reading data. The sample's demographics were very similar to the district's demographics. However, the study district had the highest free and reduced lunch rate in the state and 52% of students were classified as belonging to a minority group. Thus it is important to remember as a study limitation that the sample represented a higher-than-normal risk group. With more students in the lower part of the academic performance distribution curve, the power to detect variable interactions may have been statistically enhanced.

Instrumentation Limitations

All instrumentation used for growth and outcome measurement was and is commonly used in practice: repeated reading passages, the DORF benchmark assessment and the state high-stakes AYP criterion referenced test. However, it was important to note that all measures, including those used, have reliability limitations and therefore the interpretations made concerning student growth were correspondingly limited. The DIBELS Benchmark Assessment recently had concerns raised about its validity (Valancia et al., 2010).

The calculation of WRCM on repeated reading passages was self-reported by students. Therefore, the accuracy and reliability of scores could be questioned. Historically, self-reporting has been the common practice in the repeated reading procedure (Samuels, 1979) and has been considered a necessary part of the strategy to promote student motivation. Therefore, the inconsistency in reporting was tolerated as it exists in normal practice.

In order to control for validity and reliability, the DORF Benchmark assessments were administered by a district-level team that had been trained and had regular reliability checks. Due to these controls, the only limitations in DORF scores were those inherently part of the validity of the measure and administration issues were minimalized.

Likewise, the high-stakes AYP state criterion referenced test was highly monitored during administration. All proctoring teachers completed a testing ethics review before administration. Directions were scripted for teachers and students. Administration was monitored by the building administrator. The results were statistically analyzed and equated at the state level so that while the test was different for each grade level, the test score could be compared across grade levels. Due to the implemented standardization procedures, no unexpected factors were seen to limit the interpretation of the high-stakes AYP results.

Instructional Limitations

Although some variance in the repeated reading process was reported between teachers, the basic model proposed by Samuels was present in each classroom. However, many classroom level variables could not be controlled. For instance, the level of tier-one reading instruction received by each student varied due to the teacher’s knowledge, experience, and skill. Also the quality and quantity of interventions varied and could not be controlled between classes. Teacher knowledge varied dependent on the professional development experienced individually and at the school. In order to provide a measure of control for these variables, student growth data (level one) was analyzed using multi-level modeling clustered by teacher (level two). This hierarchical analysis prevented student growth from being masked by averaging student and teacher level variables.

Recommendations

As a result of the findings of this study, it was found that the HRACM provided the best model fit, and therefore, some form of hot read measurement provided the best indicator of when a student was ready to move to a new practice passage. This finding combined with the finding that student slope or rate of growth during repeated reading was constant between students, provided an important clue to efficient repeated reading practice. Two popular hot read criteria were identified in practice. In the Read Naturally (2010) program, students read until they have increased their reading score by 30 WRCM. In the Six-Minute Solution program (Sopris West Educational Services, 2005), students read the passage once per day for five days. Often in the Read Naturally practice, all repeated reading is completed in one sitting—a practice not recommended as per Durgunoglu (1993) and Krug (1990). Since student growth in repeated reading was shown in this study to be constant, the criteria of increasing 30 WRCM on a passage before moving to the next passage was not found to be the most efficient and therefore could not be recommended. Since student growth in repeated reading is constant, the number of times a passage is repeatedly read may be the best criteria for moving to a new passage—the criteria used in the Six Minute Solution program and recommended by Rashotte and Torgesen (1985). The Six Minute Solution program had the additional advantage of distributed practice over multiple days (Durgunoglu, 1993; Krug, 1990).

It was also found that student demographic variables had little effect on the fluency growth rate of individual students. Consequently it was found that repeated reading evidenced an equal effect for all students reading fluency growth as measure by the midyear DIBELS Benchmark assessment and Utah's end of year high-stakes AYP state criterion-referenced test. There were two exceptions to this finding: (a) students who read less than 25 WRCM initially did not show any growth from repeated reading practice, and (b) students in the non-White subgroup increased their fluency (WRCM) scores slightly faster than the White students using a repeated reading practice routine.

Future Studies

Since repeated reading was not found to be an effective strategy for students reading less than 25 WRCM, future investigation is needed to determine the best strategies for developing reading automaticity in these emerging readers. At their current level of ability and based on the study findings, repeated reading could not be recommended for these students. It can be recommended that future studies should focus on determining what reading skills need to be attained and practiced for these students to reach a level (25 WRCM) where repeated reading might be effective.

A second question emerged from the study to be addressed by future studies. Since a minimum threshold was identified for the effectiveness of repeated reading, is there a maximum threshold identifying when repeated reading is no longer effective and practice for automaticity should be discontinued in favor of other reading strategies? A growth curve analysis showing when practice trials have no appreciable effect would be of value to educators.

REFERENCES

Allington, R. (1983). Fluency: The neglected reading goal. *Reading Teacher*, *36*, 556-561.

August, D., & Shanahan, T. (Eds). (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth*. Mahwah, NJ: Erlbaum.

Bell, D. (1990). *The effect of repeated readings on elementary students' attitude toward reading* (Technical Report). Bowling Green, OH: Bowling Green State University.

Blachowicz, C. L. (2000). *Everybody reads fluency development program.* Springfield, IL: Illinois State Board of Education.

Chall, J. (1976). *The great debate: Ten years later, with a modest proposal for reading stages*. Pittsburgh, PA: Pittsburgh University Research and Development Center.

Champagne, R. A. (2008). *The impact of student mobility on Texas public school accountability ratings* (Unpublished doctoral dissertation). Tarleton State University, Stephenville, TX.

Chard, D. J. (2009). Repeated reading intervention for students with learning disabilities: Status of the evidence. *Exceptional Children, 75*, 263-281.

Chard, D. J., Stoolmiller, M., Harn, B A., Wanzek, J., Vaughn, S., Linan-Thompson, S., & Kame'enui, E. (2008). Predicting reading success in a multilevel schoolwide reading model. *Journal of Learning Disabilities, 41,* 174-188.

Chard, D. J., Vaugh, S., & Tyler, B. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities*, *35,* 386-407.

Chomsky, C. (1976). After decoding: What? *Language Arts, 53,* 288-296.

Clements, M., Reynolds, A., & Hickey, E. (2004). Site-level predictors of children's school and social competence in the Chicago child-parent centers. *Early Childhood Research Quarterly*, *19*(2), 273-296.

Cummings, K. D., Kennedy, P. C., Otterstedt, J., Baker, S. K., & Kame'enui, E. J. (2011). *DIBELS data system: 2010-2011 percentile ranks for DIBELS next benchmark assessments* (Technical Report 1101). Eugene, OR: University of Oregon.

Dahl, P. (1974). *An experimental program for teaching high speed word recognition and comprehension skills* (Final Report). Bloomington, MN: Bloomington Public Schools.

Dowhower, S. (1987). Effects of repeated reading on second-grade transitional readers' fluency and comprehension. *Reading Research Quarterly*, *22*, 389-406.

Durgunoglu, A.Y. (1993). Effects of repeated readings on bilingual and monolingual memory for text. *Contemporary Educational Psychology, 18*, 294-317.

Eldredge, J., Reutzel, D. R., & Hollingsworth, P. M. (1996). Comparing the effectiveness of two oral reading practices: Round-robin reading and the shared book experience. *Journal of Literacy Research*, *28*, 201-225.

Elley, W. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, *24*, 174-187.

Ehri, L. (1985, March). *Learning to read and spell*. Paper presented at the 69th Annual Meeting of the American Educational Research Association, Chicago, IL.

Fitzgerald, J., Amendum, S. J., & Guthrie, K. M. (2008). Young Latino students' English-reading growth in all-English classrooms. *Journal of Literacy Research, 40,* 59-94.

Folley, C. L. (1987, May). *Four longitudinal studies of the starter approach, a beginning reading strategy for nonreaders*. Paper presented at the annual meeting of the International Reading Association, Anaheim, CA.

Freeland, J., Skinner, C., Jackson, B., McDaniel, C., & Smith, S. (2000). Measuring and increasing silent reading comprehension rates: Empirically validating a repeated readings intervention. *Psychology in the Schools*, *37*, 415-29.

Fuchs, D., Fuchs, L., Yen, L., McMaster, K., Svenson, E., Yang, N., … King, S. (2001). Developing first-grade reading fluency through peer mediation. *Teaching Exceptional Children*, *34,* 90-93.

Gamse, B., Bloom, H., Kemple, J., Jacob, R., & Institute of Education Sciences. (2008). Reading First impact study: Interim report. NCEE 2008-4016. *National Center for Education Evaluation and Regional Assistance*. Washington, DC: Institute of Education Sciences.

Goode, R., Simmons, D., & Kame'enui, E. (2001). The importance and 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

Goodman, K. S. (2006). *The truth about DIBELS: What it is, what it does.* Portsmouth, NH: Heinemann.

Greene, E. C. (1988, January). Effects of new computer technology on increases in children's word recognition automaticity. *Paper presented at the Annual Meeting of the Association for Educational Communications and Technology,* New Orleans, LA.

Hindon, A., & Paratore, J. R. (2007). Supporting young children's literacy learning through home-school partnerships: The effectiveness of a home repeated-reading intervention. *Journal of Literacy Research, 39*, 307-333.

Hintze, J. M., Ryan, A. L., & Stoner, G. (2003). Concurrent validity and diagnostic accuracy of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) and the Comprehensive Test of Phonological Processing. *School Psychology Review*, *32*, 541-556.

Huey, E. B. (1908). *The psychology and pedogogy of reading.* Cambridge, MA: MIT Press.

Kameenui, E. J. (1993). Diverse learners and the tyranny of time: Don't fix blame; fix the leaky roof. *The Reading Teacher, 46*, 376-83.

Kaminski, R., Cummings, K. D., Powell-Smith, K. A., & Good, R. H., III (2008). Best practices in using Dynamic Indicators of Basic Early Literacy Skills for formative assessment and evaluation. In A. Thomas & J. Grimes (Eds.) *Best practices in school psychology V* (pp. 1181-1204). Bethesda, MD: National Association of School Psychologists.

Kato, K., Albus, D., Liu, K., Guven, K., & Thurlow, M., (2004). *Relationships between a state-wide language proficiency test and academic achievement assessments* (LEP Projects Report 4). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Koskinen, P. S., & Blum I. H. (1984). Repeated oral reading and the acquisition of fluency. In J. Niles & L. Harris (Eds.), *Changing perspectives on research in reading/language processing and instruction* (pp. 183-187). Rochester, NY: National Reading Conference.

Krug, D. (1990). Massed versus distributed repeated reading: A case of forgetting helping recall? *Journal of Educational Psychology, 82*, 366-371.

Kuhn, M. R. (2000). *Fluency: A review of developmental and remedial practices.* Ann Arbor, MI: Center for the Improvement of Early Reading Achievement.

Kuhn, M. R. (2005). A comparative study of small group fluency instruction. *Reading Psychology, 26*(2), 127-146.

Kuhn, M., & Stahl, S. (2000). *Fluency: A review of developmental and remedial practices*. Ann Arbor, MI: Center for the Improvement of Early Reading Achievement.

LeBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293-323.

Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading and Writing Quarterly*, *13*(2), 123-146.

Marcell, B. (2012). Putting fluency on a fitness plan: Building fluency's meaning-making muscles. *The Reading Teacher, 65,* 242-249.

Mastropieri, M., Leinart, A., & Scruggs, T. (1999). Strategies to increase reading fluency. *Intervention in School & Clinic*, *34*(5), 278-283.

Mercer, C., Campbell, K., Miller, M., Mercer, K., & Lane, H. (2000). Effects of a reading fluency intervention for middle schoolers with specific learning disabilities. *Learning Disabilities: Research & Practice*, *15*(4), 179-189.

Moskal, M. K. (2006). *Partnering for fluency.* New York, NY: Guilford.

Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Washington, DC: National Institute of Child Health and Human Development.

Nations Report Card. (2008). *2008 long-term trends.* Retrieved from http://nationsreportcard.gov/ltt_2008/

Nichols, W. D., Rupley, W. H., & Rasinski, T. (2009). Fluency in learning to read for meaning: Going beyond repeated readings. *Literacy Research & Instruction, 48*(1), 1-13.

O'Connor, R. E., White, A., & Lee Swanson, H. (2007). Repeated reading versus continuous reading: Influences on reading fluency and comprehension. *Exceptional Children, 74*(1), 31-46.

O'Shea, L. J., Sindelar, P. T., & O'Shea, D. J. (1985). The effects of repeated readings and attentional cues on reading fluency and comprehension. *Journal of Reading Behavior, 17*(2), 129-142.

O'Shea, L. J., Sindelar, P. T., & O'Shea, D. J. (1987). The effects of repeated readings and attentional cues on reading fluency and comprehension of learning disabled readers. *Learning Disabilities Research, 2*(2), 103-109.

Perfetti, C. A. (1985). *Reading ability.* New York, NY: Oxford University Press.

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

Rashotte, C. A., & Torgesen, J. K. (1985). Repeated reading and reading fluency in learning disabled children. *Reading Research Quarterly*, *20*(2), 180-188.

Rasinski, T. V. (1989). Fluency for everyone: Incorporating fluency instruction in the classroom. *The Reading Teacher*, *42*, 690-693.

Rasinski, T. V. (2006). A brief history of reading fluency. In S. J. Samuels (Ed.), *What research has to say about fluency instruction* (pp. 4-23). Newark, DE: International Reading Association.

Rasinski, T. V., Padak, N., Linek, W., & Sturtevant, E. (1994). Effects of fluency development on urban second-grade readers. *The Journal of Educational Research, 87*, 158-165.

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

Read Naturally. (2010). *Monitoring student performance*. Retrieved from http://www.readnaturally.com/howto/adjlevgoals.htm

Reeves, D. B. (2011, March). *Leaders make the difference: How leadership decisions influence teaching, students results, and community engagement*. Webinar presented at the Leadership and Learning Center, Englewood, CO.

Reutzel, D. R. (2006). "Hey, teacher, when you say 'fluency,' what do you mean?" In T. B. Rasinski (Ed.), *Fluency instruction: Research-based best practices* (pp. 62-85). New York, NY: Guilford.

Roehrig, A., Petscher, Y., Nettles, S., Hudson, R., & Torgesen, J. (2008). Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology*, *46*(3), 343-366.

Samuels, S. J. (1979). The method of repeated reading. *The Reading Teacher*, *32*, 403-408.

Samuels, S. J. (2006a). Reading fluency: Its past, present, and future. In T. B. Rasinski (Ed.), *Fluency instruction: Research-based best practices* (pp. 7-20). New York, NY: Guilford.

Samuels, S. J. (2006b). Toward a model of reading fluency. In S. J. Samuels (Ed.), *What research has to say about fluency instruction* (pp. 24-46). Newark, DE: International Reading Association.

Schreiber, P. A. (1980). On the acquisition of reading fluency. *Journal of Reading Behavior*, *12*, 177-186.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Mahwah, NJ: Erlbaum.

Simmons, D. C., Fuchs, L. S., Fuchs, D., Mathes, P., & Hodge, J. P. (1995). Effects of explicit teaching and per tutoring on the reading achievement of learning-disabled and low-performing students in regular classrooms. *The Elementary School Journal, 95*, 387-408.

Sindelar, P., Monda, L., & O'Shea, L. (1990). Effects of repeated reading on instructional- and mastery-level readers. *Journal of Educational Research, 83*(4), 220-226.

Sopris West Educational Services. (2005). *The six-minute solution: A reading fluency program.* Retrieved from: http://store.cambiumlearning.com/Resources/Program/Overview/pdf/sw_Overview_SixMinute_Pri_01.pdf

Stanovich, K. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, *16*(1), 32-71.

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading. *Remedial & Special Education, 25*, 252-261.

Topping, K. J. (1987). Paired reading: A powerful technique for parent use. *The Reading Teacher, 40*, 608-614.

Topping, K. J. (2006a). Building reading fluency: Cognitive, behavioral, and socioemotional factors and the role of peer-mediated learning. In S. J. Samuels (Ed.), *What research has to say about fluency* (pp. 106-129). Newark, DE: International Reading Association.

Topping, K. J. (2006b). Paired reading: Impact of a tutoring method on reading accuracy, comprehension, and fluency. In T. B. Rasinski (Ed.), *Fluency instruction: Research-based best practices* (pp. 173-191). New York, NY: Guilford.

Trachtenberg, P., & Ferruggia. A. (1989). Big books from little voices: Reaching high risk beginning readers. *The Reading Teacher, 42*, 284-289.

Valencia, S., Smith, A., Reece, A., Li, M., Wixson, K., & Newman, H. (2010). Oral reading fluency assessment: Issues of construct, criterion, and consequential validity. *Reading Research Quarterly, 45*, 270-291.

Walpole, S., Chow, S., & Justice, L. (2004). Literacy achievement during kindergarten: Examining key contributors in an at-risk sample. *Early Education and Development*, *15*, 245-264.

APPENDICES

Appendix A

Statistical Analysis Data

Table A1

Intercept or Difference in Beginning WRCM by Latent Variables

Variable	Estimate	*SE*	Estimated *SE*	Two-tailed *p* value
Using cold reads as the WRCM measure				
DORFB	.838*	.036	23.172	.000
Gender	1.646	2.183	.754	.451
Race/ethnicity	5.435*	2.190	2.482	.013
Economic	-3.718	2.653	-1.402	.161
ELL	-12.279*	4.604	-2.667	.008
Grade	1.448	1.411	1.026	.305
SpEd	-.890	3.064	-.291	.771
Using hot reads as the WRCM measure				
DORFB	.759*	.043	17.577	.000
Gender	5.048*	2.543	1.985	.047
Race/ethnicity	1.002	2.659	.377	.706
Economic	-6.990	3.493	-2.001	.045
ELL	-12.496*	4.768	-2.621	.009
Grade	4.610*	1.605	2.873	.004
SpEd	-4.307	4.847	-.889	.374
Using gain scores from cold to hot as the WRCM measure				
DORFB	.034	.084	.410	.682
Gender	1.273	1.731	.735	.462
Race/ethnicity	-2.643	1.958	-1.350	.177
Economic	-3.246	2.346	-1.384	.166
ELL	-1.367	2.686	-.509	.611
Grade	3.435*	1.068	3.215	.001
SpEd	-2.955	3.140	-.941	.347

$N = 400$.
$^*p < .05$, two-tailed.

Table A2

Regression Slope or Differences in the Increase in WRCM over Time Based by Latent Variables

Variable	Estimate	*SE*	Estimated *SE*	Two-tailed *p* value
Using cold reads as the WRCM measure				
DORFB	-.153	.113	-1.361	.174
Gender	.363	.283	1.283	.199
Race/ethnicity	.724*	.322	2.249	.025
Economic	.304	.382	.797	.426
ELL	.936	.560	1.671	.095
Grade	-.157	.203	-.771	.440
SpEd	-.402	.320	-1.254	.210
Using hot reads as the WRCM measure				
DORFB	-.086	.100	-.864	.387
Gender	.098	.375	.261	.794
Race/ethnicity	1.029*	.409	2.518	.012
Economic	.423	.496	.853	.394
ELL	1.014	.687	1.475	.140
Grade	-.319	.217	-1.467	.142
SpEd	-.252	.489	-.515	.607
Using gain scores from cold to hot as the WRCM measure				
DORFB	.159	.145	1.096	.273
Gender	-.048	.260	-.186	.852
Race/ethnicity	.181	.288	.629	.529
Economic	.113	.350	.324	.746
ELL	-.030	.398	-.075	.940
Grade	-.165	.168	-.977	.329
SpEd	.091	.486	.187	.852

N = 400.
$*p < .05$, two-tailed.

Table A3

Regression Slope of Midyear DIBELS Oral Reading Fluency Score on Demographics by Latent Variable

Variable	Estimate	*SE*	Estimated *SE*	Two-tailed *p* value
Using cold reads as the WRCM measure				
DORFB	.646*	.063	10.201	.000
Gender	-.022	.025	-.888	.375
Race/ethnicity	-.039	.028	-1.402	.161
Economic	-.023	.028	-.816	.414
ELL	-.032	.028	-1.137	.255
Grade	-.223*	.028	-7.917	.000
SpEd	-.056*	.021	-2.672	.008
Using hot reads as the WRCM measure				
DORFB	.711*	.051	13.910	.000
Gender	-1.471	1.677	-.877	.381
Race/ethnicity	-.290	1.841	-.157	.875
Economic	-1.083	2.345	-.462	.644
ELL	-3.966	3.313	-1.197	.231
Grade	-8.015*	.958	-8.365	.000
SpEd	-4.784*	1.726	-2.772	.006
Using gain scores from cold to hot as the WRCM measure				
DORFB	.946*	.025	37.293	.000
Gender	-.163	1.870	-.087	.931
Race/ethnicity	1.449	1.908	.759	.448
Economic	-2.202	2.467	-.892	.372
ELL	-6.246*	3.557	-1.756	.079
Grade	-7.473*	1.147	-6.514	.000
SpEd	-5.926*	2.088	-2.839	.005

$N = 400$.
*$p < .05$, two-tailed.

Table A4

Regression Slope of High-Stakes State Accountability Measure (CRT) on Demographic by Latent Variable

Variable	Estimate	*SE*	Estimated *SE*	Two-tailed *p* value
Using cold reads as the WRCM measure				
DORFE	.467*	.080	5.815	.000
DORFB	.062	.108	.574	.566
Gender	.985	.734	1.343	.179
Race/ethnicity	-1.550	.810	-1.914	.056
Economic	-2.828*	1.029	-2.749	.006
ELL	-2.369	1.653	-1.433	.152
Grade	-2.892*	.466	-6.206	.000
SpEd	-1.347	1.121	-1.202	.229
Using hot reads as the WRCM measure				
DORFE	.437*	.077	5.662	.000
DORFB	.010	.098	.104	.918
Gender	.858	.728	1.180	.238
Race/ethnicity	-1.454	.789	-1.843	.065
Economic	-2.587*	1.050	-2.463	.014
ELL	-2.052	1.558	-1.317	.188
Grade	-3.115*	.490	-6.354	.000
SpEd	-1.282	1.089	-1.178	.239
Using gain scores from cold to hot as the WRCM measure				
DORFE	.541*	.077	7.017	.000
DORFB	.042	.090	.465	.642
Gender	1.018	.752	1.353	.176
Race/ethnicity	-1.095	.796	-1.377	.169
Economic	-2.476*	1.050	-2.358	.018
ELL	-2.071	1.491	-1.390	.165
Grade	-3.000*	.545	-5.505	.000
SpEd	-1.079	1.296	-.832	.405

$N = 400$.
*$p < .05$, two-tailed.

Table A5

R^2 *Results*

Variable	Estimate	*SE*	Estimated *SE*	Two-tailed *p* value
Cold read model				
End DORF score	.787	.020	39.563	.000
High-stakes AYP scores	.432	.037	11.600	.000
Initial WRCM (intercept) score	.758	.031	24.607	.000
Increase (slope) in cold score	.169	.064	2.633	.008
Hot read model				
End DORF score	.717	.021	36.430	.000
High-stakes AYP score	.443	.037	11.926	.000
Initial WRCM (intercept) score	.733	.032	22.721	.000
Increase (slope) in hot score	.122	.048	2.564	.010
Gain score model				
End DORF score	.743	.024	31.466	.000
High-stakes AYP scores	.464	.039	12.044	.000
Initial WRCM (intercept) score	.134	.050	2.691	.007
Increase (slope) in cold score	.033	.048	.687	.492

Appendix B

Mplus® Version 5 Syntax

Mplus® version 5 Syntax for Analysis

Using Cold Read Scores as Advancement Criterion

```
i1 s1 | Cold 1 @ 0 Cold2@1 Cold3@2 Cold4@3 Cold5@4 Cold6@5 Cold7@6
Cold8@7 Cold9@8 Cold10@9 Cold11@10 Cold12@11 Cold13@12 Cold14@13
Cold 15@14;

i1 s1 ON DORFB Gender Race Economic ELL Grade SpEd;

CRT ON DORFE DORFB i1 s1 Gender Race Economic ELL Grade SpEd;

DORFE ON DORFB i1 s1 Gender Race Economic ELL Grade SpEd;

i1 WITH s1;
```

Using Hot Read Scores as Advancement Criterion

```
i1 s1 | Hot 1@0 Hot2@1 Hot3@2 Hot4@3 Hot5@4 Hot6@5 Hot7@6
Hot8@7 Hot9@8 Hot10@9 Hot11@10 Hot12@11 Hot13@12 Hot14@13
Hot 15@14;

i1 s1 ON DORFB Gender Race Economic ELL Grade SpEd;

CRT ON DORFE DORFB i1 s1 Gender Race Economic ELL Grade SpEd;

DORFE ON DORFB i1 s1 Gender Race Economic ELL Grade SpEd;

i1 WITH s1;
```

Using Hot Read Scores as Advancement Criterion

```
i1 s1 | y1@0 y2@1 y3@2 y4@3 y5@4 y6@5 y7@6 y8@7
y9@8 y10@9 y11@10 y12@11 y13@12 y14@13 y15@14;

i1 s1 ON DORFB Gender Race Economic ELL Grade SpEd;

CRT ON DORFE DORFB i1 s1 Gender Race Economic ELL Grade SpEd;

DORFE ON DORFB i1 s1 Gender Race Economic ELL Grade SpEd;

i1 WITH s1;
```

CURRICULUM VITAE

GREGORY PAUL LEWIS

Office: Ogden School District, 1950 Monroe Blvd., Ogden, UT 84401
Home: 827 W. 1700 S., Marriott-Slaterville, UT 84404

Telephone: Office (801)737-7288
Home (801)334-9303

ECUCATION

University	Fields of Study Major and Minor	Degree	Year
Utah State University	Curriculum & Instruction/	PhD	2012
Utah State University	Level I & II Reading Endorsement		2006
Utah State University	Educational Administrative/ Supervisory Endorsement		1991
Weber State University	Secondary Education	MS	1990
Brigham Young University	Composite Biology	BS	1982

Educational Experience

Executive Director of Elementary Education, Ogden School District (OSD), 2011-Present
Responsible for K-6 schools, technology, and assessment
Executive Director of Curriculum, Ogden School District (OSD), 2008-2011
Responsible for K12 curriculum, technology, and assessment
Reading First Coordinator OSD, 2003-2008
Coordinate the Reading First Grant in five elementary schools
Support for District K-3 Reading Initiative "Performance Plus" Grant
Coordinate district DIBELS assessment and records
Bonneville Elementary School, OSD
Elementary Principal, 1994-2003
Mound Fort Middle School, OSD
Middle School Assistant Principal, 1991-1994
Central Middle School, OSD
Science Teacher, 1982-1991

Professional Memberships

International Reading Association
Utah Elementary Principal's Association
Ogden Administrator's Association

Service

Board of Directors, Elizabeth Stewart Treehouse Children's Museum
Marriott-Slaterville City, Parks and Recreation Assistant Chair

Honors and Awards

Reading First project honored as Exemplary Reading First Project for the state of Utah, 2006
Curriculum Innovator of the Year, Utah Elementary Principal's Association, 2002

References

The following persons have written letters of recommendation on my behalf and have consented to reference inquiry:

- D. Ray Reutzel, Ph.D.
  Emma Eccles Jones Endowed Chair
  Utah State University, Logan, UT 84322-6515
  435-797-8631
- Parker Fawson, Ph.D.
  Professor and Chairperson, Department of Curriculum and Instruction
  University of Kentucky
  335 Dickey Hall
  Lexington, KY 80406-0017
  859-257-0767
- John A. Smith, Ph.D.
  Professor and Chairperson, Department of Curriculum and Instruction
  Box 19167
  Arlington, TX 76019
  817-272-0116
- Rebecca Donaldson, Ph.D.
  Title One Specialist, USOE
  250 Ease 500 South
  PO Box 144200
  Salt Lake City, UT 84114-4200
  801-538-7869
- Rich Moore, Ph.D.
  Superintendent, Livingston School District
  133 S. 5th Street
  Livingston, MT 59047
  406-222-0861

Professional Presentations

National

1. "Take'n the bus…to achievement: Effective Leadership Teams." Presented with Leslie Law and Scott Robertson at the National Reading First Conference, Reno, NV, July 2006.

Regional

1. "Template Instruction." Featured Speaker, State Curriculum Directors, SLC, UT, Nov. 2005
2. "Explicit Instruction." Featured Speaker, Student Teacher Workshop sponsored by Ogden Chapter of UIRA, Ogden, UT, Nov. 2005
3. "Explicit Instruction." Featured Speaker, Reading First New Teacher State Conference, SLC, UT, Aug 2006.
4. "Vocabulary." Presenter at regional Words Their Way Conference, SLC, UT, Nov. 2006.
5. "Finding Your Way through the Letter-Name Alphabetic Stage." Presenter at regional Words Their Way Conference, SLC, UT, Feb 2007.
6. "Teaching to Mastery." Presenter at Utah Principal's Literacy Academy, Ogden, UT, Mar. 2007.
7. "Teaching to Mastery." Presenter at State Literacy Coaches Conference, Ogden, UT, Apr. 2007.
8. "Word Games." Presenter at regional Words Their Way Conference, SLC, UT, Nov. 2007.
9. "Helping Struggling Readers in the Early Stages." Presenter at regional Words Their Way Conference, SLC, UT, Nov. 2007.
10. "Data for Coaches." Presenter with Leslie Law at Utah Advanced Coach Training, Provo, UT, Feb 2008.

Consulting and Professional Service

1. "Using the Basal Effectively." Workshop for Wasatch Elementary School, Ogden, UT, Sep 2006.
2. "A Literacy Environment." Workshop for T. O. Smith Elementary School, Ogden, UT, Feb 2007.
3. "Differentiated Instruction." Workshop for Grandview Elementary School, Ogden, UT, Aug 2007.
4. "Differentiated Instruction." Workshop for T.O. Smith Elementary School, Ogden, UT, Aug 2007.
5. "Fluency." Workshop for Liberty Elementary School, Murray, UT, Aug, 2007.
6. "Effective Basal Instruction." Workshop for Hillcrest Elementary School, Ogden, UT, Aug 2007.
7. "Fluency and the Basal." Workshop for Lincoln Elementary School, Ogden, UT, Aug 2007.

8. "Words Their Way." Workshop for Ogden City Schools Teachers, Ogden, UT, Aug 2007.
9. "Instructional Change." Workshop for Grandview Elementary School, Ogden, UT, Jan 2008.
10. "Differentiated Instruction." Workshop series for Wasatch Elementary School, Ogden, UT, Jan-Feb 2008.
11. "Explicit Teaching." Workshop for Grandview Elementary School, Ogden, UT, Jan 2008.
12. "Vocabulary." Workshop for Horace Mann Elementary School, Ogden, UT, Feb 2008.
13. "SBRR Reading Instruction." Multiple day workshop for Quest Academy, Hooper, UT, 2008-09.
14. "Turnaround School Leadership." Multiple day workshop for administrators, Ogden, UT, 1011-12.

Adjunct Class Instructor

1. Theories and Models of Reading TEAL/ELED 6230 (Foundations course for reading endorsement), Spring 08, Fall 08, Spring 10.
2. Supervision of Reading, TEAL 6590 (Level II reading supervision endorsement), Summer 09.