

12-22-2009

Reading Regression and Recoupment Among Students With and Without Learning Disabilities

Nicole E. Musil

Indiana University of Pennsylvania

Follow this and additional works at: <http://knowledge.library.iup.edu/etd>

Recommended Citation

Musil, Nicole E., "Reading Regression and Recoupment Among Students With and Without Learning Disabilities" (2009). *Theses and Dissertations (All)*. 548.

<http://knowledge.library.iup.edu/etd/548>

This Dissertation is brought to you for free and open access by Knowledge Repository @ IUP. It has been accepted for inclusion in Theses and Dissertations (All) by an authorized administrator of Knowledge Repository @ IUP. For more information, please contact cclouser@iup.edu, sara.parme@iup.edu.

READING REGRESSION AND RECOUPMENT AMONG STUDENTS
WITH AND WITHOUT LEARNING DISABILITIES

A Dissertation

Submitted to the School of Graduate Studies and Research
in Partial Fulfillment of the
Requirements for the Degree
Doctor of Education

Nicole E. Musil

Indiana University of Pennsylvania

December 2009

Indiana University of Pennsylvania
The School of Graduate Studies and Research
Department of Educational & School Psychology

We hereby approve the dissertation of

Nicole Elise Musil

Candidate for the degree of Doctor of Education

Joseph F. Kovalski, D.Ed.
Professor of Educational & School
Psychology, Advisor

William Barker, Ph.D.
Professor of Educational & School
Psychology

Becky Knickelbein, Ph.D.
Associate Professor of Education of
Exceptional Persons

Lynanne Black, Ph.D.
Assistant Professor of Educational &
School Psychology

ACCEPTED

Timothy P. Mack, Ph.D.
Dean
The School of Graduate Studies and Research

Title: Reading Regression and Recoupment Among Students With and Without Learning Disabilities

Author: Nicole Musil

Dissertation Chair: Joseph F. Kovalesski, D.Ed.

Dissertation Committee Members: William Barker, Ph.D.
Becky Knickelbein, Ph.D.
Lynanne Black, Ph.D.

The purpose of the current study was to compare students' oral reading fluency (ORF) skills before and after summer break. This study examined whether students whose ORF regressed would regain those skills within a month of returning to school. In particular, the study examined differences between students with and without specific learning disabilities (SLDs) in reading. Additionally, the study analyzed the effects of age, sex, summer program attendance, and initial low achievement had on regression and recoupment.

Examiners administered curriculum-based measurement (CBM) probes measuring ORF to 137 students in May and September 2008. Students whose scores dropped ten percent or more received a follow-up administration in October 2008. Several repeated measured analyses of variance (ANOVA-RM) were conducted to determine whether students' age, sex, SLD status, achievement level, and summer program attendance impacted ORF regression and recoupment. Given that only four students with SLDs participated, inferential statistics were inappropriate in analyzing the impact SLD status on regression. Overall, the analysis yielded insignificant results when comparing relative ORF among different groups of students. Analyses also indicated that students did not regress as a whole group. However, significant recoupment occurred among those students whose ORF regressed from pre-test to post-test.

The use of a convenience sample limited the generalizability of the current study's results to that of other samples. In particular, many groups were too small to have adequate statistical power to meet the

assumptions of the analysis used. Including more students with SLDs would have led to a stronger analysis to determine whether this group differs significantly from the general population. Additionally, the current sample lacked students eligible for free or reduced lunch. Having a more economically diverse sample would help to determine whether this factor impacted regression and recoupment. Finally, parent survey information describing students' access to summer reading practice and activities was inaccessible. Thus, it is recommended that further studies collect parent information regarding summer reading practice.

TABLE OF CONTENTS

CHAPTER		Page
I	THE PROBLEM.	1
	Introduction	1
	The Impact of Summer Break	3
	SLD vs. Low Achievement.	8
	How Much Regression is "Typical"?.	10
	Sex Differences and Reading Achievement.	13
	Statement of the Problem	14
	Summary.	15
	Research Questions and Hypotheses.	16
II	REVIEW OF RELATED LITERATURE	19
	ESY Background	19
	Case Law and ESY	20
	Armstrong v. Kline	20
	Other Relevant ESY Cases	28
	Nature and Definition of Regression.	29
	Recoupment and Other Factors to Consider	29
	Beliefs About ESY.	30
	The SLD Construct.	30
	Current SLD Regulations.	34
	Family Income and Summer Reading Loss.	37
	Impact of Summer Reading Intervention Programs	39
	Sex Differences.	41
	Curriculum-Based Measurement	44
	Developmental Status	46
	Readability.	49
	Summary.	51
III	METHODS.	53
	Introduction	53
	Design	54
	Population	56
	Sample	56
	Measurement.	57
	Predictor Variables.	57
	Special Education Status	57
	Sex.	58
	Grade.	59
	Achievement Level.	59
	Summer Program Attendance.	60
	Developmental Status	60
	Dependent Variables.	61
	Pre-test Oral Reading Fluency.	61
	Post-test Oral Reading Fluency	61
	Follow-up Oral Reading Fluency	61
	Instrument	61
	Description and Development of Probes.	62
	Readability.	63

CHAPTER
III

METHODS (Continued)

	Validity	66
	Reliability.	68
	Standard Error of Measurement.	68
	Alternate-Form Reliability	69
Procedures		69
	Examiner Training	69
	Pre-test	69
	Creation of Achievement Level Groups	70
	Post-test.	70
	Follow-up Assessment	70
	Analysis	71
	Statistical Techniques	73
	Computer Programs.	73
	Parent Survey.	74
Summary.		74
IV	RESULTS.	76
	Introduction	76
	Did Students Regress?.	77
	Impact of Standard Error of Measurement.	78
	Summer Regression.	79
	Summer Improvement	79
	Minimal Changes.	80
	Research Question 1: Specific Learning Disabilities.	80
	Research Question 2: Sex Differences	81
	Regression	82
	Recoupment	83
	Research Question 3: Low Achievement	84
	Regression	84
	Recoupment	86
	Research Question 4: Summer Reading Program Attendance	87
	Research Question 5: Grade Level	89
	Research Question 6: Overall Recoupment.	91
	Summary.	92
V	DISCUSSION	95
	Overview.	95
	Current Findings on Regression and Recoupment	95
	Research Questions and Hypotheses	100
	Specific Learning Disabilities.	100
	Sex Differences	102
	Low Achievement	103
	Summer Reading Program Attendance	105
	Developmental Status (Age and Grade Level).	107
	Recoupment.	109
	Limitations	110
	Implications for Research	113
	Implications for the Practice of School Psychology.	115
	Summary	117
REFERENCES.		120

LIST OF TABLES

1	Research Questions, Latent Variables, Observed Variables, Instrument/Source, Validity and Reliability.	58
2	AIMSweb Benchmark Scores as of May 2008.	59
3	Grade Equivalent Readability Estimates for R-CBM Passages using OKAPI! and Microsoft Word Software.	65
4	Alternate-form Reliability of AIMSweb Standard Benchmark Reading Assessment Passages.	66
5	Standard Error of Measurement and Alternate Form Reliability for R-CBM Passages	68
6	Research Questions, Hypotheses, Variables, Statistical Analyses, and Statistical Assumptions for the Reading Regression-Recoupment Project.	72
7	Cross-tabulation of Students Based on Achievement level, Sex, Grade, Learning Disability Status, and Summer School Attendance	76
8	Descriptive Statistics Regarding Whole Group Pre-test to Post-test Changes in ORF scores.	77
9	Analysis of Variance with Repeated Measures for Variables Impacting ORF.	78
10	Pre-test to Post-test Changes in ORF Scores.	79
11	Descriptive Statistics Regarding SLD Status and Oral Reading Fluency Measures.	81
12	Descriptive Statistics Regarding the Impact of Sex Differences on Oral Reading Fluency Measures.	82
13	Analysis of Variance with Repeated Measures for Effect of Sex Differences on Oral Reading Fluency Recoupment.	83
14	Descriptive Statistics of the Impact of Low Achievement on Summer Regression of Oral Reading Fluency Measures.	85
15	Descriptive Statistics Regarding the Impact of Low Achievement on ORF Recoupment	87
16	Descriptive Statistics Regarding the Impact of Summer Reading Program Attendance on Oral Reading Fluency Measures.	88
17	Descriptive Statistics of the Effect of Grade Level on ORF Measures.	90

LIST OF TABLES (continued)

18	Analysis of Recoupment. Descriptive Statistics by Grade Level on Post-Test and Follow-Up Oral Reading Fluency Measures.	91
19	Descriptive Statistics Regarding ORF Recoupment	92
20	Analysis of Variance with Repeated Measures for Recoupment.	92

LIST OF FIGURES

1	Research Path Diagram of the Latent Variables.	18
2	Research Design Diagram for Regression-Recoupment Project. .	54
3	Research Path Diagram of the Regression Recoupment Project .	55
4	Interaction between Sex and Achievement Level for ORF Regression	86
5	Interaction of Sex and Summer Program Attendance on ORF Regression	89

CHAPTER I

THE PROBLEM

Introduction

"Data-based decision-making and accountability" comprises one of the key roles and job functions of school psychologists. In creating their second *Blueprint*, or model for training and practice, the National Association of School Psychologists (NASP) listed this as the first of ten domains of school psychology leadership and functioning (National Association of School Psychologists, 1997). At that time, NASP reported that data-based decision-making and accountability should be "the organizing theme for school psychology training and practice," stating that it should "permeate every aspect of the practice."

School psychologists traditionally have focused their practice on collecting data to make decisions about individual students. A more recent revision to the *Blueprint* recommend that school psychologists extend their data collection beyond the individual student level, focusing on whole classrooms, school systems, and programs as well (National Association of School Psychologists, 2006). Collecting group data allows school psychologists to determine the efficacy of programs, class interventions, and system policies. In addition, this practice provides a normative comparison for individuals or groups of individuals. For example, collecting and interpreting data about students in the general education population provides a comparison for those with disabilities.

School psychologists serve on school teams that determine whether students with disabilities need extended school year (ESY) services. Federal law mandates that school teams make this decision on an annual basis. ESY services aim to prevent regression, or a loss of skills, during breaks in school. More specifically, students with disabilities may

require ESY services when data suggests that it will take them longer than a "typical" peer to "recoup," or regain any skills they lost during educational breaks.

Some school systems regularly collect data regarding regression and recoupment of skills for their students with disabilities. However, at this time there is a dearth of data regarding academic skill regression and recoupment among general education students. School psychologists often benefit from normative comparisons when making educational decisions. Both local and national norms can inform educational decisions such as whether a specific child (or group of children) requires ESY services. The current study attempts to provide such information for a single school district.

In addition to comparing the academic performance of students with and without disabilities before and after summer break, the current study focuses specifically on students with specific learning disabilities (SLD). According to the National Center for Learning Disabilities (2008) students with SLD comprise the largest category of students receiving special education services under the Individuals with Disabilities Education Act (IDEA). In addition, there is a lack of data suggesting that students with SLD are as likely to experience regression and students with more severe disabilities.

School psychologists play an important role in the assessment of students' educational needs. Furthermore, many school psychologists spend a considerable amount of time assessing whether students have SLD as well as how to educate students with poor reading skills regardless of disability. Thus, this study examined the pre and post-summer reading performance of students with SLD compared to their peers. Such information has value for the field of school psychology due to the professional roles

of using normative comparisons when determining eligibility for SLD, monitoring progress, and making decisions about ESY eligibility.

The Impact of Summer Break

Why do most American students attend school nine months out of the year? When formal education was in its infancy, community needs determined the length and timing of school calendars. During that time, the school year varied from 11 to 12 months per year in urban communities and five to six months in agricultural areas. In an effort to standardize school calendars across the nation at the turn of the century, a compromise yielded the current school calendar. Although 85% of the population worked in agriculture at that time, less than 3% do today (Cooper, Nye, & Charlton, 1996). Thus, the current practice of a summer "vacation" from school presently serves as more of a welcome respite than an economic essentiality. Having this extended time away from school intuitively permits children with time to relax, process learned information, and possibly participate in activities that may potentially enhance their future academic achievement. However, despite its seemingly favorable effects, one must consider *which* students benefit from a three-month break.

Children and youth with disabilities comprise an important group of students for whom a three-month repose from formal educational experiences may hinder academic progress. Due to various learning problems and accompanying adverse educational effects experienced by students with disabilities, they often have slower rates of skill acquisition and generalization. This can make extended breaks from school and special education services particularly detrimental for these children and youth. In addition, long-term outcomes are less favorable among students with disabilities compared to their general education peers. For example, nation-wide data on school dropout rates have shown that students with

various disabilities are about twice as likely as their peers to drop out of school (U.S. Department of Education, 2007).

Students with different disabilities vary in terms of dropout rates, with students with emotional disturbance (ED) showing the poorest outcomes. According to the U.S. Department of Education, Office of Special Education and Rehabilitative Services, Office of Special Education Programs (2007), in 2005 dropout rates were about 56% for students with ED, 32% for students with specific learning disabilities (SLD), and 29% for students with mental retardation (MR). Similarly, when comparing graduation rates of students with and without disabilities, only 51.9% of students receiving special education completed high school compared to 73.9% of their peers in general education (National Center for Education Statistics, 2007).

Educators and other professionals have generated a variety of ideas regarding the mission of education when considering long-term outcomes for students with disabilities. Siegel and Sleeter (1991) cited preparing students with disabilities to participate in the adult world to the maximum extent possible as one such goal. Court cases such as *Armstrong vs. Kline* (1979) and the appellate cases that followed considered the attainment of the highest levels of self-sufficiency and individual potential as desired educational outcomes.

While excluded from statistics regarding graduation and dropout rates, children with severe and profound MR also have notable post-educational disadvantages. Members of this population sometimes are referred to as "severely and profoundly impaired" (SPI) due to the likelihood that they may have other impairments. Due to the severity of their disability, students with SPI often require intensive rehabilitative services throughout their life spans. For students with SPI as well as other those with other disabilities, extended breaks from their

educational environment and related services may hinder the goal of maximizing their individual potential and participation in society. The rationale of ESY includes allowing students to sustain the development and progress they made during the school year. Thus, ESY serves as a preventive measure for those students for whom such gains would be lost during extended breaks without access to their special education and related services.

Of these three disability categories recognized as considerably more likely to drop out of school than peers in general education, ED and MR are considered to be "severe" disabilities, while SLD often is considered to be a "mild" condition. Court cases such as *Armstrong vs. Kline* (1979) and *Battle vs. Commonwealth of Pennsylvania* (1980) set the precedent that breaks in schooling had detrimental effects on some students with "severe" disabilities such as MR and ED.

The court found that some disabled students showed significant losses of critical skills over school breaks as well as difficulty regaining lost skills in a timely manner upon returning to school. These cases did not determine how much regression was "typical" among students with severe disabilities. Such regression may be referred to as an absolute change in skill level during educational breaks. They also failed to assert that students without disabilities did not experience skill regression over breaks in education. Instead, data on the individual plaintiffs and testimony by expert witnesses led to the decision that children with certain "severe" disabilities may be eligible to receive schooling and special education services outside the duration of the regular school year. The court decided that limiting these students' access to education beyond the standard 180 days violated their right to a "free, appropriate public education" (FAPE) promised to individuals with disabilities by Section 504 of the Rehabilitation Act of 1973.

The court decree from *Armstrong v. Kline* deemed the following disabilities to be "severe," thus warranting the potential need for Extended School Year (ESY) services: autism/pervasive developmental disorders, ED, severe MR, degenerative impairments with mental involvement, and severe multiple disabilities. However, the *Armstrong* and later cases made no mention of the potential skill losses that students with *mild* disabilities may potentially experience. Just as these cases lacked data regarding the actual regression shown by groups of students with various disabilities, little is known about how much skill loss is shown by students with mild disabilities. Nor is much known regarding normative loss shown by the comparison group of these students' general education peers.

Nevertheless, due to *Armstrong v. Kline* and the related class-action lawsuits that followed, IDEA currently mandates the consideration of ESY services for all students with disabilities. According to Section §300.106 of IDEA:

1. Each public agency must ensure that extended school year services are available as necessary to provide FAPE, consistent with paragraph (a)(2) of this section.
2. Extended school year services must be provided only if a child's IEP Team determines, on an individual basis, in accordance with Sec. 300.320 through 300.324, that the services are necessary for the provision of FAPE to the child.
3. In implementing the requirements of this section, a public agency may not—
 - a. Limit extended school year services to particular categories of disability; or
 - b. Unilaterally limit the type, amount, or duration of those services.

Thus, schools may deem Extended School Year (ESY) services necessary for any student identified as having a disability. What makes a student with special needs eligible for ESY? According to the IDEA, any student receiving special education services may be eligible for ESY. Two major factors impacting a student's potential need for ESY include regression, defined as a significant loss of skills over time, and recoupment, or the time it takes to regain those skills.

Students with SLD constitute one of the largest groups of school-aged children receiving special education services, and approximately half of all students receiving special education have SLD as their primary disability category (U.S. Department of Education, 2006). In addition, most students identified as having SLD have reading difficulties; it has been estimated that 80% or more of the approximately 2.9 million school-aged children identified as having this SLD have reading disabilities (Lichtenstein, 2008; U.S. Department of Education, 2006).

Despite the prevalence of SLD, especially in reading, few studies have assessed the extent to which these students show skill regression over school breaks. In particular, students with "mild" disabilities adversely impacting reading skills, such as children identified as having SLD have yet to receive adequate attention in the literature regarding the need for ESY. Similarly, few studies have measured the severity of regression and (lack of) recoupment affecting any students with disabilities. Finally, little is known about how much skill regression is present among students in the general education population.

Case law has deemed students with SLD as potentially eligible for ESY services given academic skill regression-recoupment as well as other individual factors (*Reusch v. Fountain*, 1994). At least in theory, having a longer break from academic instruction can have more significant detrimental effects on students with SLD compared to their peers

(Katsiyannis, 1991). However, there are little data regarding the prevalence of these students showing regression-recoupment rates warranting the need for ESY.

Another issue in determining eligibility for ESY for students with SLD relates to the definition of SLD. Based on one historically prevalent definition, students identified using the ability-achievement discrepancy model have higher IQs than some of their peers with similarly poor academic achievement. Thus, it is possible that some low-achieving students without SLD may show similar levels of regression of reading skills during summer months.

SLD vs. Low Achievement

When making a determination regarding a child's potential disability status (eligibility for SLD), school teams must rule out several factors that may contribute to a child's lack of achievement. According to IDEA 2004, the definition of SLD, found in United States Code, follows:

The term 'specific learning disability' means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which disorder may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations.

Such term includes such conditions as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia.

Such term does not include a learning problem that is primarily the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage. (20 U.S.C. §1401 [30])

In addition to the federal definition, a historic and pervasive conceptualization of "learning disability" stipulates that students meeting the SLD criteria have both average intelligence and deficient achievement. Thus, given two poor readers, one with an IQ in the "average" range and one with a "below average" score, the student with "average" intelligence is more likely to be eligible for special services

(when an ability-achievement discrepancy is considered when making this determination).

What, besides performance on tests of cognitive abilities or general intelligence, differentiates these two students? Several studies have found minimal or nonexistent differences when comparing deficient readers with and without learning disabilities. Fletcher et al. (1994) found that despite the assumption that cognitive processing deficits distinguish students with SLD and poor readers, one study failed to find any significant differences between these two groups. Another study found no significant differences in the developmental trajectories of precursor and reading-related skills when comparing low achieving readers and students with learning disabilities (Shaywitz, Fletcher, Holahan, & Shaywitz, 1992). In addition to examining student data related to cognitive abilities and achievement, the authors of this study surveyed parents of both groups.

Parents of children with low reading achievement and those with SLD reported similarities in their children's self-perceptions, health histories, and behavior. Shaywitz et al. also found that students with SLD showed better reading achievement in second grade compared to other low-achieving students. Furthermore, the study found that compared to their similarly low-achieving peers, students with SLD showed better rates of reading improvement between second and fifth grade. Overall, when comparing low achieving students with and without identified SLDs, Shaywitz et al. found more similarities than differences between the two groups. Other authors comparing children with SLD to poor readers have drawn similar conclusions regarding the minimal differences between these two groups of students (Hoskyn, & Swanson, 2000; Spear-Swerling & Sternberg, 1998; Fuchs, Fuchs, Mathes, Lipsey, & Roberts, 2001).

How Much Regression is "Typical"?

While researchers have studied summer decline in reading skills for over 100 years, many of these studies did not use statistical analyses. Additionally, most research on reading skill loss made comparisons based on often-misunderstood grade equivalent scores. Assessors (such as teachers and researchers examining skill regression) frequently misinterpret grade equivalent scores as well, assuming that these scores indicate a student's instructional level. Only one study using curriculum-based measurement (CBM), a measurement tool designed to assess incremental progress over time, could be located.

When asking how much regression a "typical" student shows, one must consider the definition of "typical." Taking one reading teacher's definition of "average" to mean typical among her students, Ross (1974) found that such students improved over summer months. In her study, Ross compared 119 sixth-grade students' different levels of reading skill change over the summer months. Data analysis included a comparison of poor, average, and skilled readers' regression rates. Results indicated that average and skilled readers showed improvements, while poor readers regressed.

More recently, Mraz and Rasinski (2007) measured the changes in reading decoding and fluency skills among 116 middle-class first, second, and third grade students during summer break. Their results yielded less optimistic results than Ross' data; they found that between May and September, 45% of the sample's reading decoding skills declined, and 25% of the students' experienced a decrease in reading fluency. Mraz and Rasinski also found that lower achieving students' skills declined more than those of their "average" peers.

Looking at multiple studies often is helpful when considering how much summer skill regression is "typical" outside of special education

populations. In order to assess overall findings of summer skill regression, Cooper, Nye, and Charlton, (1996) conducted a meta-analysis of 39 summer regression studies conducted between 1975 and 1996. As a whole, their findings indicated that "at best" students showed no academic growth on repeated measures of standardized tests. At worst, they lost about 1/10 of a standard deviation, or one month grade equivalent, of reading skills.

Cooper et al. described these estimates as conservative due to the studies' unclear descriptions of how much instructional time was included between measures. For example, school districts included in the studies failed to specify the length of their summer vacations. The authors also presumed that assessment took place before the last day of school for the ending school year, and after the first day of the following year. Thus, the unknown differences in instructional time included in the studies impacted the ability to make an accurate assessment of actual skill regression over summer months. Somewhat surprisingly, results of the analysis indicated longer summer intervals in their meta-analysis yielded less significant losses in test scores. The authors hypothesized that rather than actual longer summer breaks, more instructional time between pre- and post-tests led to the observed differences in skill regression.

A student's grade level also has been found to impact the degree of skill loss observed over summer months. For example, results across studies found that students in lower elementary grades showed minimal regression, while those in grades 4 and up showed more significant losses (Cooper, Nye, & Charlton, 1996). The authors contended that these results likely arose from "floor effects" of the standardized academic tests used in these studies' repeated measures designs.

Several studies have aimed to determine whether students with SLD showed reading skill regression during the summer months. Cornelius and Semmel (1982) found that students with SLD showed significant reading

skill regression. They also determined that this skill loss was preventable given access to a five-week summer program. Shaw (1982, as cited in Allinder & Eicher, 1994), used the Wide Range Achievement Test (WRAT) to measure reading skills before and after summer break, also finding a statistically significant decrease in reading skills at the end of the break.

One study comparing reading achievement over the summer months among students with "mild handicaps" yielded contrasting results. Cook and Schwartz (1969), who used the WRAT to measure pre- and post-summer reading achievement among students with disabilities, found that the students showed no significant regression in reading skills over the summer months. Thus, in the few studies examining this population, students with mild disabilities such as SLD have either been found to regress in reading progress or experience no change in their skill level during summer months.

A later study conducted by Allinder and Eicher (1994) had two strengths compared to earlier studies on regression among students with mild disabilities. First, they considered both regression and recoupment; second, they used curriculum-based measurement (CBM) to measure change over time. Using a sample of 75 elementary school students identified as having mild disabilities impacting academic skills in reading and math, the authors found that the students' reading skills significantly declined over the summer. However, six weeks after their return to school, the students not only recouped their previously lost reading skills: they also showed significant improvements compared to their previous scores.

Allinder and Eicher commented that practitioners could interpret the results of their study regarding student recoupment in two ways. The fact that students recouped their skills within six weeks might question the need for services such as ESY. On the other hand, educators might consider

the fact that students took six weeks to regain their lost skills as evidence that IEP teams should consider such a service in order to prevent similar skill losses.

Sex Differences and Reading Achievement

While no data currently exist regarding males' and females' regression and recoupment of reading skills, nation-wide data has indicated that about two thirds of students receiving special education services are male (US Dept of Education, 1998). Sex differences are most pronounced in the area of SLD. Researchers have attributed this phenomenon to the "three B's": biological and behavioral differences between boys and girls, and bias in special education referral procedures (Tschantz & Markowitz, 2003). Additionally, Shaywitz, Shaywitz, Fletcher, and Escobar (1990) determined that the method of ascertainment of SLD in reading and the statistical methods employed impacted whether children received a reading disability diagnosis.

Despite some contention that males' vulnerability to reading disability is a "myth," other research has supported this claim (Flannery, Liederman, Daly, & Schultz, 2000; Liederman, Kantrowitz, & Flannery, 2005). For example, when examining a large sample (n = 32,223) of children with learning disabilities, Flannery, Liederman, Daly and Schultz (2000) found a significant male to female ratio (about 2 to 1). Furthermore, even when eliminating ascertainment bias, Liederman, Kantrowitz, and Flannery (2005) still found a "significant preponderance" of boys with reading disabilities. However, a meta-analysis examining 39 studies of summer setback found no significant differences between male and female students (Cooper, Nye, & Charlton, 1996).

When considering regression and recoupment, it is important to note that none of the aforementioned studies used progress-monitoring data such as CBM to assess improvement or decline in reading skills. Given its

technical validity and capability for measuring small increments of change over time, using CBM to compare boys' and girls' reading progress would likely provide a valid basis for comparison.

Taken together, despite biases and behavior differences between the sexes, several studies have implied that boys may truly have more vulnerability to reading disabilities.

Statement of the Problem

There is a dearth of research examining the prevalence of academic skill regression and delayed recoupment among students identified as having so-called "mild" disabilities (e.g. SLD). At the same time, students with SLD constitute about half of all students receiving special education services under IDEA. Decisions regarding ESY often rely on data regarding students' regression and recoupment of academic skills; however, such data rarely is collected and/or reported by school systems. Furthermore, there are no federal definitions of "regression" or "recoupment," nor has case law set any quantifiable standards defining significant regression and recoupment.

Students with SLDs account for 45% of students aged 6-21 receiving special education services under IDEA, comprising the largest of all the disability categories serviced according to this law (National Center for Learning Disabilities, 2008). Thus, students with SLDs make up a significant proportion of special education students who may be eligible for ESY.

However, few studies have examined the degree of skill regression experienced by this group of students. Of the few studies documenting skill regression, only one considers recoupment rates (Allinder & Eicher, 1994).

Studies considering skill regression of students with SLD have used global, standardized tests which lack sensitivity to change (Cook &

Schwarz, 1969; Cornelius & Semmel, 1982; Ross, 1974). Only one study (Allinder & Eicher, 1994) used curriculum-based measurement (CBM), a metric intended to measure student progress over time. This study found that although students with SLD did show regression in reading fluency skills, they not only recouped but significantly improved their performance upon returning to school.

The current study will expand on Allinder and Eicher's work by examining how much regression takes place among students with SLD as well as their peers in the general population. In addition, this study will assess how much recoupment takes place among both groups of students. Thus, while the literature currently lacks information comparing summer regression and recoupment rates of students with learning disabilities to their peers, the current study will examine this potential difference. Finally, due to empirically observed vulnerabilities to reading disabilities experienced by males, the current study will examine whether sex differences impact regression and recoupment of reading fluency skills. The study's analysis will compare males' and females' performance among students receiving special education services as well as those in the general education population.

Summary

School psychologists spend a considerable amount of time collecting data to support quality decision-making and accountability. Such decisions include, yet are not limited to, assisting in educational planning for individual students with disabilities. Having normative data such as students in general education often is helpful in informing large-scale educational decisions. Such data also serves as a comparison group when considering the educational needs of students with disabilities.

School psychologists participate in school teams that make decisions about individual students with disabilities, such as whether ESY services

are warranted. Factors such as anticipated regression and lack of recoupment over extended school breaks may warrant such services. Which students receive ESY? Students with severe physical, cognitive, and emotional disabilities usually have more pronounced needs for ESY compared to those with more mild disabilities such as SLDs.

Students with SLDs comprise the largest group of students with disabilities, with most of them struggling in reading. However, little information currently is available regarding the extent to which students' reading skills decline over the summer. The current study assesses students' oral reading fluency (ORF) skills using CBM, a validated measurement of overall reading competence for the age group used. This type of data has strong technical properties and also is able to be administered in a quick and inexpensive manner.

The dearth of quality data regarding reading skill regression and recoupment among students with and without disabilities makes ESY decisions for this group difficult. Such data also would be beneficial in informing schools about the efficacy of summer reading practice or specific programs. Finally, collecting regression and recoupment data will help determine whether differences occur in subgroups of students such as different sexes and those with initial low achievement. Overall, collecting CBM data before and after summer will assist in data-based decision-making regarding students with and without disabilities. Such data may serve in educational planning, program evaluation, and enhancing general knowledge about regression and recoupment of reading fluency skills.

Research Questions and Hypotheses

1. How much oral reading fluency regression/recoupment do students with SLD show compared to their peers in general education during summer break?

It is hypothesized that students with SLD will show more regression and less recoupment than their peers in general education.

2. How do male and female students differ in terms of oral reading regression/recoupment following summer break?

It is hypothesized that male students will show more regression and less recoupment than their female peers.

3. What relationship will exist between pre-, post- and follow-up R-CBM measures?

It is hypothesized that students with very deficient skills will be more likely to show regression and lack of recoupment than those scoring closer to average/normal limits on the pre-test.

4. How does summer reading program attendance impact ORF regression and recoupment?

It is hypothesized that attending a summer reading program will lead to students either maintaining (no change) or improving their ORF scores from pre-test to post-test.

5. Is there any association between age / grade (developmental status) and ORF regression and recoupment?

It is hypothesized that ORF will increase with age/grade. However, no significant differences are anticipated in terms of regression/recoupment for the age groups assessed in the current study.

6. Will one month back in school improve ORF for students who regressed from spring to fall (pre-test to post-test)?

It is hypothesized that returning to school for one month will significantly increase ORF scores for students who regressed from pre-test to post-test.

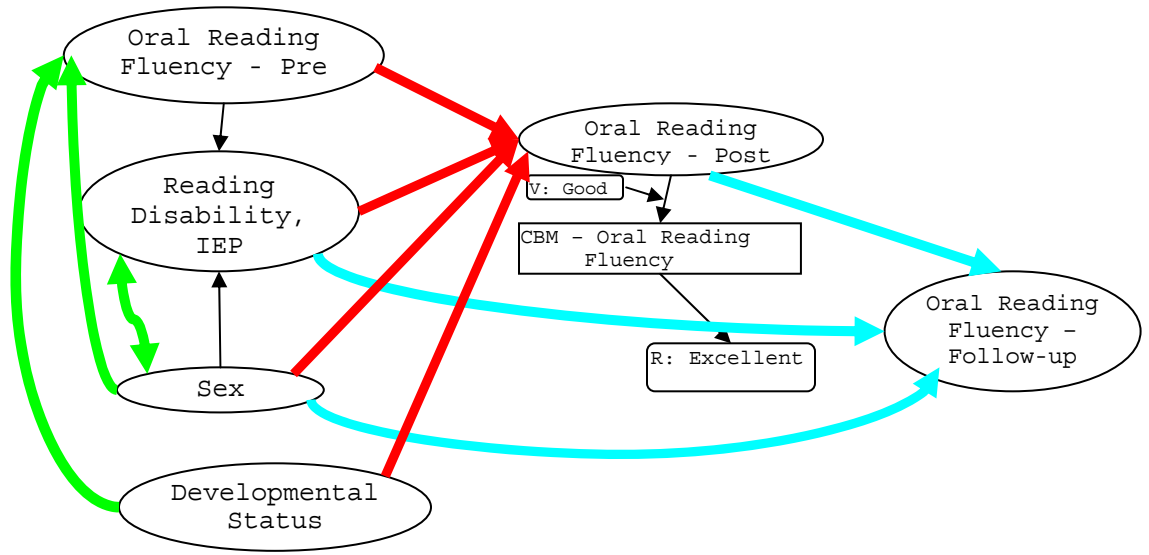


Figure 1. Research path diagram of the latent variables.

CHAPTER II

REVIEW OF RELATED LITERATURE

ESY Background

According to the Education for All Handicapped Children Act of 1975 (EHA), now called the Individuals with Disabilities Education Act (IDEA), students with disabilities are entitled to a "free, appropriate public education" (FAPE). Given extended breaks from school, some students may show regression in the areas of academic, behavioral, social, and communication skills. Certain students with disabilities may be especially susceptible to such losses, which may be determined to be severe enough to warrant a more continuous access to the services they received in school. Thus, according to section §300.106 of IDEA, such students may be eligible to receive Extended School Year (ESY) services. Such determinations are made by the Individual Education Plan (IEP) team for each child receiving special education services.

ESY initially began due to noted regression among students with disabilities displaying a regression in skills that took an inordinate amount of time to regain or "recoup" following instructional breaks. Currently, ESY services must be considered for all students with disabilities. For example, according to the current definition in §300.106(a)(3) of IDEA, public agencies may not:

- (i) Limit extended school year services to particular categories of disability; or
- (ii) Unilaterally limit the type, amount, or duration of those services.

Case Law and ESY

According to Prasse (2002), case law and other legal sources have influenced the provision of special education services. As a result of such litigation, ESY services are guaranteed under the free, appropriate public education (FAPE) clause of IDEA. ESY eligibility issues such as length of programming and using regression/recoupment to make eligibility decisions also emanated from legal sources (Olmi, Walker, & Ruthven, 1995).

Armstrong v. Kline

Of all the court cases pertaining to ESY, *Armstrong v. Kline* (1979) set a precedent by questioning the appropriateness of the traditional school year. This case determined that denying programming beyond 180 days to children with severe and profound mental retardation and severe emotional disturbance violated the "appropriateness" clause of EHA. When considering appropriateness, this case determined that decisions must be made based on the individual child's needs. Thus, if an individual student demonstrated a need for services beyond the typical school calendar, the school district must provide those services.

In addition to challenging the "180 day rule" formerly in effect in the state of Pennsylvania, *Armstrong v. Kline* (1979) began a series of class action lawsuits. In fact, Armstrong and four handicapped children and their families began three class action lawsuits in 1978. Due to the overlapping nature of the issues presented in their cases, the three lawsuits were consolidated and held in the Circuit Court the following year. The plaintiffs included Gary Armstrong, an eight year old classified as severely and profoundly impaired (SPI); Richard H., an 18-year-old with a severe emotional disturbance (ED); Patricia Sue Battle, a 20-year-old

with ED as well as brain injury; and Natalie Bernard, 17 years old, with mental retardation (MR) and an orthopedic impairment.

According to the Federal Rules of Civil Procedure, these five individual cases formed a class composed of "All handicapped school-aged persons in the Commonwealth of Pennsylvania who require or may require a program of special education and related services in excess of 180 days and the parents or guardians of such persons." All three actions filed suit against Caryl Kline, who served as Pennsylvania's Secretary of Education at that time. As Secretary, Kline was the chief official of the Pennsylvania Department of Education (PDE), the state education agency responsible for providing a FAPE to all school-aged children with disabilities in Pennsylvania. The local education agencies (LEAs) or school districts that the students attended also were defendants in the case. Finally, the "approved private schools" (APS) attended by two of the plaintiffs also were sued given their status as sites approved to provide services to handicapped children.

The defendant school districts, PDE, and APSs all were eligible recipients of federal funds for the provision of services to students with disabilities. Those funds included federal financial assistance, some of which was received under EHA. At that time, the PDE policy and practice did not require the schools to provide any child with an education or educational services in excess of 180 days per year. The PDE also informed hearing officers that they lacked the authority to order special education programming for greater than 180 days.

Prior to the lawsuit(s), the plaintiffs attempted to challenge the 180 day rule. For example, in Gary Armstrong's case, the hearing officer recommended as much programming as possible to prevent regression. The examiner had made that decision due to the degenerative nature of his condition as well as the expert opinion that "a 12 month program for Gary

was important to prevent regression during the summer months." The other four families included in the *Armstrong v. Kline* case did not request special education due process hearings. Had they done so, however, their efforts likely would have been fruitless considering the hearing officers' orders that they could not recommend programs in excess of 180 days.

Regarding characteristics of the plaintiffs and the educational needs of those in their class, their disabilities were considered to be "two separate, occasionally overlapping, categories." These categories included SPI with MR and other handicaps and ED. Students with severe retardation were described as those with IQs below 30, often experiencing difficulty moving, and often entering the school setting lacking many basic language and self-help skills. Those with profound retardation were described as likely to be unable to speak or walk, with minimal means of communication.

Because students with more severe forms of retardation are more likely to have other impairments, children with severely and profound MR often were grouped together and referred to as SPI. The findings of the court recognized that these children usually learned at a much slower rate than their peers. They also had difficulty remembering learned information and generalizing it to multiple settings. One of the defendants, Natalie Bernard, had Down's Syndrome and was considered SPI due to her delayed intellectual and social abilities. She also had developed double scoliosis, requiring her to wear a brace due to two sideways curvatures of her spine.

Gary Armstrong, who exhibited the signs of a normal child until he was about two years old, was diagnosed with San Filippo, Type A Syndrome, a genetic disease. Due to his condition, he began showing signs of hyperactivity, his rate of learning decreased, and his parents were unable to toilet train him. Symptoms of San Filippo, Type A Syndrome included

progressive physical, emotional, and mental deterioration as well as joint contractures and seizures. Almost all children with this condition start acting aggressively as the disease progresses, making it difficult for their parents to care for or control them.

At the time of the trial, Gary Armstrong had SPI including physically disabling progressive joint contractures and deafness. At the time of *Armstrong v. Kline*, certain students with ED were considered to have "severe" disabilities. The court considered children with autism spectrum disorders and those with schizophrenia as having severe ED. In addition, so-called "symbiotic" children who showed extreme attachment and exhibited high levels of anxiety and bizarre behavior if their routines or demands for sameness were not met were included in the group of children considered to have ED. The case described children with ED as having extreme difficulties learning due to poor frustration tolerance, impulsivity, inattentiveness, and lack of self-control. Similar to SPI children, those with ED reportedly had difficulty generalizing skills they had learned.

Three of the plaintiffs in the *Armstrong* case had ED, resulting in symptoms such as temper tantrums, compulsive behavior, self-abuse, and hyperactivity. Regarding the purpose of education for children with SPI and ED, expert witnesses in the *Armstrong* case concurred that goals such as reaching the child's individual potential, or the highest level of self-sufficiency possible were appropriate. For children with degenerative diseases such as Gary Armstrong, providing an education that allows the child to be mobile, at home, and increasing life expectancy were judged to be appropriate goals.

The educational programs of children with disabilities severe enough to warrant serious consideration of ESY services such as those mentioned in the *Armstrong v. Kline* case often focused on functional rather than

"traditional" academic skills such as reading and mathematics. For example, Natalie Bernard's educational program included basic functional academics such as basic language and arithmetic. Due to the goals of her education, Natalie also learned self-help skills such as getting dressed and pre-work skills. The aim of her program included providing her with skills that eventually would allow her to live in a group home or sheltered workshop rather than an institutional setting. In Gary Armstrong's case, due to his syndrome and the behaviors it caused, his programming aimed to keep him living with his family as long as possible. The skills he learned were even more basic, focusing on allowing his parents to manage him. Gary learned self-control skills to manage his violent and aggressive behavior, sign language to cope with his hearing loss, and basic skills such as coming when called, feeding, and toilet training. Related services such as psychotherapy often are included in the curricula of students with severe ED (SED). Overall, the educational programs of students included in the *Armstrong v. Kline* case class action lawsuit differed significantly from those of students with "mild" disabilities such as SLD.

Based on the expert testimony provided in *Armstrong v. Kline*, the court decided upon several factors in the educational environment impacting "whether and at what rate" a child learned. They identified child's disability as an important limiting factor on how much a child could learn. Teacher competency and opportunities to practice outside of the classroom also were considered to be significant. Finally, expert testimony concluded that allowing a child more time to practice a skill increase the likelihood of mastery. Thus, the plaintiffs contested that during breaks in programming, their children's skills and development regressed so much that it rendered progress "impossible." They claimed that as a result, the 180 day rule deprived children with severe

disabilities of educational programming. This required these students to spend a greater amount of time recouping lost skills compared to the time that they actually received services.

Despite the lack of empirical studies supporting or contradicting these claims, the expert witnesses agreed that in the general opinion of educators and others who worked with students with SED and SPI, these children suffered significant skill losses due to breaks. The plaintiffs' expert witnesses included professionals with experience working with children with severe disabilities as well as advanced degrees in areas such as special education and psychology. Considering case histories, Gary Armstrong lost most of his previously acquired progress in eye contact, feeding, dressing, obeying simple commands, motor coordination, and communication during a summer break. Natalie Bernard regressed in functional skill level and emotional development when coming home for weekends. A neurology professor with considerable experience assessing students with SPI predicted that were Natalie deprived of her year-round program, she likely would experience severe regression.

Overall, the court's findings concluded that breaks in educational programming served as a catalyst, if not a cause, for regression. Defendants claimed regression was not caused by educational breaks, and merely coincided with them. They asserted that unqualified teachers and parents' failure to give their children proper practice in the skills they had learned also were to blame for skill loss during breaks. However, the court considered parents' lack of expertise to maintain their child's skills level, as well as the insufficient time available due to other family responsibilities. Due to these considerations as well as evidence from individual children such as Gary Armstrong, who regressed when deprived of summer programming and did not regress when it was provided, parental factors were dismissed as the main cause of regression.

Among the serious ED population, increases in inappropriate behaviors and emotional problems were observed during interruptions in educational programming. The court declared that students with SED and those with SPI both had difficulties generalizing skills they had learned, making a change of environment especially detrimental for some children with these disabilities. Furthermore, through case study of the plaintiffs' histories, the court determined that some children with SED and SPI regressed significantly during breaks. This regression also was estimated to be greater than that of their "typical" peers. Next, the court needed to consider how much time such students needed to recover from the regression due to breaks in educational programming.

Similar to the present time, no empirical studies regarding recoupment of lost skills were available to the court of *Armstrong v. Kline*. Thus, the court relied on expert testimony, which found that the rate of relearning "for all types of children SED, SPI and non-handicapped" differed for each individual child. Case examples found that some children with SED and SPI returned to their pre-break skill level within two weeks, while others took over nine months to recoup their lost progress. Experts' opinions also established that if non-disabled children had a loss in skills, they would return to their previous level in a month or less. Thus, regarding recoupment, the court decided that although some students with SPI and SED showed similar patterns of regression and recoupment as their non-disabled peers, others showed significant regression with a lack of recoupment.

Next, the court needed to decide which children should have access to an educational program lasting more than 180 days. One distinguished expert, Terrence J. Piper, held a Ph.D. in Special Education and served as an Associate Professor and Chairperson of the Department of Special Education at Temple University. He testified that students with SPI

requiring two months or more to recoup lost skills needed year-round schooling. Another expert, Bertram A. Ruttenberg, M.D., Ruttenberg of the Elwyn Institute, declared that children with SED who had not "developed a stable relationship, who have not developed a real image of themselves, who have not developed impulse control" needed a twelve month program. Given these testimonies, the court finally decided that neither all students with SPI and SED nor only these categories of students needed more than 180 days of educational programming. Thus, it was ruled that because each child had unique needs, decisions regarding school year length would be made on individual bases by those who were familiar with the child.

Finally, the court found Pennsylvania's "180 day rule" in violation of Section 504 of the Rehabilitation Act of 1973 as well as EHA, which defined FAPE as:

". . . special education and related services which (A) have been provided at public expense, under public supervision and direction, and without charge, (B) meet the standards of the State educational agency, (C) include an appropriate preschool, elementary or secondary school education in the State involved, and (D) are provided in conformity with the individualized education program . . ." 20 U.S.C. § 1401(18).

In addition, the Act's definition of "special education" further clarified the terms of an "appropriate" education:

"The term 'special education' means specially designed instruction, at no cost to parents or guardians, to meet the unique needs of a handicapped child, including classroom instruction in physical education, home instruction, and instruction in hospitals and institutions." 20 U.S.C. § 1401(16).

The 180 day rule was found to prevent school districts and other agencies receiving public funds must adequately from meeting these "unique needs" of some students with disabilities. Thus, in its final decision, the court in *Armstrong v. Kline* deemed that the rule needed to fall. Although this decision is not binding upon other states, it has set a judicial precedent for other U.S. Circuit Courts. In addition, IDEA

mandates that a student's individual needs, rather than the services that an educational agency provides, drive that student's educational program.

There currently is a federal requirement by IDEA that IEP teams consider whether ESY services are appropriate for all individual students with disabilities. Other than data on individual students, little empirical data has been published to date regarding regression/recoupment in groups of students with disabilities. In particular, there is a dearth of such data pertaining to students with specific learning disabilities (SLD).

Other Relevant ESY Cases

Following the decision in *Armstrong v. Kline*, a later case made ESY eligibility decisions applicable to a wider variety of children with disabilities. The Supreme Court case *Board of Education v. Rowley* (1982) determined that individual decisions must consider multiple factors (as opposed a single criteria such as the child's disability category).

In addition three federal judicial cases found that limiting a child's access to special education services to 180 days when more service was needed violated EHA (*Crawford v. Pittman*, 1983; *Georgia Association of Retarded Citizens v. McDaniel*, 1983; *Yaris v. Special School District of St. Louis County*, 1984).

Another case determined that ESY was warranted in order to prevent skill regression during a break from school, and that transportation could be included as a related service (*Alamo Heights v. Texas State Board of Education*, 1986). Later, federal court case *Cordrey v. Euckert* (1990) also dealt with ESY as a preventive measure against regression. In that case, because the child's parents (plaintiffs) could not show proof of regression, the court's decision favored the school district.

Similarly, the decision in *Johnson v. Bixby Independent School District* (1990) determined that both "retrospective" and "predictive"

data, or regression and recoupment, should be considered when making decisions related to ESY (Olmi, Walker, & Ruthven, 1995).

Nature and Definition of Regression

At this time, no federal statutes define "regression" and "recoupment". Based on section 34 CFR §300.309 of IDEA, state education agencies may define these terms. In addition to the *Armstrong v. Kline* decision, additional case law has determined regression to mean the likelihood that a child will lose critical skills (*Reusch v. Fountain, 1994*). One example of a state definition of regression, coming from Iowa, includes the inability of a student to maintain an acquired skill during school breaks (Iowa State Department of Education, 2001).

Some regression of skills over extended school breaks is expected even for non-disabled students. Students likely vary in the severity of their skill loss, as well as the amount of time it takes to regain previous skill levels. To date, no studies examining the regression of reading fluency skills among the general population could be found.

Recoupment and Other Factors to Consider

According to court findings such as those determined by *Reusch v. Fountain* (1994), when skill regression occurs, the IEP team also should consider "recoupment," or whether the student is able to recover those skills within a reasonable time. The court also named five other factors that the IEP should consider when making ESY eligibility determinations. In addition to regression and recoupment, they included the following: the degree to which the child progressed toward his or her IEP goals, the nature and severity of the disability, emerging skills that might decline during a lengthy break in education, potential interfering behaviors, and any other special circumstances pertaining to the child. In addition, case law has determined that predictive data also should include predictions of future regression (*Johnson v. Bixby Independent School District, 1990*).

Beliefs About ESY

Edgar, Spence, and Kenowitz (1977) found that special education teachers reportedly believe that more severely disabled students demonstrate greater skill loss during school breaks. However, little evidence supported this claim in the literature. This may have been the case due to the general lack of research literature regarding ESY.

Few studies have examined the efficacy of ESY, and case law provides nearly as much information regarding eligibility criteria as IDEA. One reason for the lack of clarity regarding the provision of ESY services is that IEP teams make eligibility determinations on a case-by-case basis. One study found that responsibility for ESY often falls at the district level as opposed to the state level (Ahearn, 2000).

When surveying all 50 states' collection of ESY data, only 13 states collected such data. Furthermore, only nine of those states compiled and reported their collected data pertaining to ESY. Thus, few specific criteria aid IEP teams in determining the need for ESY. Individual school districts generally are perceived as responsible for setting their own criteria guiding ESY-related decisions for their students.

The SLD Construct

Beginning with the EHA of 1975, federal special education law has provided a definition of SLD that has experienced little change to this day (Lichtenstein & Klotz, 2007). This definition stated that SLD was:

A disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which disorder may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations, (20 U.S.C. §1401 [30]).

Due to the dearth of evidence regarding these psychological processes as well and the lack of guidelines specifying how school personnel should interpret assessment information, SLD rates increased

rapidly given this definition. Two years later, in 1977, federal regulations mandated an "ability-achievement discrepancy" as part of the SLD identification. This model's lack of validity as well as studies demonstrating that "IEP teams, more often than not, disregarded federal and state mandates for making SLD identification decisions," (Lichtenstein & Klotz, 2007) led to substantial variation among students with SLD. Studies examining decision-making procedures have found that teams often made eligibility decisions in the absence of formal criteria and in spite of the available data (Ysseldyke, Algozzine, Richey, & Graden, 1982).

The vagueness of the SLD definition as well as the application of wide variation among LEAs in test selection, criteria, and interpretation of data led to much differentiation in the amount of students identified. McLeskey and Waldron (1991) explored the issue of changing guidelines leading to differences in SLD identification rates and types of students identified when comparing reports before and after statewide implementation of specific SLD guidelines.

In a review of over 1,500 multidisciplinary team reports on students referred for assessment and labeled as SLD, they found a great deal of variation among the students identified as SLD. For example, 20% of the students identified as having SLD had low achievement, yet failed to meet SLD criteria. The authors hypothesized that professionals knew that without the SLD label, whether appropriate or not, these struggling students would not receive any supportive services. These professionals appeared to be driven by what Hewett & Forness (1974) called a "service motivation" hoping to provide assistance to needy students.

The desire to provide services regardless of the data led teams to dismiss the "scientific motivation" of making accurate decisions using systematic methods. McLeskey and Waldron posited that such motivation to identify (or misidentify) low-achieving students as having SLD arose when

alternative program options were unavailable to these students. In addition, they noticed another group of students who identified as having SLD had scored above or near grade level on measures of academic achievement. These students represented about 16% of the sample in their study.

Other studies looking at SLD identification such as Shepard and Smith (1983) found similar levels (about 20-25%) of students whose academic achievement levels fell close to their grade level. Together, authors postulated reasons such as desires to remove these students from their general classroom settings as well as parental pressure for placement for labeling these students as having SLD. Alternatively, the "fundamental problem" with the SLD construct was cited as another possible reason for the misapplication of its criteria when identifying students as having SLD. The view posed by the "medical model," claiming that the so-called disability lies within the child lies in opposition to the idea that SLD may be a socially (and environmentally) constructed phenomenon (Ho, 2004).

Several studies examining SLD have determined that school teams identify students as having this disability when data indicate that they do not meet federal or state-defined guidelines. For example, when examining test scores of students referred to school teams, MacMillan, Gresham, and Bocian (1998) found that less than half of those students eventually classified as SLD met the defined criteria. In their analysis of 150 student records, some of the students identified as SLD made up a heterogeneous group. Some of the students in the sample fell within the range to qualify for special education services as having mental retardation (MR), while others did not meet criteria to receive *any* special education services. Looking at the identification patterns in these schools, MacMillan et al. concluded that the school teams' SLD

identification practices rarely complied with federal and state regulations. Instead, the observed pattern consisted of unexpected underachievement, a generally accepted marker of SLD.

Another study examining SLD identification procedures also found students identified as SLD to be a heterogeneous group with many identified students failing to meet eligibility criteria. Gottlieb, Alter, Gottlieb, and Wishner (1994) conducted two sets of studies looking at IQ and academic achievement scores among urban students identified as having SLD. First of all, regarding heterogeneity in the SLD population, they found that children classified as SLD in urban school districts had IQ scores about 1.5 standard deviations lower than their suburban peers. This difference was about twenty IQ points, with 81.4 as the mean IQ for urban samples while the suburban mean was 102.8. This may be attributed in part to the likelihood that children in urban environments are more likely to experience poverty, live in less cognitively enriching households, and have more frequent absences. In terms of schooling, poor children are more likely to attend schools with higher levels of violence, a higher teacher turnover rate, less qualified teachers, and fewer resources than their more affluent counterparts (Whitehouse, 2006).

When examining urban students with SLD in a more in-depth manner, Gottlieb et al. found that some students with IQs in the MR or borderline MR ranges were classified as SLD. In 1984, Gottlieb et al. found that 41% of their sample of students identified as SLD had IQ scores between 70 and 85, while 7.5% had IQ scores below 70. In their 1992 follow-up study, they found that 16.6% of the children labeled as SLD had IQs less than 70. Considering their finding that 1 out of 6 students with SLD had IQ scores in the MR range led the researchers to question the SLD construct itself.

School teams' desire to grant students who might not otherwise qualify for special education services likely drives this practice.

Classifying students with low IQ scores as SLD rather than MR allows school teams to avoid using this "pessimistic" term and obviates the adaptive skills requirement of MR classification. In addition, labeling students with IQs well below normal limits as SLD allows students access to educational support services that they might not otherwise receive.

Many researchers, including Gottlieb et al. (1994) concluded that both the problem and solution to such issues of SLD identification rests within general education. Either way, research has confirmed that students with SLD represent a widely varied group of students. The main characteristic these students share is poor academic achievement regardless of data indicating the presence SLD according to federal and state standards and guidelines. This makes the question of differentiating low achievers from students with SLD (and students with "true" SLD according to guidelines versus those identified as such in schools) increasingly complicated. Due to differential practices among LEAs regarding identification of students with low ability and low achievement as opposed to those meeting SLD makes looking at comparisons within and between these groups difficult.

Current SLD Regulations

The current version of IDEA, implemented in 2004, governs the identification of students with SLD. The U.S. Department of Education (2006) published a commentary to the 2004 regulations in 2006. Low achievement still is the first requirement for students to receive special education services if found to have this disability. The standards define low achievement as follows:

The child does not achieve adequately for the child's age or to meet State-approved grade-level standards ... when provided with learning experiences and instruction appropriate for the child's age or State-approved grade level standards (§300.309(a)(1)).

According to Lichtenstein and Klotz (2007), determining whether a child has low achievement involves measuring the child's performance based on criterion-based curricular expectations as opposed to "comparing the child to others in the class or in the school district." The legislation also indicates that using a local norm as a guide for low achievement is inappropriate if most of the child's peers fail to meet standards. In addition, current grade-level standards represent an inappropriate gauge of achievement if the referred student has been retained in their current grade (Lichtenstein & Klotz, 2007). Complying with these standards poses a potential challenge for school teams making SLD decisions in schools where many students fail to meet academic standards. For example, inner-city schools with few resources such as those described by Gottlieb, Alter, Gottlieb, and Wishner (1994) often have many students qualifying as having "low achievement" according to IDEA's definition. Thus, IDEA mandates that a comprehensive evaluation also is necessary to rule out causes for low achievement other than a disability intrinsic to the child.

Once a student shows underachievement, the next criteria that must be met in order to determine a child's eligibility for SLD must include one of the two following indicators of SLD. The first requires that the child "exhibits a relevant pattern of strengths and weaknesses in performance and/or achievement, relative to age, grade, or intellectual development," (§300.309). This data typically is collected via an individual psychological and educational assessment.

Alternatively, IDEA (2004) regulations indicate that schools may use a response to intervention (RTI) within the context of a multi-tiered service delivery system. The RTI approach assumes that the student is receiving appropriate instruction and interventions within general education. This method involves the monitoring of a child's change in academic performance over time given evidence-based intervention

(Lichtenstein, 2008). In order to meet SLD criteria according to this model, the regulations state that:

"[The] child does not make sufficient progress to meet age or grade standards in response to scientific, research-based intervention." (§300.309).

In addition, the regulations indicate that when using the RTI approach, schools must include "documentation of instructional strategies used and the student-centered data collected," (§300.309). Once one of these two methods is used to determine either a child's pattern of strengths and weaknesses or RTI given quality instruction and interventions, several other factors must be examined when making a possible SLD determination.

As a third step in the SLD determination process, the findings of the first two areas (underachievement and either lack of progress using RTI or a relevant pattern of strengths and weaknesses), several other factors must be ruled out as primary causes of the learning difficulty. These areas include visual, motor, or hearing disabilities, mental retardation, emotional disturbance, cultural factors, environmental or economic disadvantage, and limited English proficiency. In addition, the child's underachievement must not be due to a lack of appropriate instruction in reading or math (§300.309). The mandated comprehensive evaluation to determine the whether a child's underachievement results from one of these factors. Thus, if appropriate, other areas to assess when evaluating whether a student may qualify for services as a child with SLD may include health, vision, social and emotional status, general intelligence, academic performance, communicative status, and motor abilities (IDEA §300.304(c)(4)).

Once a comprehensive evaluation is completed, a group or team decides whether the student meets SLD criteria. The team determining eligibility for special education services must include the child's parent(s), teacher, and at least one individual qualified to conduct

individual diagnostic examination of children, such as a school psychologist or remedial reading teacher (§300.308). OSEP has allowed LEAs to decide which specific support services personnel must take part in the eligibility process. According to Lichtenstein and Klotz (2007), the group's composition may vary based on the child's suspected disability, expertise of local staff, and other factors. Thus, SLD identification currently requires low achievement as well as the collection of comprehensive data ruling out extrinsic reasons for such low achievement. However, the final decision about how to interpret that data rests with school team members. As indicated earlier, these data often are overlooked by teams when making SLD decisions.

Family Income and Summer Reading Loss

School-aged children spend the majority of their time outside of school. In fact, while enrolled in school, these children spend less than one-third of their waking hours in school. Thus, activities outside of school play an important role in school aged children's academic skill development. Community and family resources play a significant part in determining the amount of time children spend in academically-enriching activities outside of school. In particular, the amount of time a student spends reading has been well documented as the best predictor of reading achievement (Allington, 2006). More specifically, the reading of books predicted several measures of children's reading proficiency, including cumulative gains in reading progress between 2nd and 5th grade (Anderson, Wilson, & Fielding, 1988).

Across studies of American elementary school children, Anderson et al. found that on average, students read for only 8-12 minutes per day, and only 4-5 of those minutes were spent reading books. Why did students read so little? The authors found that a lack of access to books, as well as spending time engaged in other activities such as watching television

played a part. In a related study, Kim (2006) found that giving low-income children access to books over the summer improved their reading skills. Exposure to enriching out-of-school activities such as summer camps differentiated between children of different income levels, further (though sometimes indirectly) impacting their academic skills (Miller, 2007).

A number of other researchers also have determined that differences in summer activities and opportunities were important. These differences contributed significantly to academic achievement differences between affluent children and their more disadvantaged peers. A study by Heyns (1978) demonstrated that students' reading and mathematics skills improved during the school year regardless of race and socio-economic status (SES). However, while middle-class children maintained or improved their skills during the summer, those from poor families experienced a regression in their academic skills. In 1981, Ginsburg conducted a replication of Heyns' study, finding that although SES impacted achievement, it did not differentiate between academic skill learning during the school year or summer months.

Twenty years later, Alexander, Entwisle, and Olson (2001) conducted a similar longitudinal study looking at reading losses among children from middle and low-income families. They used a representative sample of urban students (n = 368), two-thirds of whom were considered "low income" based on free or reduced meal status. Using repeated measures of the California Achievement Test (CAT) Reading Comprehension subtest in the fall and spring, they found that family income had a profound impact on summer regression. In fact, Alexander et al. determined that summer setbacks accounted for "three grade levels" of difference between income groups by sixth grade.

These findings were consistent with Heyns' (1981) results. Similarly, a meta-analysis of 39 summer regression studies found that while middle-class children's reading skills remained the same or improved, lower-income students' skills declined (Cooper, Nye, & Charlton, 1996). By the end of elementary school, the effect of accumulated summer reading loss by low-income students widened. In fact, researchers estimated that summer regression accounted for 66 to 80% of the "achievement gap" between income groups (Hayes & Grether, 1983; Miller, 2007).

Despite what is known about the environmental variables serving as detriments to some low-income students' academic achievement (Whitehouse, 2006) combined with IDEA's exclusionary factor of "environmental or economic disadvantage" §300.309(a)(3) when identifying SLD, low-income students still receive this label. As mentioned earlier, low income students in urban areas often receive SLD services despite lower cognitive scores than the "typical" achievement compared to their higher income peers with SLD (Gottlieb, Atler, Gottlieb, & Wishner, 1994). The practice of overlooking or ignoring poverty as an "environmental or economic disadvantage" in order to give low-achieving urban students access to services makes the SLD group increasingly diverse in terms of ability and achievement level.

Impact of Summer Reading Intervention Programs

Several research studies have indicated that summer school programs have helped at-risk students to make academic gains. Lauer, Akiba, Wilkerson, Apthorp, Snow, and Martin-Glenn (2006) conducted a meta-analysis of several such programs taking place during "out-of-school-time" (OST). Basing the study upon the No Child Left Behind (NCLB, 2001) mandate that "children in schools that fail to help all children reach proficiency are eligible to receive supplemental educational services," they assessed

programs designed to remediate students' academic skill deficits. Some of the programs took place during after school hours, while others occurred during the summer.

When analyzing the effects of 35 studies of OST programs using control or comparison groups, Lauer et al. found that as a whole, these programs effected small yet significant improvements in academic skills. The individual attention provided to students via tutoring was found to be especially beneficial in enhancing those students' reading skills. The timing of interventions (during the summer versus after school) had no impact on the relative effect sizes of improvement in the examined studies.

While Lauer et al. examined the types of academic gains made by OST programs taking place at different times, other studies have taken this analysis further. For example, when comparing purely academic programs to "hybrid" programs combining academics and youth development, findings across such studies have demonstrated that the latter had higher efficacy in improving academic skills (Chaplin & Capizzano, 2006; Miller, 2007). Characteristics of these more successful programs included engaging experiential activities that allowed students to form close relationships with staff. Unlike traditional summer schools, the hybrid programs were less punitive in nature (Miller, 2007). As a result, participating students gained an appreciation for learning embedded in the context of improving specific skills.

Summer reading interventions programs need not always include specific instruction to yield improvements in reading skills. For example, Kim (2006) found that providing fourth-grade students with books for independent summer reading led to significant skill gains for at-risk populations. Families received books over summer vacation, accompanied with instructions encouraging students to use comprehension strategies and

practice silent and oral reading. An analysis of repeated measures of the Iowa Tests of Basic skills (ITBS) demonstrated positive effect sizes for the at-risk populations including minority students, children with few books in the home, and dysfluent readers.

Sex Differences

Current statistics indicate that boys are about twice as likely as girls to receive special education services, with particular over-representation in the SLD category (Lichtenstein, 2008). Boys also are more likely than girls to experience slower brain maturation, resulting in poorer school readiness skills such as language and vocabulary development, self-control, and fine motor skills (Raffaele Mendez, Mihalis, & Hardesty, 2006).

During the early elementary school years, boys' higher activity levels lead them to have more behavior problems than their female counterparts. Over the past ten years, girls have consistently outperformed boys on reading and writing measures of the National Assessment of Educational Progress (NAEP), which is given in grades 4, 8, and 12 (Freeman, 2005). These behaviors and poorer academic performance co-occur with boys' greater likelihood to repeat one or more grades. In 1999, retention statistics demonstrated that while "8% of boys ages 5 to 12 had repeated at least one grade compared to 5% of girls the same age (Raffaele Mendez et al., 2006).

Factors such as the aforementioned biological differences have been recognized in the literature as contributing to the higher prevalence of SLD in boys as opposed to girls (Tschantz & Markowitz, 2003). Boys' noted behavior problems and higher activity levels compared to their female peers may also lead to their over-identification as SLD. For example, high activity levels could make attending to instruction and doing independent seatwork more difficult for male students.

Boys' propensity toward disruptive behavior also could make them more likely to be misidentified as having SLD despite having academic skills in the "average" range academic skills. For example, when looking at over 1500 evaluation reports of students identified as SLD, McLeskey & Waldron (1991) found that about 16% of their sample had academic skills near or above grade level.

Similarly, Shepard and Smith (1983) found that close to 1 out of 4 students in their representative sample of 1,000 students with SLD had close to average levels of academic achievement. Speculations about reasons behind these students being identified as SLD included "a general desire to remove students from general classrooms who were troublesome, difficult-to-teach, or manifested minor behavior problems" (McLeskey & Waldron, 1991). Based on the literature's conclusions about boys, it appears that more male students *without* academic deficiencies would be likely to be identified as SLD compared to their female classmates.

Studies looking at referral and bias for male versus female students have yielded mixed conclusions regarding the higher prevalence of SLD among boys. In particular, some studies have considered ascertainment bias, distorting the true frequency of SLD due to data collection methods, as a reason for higher SLD prevalence rates among males. For example, Shaywitz, Shaywitz, Fletcher, and Escobar (1990) examined the extent to which referral bias contributed to the finding that about twice as many boys receive SLD services compared to girls. In their study, they found that using objective measures such as standardized tests to determine which students would be referred for student support led to comparable levels of boys and girls receiving referrals.

On the other hand, Shaywitz et al. also determined that students referred by subjective methods such as teacher referral led to more boys than girls receiving referrals for student support. Boys' higher activity

level and greater likelihood to have behavior problems have been used to explain that reasons that more boys are referred using such subjective methods. However, this claim was invalidated by Mirkin (1982), who found negligible differences when comparing students referred subjectively via teacher referral and those chosen due to performance on objective measures such as academic screenings. Academic failure was found to be the main reason for referral using both methods.

Another study examining the ratio between males and females with SLD in reading found a ratio of 2:1 when analyzing a sample of over 32,000 participants (Flannery, Liederman, Daly & Schultz, 2000). These results suggested that male bias for SLD may have been due to more than ascertainment bias in research. In order to further assess this claim, Liederman, Kantrowitz, and Flannery (2005) also studied the impact of bias as it pertained to males and females relative identification as having SLD. When analyzing sex differences in SLD across studies using statistical methods to minimize ascertainment bias, they still found a "significant preponderance" of boys suggest had SLD in reading. This finding suggested that the greater frequency SLD among males was "not a myth but a reliable phenomenon" (Liederman, Kantrowitz, & Flannery 2005).

Despite well documented studies regarding the greater likelihood and possible vulnerability of males in terms of SLD, no data to date indicates whether the sexes differ in terms of reading progress over summer. In a meta-analytic review of 39 studies of academic skill decline over summer break, no significant differences were found between girls' and boys' academic decline. Instead, it was found that skill regression over summer was the rule for most students regardless of sex (Cooper, Nye, & Charlton, 1996). However, the assessments used to measure academic progress in the studies included in this meta-analytic review often were not designed for that purpose. Instead, most of the included studies used repeated measures

of standardized, norm-referenced tests. Results were measured using grade equivalent scores, which often are misinterpreted, and sometimes misunderstood by professionals as well as parents. Studies using Curriculum-Based Measurement (CBM), which was designed to measure incremental progress over time, have yet to compare boys' and girls' reading growth over summer break.

Curriculum-Based Measurement

When measuring health, economic growth, behavior, and other indices of status, progress, and change over time, certain indicators are considered to be especially important. These indicators, which are "simple, accurate, and reasonably inexpensive in terms of time and materials", are "collected on an ongoing basis over time", and "shape a variety of decisions" (Shinn, 2002) all fall under the category of General Outcome Measures (GOMs). In education, a set of test procedures referred to as CBM are GOMs that indicate success in basic academic skills. Like other GOMs, CBM has come to be considered so important that they currently are routine in many educational settings.

More specifically, Curriculum-Based Measurement in Reading (R-CBM) has a strong evidence base supporting its use as a valid measure of basic reading skills for elementary students in general and special education environments (Deno, 1985; Deno & Fuchs, 1987; Shinn, 2002). Data provided using CBM has proven useful in monitoring progress, developing local norms, setting goals, evaluating interventions, and as a universal screening tool for assessing basic skills.

Using R-CBM has several advantages, many of which are relevant for measuring academic progress and summer regression/recoupment for students with SLD. First, using CBM procedures allows for a quick assessment of students' current level of functioning within the curriculum. Second, this procedure helps to determine the appropriateness of the instructional

program or curriculum for individual students or classes. Third, this data may aid determining whether that student requires a change in educational placement. Finally, assessment information derived from CBM provides meaningful information to parents, school personnel, and students themselves.

Several school districts and other LEAs nationwide regularly collect R-CBM data to monitor academic progress among general and special education students. R-CBM data also are used as a screening measure to determine which students may need additional support in order to meet recognized standards or reading progress. These data also measure how well students in general and special education respond to changes in their instructional program.

R-CBM's quick, efficient nature, empirical validation, and sensitivity to change over time make it an ideal tool for measuring how much skill loss students experience during the summer (or other breaks in educational programming). While such simple, direct CBM measures such as oral reading fluency have been validated for elementary students, these procedures have less validity for secondary students (Espin & Tindal, 1998). However, using R-CBM at the secondary level has received support for students with low levels of reading achievement compared to their peers. On average, high school students with disabilities read at a fourth-grade level (Deshler, Shumaker, Alley, Warner, and Clark, 1982) and made few gains in reading beyond the fourth or fifth grade level (Warner, Schumaker, Alley, & Deschler, 1980). Additionally, many of these students had trouble with simple word identification, hindering their reading fluency and comprehension (Lenz & Hughes, 1990).

Evidence also has supported a relationship between oral reading fluency and scores on a standardized test of reading achievement for secondary students with below average reading skills. Thus, R-CBM can be

done in a quick, efficient manner, has a strong sensitivity to change, and has documented validity for measuring progress in overall reading development for struggling readers in both elementary and secondary school.

These characteristics make R-CBM a useful tool for measuring regression and recoupment of reading skills for special education students such as those with SLD in reading. In addition, due to its validity in measuring reading growth among large groups of elementary students, using R-CBM as a screening measure to assess changes in reading skills over the summer has the potential to provide a normative comparison group for students with SLD to general education students.

Developmental Status

Children often make improvements in a variety of areas as they get older. For example, most typically developing children show improvements in cognitive abilities, language development, attention and behavior, communication skills, adaptive skills, and social skills as they mature. Given this development as well as exposure to school, children in kindergarten and beyond also are expected to make gains in reading skills. Research literature suggests differences in reading development when comparing age (or grade) levels. In addition, reports from AIMSweb, a large database with student R-CBM scores from multiple years of data collection shows improvement in reading fluency scores as students mature.

Do certain grade levels show different rates of improvement in reading skills? When considering normative language development of younger elementary children, one study compared reading skill growth rates of kindergarteners to those of first graders. That study's sample was drawn from Early Childhood Longitudinal Study-Kindergarten cohort (ECLS-K) study, a federally sponsored database that documented kindergarten students' academic growth over two school years.

The ECLS-K obtained a nationally representative sample by using a multistage probability sample design (McCoach, O'Connell, Reis, & Levitt, 2006). The ECLS-K reading assessment measured the children's early literacy skills, word reading, vocabulary, and reading comprehension skills using selected items from the Peabody Individual Achievement Test-Revised (PIAT-R). Results indicated that the first grade students made considerably more reading progress than did kindergarten students.

Why did first graders show more growth than their slightly younger counterparts? The authors hypothesized that this difference arose due to kindergarten students' entering school with different pre-reading skills such as letter naming and phonological awareness (McCoach, O'Connell, Reis, & Levitt, 2006). In other words, kindergarten students entered school with a variety of different backgrounds and skill levels in terms of pre-reading skills. On the other hand, most first graders had exposure to early academic skills through school. This difference between grade levels provided a possible justification for the greater gains observed among first grade students compared to their kindergarten peers.

No research could be found comparing students' reading achievement by grade level beyond first grade. However, the online database AIMSweb has data for oral reading fluency scores of students in grades 1-11. This database includes R-CBM scores for each grade level three times per school year: fall, winter, and spring. In fact, when describing the development and technical properties of the R-CBM passages, Howe & Shinn (2002) commented that, "Developmental trends, as expected, confirmed an increase in mean WRC across grades."

The reading fluency rate of improvement (ROI) provides another gauge of grade level oral reading fluency comparisons. This metric is derived by comparing benchmark scores. The publisher of the AIMSweb passages defined

ROI as the "Spring Score minus Fall Score (or Winter minus Fall) divided by 36 weeks (or 18 weeks)" (NCS Pearson, Inc., 2008).

When comparing R-CBM scores across grade levels in the AIMSweb sample ($n > 60,000$ for each grade level), ROIs dropped slightly each school year from first through eleventh grade. For example, among students at the 50th percentile for each grade, the ROI dropped by about 10-20% (0.1 - 0.2) each school year (NCS Pearson, Inc., 2008). These scores were based on students reading passages aloud at their current grade level. Grade level probes increased in difficulty as students advanced in age. For example, 5th grade reading passages were slightly more difficult than 4th grade probes. It is important to note that although the rate of improvement decreased slightly between grade levels, students still made improvements.

Because AIMSweb computes scores for each grade level three times per year, this allows for a comparison between students' ORF performance at the end of one school to their ORF at beginning of the next school year. When making this comparison, each grade level showed similar levels of regression. More specifically, when comparing students' spring ORF to their fall ORF the following the school year, similar gaps in WRC were observed. Drops in ORF were noted from spring to fall between grade levels.

It is important to note that unlike the current study, students included in the AIMSweb sample were assessed at their current grade level. Thus, at the end of 3rd grade, students read 3rd grade probes. Then, at the beginning of 4th grade, students read the more difficult 4th grade probes. The spring and fall assessments therefore compared reading fluency using different test probes, limiting the validity of this comparison.

One research study comparing reading fluency between students of different age groups could be found. Tressoldi, Lorusso, & Brenbati (2008) assessed whether older children identified as having "dyslexia" or SLD

made comparable gains to their younger peers when both had access to remediation. The study compared reading fluency progress of students in third and fourth grade to gains made by students in sixth through eighth grade. Two treatments were used. The first was based on Bakker's Balance model, which focused on activating different brain areas. The other focused on sublexical syllable recognition, focusing on word-attack skills. The only statistically significant difference between age groups included younger students' having greater improvements in accuracy given the sublexical treatment. However, overall results indicated that the older students made comparable gains to their younger peers in terms of fluency and accuracy. The authors contended that it was "never too late" to remediate deficits in these areas.

Given what research studies and AIMSweb database information indicate, students' normative reading fluency improves as students advance in age and grade in school. Also, some minimal differences exist between older and younger students, with the latter group showing a slightly greater likelihood of improving reading accuracy given a specific intervention. The current study examined students in 2nd through 5th grade. This age range was chosen given the validity of R-CBM as a measure of overall reading competency for this group. Thus, it is hypothesized that no significant differences will be found between grade levels in terms of regression and recoupment. Choosing a sample with a larger age range may have more likelihood to yield differences between groups.

Readability

Just as students' reading competency typically improves with age, CBM passage difficulty also increases with grade level. Readability formulas (RFs) provide one means with which to match a passage's difficulty to a reader's grade level.

According to Dale and Chall (1949), a reader's success with a text includes their ability to read it fluently, comprehend it, and find it interesting. Educators typically use data from RFs to match a text to the reader's reading or language abilities. Two measures by which to measure readability include vocabulary and syntax (Oakland & Lane, 2004). Such factors have been applied when analyzing the readability of R-CBM probes.

Several studies have examined and critiqued these formulas and their validity. Some criticisms include their low reliability, and over-reliance on "surface" features such as vocabulary and syntax. In focusing on these surface features, RFs neglect "structure level" features such as inference load and story structure (Oakland & Lane, 2004).

Another study found that RFs inaccurately measured increases in difficulty among passages. In that study, the researchers compared various readability estimates to students' actual reading performance. They found low correlations between these formulas. Additionally, the RFs poorly estimated the students' performance (Fuchs, Fuchs, & Deno 1983). As a result, the authors concluded that these formulas had limitations in terms of predicting passage difficulty. Fuchs et al. also noted that RFs ignored characteristics such as a student's background knowledge and familiarity with a text.

Despite the use of RFs to estimate difficulty and grade level appropriateness, these methods have limitations when applied to R-CBM. For example, Ardoin, Suldo, Witt, Aldrich, & McDonald (2005) compared the relationship between five RFs estimates and students' actual performance on R-CBM measures. They found a lack of relationship between 4 out of the 5 RFs typically used in the R-CBM literature. In addition, several of the RFs used overestimated passage difficulty. Ardoin et al. also found that only two of the variables commonly assessed by RFs had significant relationships with ORF. These components included syllables per 100 words

and number of words not included in the Dale-Chall word list. However, RFs often look at variables such as sentence length, which did not relate to ORF. Overall, the authors concluded that RFs lacked support when attempting to estimate passage difficulty.

Due to concerns regarding readability, Ardoin & Christ (2008) proposed a solution for ensuring similarity between probes. They recommended using the same three probes for universal screenings and benchmarking. Both of these assessments took place about three months apart. Spacing sessions at intervals of one month or more was recommended to avoid practice effects.

Several studies have found limitations of RFs in reliably and accurately predicting text difficulty. Thus, for the purpose of this study, their use served as a means by which to *estimate* of the grade-level appropriateness for the passages used in this study.

Summary

At the present time, federal law requires that IEP teams consider whether students with disabilities require ESY services. Anticipated regression and recoupment serve as major factors considered when making this decision. However, no federal statutes currently defines either regression or recoupment, leaving state education agencies the task of determining any skill losses students may experience.

No studies to date examine how much students with and without disabilities regress and recoup skills over the summer, specifically in reading skills. Instead, ESY decisions are made on a case-by-case basis. This lack of data makes the justification of ESY decisions for students with disabilities such as SLD difficult. The nebulous nature of the SLD construct and inconsistencies in SLD identification further complicate these decisions. For example, the literature provides no comparison of

students with SLDs to either their "typical" or low-achieving peers in general education.

Several other factors likely lead to differences in regression and recoupment of different groups. For example, family income level, out-of-school programs, availability of reading materials, and summer reading practice also have been found to influence reading achievement. In addition, boys have been found to have a greater risk of having SLDs and low achievement compared to their female peers. Finally, developmental differences in ROI have shown that students' academic maturation may vary on different measures as students mature.

Although some studies have assessed change in reading skills over summer break, few have used CBM, which has been shown to measure incremental reading progress over time. CBM offers benefits such as speed and efficiency of administration and scoring paired with validity and cost-effectiveness. Current data indicates that students increase in ORF as they age; however readability formulas lack consistency in measuring relative difficulty of passages. This makes using identical CBM passages for repeated measures ideal for comparing students' change in reading skill over the summer. Thus, the current study aims to examine regression and recoupment among students with and without learning disabilities. Factors such as sex, grade level, summer program attendance, and initial low achievement also will be examined to determine any differences that may occur within or between groups.

CHAPTER III

METHODS

Introduction

The current study examined summer regression of oral reading fluency (ORF) among 136 elementary school students. At the time of the pre-test (May 2008), the students were in grades 2, 3, and 4. During the summer, all of the students advanced one grade level. At the time of the post-test in September 2008, the sample of students had moved on to grades 3, 4, and 5.

Curriculum-based measurement of ORF(R-CBM) is an empirically validated, standardized procedure measuring reading achievement through ORF (Howe & Shinn, 2002). A median score of number of words read correctly per minute (WPM) was calculated for every student at each data collection point. Student reading achievement was measured at the end of the 2007-2008 school year using R-CBM probes from the AIMSweb system.

The students' ORF was measured again at the beginning of the following school year (September 11, 2008). Finally, a follow-up assessment was done for those students who regressed. Students who lost 10% or more of their May score during the post-test were reassessed four weeks after the September assessment. This follow-up assessed whether students whose skills regressed were able to "recoup" fluency skills they had lost. The follow-up measurement took place on October 7, 2008.

During the pre-test, students were assessed using probes from the grade they entered in the fall. For example, during the pre-test assessment in May, 2nd grade students received 3rd grade probes, and 3rd graders were assessed using 4th grade probes. The same three benchmark

probes were used for the post-test and follow-up measures in the fall, at which time they represented the students' previous grade level.

Before the pre-test, the sample was stratified according to three different pre-existing predictor variables: grade, sex, and special education status. School records provided information regarding students' status for each of these factors.

In addition to the three factors above, two more predictor variables were created after the pre-test. Students were categorized as having "low achievement" based on their pre-test scores. Finally, students were divided into groups of those who had or had not attended a school-sponsored summer reading program.

Design

This study used a quasi-experimental design with non-equivalent groups using a pre-test and post-test measure for all participants. The follow-up measure was done for a selected subset of students who had regressed from pre-test to post-test.

The participants were grouped by grade, sex, achievement level (at or below benchmark), summer school attendance, and SLD status. All students' ORF skills were assessed in May and September 2008. Students showing a drop in reading fluency (words per minute) of 10% or greater also were re-assessed in the follow-up measurement in October 2008. Figures 2 and 3 depict the research design and path diagrams, including variables, relationships, measurements, etc.

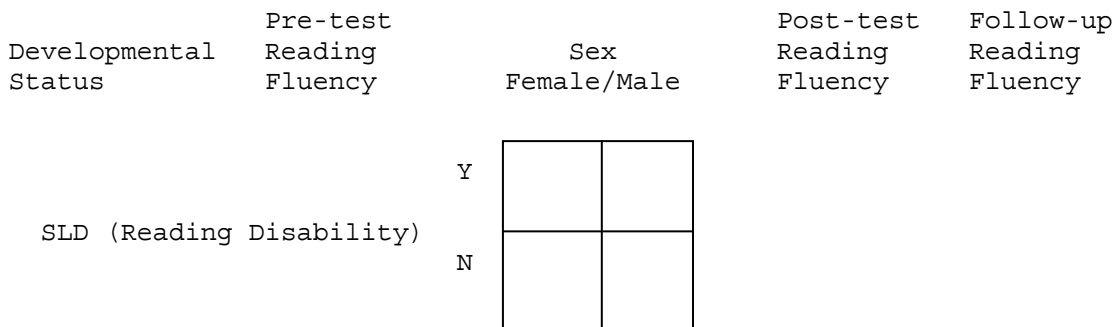


Figure 2. Research design diagram for regression-recoupment project.

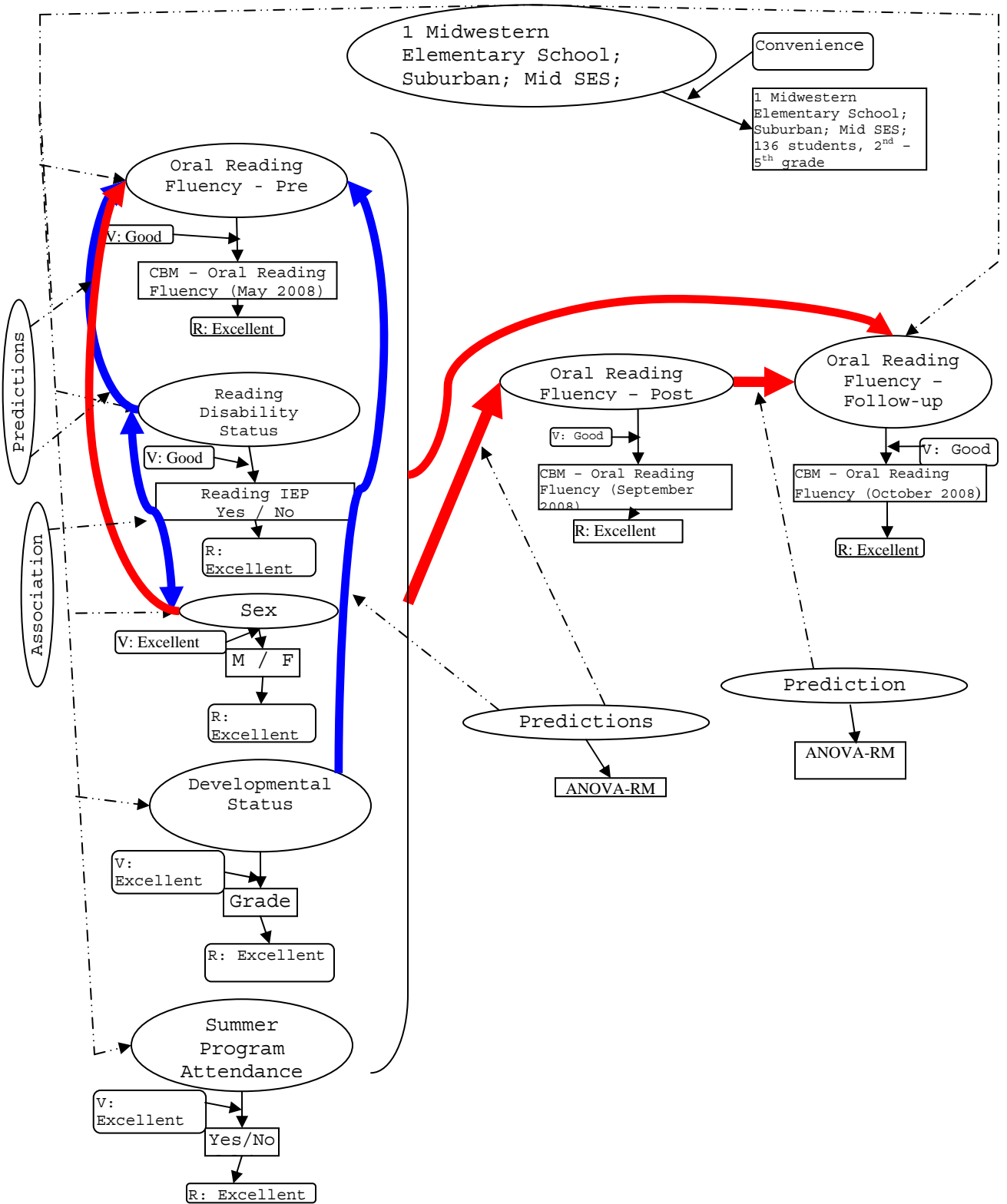


Figure 3. Research path diagram of the regression recoupment project.

Population

The population of interest included elementary school students in grades 2-4 (at the time of the pretest). This age group was chosen given the validity of the ORF construct as an overall measure of reading competency. In addition, the study aimed to assess reading skills of students within general education as well as those with SLD in order to compare the performance of the two groups in terms of reading regression and recoupment.

Sample

The sample consisted of 136 students in 3rd, 4th, and 5th grade (as of the September 2008 post-test date). 72 males and 66 females were assessed. Four students in the sample were identified as having SLD, and 16 students attended a summer program through the school during the summer of 2008. Students ages ranged from 8-9 years old for entering 3rd graders, 9-10 years for 4th graders, and 10-11 for 5th graders. All of the present and eligible students enrolled in the third through fifth grades at the study site participated.

All participants were enrolled in an elementary school located about five miles outside of a large city in the Midwest. As of July 2007, the population of the town where the study took place was 8,079. The town's estimated median household income as of 2006 was \$47,800 (Illinois Profile, 2008).

According to the Illinois Interactive Report Card (IIRC, 2008), this school had an enrollment of 620 students in 2008. The student attendance rate was 96.5%, compared to the average of 93.3% for the state of Illinois. The IIRC also described the school's population as 77.9% white, 14.5% Hispanic, 7.4% Asian, and 0.2% Black.

As of 2008, the school's teachers had an average of 13.9 years of teaching experience, with 60% of them holding masters' degrees. The remaining 40% of teachers held bachelor's degrees.

Measurement

Predictor Variables

Special Education Status

This factor compared students with and without specific learning disabilities (SLD) in reading. Students were placed in one of two groups: SLD and no SLD. For the purpose of this study, students identified as having SLDs in reading were defined and identified based on having an Individual Education Plan (IEP) for SLD that included goals addressing reading skills.

Students' SLD status was obtained from school records in order to differentiate students with and without disabilities. The school district identified students with SLD based on comprehensive assessments using R-CBM progress monitoring data as well as data from individual norm-referenced academic achievement measures.

Students with IEPs for SLDs in reading were identified and coded into the data. Having such a small sub-sample limited the statistical power of the analysis. The fact that such a small percentage of the sample had SLD also limited generalizability to other samples. The school provided information regarding SLD status based on school records. This source of information was considered to have good validity and very good reliability.

Table 1 indicates further reliability and validity information for all of the current study's six main research questions. It also includes information regarding latent and observed variables, instruments used, and others source of information if applicable.

Table 1.
*Research Questions, Latent Variables, Observed Variables,
 Instrument/Source, Validity and Reliability*

Research Questions	Latent Variable	Observed Variables	Instrument /Source	Validity	Reliability
1. How much ORF regression and recoupment do students with SLD show compared to their peers in general education during summer break?	SLD	Reading IEP (SLD)	School Records	Good	Very Good
	Reading Skill	ORF score	AIMSweb R-CBM	Very Good	Excellent
2. How do male and female students differ in terms of ORF regression and recoupment following summer break?	Sex	Male/Female	School Records	Excellent	Excellent
	Reading Skill	ORF score	AIMSweb R-CBM	Very Good	Excellent
3. What relationship will exist between pre-, post- and follow-up R-CBM measures?	Reading Skill	Pre, Post, & Follow-up ORF scores	AIMSweb R-CBM	Very Good	Excellent
4. How does summer reading program attendance impact ORF regression and recoupment?	Summer Reading Practice	Summer Program Attendance	School Records	Excellent	Excellent
	Reading Skill	ORF score	AIMSweb R-CBM	Very Good	Excellent
5. Is there any association between age/grade and ORF regression and recoupment?	Developmental Status	Grade	School Records	Excellent	Excellent
	Reading Skill	ORF score	AIMSweb R-CBM	Very Good	Excellent
6. Will one month back in school improve ORF for students who regressed from spring to fall?	Reading Skill	Pre-post ORF difference	AIMSweb R-CBM	Very Good	Very Good
		ORF score-Follow-up			

Sex

The assessment compared boys' and girls' ORF scores. Information regarding each student's sex was obtained from school records. These records were considered to have excellent reliability and validity in accurately reporting each student's sex. Table 1 includes further information.

Grade

Students were divided by grade level, and school records and class lists provided information about individual students. All 136 of the students in the study advanced to the next grade level at the end of the 2007-2008 school year. The school records were considered to have excellent reliability and validity in reporting students' grade levels. Table 1 includes further information regarding this variable.

Achievement Level

Students' scores were stratified according to whether their pre-test scores met or exceeded the benchmark level provided by AIMSweb. At the time of the study, AIMSweb provided an aggregate norm table indicating scores in correct words per minute for each grade level. Spring 2008 benchmarks are listed in Table 2. Students whose pre-test scores fell below the benchmark score for their grade level were classified as having "low achievement." All students were grouped according to whether or not they met this criterion.

Table 2.
AIMSweb Benchmark Scores as of May 2008 (AIMSweb, 2008)

Grade Level	AIMSweb Spring 2008 Benchmark (50 th percentile)
2 nd	100
3 rd	118
4 th	131

Using benchmark scores determined by AIMSweb was considered to have very good validity and excellent reliability in differentiating students with low versus "benchmark" ORF skills. Table 1 includes further information regarding reliability and validity related to achievement levels.

Summer Program Attendance

Students were grouped into those who had or had not attended a school-sponsored reading program during the summer. This program was separate from Extended School Year (ESY) for special education, for which none of the participants qualified. Students were selected as eligible for the program based on below-average achievement; however, participation was voluntary. The program ran from July 21 to August 15, 2008. The sessions took place from 8:30-11:30 am Monday through Friday, and taught by a reading specialist.

The school provided information regarding which students attended a summer reading enrichment program through the school. The summer program took place for 20 sessions at three hours per session over a four week period. Students attended the program six weeks prior to the beginning of the school year, from Monday through Friday from July 21, 2008 to August 15, 2008.

Developmental Status

Students' reading scores were compared based on grade level as a measure of developmental status. Students' grade levels were provided by the school district. The sample assessed reading skills of students over a small age range. Based on what research studies and data from the AIMSweb R-CBM database demonstrate regarding ORF development among students in 2nd through 4th grade, it was hypothesized that no significant differences in regression or recoupment would arise between grade levels. However, increases in fluency were expected as students advanced in grade levels. For example, even given slightly more difficult passages, large data sets demonstrate growth in ORF as students mature (NCS Pearson, Inc., 2008).

Dependent Variables

Pre-test Oral Reading Fluency

Students' reading fluency scores were reported in median WRC per minute among the three standardized AIMSweb benchmark probes at each administration.

Post-test Oral Reading Fluency (Pre-test to Post-test Difference)

Post-test data were collected and recorded using the same probes and procedures the pre-test data. Post-test scores held the most meaning in comparison to pre-test scores. This difference determined whether students had improved their ORF scores over the summer break, maintained the same scores, or regressed in ORF skills. The difference between pre-test and post-test scores was determined by subtracting the median pre-test score from the median post-test score. Students whose scores improved over the summer had a positive number indicating how many WRC their ORF score had improved over the summer. Those who regressed had a negative number for this difference.

Follow-up Oral Reading Fluency (Post-test to Follow-up Difference)

Follow-up data were collected and recorded using the same probes and procedures the pre-test and post-test data. Follow-up assessments were conducted with those students who had lost at least 10% of their ORF from pre-test to post-test. In order to ascertain whether students improved, remained the same, or recouped their post-test scores after returning to school for a month, the difference between post-test and follow-up scores was computed.

Instrument

R-CBM was used to measure students' ORF for pre-test, post-test, and follow-up measures. The three grade-level benchmark probes from the commercially-available AIMSweb system (AIMSweb, 2002) assessed students'

reading performance and progress in words read correctly per minute. The standard AIMSweb benchmark probes for each grade level were used.

The reading passages were presented to students on 8 by 10 inch paper, and text was double-spaced in 12-point Arial font. Examiners received nearly identical copies of the probes. However, examiners had access to numbers at the end of each line of text corresponding to the total number of words in the text up to that point. Each probe had about 350 words of text to be read.

Description and Development of Probes

Howe & Shinn (2002) described the rationale and procedures used in the creation of the R-CBM benchmark probes. In order to be used in multiple school districts, the probes were designed to be standardized, grade appropriate and "curriculum independent." Readability formulas (RFs) determined the extent to which passages had an appropriate level of difficulty. Curriculum independence meant that passages were not tied to any published basal texts or trade books. Instead, they provided a common standard by which multiple school districts could compare their students' performance to a national norm.

Thirty teachers and 10 paraprofessionals from a medium-sized suburban/rural education district in the Midwest participated in the writing of the AIMSweb graded reading passages, including the benchmark probes (Howe & Shinn, 2002). Authors were recruited based on their familiarity with reading instruction and how students learn to read; they also possessed knowledge of "the kinds of literature students at various grades encountered." The R-CBM probe authors also needed to write passages that fit parameters of commonly used RFs. In addition, potential authors received specifications regarding the number of words per passage to be written for different grade levels.

Next, authors were given directives for the number of syllables and sentences per 100 words per grade according to the Fry readability formula (Fry, 1968). Finally, stories needed to have beginnings and endings, and authors received instructions regarding how to include proper names in the passages. All of the passages used in this study were fictional narrative texts.

Readability

Several RFs quantitatively assessed the difficulty of R-CBM passages used in this study. These methods estimated the age and grade appropriateness of the instrument as a measure of reading fluency for each grade level. Although a multitude of factors contribute to a text's level of difficulty, readability broadly refers to "the ease with which a reader can read and understand text," (Oakland & Lane, 2004, p. 244). More specifically, readability encompasses elements of language as well as factors such as interest.

Howe and Shinn (2002) described the various ways in which the R-CBM probes' readability was determined during their development. First, passages that failed to meet passage difficulty criteria using the Lexile-graded standards method were excluded from the initial set of passages during development. Once the final R-CBM passages had been chosen for publication, four more RFs measured the passages' readability: Flesch, Powers, Fry, and Spache. The computer software program *Readability Calculations*, developed by Micro Power and Light (*Readability Calculations*, 1999), compared readability estimates given by the aforementioned formulae.

Readability data yield whole number scores pertaining to grade levels. Thus, using the RFs listed above, each set of passages was assigned a number in order to calculate correlations between readabilities for each grade level. For example, second grade probes were assigned the

value of two, third grade probes had a value of three, etc. Correlations between grade level and each of the RFs ranged from 0.78 to 0.99, with a median correlation of 0.95 (Howe & Shinn, 2002). Across grade levels, these correlations indicated that multiple RFs confirmed solid grade-level appropriateness for the R-CBM passages.

In addition to the RFs used in the development and publication of the AIMSweb R-CBM probes, the researcher conducted additional readability assessments of the nine passages used in this study. These passages included three AIMSweb R-CBM benchmark probes each for grades 2, 3, and 4. The researcher used the *OKAPI!* tool (Wright, 2009), an online instrument that calculates the readability of written material. The *OKAPI!* tool uses the Spache formula, developed by Spache (1953) to evaluate primary texts. This index uses two main factors to approximate a reading passage's grade level. First, the Spache considers the average sentence length in words per sentence. Next, it calculates the percentage of "difficult" words, or those not found on the *Spache Revised Word List* (Spache, 1953). Table 3 lists the readability estimates generated using the *OKAPI!* program to compute the Spache grade level.

Table 3.
Grade Equivalent Readability Estimates for R-CBM Passages using OKAPI! and Microsoft Word Software

Grade	Passage	Spache	Flesch-Kincaid
2	1	2.4	4.7
	2	2.5	2.7
	3	2.7	3.9
3	1	3.2	4.3
	2	2.9	2.9
	3	2.8	3.6
4	1	3.5	5.2
	2	3.9	5.2
	3	3.7	5.2

The Flesch-Kincaid formula also was used to compute readability for each of the benchmark passages used in this study. The researcher accessed this formula using the spelling and grammar feature of the software program *Microsoft Office Word 2003*, a Windows®-based software application by Microsoft Corporation.

The Flesch-Kincaid grade level test considers two factors in determining an estimated grade level. First, it assesses average sentence length, or the number of words divided by the number of sentences. Second, it calculates the average number of syllables per word, or the number of syllables divided by the number of words (Microsoft Corporation, 2009). Table 3 includes the grade level approximations generated using this formula.

Several studies have assessed the validity and utility of RFs. Overall, RFs have been found to lack both consistency and validity, particularly for use with R-CBM (Ardoin & Christ, 2008; Ardoin, Suldo, Witt, Aldrich, & McDonald, 2005; Fuchs, Fuchs, & Deno, 1983). Thus, this

data should be viewed in light of this limited utility as a precise measurement of passage readability. Instead, these RFs provide estimates of relative reading difficulty levels. Table 4 shows inconsistencies found when using different quantitative RFs to evaluate a text.

Table 4.
Alternate-form Reliability of AIMSweb Standard Benchmark Reading Assessment Passages

Grade	Passage	Alternate-Form Reliability
2	1	.81
	2	.80
	3	.85
	Mean	.82
3	1	.85
	2	.83
	3	.87
	Mean	.85
4	1	.87
	2	.82
	3	.86
	Mean	.85

Validity

According to Fuchs and Fuchs (1992), research examining correlations between ORF) and widely used commercial reading tests has established ORF's criterion validity. The authors also indicated that ORF scores correlate with ages and discriminate between student categories, supporting its well established discriminative validity. Marston (1989) conducted a comprehensive analysis of different aspects of reliability and validity for CBM ORF. In reviewing research studies examining the technical properties of ORF, he found strong correlations between having

students read aloud for one minute and a variety of commercial reading tests.

The samples of the studies reviewed by Marston included students in general education as well as those with "mild" disabilities, ranging from grades 1-6. Correlation coefficients ranged from .63 to .90, with most falling about .80, supporting the claim that ORF measures had strong criterion validity. Marston also found that ORF measures correlated with teachers' ratings of students' reading skill ($r = .86$); the correlation between teacher judgments and ORF was stronger than that between teacher ratings and published tests. Thus, research indicates that ORF has strong criterion validity as a measure of overall reading proficiency.

Good and Jefferson (1998) reviewed five studies examining the concurrent, criterion-related validity for CBM reading probes. They used a set of measures for grades 2-6 from The Test of Reading Fluency (Children's Educational Services, 1987). Correlations compared R-CBM scores to a variety of norm-referenced, commercially published reading tests such as the Woodcock-Johnson Psycho-Educational Battery (Woodcock & Johnson 1977).

When considering each grade level individually, the authors obtained median validity coefficients across studies for each grade. Among those median scores for grades 2-6, 4th grade had the strongest validity (.73) and 5th grade had the weakest validity (.62). However, these scores were close and indicated moderate to strong validity.

When examining validity coefficients for each grade level, Good and Jefferson (1998) examined the validity of R-CBM across studies. Findings indicated that overall, all of the coefficients for CBM reading measures ranged from .60 to .80, concluding that R-CBM had strong construct validity.

Reliability

Marston (1989) also examined reliability for CBM reading procedures. In reviewing the literature, he found test-retest reliability coefficients over a ten week period ranged from .82 to .97, with most falling above .90. Marston also found parallel form reliability estimates to be strong; they ranged from .84 to .96. Finally, he found that R-CBM had excellent inter-rater reliability, with a correlation of .99. Overall, these studies of R-CBM ORF tasks support technical adequacy on a wide range of reliability and validity measures.

Regarding reliability of the specific AIMSweb standardized benchmark probes used, Howe and Shinn (2002) determined alternate-form reliability coefficients of ranging from .80-.87. These coefficients represented the average correlation for each alternate-form probe used. Table 4 lists the coefficients for each benchmark passage used.

Standard Error of Measurement

According to Howe and Shinn (2002), the R-CBM passages used in the current experiment had standard errors of measurement (SEM) ranging from 9.7 to 10.5 words per minute. Table 5 lists the SEM for passages at each grade level studied.

Table 5.
Standard Error of Measurement and Alternate Form Reliability for R-CBM Passages. Adapted from Howe and Shinn (2002)

Grade	Standard Error of Measurement	Alternate Form Reliability
2 nd	9.9	.82
3 rd	10.5	.85
4 th	9.7	.85

Alternate Form Reliability

The alternate form reliability ranged from .82 to .85 when comparing the three passages for each grade level, which is considered to be high (Howe & Shinn, 2002). This metric examined the correlation between each of the three benchmark assessment passages given to each student during the repeated administrations. Table 5 lists the correlation for grades 2, 3, and 4 as used in the study.

Procedures

Examiner Training

Eleven examiners, including teachers and paraprofessionals at the school, administered assessments. The district school psychologist trained the examiners by presenting AIMSweb's standard PowerPoint® presentation and having them practice using the associated practice exercises. The power point and practice exercises are publicly available through the AIMSweb website training and support section (<http://www.AIMSweb.com/support-training/training/training-materials>). The practice exercises include eight videotaped sessions of R-CBM with students.

Trainees practiced administering and scoring these exercises during training. Next, they checked their responses using the answer key provided by the AIMSweb® *Training Workbook* (Shinn & Shinn, 2002). The May 2008 administration was the examiners' third time doing benchmarking with students. Thus, the examiners participated in benchmarking for about 50 students each in September 2007 and January 2008 prior to the May 2008 pre-test assessment.

Pre-test

In May 2008, all students in 2nd, 3rd, and 4th grade were assessed using their grade-appropriate R-CBM probes as part of the school's benchmarking procedure. The probes assessed each student's ORF rate in WRC per minute.

Three probes were administered, scored, and WRC scores were recorded. The median of the three scores, the metric advocated by AIMSweb as the standard comparison metric for benchmarking, also was identified and recorded for data analysis purposes.

Creation of Achievement Level Groups

Data from the pre-test were coded to protect students' anonymity and entered into a Microsoft Excel spreadsheet. The students' grade level, sex, pre-test score, summer school attendance, and disability status were entered into the spreadsheet. During the summer of 2008, the school district provided the researcher with the pre-test data.

Using these data, the researcher compared each of the students' scores to divide students into one of two groups. Students who scored below the 50th percentile benchmark established by AIMSweb (2008) for their grade level during the spring 2008 benchmarking session were grouped together as having "low achievement." The students who met or exceeded the predetermined benchmark scores were included in one group showing average or above-average achievement.

Post-test

In September 2008, one week after the beginning of school, the assessment team re-administered the same three probes used during the pre-test (the previous grade level) to reassess students' ORF. Once again the WRC from all three probes as well as median WRC per minute were recorded. During the post-test, the researcher was given the probes with random identification numbers and demographic information (grade, sex, LD status, etc.) to enter into spreadsheet software. All students participating in the regular CBM benchmarking procedure participated.

Follow-up Assessment

In October 2008, four weeks after the initial assessment, a follow-up assessment was done with students who had shown "significant"

regression, which was defined as losing 10% or more of the initial words read correctly (WRC) during the summer break. For example, if a student had a median of 80 WRC per minute in the summer and 72 or fewer WRC in the fall, they received this follow-up assessment.

Analysis

Means, standard deviations, number of students in each group, and ranges (if applicable) were calculated to provide informal descriptive analysis for each variable.

A repeated measures analysis of variance (ANOVA-RM) was used to analyze the impact of age/developmental status (grade level), sex, initial low achievement, summer school attendance, and SLD status on ORF regression.

Assumptions. Procedures were followed to establish whether assumptions specific to each of the statistical procedures used were met. This helped determine the appropriateness of using each statistical technique for the available data. For ANOVA-RM, these assumptions include the use of interval or ratio data, normality of data, and equal variances of dispersion.

The first assumption to be verified included the use of interval or ratio data. These types of data possess both nominal (naming) and ordinal (ordering) properties. Interval data has equal quantities represented between numbers. Ratio data shares the properties of interval data in addition to a zero point.

The second assumption to be met required that scores be normally distributed. Checking this assumption can be accomplished using visual inspection by plotting a frequency distribution accompanied by a normal curve.

A third assumption included ensuring equal variances between data points. Equal variances are a measure of each point from its mean value.

Thus, measures of dispersion and standard deviations are equal. Table 6 lists the statistical techniques used for each research question, accompanied by their appropriateness based on meeting the required assumptions for each measure.

Table 6.
Research Questions, Hypotheses, Variables, Statistical Analyses, and Statistical Assumptions for the Reading Regression-Recoupment Project

<u>Research Questions</u>	<u>Hypotheses</u>	<u>Variables</u>	<u>Statistic</u>	<u>Assumptions</u>	<u>Assumptions Appropriateness</u>
1. How much ORF regression and recoupment do students with SLD show compared to their peers in general education?	Students with SLD will regress more and recoup less than peers.	SLD status and ORF score	ANOVA-RM	1. Interval or Ratio Data 2. Normality 3. Equal Variances	1. Instrument uses ratio data 2. Examine histogram within a normal curve 3. Descriptive statistics
2. How do male and female students differ in ORF regression and recoupment?	Males will regress more and recoup less than females.	Sex and ORF score	ANOVA-RM	1. Interval or Ratio Data 2. Normality 3. Equal Variances	1. Instrument uses ratio data 2. Examine histogram within a normal curve 3. Descriptive statistics
3. What relationship will exist between pre, post, and follow-up R-CBM measures?	Deficient readers will regress more and recoup less than peers.	Pre, Post, & Follow-up ORF scores	ANOVA-RM	1. Interval or Ratio Data 2. Normality 3. Equal Variances	1. Instrument uses ratio data 2. Examine histogram within a normal curve 3. Descriptive statistics
4. How does summer reading program attendance impact ORF regression and recoupment?	Program attendance will lead to the same or better ORF.	Summer Program and ORF scores	ANOVA-RM	1. Interval or Ratio Data 2. Normality 3. Equal Variances	1. Instrument uses ratio data 2. Examine histogram within a normal curve 3. Descriptive statistics
5. Is there any association between grade and ORF regression and recoupment?	ORF will increase with age/grade. No differences anticipated in terms of regression and recoupment.	Grade and ORF score	ANOVA-RM	1. Interval or Ratio Data 2. Normality 3. Equal Variances	1. Instrument uses ratio data 2. Examine histogram within a normal curve 3. Descriptive statistics
6. Will one month back in school improve ORF for students who regressed from pre-test to post-test?	Returning to school for one month will increase regressors' ORF scores.	Post-test ORF scores Follow-up ORF scores	ANOVA-RM	1. Interval or Ratio Data 2. Normality 3. Equal Variances	1. Instrument uses ratio data 2. Examine histogram within a normal curve 3. Descriptive statistics

Statistical Techniques

Question One: An ANOVA-RM was planned to compare how much ORF regression and recoupment do students with SLD showed in comparison to their peers in general education. However, given that only four students in the sample were identified as having SLDs, this sample size was too small for such analysis.

Question Two: An ANOVA-RM was used to compare how boys and girls compared in terms of regression of ORF skills. An additional ANOVA-RM was used to determine relative recoupment between the sexes.

Question Three: An ANOVA-RM was used to compare regression between students with low versus benchmark achievement levels. Too few students with low achievement regressed significantly ($n = 3$), making statistical analysis inappropriate for this group in terms of recoupment.

Question Four: An ANOVA-RM was used to compare how much regression occurred among students who attended the school-sponsored summer reading program compared to their peers. Because only 3 students who attended the summer program regressed, no statistical procedure was appropriate in determining their recoupment relative to other regressors.

Question Five: An ANOVA-RM was used to determine whether students of different grade levels differed in terms of regression. Descriptive statistics as well as ANOVA-RM was used to determine whether ORF increased significantly with increasing grade levels.

Question Six: An ANOVA-RM was used to determine the degree to which regressors recouped their losses in ORF. This analysis compared post-test and follow-up R-CBM measures.

Computer Programs

Individual student scores were coded by the school and entered into a Microsoft Excel spreadsheet. Next, these data were prepared and entered into a spreadsheet using the Statistical Package for the Social Sciences

(SPSS) software for statistical analysis. SPSS was used to check assumptions for each of the analyses used and analyze the data.

Parent Survey

Following the post-test, a parent survey was given to the school district to send home to parents. One component included a permission letter explaining that responses would be coded and confidential, and were intended for use in dissertation research. This survey was intended to ascertain the amount of reading practice and instruction students received over the summer. Other than participation in a school-sponsored summer reading program, little was known regarding how much reading practice students had over the summer break. In addition, little was known regarding students' access to summer reading instruction via a program other than their school. Thus, the researcher attempted to obtain such information in a survey sent home to parents. This parent survey asked the following questions:

1. Approximately how many hours of reading did your child do per week this summer?
2. Did your child attend an academically focused reading program this summer?
3. If so, which program?

Unfortunately, after some consideration and review, the school district's administration decided that sending a permission letter and survey home to parents might cause confusion regarding the standard benchmarking procedure already in place. Therefore, parent survey was not administered as originally planned.

Summary

The current study measured summer regression of ORF skills among 136 elementary school students in grades 2, 3, and 4 (during the pre-test). Repeated measures of identical R-CBM probes, a measurement system

validated for this age group, were administered to all of the participating students in a pre-test and post-test. Students whose ORF skills regressed by 10% or more during the summer received a follow-up measure using the same probes one month later.

In addition to examining reading regression and recoupment, this study specifically focused on comparing different groups of students. For example, data compared students based on sex, initial achievement levels, grade level, learning disability status, and summer program attendance. An ANOVA-RM was used to analyze several of these factors, and assumptions were checked to determine appropriateness of these techniques for the data set. The computer program SPSS was used to analyze statistical data.

CHAPTER IV

RESULTS

Introduction

One main focus in this study was summer regression, or the change in oral reading fluency (ORF) from May to September. Thus, pre-test to post-test changes were a main focus of the analysis. Additional variables analyzed included specific learning disability (SLD) status, age, sex, achievement level, and summer program attendance. Table 7 lists a cross-tabulation of different groups sub-divided by their overlapping characteristics in terms of these factors.

Table 7.
Cross-tabulation of Students Based on Achievement Level, Sex, Grade, Learning Disability Status, and Summer School Attendance

Low Achievement	Sex	Diagnosed with Learning Disability			Grade			Total
		No	Attended Summer Program	Yes	2	3	4	
No	Female	No	Attended Summer Program	No	13	14	18	45
				Yes	0	1	0	1
			Total		13	15	18	46
		Yes	Attended Summer Program	No	0	0	0	0
				Yes	0	1	0	1
			Total		0	1	0	1
	Male	No	Attended Summer Program	No	10	16	16	42
				Yes	1	0	2	3
			Total		11	16	18	45
		Yes	Attended Summer Program	No	0	0	0	0
				Yes	0	0	0	0
			Total		0	0	0	0
Yes	Female	No	Attended Summer Program	No	3	2	6	11
				Yes	2	1	3	6
			Total		5	3	9	17
		Yes	Attended Summer Program	No	0	0	2	2
				Yes	0	0	0	0
			Total		0	0	2	2
	Male	No	Attended Summer Program	No	5	6	8	19
				Yes	1	0	4	5
			Total		6	6	12	24
		Yes	Attended Summer Program	No	0	0	0	0
				Yes	0	0	1	1
			Total		0	0	1	1

A repeated-measures analysis of variance (ANOVA-RM) was computed to determine the impact of age, sex, achievement level, and summer program attendance on pre-test to post-test change in ORF. Given that only four students in the sample were found to have specific learning disabilities (SLD), this group was not analyzed using ANOVA-RM.

Did Students Regress?

R-CBM scores were calculated as the median words read correctly per minute (WRC) among three passages for each assessment period. Table 8 lists descriptive statistics regarding the whole group's change in ORF from spring to fall.

Table 8.
Descriptive Statistics Regarding Whole Group Pre-test to Post-test Change

	N	M	SD	Range
Pre-test	136	127	38	16-203
Post-test	136	130	39	20-221
Difference	136	2.85	13	16-221

An ANOVA-RM indicated that the whole group's change in ORF from spring to fall was not significant. Table 9 lists the results of this analysis.

Table 9.
Analysis of Variance with Repeated Measures for Variables Impacting ORF

Source	df	F	η	p	power
Between subjects					
Grade (G)	2	9.12*	.14	<.01**	0.97
Achievement Level (Ach)	1	58.37*	.34	<.01***	1.00
Summer Program Attendance (Sum)	1	4.99	.04	.03	0.60
Sex	1	2.65	.02	.12	0.37
G x Ach	2	1.46	.03	.23	0.31
G X Sum	2	0.31	.01	.73	0.10
G x Sex	2	0.56	.01	.57	0.14
Ach x Sum	1	0.23	.00	.63	0.08
Ach X Sex	1	4.52	.04	.04	0.56
Sum X Sex	1	4.17	.04	.04	0.53
G x Ach X Sum	1	0.13	.00	.71	0.07
G x Ach X Sex	2	0.06	.00	.93	0.06
Grade X Sum X Sex	1	0.98	.01	.32	0.17
Residual	116	(990.93)	.00	<.01	0.00
Within subjects					
Time	1	2.56	.02	.11	0.36
Grade (G)	2	0.08	.00	.91	0.06
Achievement Level (Ach)	1	0.01	.00	.92	0.05
Summer Program Attendance (Sum)	1	0.01	.00	.92	0.05
Sex	1	0.33	.00	.57	0.09
G X Ach	2	0.07	.00	.94	0.06
G X Sum	2	0.40	.01	.67	0.11
G X Sex	2	0.61	.01	.54	0.15
Ach X Sum	1	0.08	.00	.78	0.06
Ach X Sex	1	0.25	.00	.62	0.08
Sum X Sex	1	0.43	.00	.51	0.10
G X Ach X Sum	1	0.08	.00	.78	0.06
G X Ach X Sex	2	1.04	.02	.35	0.23
G X Sum X Sex	1	0.38	.00	.56	0.09
Residual	116	(88.49)	.00	<.01	0.00

Note. Values enclosed in parenthesis represent mean square errors.

* $p < .001$, ** $p < .0001$

Impact of Standard Error of Measurement

As mentioned earlier, the standard error of measurement (SEM) for the probes used ranged from 9.5 to 10 words per minute, differing slightly

with each grade level probe used. Because test scores lack precision, each student had a range in which the actual score reflected his or her "true" score. Due to the specific SEM of R-CBM passages used in this study, the range of the actual score +/- 10 theoretically included the students' true scores.

Summer Regression

Overall, most students' ORF scores did not decrease from pre-test to post-test. Of the 136 participating students, only 44 students showed decreases in R-CBM score during the summer. Among those who regressed, the majority (n = 27) showed decreases of ten or fewer words from pre-test to post-test. Thus, only 17 students showed regression equal to or greater than that caused by error according to the SEM. Only 17 of these showed a loss of 10% or more of their pre-test score during the post-test. For the purposes of this study, such an increase was considered to be "significant," warranting the follow-up assessment which took place one month later. Table 10 includes further information regarding the number of students who scored within, greater than, or less than the SEM.

Table 10.
Pre-test to Post-test Changes in ORF Scores

<u>Regression</u>	<u>n</u>
Within SEM	27
Greater than SEM	17
Total	44
<u>Improvement</u>	<u>n</u>
Within SEM	36
Greater than SEM	36
Total	72
Minimal Change*	21

*defined as 1-2 WRC variation from pre-test to post-test

Summer Improvement

As indicated in Table 10, some students regressed, even more of the participating students improved their ORF scores after summer break. A

total of 72 students had higher post-test scores compared to their pre-test scores. Of the students who improved, 36 of them improved by 10 words or less, falling within the range of the SEM. Thus, their improvements may have happened due to measurement error as opposed to true improvement. The remaining 36 students whose scores improved had increases of more than 10 WRC per minute, reflecting improvements caused by factors other than chance.

Minimal Changes

While some increases and decreases took place when comparing pre-test and post-test scores, other students showed very little change in ORF over the summer. As indicated in Table 10, 21 of the 136 students had scores that were either identical to or 1-2 words greater or less than their pre-test scores. Combining this group with students who scored slightly higher ($n = 36$) or slightly lower ($n = 27$) than their pre-test scores as defined by the SEM limits yielded a total of 120 students scoring at or close to their pre-test score.

Research Question 1: Specific Learning Disabilities

One major research question that this study aimed to answer was whether students with SLDs showed more ORF regression over the summer than their peers. The changes in ORF from pre-test to post-test were examined in order to answer this question. However, due to the small size of this subsample, no statistical analysis procedure was used to determine these differences.

Overall, students with SLD had lower mean pre-test and post-test scores compared to their peers. For example, students with SLD had mean scores of 102 WRC for both the pre-test and post-test. The rest of the sample had an average pre-test score of 128 and an average post-test score of 131 WRC. Table 11 provides more descriptive information.

Table 11.
Descriptive Statistics Regarding SLD status and Oral Reading Fluency Measures

SLD Status	Pre-test			Post-test			
	n	M	SD	Range	M	SD	Range
No SLD	133	128	38	16-123	131	39	20-221
SLD	4	102	32	64-142	102	29	63-129

Of the four students with SLD in the sample, two attended the school sponsored summer reading program. The two students who did not attend the program achieved nearly identical pre-test and post-test scores; one showed a decrease of one WRC while the other dropped two WRC. Both of these changes were within the realm of the SEM. Of the two students with SLD who attended the summer program, one increased ORF from pre-test to post-test by 30 WRC. The other student showed a decrease of about 20 WRC.

Regarding recoupment, only 1 of the 4 students with SLD regressed in ORF score during the summer break. This female student also attended the summer reading program sponsored by her school. With a pre-test score of 142 WRC and post-test score of 124 WRC, this student regressed by 17 WRC from May to September. At the time of the follow-up measure, however, she improved her median score to 141 WRC. Further investigation indicated that this student had attention-deficit disorder, wore her glasses inconsistently, and showed varied academic performance on a regular basis. Given that only one student with SLD regressed over the summer, the researcher was unable to use a statistical analysis with adequate power to draw conclusions about recoupment of students with SLD as a group.

Research Question 2: Sex Differences

A second research question this study concerned whether male and female students differed in terms of regression/recoupment. It was

hypothesized that male students would regress more and recoup less compared to their female peers.

Regression

The ANOVA-RM indicated that boys' and girls' scores showed similar changes from pre-test to post-test, $F(1,116) = 0.33, p > .05$. In addition, boys' and girls' scores did not differ significantly from each other, $F(1,116) = 2.65, p > .05$. However, based on the sample size, the observed power for both between group (0.37) and within group (0.09) differences were low. This means that there was only a 37% chance of detecting sex differences in scores, and a 9% chance of detecting differences in ORF regression if those differences existed. Thus, the observed power fell below the generally accepted minimum of 0.8 (Cohen, 1988). This led to a strong possibility of failing to detect a significant effect that may have occurred.

Despite a lack of significance as a main effect, sex was a contributing factor in its interactions with low achievement and summer program attendance. More information regarding these interactions will be described below.

As indicated in Table 12, group pre-test and post-test means for males versus females demonstrated that males' ORF rates were slightly lower than those of their female peers for both pre-test and post-test measures.

Table 12.
Descriptive Statistics Regarding the Impact of Sex Differences on Oral Reading Fluency Measures

Sex	n	M	Pre-test		Post-test		
			SD	Range	M	SD	Range
Female	66	133	40	16-203	136	40	20-221
Male	71	121	35	42-186	124	37	42-197

Overall, 66 females and 71 males participated in the study. Females attained higher scores than males for both pre-test and post-test measures. More specifically, females had an average score of 133 WRC for the pre-test and 136 for the post-test, while males' had average pre-test score of 121 and an average post-test score of 124 WRC. Thus, both groups showed similar levels of improvement of about three WRC from pre-test to post-test measurement.

Recoupment

An additional ANOVA-RM was computed to determine the whether sex significantly impacted recoupment. This factor was not found to be significant; males and females did not differ regarding recoupment. However, the observed power was low for both between (0.25) and within-subjects factors (0.07). Table 13 lists descriptive statistics and ANOVA-RM data for the post-test to follow-up ORF differences.

Table 13.
Analysis of Variance with Repeated Measures for Effect of Sex Differences on Oral Reading Fluency Recoupment

Sex	n	Post-Test			Follow-up		
		M	SD	Range	M	SD	Range
Male	8	109	34	73-163	128	38	123-206
Female	6	133	28	101-176	156	32	72-189
Total	14	119	33	73-176	140	37	72-206

Analysis of Variance with Repeated Measures

Source of Variation	df	<i>F</i>	η	<i>p</i>	power
Between Students	1	1.905	.14	.193	0.25
Within Students	1	0.191	.83	.670	0.07
Time x Sex	1	0.191	.02	.670	1.00
Residual	12	48.87			

A total of 8 males and 6 females received the follow-up measure to assess recoupment. The researcher selected this group of students due to their showing "significant" regression, operationally defined as an ORF decline of 10% or more from pre-test to post-test.

As a whole group, students' mean WRC scores improved, increasing from 119 to 140 WRC from post-test to follow-up. Among the regressors, male students had a post-test mean of 109 with a follow-up mean of 128. Their female peers had slightly higher scores for both post-test (133 WRC) and follow-up measures (156 WRC).

Most of the males (7 out of 8) and all of the females showed improvement from post-test to follow-up. The one male student who failed to recoup his lost ORF skills over the summer stayed almost the same, scoring one WRC less on the follow-up measure.

Research Question 3: Low Achievement

Another research question in this study sought to determine the relationship between pre-test, post-test and follow-up R-CBM measures among students with different initial achievement levels. It was hypothesized that students with very deficient skills will be more likely to show regression and lack of recoupment than those scoring closer to average/normal limits on the pre-test.

As described earlier, students' pre-test scores determined whether they qualified as having "low achievement," defined as having below-benchmark ORF scores during the summer. Given this criterion, 44 students in the sample qualified as having low achievement. The 93 other students were classified as having "benchmark" achievement, meaning they scored at or above the benchmark during the May 2008 pre-test.

Regression

Students in the low achievement group had lower pre-test and post-test scores than their peers, which occurred largely due to the creation

of this group using benchmark cut scores. Low achievers had an average pretest score of 87 WRC, with an average post-test score of 90 WRC. The rest of the sample (benchmark achievement group) had mean pre-test scores of 147 WRC, improving to 149 WRC at the post-test measure. Both groups showed similar levels of improvement despite their initial score differences. More descriptive statistics for this comparison are included in Table 14.

Table 14.
Descriptive Statistics of the Impact of Low Achievement on Summer Regression of Oral Reading Fluency Measures

Low Achievement	n	Pre-test			Post-test		
		M	SD	Range	M	SD	Range
Low Ach	44	87	24	16-129	90	27	20-139
Reg Ach	93	147	26	93-203	149	28	88-221

An ANOVA-RM examined the effect of achievement level on ORF regression; results are included in Table 9. Achievement levels did not differ in terms of ORF change from pre-test to post-test, $F(1,116) = 0.01$, $p > .05$. The small sample size led to an observed power of 0.05, or a 5% possibility of detecting a significant difference between these groups. This power level fell below the generally accepted minimum established by Cohen (1988).

Despite the lack of significant differences found in terms of relative regression, the low and benchmark achievement groups differed significantly in terms of overall ORF scores: $F(1,116) = 58.37$, $p < .01$. In addition, the overall achievement levels differed in terms of sex differences, for which a significant interaction effect was observed $F(1,116) = 4.52$, $p < .05$.

In order to do a post-hoc analysis of this interaction, each individual student's pre-test and post-test score were averaged to create

an overall ORF mean for each student. Next, group means were computed for boys and girls. Finally, students' scores were grouped by achievement level.

Results indicated that girls with benchmark achievement had higher ORF scores than boys with benchmark achievement. On the other hand, boys and girls with low achievement had very similar ORF scores. This interaction is depicted in Figure 4.

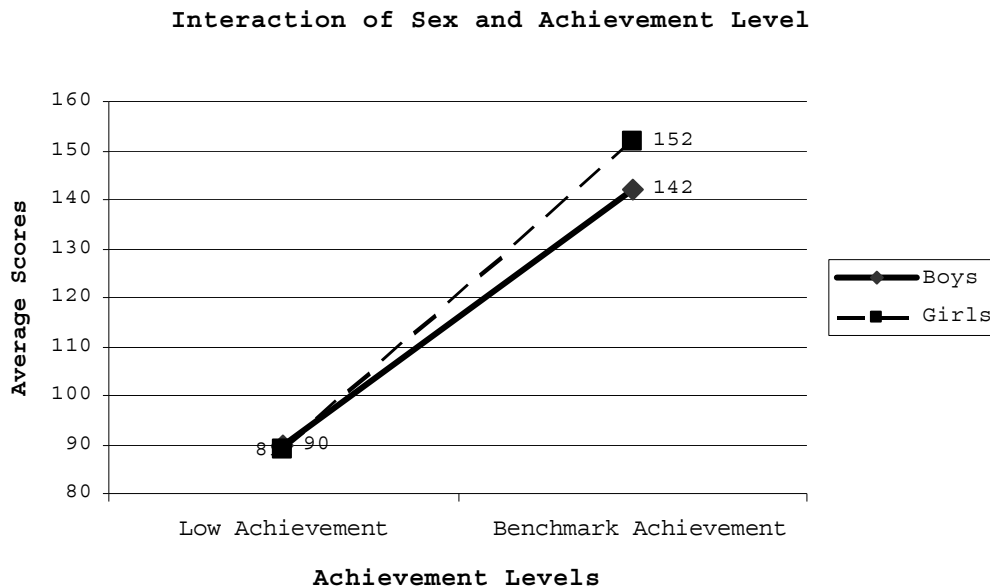


Figure 4. Interaction between sex and achievement level for ORF regression.

Recoupment

As a group, students with initially low scores who regressed from pre-test to post-test had an average post-test score of 87 WRC with an average follow-up score of 102. The rest of the group identified as having initial benchmark achievement showed an average pre-test score of 127, with an average follow-up score of 150. It is important to note that only three students identified as having low achievement showed a "significant"

amount of regression from pre-test to post-test. Table 15 includes more descriptive information regarding this comparison.

Table 15.
Descriptive Statistics Regarding the Impact of Low Achievement on ORF Recoupment

Achievement Level	n	Post-Test			Follow-up		
		M	SD	Range	M	SD	Range
Benchmark Achievement	11	127	31	88-176	150	33	106-206
Low Achievement	3	87	14	73-101	102	27	72-110
Total	14	119	33	73-176	140	37	72-206

Of the three "low achieving" students participating in the follow-up measure, two attained post-test ORF scores that exceeded their pre-test scores. The third student's score dropped 10 WRC from pre-test to post-test, with a follow-up score one WRC lower than the post-test score.

Overall, only 1 of the 3 students with initial "low achievement" regressed and did not recoup ORF skills lost during the summer after one month's return to school. Because so few participants regressed, no inferential statistical procedure could appropriately determine whether differences between achievement level groups took place.

Research Question 4: Summer Reading Program Attendance

Given an extended absence from school, reading practice and instruction likely impacts changes between pre-test and post-test measures. It was hypothesized that attending a summer reading program (SRP) would help students maintain or improve their ORF scores from pre-test to post-test.

The students attending the SRP had an average pre-test score of 93 WRC, improving to 98 WRC over the summer. Those not participating in the summer reading program had a higher average pre-test score (132 WRC) and

also showed some improvement over the summer, attaining a post-test score of 134 WRC. Table 16 includes more descriptive information comparing students who attended the school's SRP to those who did not participate.

Table 16.

Descriptive Statistics Regarding the Impact of Summer Reading Program Attendance on Oral Reading Fluency Measures

Summer Program	n	Pre-test			Post-test		
		M	SD	Range	M	SD	Range
No	120	132	37	16-203	134	38	20-221
Yes	17	93	30	42-150	98	32	42-172

About half (9 out of 17) of the students who attended the summer reading program improved their ORF score from pre-test to post-test. Five others maintained their scores, having attaining the same or 1-2 word difference from pre-test to post-test. Thus, the majority of summer reading program attendees (14 out of 17) performed as predicted. However, three students showed a decline in ORF scores from pre-test to post-test despite their access to summer reading instruction. All three of these students showed "significant" increases of 12 WRC or more, or 10% of their pre-test scores. Interestingly, these students all recouped their losses after a month back in school as measured by the follow-up assessment.

An ANOVA-RM analyzed the impact of SRP attendance on ORF measures. As shown in Table 9, students who attended the SRP did not differ from those who did not in terms of ORF regression $F(1,116) = 0.01$, $p > .05$. Due to the small sample size of this group, the observed power for this effect was 0.05, which fell below Cohen's (1988) minimum standard of 0.80 for adequate power. Thus, there was a strong likelihood of failing to find an effect if one was actually present.

On the other hand, a significant main effect was observed when comparing ORF scores between groups of students who attended the SRP as

opposed to those who did not: $F(1,116) = 4.99, p < .05$. This effect depended on the students' sex, and the interaction between sex and SRP attendance also was found to be significant $F(1,116) = 4.17, p < .05$.

To perform a post-hoc analysis of this interaction, individual means were calculated for each student's ORF, combining pre-test and post-test scores. Next, group means were computed for each sex and plotted in terms of SRP attendance. Figure 5 depicts this interaction.

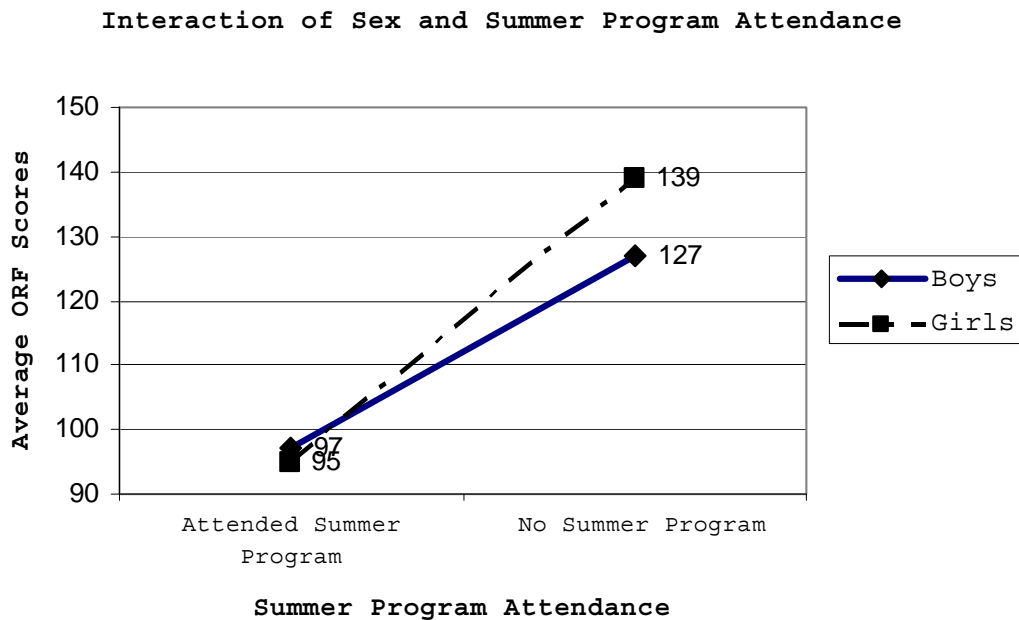


Figure 5. Interaction of sex and summer program attendance on ORF regression.

As shown in Figure 5, girls who did not attend the SRP had higher average ORF than boys who did not attend. Among students who attended the SRP, both sexes had similar average ORF scores.

Research Question 5: Grade Level

Another question this study sought to answer regarded whether age and grade (developmental status) had any impact on ORF regression and recoupment. Based on developmental and learning growth trends, it was hypothesized that ORF would increase with age/grade. However, no

significant differences were anticipated in terms of regression/recoupment for the age groups assessed in the current study (grades 2-4).

As expected, the grade levels differed significantly among themselves, $F(2,116) = 9.12, p < .01$. Table 17 includes further descriptive statistics regarding each grade level's performance as a group.

Table 17.
Descriptive Statistics of the Effect of Grade Level on ORF Measures

Grade (at pre-test)	n	Pre-test			Post-test		
		M	SD	Range	M	SD	Range
2 nd	35	109	39	16-193	114	42	20-199
3 rd	42	129	36	55-203	128	38	55-221
4 th	60	136	34	57-198	140	35	63-197

An ANOVA-RM was calculated to determine the effect of grade level on ORF regression. As shown in Table 17, grade levels did not differ in terms of regression $F(2,116) = 0.08, p > .05$. Due to the sample size, the observed power for this effect (0.06) was very low, falling below the accepted level (Cohen, 1988). Thus, the small size of the sample made it difficult to detect significant differences in terms of relative regression among the three grade levels.

As expected, students' ORF increased by grade level for both pre-test and post-test. Students who were in 2nd grade at the time of the pre-test had an average initial ORF score of 109 WRC and a post-test score of 114 WRC. The 3rd grade students had an average pre-test score of 129 with mean post-test scores of 128 WRC. Finally, 4th graders attained an average pre-test score of 136 with an average post-test score of 140 WRC.

Because only 14 total students regressed significantly, using an inferential statistical procedure was not an appropriate means by which to

measure recoupment. For ease of comparison, students are described using their grade level at the time of the pre-test, which also corresponds to the grade-level probe used for the repeated measures analysis. Of the 14 total students who regressed, 2 were second graders, 6 were third graders, and 6 were fourth graders. Students showed improvements in ORF from post-test to follow-up at each grade level. Table 18 includes descriptive statistics for post-test and follow-up scores at each grade level.

Table 18.
Analysis of Recoupment. Descriptive Statistics by Grade Level on Post-Test and Follow-Up Oral Reading Fluency Measures

Grade*	n	Post-Test			Follow-up		
		M	SD	Range	M	SD	Range
2 nd	2	82	12	73-90	97	35	72-122
3 rd	6	115	26	88-163	136	28	106-189
4 th	6	135	34	87-176	157	37	110-206
Total	14	119	32	73-176	140	37	72-206

*Grade reported at time of pretest

Research Question 6: Overall Recoupment

A final aim of this study included determining whether students who regressed in ORF over the summer would recoup lost skills after one month back in school. It was hypothesized that one month back in school would significantly increase ORF scores for students who regressed from pre-test to post-test. An ANOVA-RM was done to determine whether significant changes occurred from post-test to follow-up. It was found that regressors improved their ORF scores significantly ($p < .00001$). Table 19 lists descriptive statistics and analyses conducted to determine recoupment, while Table 20 indicates results of the ANOVA-RM.

Table 19.
Descriptive Statistics Regarding ORF Recoupment

	n	M	SD	Range
Post-test	14	119	33	87-161
Follow-up	14	140	37	72-206

Table 20.
Analysis Variance with Repeated Measures for Recoupment

Source of Variation	df	F	η	p
Time	1	59.93*	.83	.00*
Error	12	(48.87)		

Note. Value in parenthesis represents mean square error.
*p < .0001

A total of 14 students showed a decrease in ORF of 10% or greater from May to September. Eight males and six females comprised this group. In general, all except one of the students who regressed "bounced back," achieving scores close to (or within 1-2 words) or greater than their pre-test score. The one student who failed to regain his pre-test score was a second-grade boy who read below benchmark during all three trials. His post-test and follow-up scores were about the same: 73 and 72 words, respectively.

Of the 14 students who regressed, nine achieved higher scores during the follow-up compared to the pre-test. Two in this group attended the school's summer reading program, both attaining higher follow-up reading scores compared to their pre-test scores.

Summary

Several ANOVA-RMs were used to analyze the data on ORF pre-test, post-test, and follow-up measures to answer the research questions

regarding grade, sex, SRP attendance, low achievement. Given that only four students with SLDs participated, inferential statistics were inappropriate in analyzing the impact SLD status on regression. Similarly, given the small sample size for recoupment in regards to certain factors, inferential statistical procedures were only used when sample sizes were sufficiently large to meet the assumptions of the analyses used. Overall, the analysis yielded insignificant results when comparing relative ORF among different groups of students. Additionally, an ANOVA-RM indicated that students did not show regression as a whole group. However, power analysis indicated that the observed power levels were too low to detect differences in regression. Analyses did indicate that significant recoupment occurred among those students who regressed significantly (with ORF scores dropping 10% or more from pre-test to post-test).

The students with SLD had lower mean pre-test and post-test scores compared to the rest of the sample. Only 1 of the 4 students with SLD regressed in ORF score during the summer break. This student recouped in ORF from post-test to follow-up.

The ANOVA-RM conducted to determine the impact of grade, sex, SRP attendance, low achievement, and their interactions found no significant within-group variation for any of these factors. Significant between-group differences were found for the interactions between summer program and sex, as well as for the interaction between sex and low achievement.

As expected, students' ORF increased by grade level, though no significant differences were found in regression or recoupment. Males' ORF was slightly lower than females'; however, these differences were not found to be significant. Students with low achievement and those who attended the summer reading program had lower ORF scores than the rest of the group. These differences occurred due to the selection and creation of these groups.

Due to the small size of groups created, no statistical analysis was done to determine the effects of recoupment on grade, summer program attendance, and achievement level. An ANOVA-RM was computed to determine whether sex impacted recoupment; no significant differences were found. Regarding overall recoupment, an ANOVA-RM indicated that those who regressed showed significant recoupment.

CHAPTER V

DISCUSSION

Overview

Given what prior research has demonstrated regarding reading development and factors such as specific learning disability (SLD) status, age, sex, and other demographic factors, it was hypothesized that one or more of these factors would influence reading regression over the summer. None of these factors significantly impacted reading regression when analyzed alone or in combination.

When examining differences between groups, grade level was found to have a significant effect; oral reading fluency (ORF) scores increased with grade level. Sex interacted with both achievement level and summer reading program (SRP) attendance.

Regarding recoupment, the current study showed that those who regressed significantly recouped their lost skills four weeks after the post-test. Finally, students who attended a school-sponsored SRP maintained or improved their ORF skills from pre-test to post-test. The following sections will revisit extant information on regression and recoupment in relation to the current study, and examine each of the original hypotheses, limitations, implications for future research, and relevance to the field of school psychology.

Current Findings on Regression and Recoupment

Unlike other studies and ORF data found in large normative samples such as AIMSweb (NCS Pearson, Inc., 2008), the current study found that few students regressed during the summer. Instead, most students in the current sample showed slight improvements or minimal changes in ORF during the summer break. Important factors distinguishing the current study from those mentioned above include characteristics of the specific sample used,

lack of information about family income and summer activities, the use of repeated measures of CBM, and likely differences in elapsed instructional time between measurements. These differences may explain the varying outcomes of extant data and research compared to the current study's results, and will be discussed in turn.

Characteristics of the current study's sample likely contributed to the lack of significant regression when comparing pretest and post-test ORF measures. Most of the current study's sample included students in general education, including some students with below average achievement. Three previous studies examined summer reading regression among elementary students in general education. In a meta-analysis of summer regression studies, Cooper, Nye, and Charlton (1996) found that students either showed no academic growth or a slight loss (about 1/10 of a standard deviation) in reading skill.

Previous research has yielded mixed conclusions regarding regression of reading skills during the summer. While some studies suggest skill maintenance for "typical" students (Ross, 1967) or those with mild disabilities (Cook & Schwartz, 1969), others ranged from maintenance to skill loss (Cooper, Nye, and Charlton, 1996). As a whole, many of the studies found that low achieving students (Ross, 1967; Cooper, Nye, and Charlton, 1996) or those with mild disabilities (Allinder & Eicher, 1994; Cornelius & Semmel, 1982; Shaw, 1982) lost skills during the long academic break. The current study's finding that most students showed little regression or change in reading score from pretest to post-test corresponds with the conclusion of skill maintenance. However, findings differed in that the current study found only a small percentage showed a significant loss in reading skill. The current sample's small number of students with disabilities likely led to the observed lack of significant

differences in regression and recoupment when comparing students with and without SLDs.

Differences in measurement among studies of summer regression and recoupment may also account for discrepancies in results. Researchers have studied regression of reading skills over summer break for over 40 years. The current study's use of CBM as a repeated measure distinguished it from most of the earlier research done on summer regression and recoupment. At present only one previous study using an instrument sensitive to small incremental growth (CBM) could be identified. Instead, most of the previous studies assessing regression used repeated measures of standardized, norm-referenced tests. Unlike many norm-referenced tests, CBM was designed to measure incremental progress over time, and has sound technical properties for the grade levels used in the study. Thus, the variation among previous studies' results and those found in the current sample may have occurred due to differences in measurement tools. For example, among the research studies reviewed, only Allinder & Eicher (1994) used CBM while others used repeated measures of different norm-based instruments. Unlike CBM measures, norm-referenced tests may lack adequate sensitivity to small changes over time, such as those that take place over summer break.

Decreases in ORF over the summer break also can be observed upon examination of large databases of ORF benchmarking data such as AIMSweb (NCS Pearson, Inc., 2008). When comparing students' AIMSweb benchmark ORF from spring to fall, students appear to regress, showing fewer words read correctly per minute after summer break (NCS Pearson, Inc., 2008). However, pretest and post-test measures differ given that the ORF measures increase in difficulty as students advance in grade level. This makes post-test measures more difficult than pretest measures. For example, a second grade student would read a second grade passage to measure fluency

before summer. Once the student has advanced to the third grade in the fall, a third grade probe assesses their post-summer ORF. Thus, the increasing difficulty across measures in the AIMSweb database may lead to the appearance of more summer regression than actually takes place. In other words, such summer ORF "regression" may be as much a function of increasing difficulty between probes as much as one of actual skill loss.

The practice of using identical probes for repeated measures distinguished the current study's comparison of spring to fall ORF to that of the AIMSweb database (NCS Pearson, 2008), which collects grade-appropriate data three times per year. This difference likely accounted for the lack of regression seen in the current study as opposed to the regression observed by AIMSweb; AIMSweb's comparison using probes of varying difficulty may have led to more significant regression than a comparison using identical measures. In this way, the current study may serve as a refutation of previous findings when considering the regression/recoupment of "middle income" students similar to those included in the current study (Ross, 1974; Mraz & Rasinski 2007; Cooper, Nye, and Charlton, 1996).

The current study's sample was predominantly "middle income", based on lack of eligibility for free or reduced lunch status. The finding that students maintained or improved their reading skills over the summer concurred with previous studies regarding family income and summer regression. Previous studies have found that students from middle versus low-income families differed in terms of summer regression. When using repeated measures of academic achievement before and after summer, several of these studies concluded that students coming from low income families regressed more than their more affluent peers (Alexander, Entwisle, & Olson, 2001; Ginsburg, 1981; Heyns, 1978). Data showed that while most students showed academic improvement during the school year, low income

students showed significant losses during the summer. Additionally, a meta-analysis of 39 studies found that middle-class children's reading skills remained the same or improved during the summer; however, lower-income students showed academic losses (Cooper, Nye, & Charlton, 1996).

One reason students with different income levels show differential reading loss or gain during out-of-school time likely stems from different opportunities between these groups. The amount of time spent reading outside of school, specifically the reading of books, has been found to improve reading achievement (Allington, 2006; Anderson, Wilson, & Fielding, 1988). Kim (2006) found that providing low-income children with books helped improve their reading skills over the summer. In addition, exposure to out-of-school activities such as camps differed among children with different family income levels. Such activities often led to improvements in academic skills for those who had these opportunities (Miller, 2007).

Given the current's sample's lack of "low income" students, the observed lack of regression (for most participants) may have resulted from access to more enriching activities or environments. For example, students in the sample may have participated in summer activities found to be more accessible to more affluent students compared to less advantaged peers. The lack of parent survey information also made it impossible to ascertain how much reading practice participating students in the current study's sample had during the summer. Despite knowing that the school's summer reading program benefited participating students, any further assumptions about the sample's summer activities are merely speculative. For these reasons, the use of a convenience sample makes the current study's results applicable only to the participating students, and perhaps those with similar demographic characteristics.

Another possible reason behind the current study's finding that most students improved or did not change over the summer may stem from maturation. As most children age, they often improve physically, mentally, emotionally, and psychologically. It is possible that other studies' assumptions that students generally lose reading skills were in error. For example, it is possible that such previous results occurred as a byproduct of poor instrumentation such as repeated measures of tests not designed as progress monitoring measures. Such effects likely had little influence given the relatively short duration between assessments and the precedent that students generally improve over time. Knowing more about access to instruction or practice, which has shown to more clearly improve students' academic skills would likely assist in ruling out maturation effects.

Research Questions and Hypotheses

The current research study aimed to answer several hypotheses generated by the research questions. More specifically, analyses examined the impact of specific learning disabilities, sex, age/grade level, initial low achievement, and SRP attendance had on regression and recoupment.

Specific Learning Disabilities

The first research question involved determining how much ORF regression/recoupment students with SLD showed compared to their peers in general education. It was hypothesized that students with SLD would show more regression and less recoupment than their peers in general education. In the current study, no significant differences were found when comparing students with and without SLD in terms of regression and recoupment. The small number of students with SLD in the study's sample (only four participated) did not allow for a robust analysis of this question. The fact that too few students with SLD participated made it difficult to draw conclusions about how this sub-sample performed as a group.

Considering the four students with SLD in the sample, two attended a SRP through the school. Of these two students, one regressed and one showed improvement in ORF skills. The student who regressed was able to recoup any lost skills by the time of the follow-up measurement conducted four weeks after the post-test. The other two students with SLDs in the sample attained post-test scores that were within 1-2 WRC of their pre-test scores.

Several other previous studies on summer regression and recoupment focused on students with SLD. Cook and Schwartz (1969) used the Wide Range Achievement Test (WRAT) to measure pre- and post-summer reading achievement among students with "mild" disabilities such as SLDs. The students in their study did not show significant regression in reading skills over the summer months.

A later study conducted by Allinder and Eicher (1994) used R-CBM to measure regression and recoupment. Their sample assessed ORF skills of 75 elementary school students identified as having mild disabilities in reading and math. In this case, the authors found that the students' reading skills significantly declined over the summer. Upon returning to school for six weeks, the students not only recouped their previously lost reading skills, they also showed significant improvements compared to their previous scores.

Thus, previous information suggests that similar to low achievers, students with SLD have either been found to regress or experience no change in reading skills during summer months. There also is some extant evidence that a) summer instruction may prevent skill loss, and b) students recoup losses within about one month's return to school.

The current study's small number of participants with SLD had a strong impact on the results. For example, having only 4 out of 136 students in the sample with SLD made it difficult to reach significance

using inferential statistics. However, assuming that the two students whose scores stayed about the same had no summer reading intervention, this finding concurs with Cook and Schwartz's (1969) findings that little regression occurred among students with mild disabilities. The finding that the one student who regressed was able to recoup lost fluency skills after six weeks back in school was consistent with Allinder and Eicher's (1994) finding that students who regressed recouped lost skills. Because the student whose ORF improved over the summer had attended several weeks of reading instruction during that time, this finding also has some commonality with Allinder and Eicher's results.

Sex Differences

A second research question aimed to determine how male and female students compared in terms of ORF regression and recoupment following summer break. It was hypothesized that male students would show more regression and less recoupment than their female peers.

The current study determined that despite females having slightly higher average pre-test and post-test scores compared to their male peers, both sexes showed similar levels of ORF improvement from pre-test to post-test. In addition, no significant differences occurred when comparing boys' and girls' pre-test scores, post-test scores, regression from pre-test to post-test, or recoupment of lost ORF skills among students who regressed.

To date, no prior research has compared males' and females' ORF regression and recoupment. However, some data exist regarding differences in academic achievement between boys and girls. According to the US Department of Education (1998), about two-thirds of all students receiving special education are boys, and boys comprise an even higher majority of students with SLDs. Some studies have claimed that biological differences lead to an over-identification of boys as having disabilities such as SLD

(Tschantz & Markowitz, 2003); while others posit that it is boys' behavior resulting from these biological differences that leads to special education referral bias (Shaywitz, Shaywitz, Fletcher, & Escobar, 1990). However, other studies have found that boys have a true vulnerability to SLD even after eliminating bias (Flannery, Liederman, Daly & Schultz, 2000; Liederman, Kantrowitz, & Flannery, 2005). Nevertheless, one meta-analysis found no differences between boys' and girls' regression of reading skills over the summer (Cooper, Nye, & Charlton, 1996). It also is important to note that none of these previous studies used R-CBM or another instrument designed to assess improvement or decline in reading skills over a short period of time.

The current study's finding that boys had slightly lower overall ORF scores than girls was not found to be statistically significant. This finding regarding regression concurred with Cooper, Nye, and Charlton's (1996) findings that boys and girls did not differ in terms of summer regression in reading skills.

Low Achievement

Another research question this study sought to answer involved whether students with initially low ORF skills would regress more or recoup less compared to their peers with stronger performance. It was hypothesized that students with very deficient skills will be more likely to show regression and lack of recoupment than those scoring closer to average/normal limits on the pre-test.

The current study compared the ORF performance of students with "low achievement," defined as having below-benchmark ORF scores during the pre-test, to those who scored at or above the benchmark defined by the AIMSweb R-CBM measure. Regarding regression, students with benchmark and below average ORF skills showed similar levels of improvement over the summer. A repeated measures analysis of variance (ANOVA-RM) comparing the two groups

determined that differences between students' initial achievement levels (above versus below benchmark) had no significant impact on their level of regression. Students with low versus benchmark achievement differed significantly from each other based on their sex. In particular, girls with benchmark achievement had higher average ORF skills than boys. Students with low achievement showed similar ORF for both sexes.

Few studies have examined how students with below average achievement fared in terms of regression and recoupment of reading skills. However, some prior research has found similarities when comparing generally deficient readers to students with SLDs. For example, both groups had similar development regarding reading and related skills (Fletcher, Shaywitz et al., 1994; Shaywitz, Fletcher, Holahan, & Shaywitz, 1992). In some cases, students with simple low achievement fared worse than their peers with SLD. For example, Shaywitz et al. (1992) found that students with SLD had stronger reading achievement and showed better reading improvement between second and fifth grade compared to low achieving peers without SLD.

Other studies compared students with average versus low achievement in terms of summer regression. Ross (1974) found that average and skilled readers improved during the summer, while less skilled readers regressed. Mraz and Rasinski (2007) also found that lower achieving students' skills declined more than those of their "average" peers. Taken together, these studies assert that initial skill level (average or below average) impacted a student's likelihood of experiencing summer reading regression.

The current study's findings that students with low achievement did not differ significantly from their peers in terms of regression and recoupment were unexpected given previous research. Instead, most students in the current study's sample showed minimal differences or improvement in ORF from spring to fall, regardless of achievement level.

Summer Reading Program Attendance

A fourth research question the current study sought to answer included whether attending a school-sponsored SRP impacted ORF regression and recoupment. It was hypothesized that attending a SRP would lead to students either maintaining (no change) or improving their ORF scores from pre-test to post-test.

Students' pre-test and post-test scores were used to compare how students who attended the school's summer reading program fared relative to their peers. Students attending the SRP had lower pre-test scores, yet both groups showed similar levels of improvement in ORF. Despite gains shown among many of the SRP attendees, this factor failed to achieve statistical significance.

In the current study, students who attended the school's SRP started with lower pre-test ORF scores compared to the rest of the sample. This difference was found to be significant using an ANOVA-RM. Students who attended the SRP did not differ from the rest of the students in terms of regression. However, a significant effect was found when comparing groups for the interaction between sex and SRP attendance. More specifically, girls who attended the SRP had similar ORF to boys who attended. On the other hand, among students who did not attend the school's SRP, girls had higher ORF scores than boys.

Most students in the total sample improved or maintained their ORF scores from pre-test to post-test. Thus, it is possible that the summer program improved ORF or prevented loss among students who might have otherwise regressed. This may have led to the lack of statistical significance when comparing regression of students who attended the summer program to those who did not.

Several studies have documented that SRPs have effectively prevented reading skill loss. Cornelius and Semmel (1982) determined that access to

a five-week summer program prevented students with SLDs from declining in reading skills during the summer. When looking at a broader population in a meta-analysis, Lauer, Akiba, Wilkerson, Apthorp, Snow, and Martin-Glenn (2006) found that academic programs taking place during "out-of-school-time," whether after school or during the summer, led to significant improvements in academic skills. Other studies found that even programs combining academic skills and general youth development led to improved academic skills (Chaplin & Capizzano, 2006; Miller, 2007). In addition to summer programs, Kim (2006) found that providing at-risk and dysfluent readers with books for independent summer reading led to significant skill gains.

It is likely that at least some of the sample's students who did not attend the school's SRP had some access to reading practice, youth development programs, or reading instruction over the summer. Given what the literature suggests about summer activities and skill gains, it is likely that several students in the sample received summer reading practice that would improve their ORF skills. The exclusion of these students from the summer reading intervention subset may have led to the lack of significant findings in the current analysis. However, given that no students in the sample could be considered "low income" based on free/reduced lunch eligibility, it is likely that they had access to somewhat "enriched" environments that promoted literacy (Hart & Risley, 1995; Kim, 2006).

Several issues likely contributed to the lack of significance when analyzing the improvement of students who had attended the school's SRP. First, parent survey information regarding other students' summer reading activities was unavailable. Second, students participating in the SRP likely began with below-average reading achievement. In fact, 12 of the 17 SRP attendees had below-benchmark pre-test scores. Finally, a small

proportion of the sample attended the summer program. Together, these issues likely impacted the lack of statistical significance found for this factor.

Developmental Status (Age and Grade Level)

A fifth research question this study sought to answer regarded whether developmental status (age and grade level) had any impact on ORF regression and recoupment. It was hypothesized that ORF would increase with grade level; however, no significant differences were anticipated in terms of regression and recoupment for the age groups assessed in the current study.

As expected, students increased their ORF scores as they advanced in grade level, with older students reading more words per minute than students in lower grades. Such improvements in oral reading fluency were expected based on developmental trends by age as well as those noted in years of archival data collected by databases such as AIMSweb (NCS Pearson, Inc., 2008). Despite older students scoring significantly higher on isolated pretest and post-test measures, grade level had no impact upon reading fluency regression based on an ANOVA-RM. Students showed similar levels of ORF change from pre-test to post-test regardless of grade level, with no statistical significance when comparing the three grades.

These findings correspond with other research that demonstrates that as children mature, they often improve in developmental areas such reading and language skills. Research literature as well as data collected by AIMSweb, a large database with student R-CBM scores, confirms this statement. Reports from multiple years of data collection show improvements in ORF as students' grade levels increased (Howe & Shinn, 2002).

When comparing spring ORF to fall ORF, AIMSweb's aggregate norm tables demonstrate similar levels of regression at each grade level.

However, in contrast to the current study, students included in the AIMSweb sample were assessed at their current grade level. Thus, the norm tables show non-equivalent probes across spring and fall, with the fall probes increasing in difficulty each year. This may suggest that what appears to be regression may be at least in part a byproduct of increasingly difficult probes, whereas the current study used equivalent probes across the two different school years.

The use of equivalent probes may have led to some difficulties in determining a significant relationship between the analyzed factors and R-CBM before and after summer. The finding that students' ORF improved from pre-test to post-test was unusual given data from large samples of students indicating regression in ORF or reading skills over breaks in academic instruction. For example, as mentioned earlier, AIMSweb benchmark norm tables each grade level indicated similar levels of regression for each grade from spring to fall (NCS Pearson, Inc., 2008). Unlike the current study, AIMSweb's benchmark probes assess students at their current grade level. For example, at the end of 3rd grade, students read 3rd grade probes. Then, at the beginning of 4th grade, students read the more difficult 4th grade probes. Thus, students read increasingly difficult probes each school year, making pre-test to post-test comparisons based on different levels of difficulty. On the other hand, the current study's use of equivalent probes facilitated a comparison of students' ORF with probes of equal difficulty from pre-test to post-test. This serves as one possible explanation of the current study's observed improvements seen from pre-test to post-test.

In addition to R-CBM normative data, one research study compared ORF among students of different age groups with SLD. Tressoldi, Lorusso, and Brenbati (2008) found that when comparing older and younger students, 3rd and 4th grade students with SLD made similar gains in ORF and reading

accuracy compared to those in sixth through eighth grade. Such improvements were found to be proportional to the ORF at each grade level. The authors concluded that it was "never too late" to intervene in terms of offering reading skill remediation programs to students in upper grades.

Recoupment

A final research question the current study sought to answer was whether students whose ORF regressed from spring to fall (pre-test to post-test) would recoup their lost skills upon returning to school for one month. It was hypothesized that students who regressed would regain lost ORF skills after four weeks back in school.

In the current study, only a minority of students (14 out of 136) regressed, showing ORF decrements of 10% more from pre-test to post-test. Of this group, 13 out of 14 students improved after one month back in school. Students' follow-up scores were significantly higher than their post-test scores. All except one of the students showed improvements from post-test to follow-up. Many of them showed relatively similar levels of improvement, with a range of 5-30 WRC higher on the follow-up, with an average improvement of 20.5 WRC.

These results are in agreement with Allinder and Eicher (1994) who found significant ORF regression among the 75 elementary school students with SLDs in their sample. However, upon returning to school for six weeks, the students recouped their previously lost reading skills. In addition, their ORF performance improved significantly compared to their previous scores.

The current study's finding that students who regressed in ORF during the summer recouped their losses was comparable to Allinder and Eicher's conclusion regarding students with SLDs. Having a larger sample of students who regressed, those with SLDs, and those who received a SRP

likely would have helped to increase certainty that a variety of regressors are able to recoup lost ORF skills.

Limitations

Several factors limited the current study's ability to determine a relationship between ORF change over the summer and each of the factors analyzed. These factors included the use of a convenience sample, the current sample's demographics, the lack of information regarding students' access to summer activities promoting reading achievement, and timing of the assessment. Each of these factors likely influenced the study's results.

The use of a convenience sample strongly influenced this study's results and its lack of generalizability to other samples. The sample used in this study differed from the general population in several ways. First of all, only four of the 136 of the sample's participants were identified as having SLD. Having such a small group of students with SLD posed a limitation for achieving significance in a statistical analysis. In addition, 2 of the 4 students with SLD attended a summer reading intervention program, making it difficult to ascertain how a larger group of students with SLD would fare given no reading instruction.

The current sample also differed from the general population in terms of access to school-based support to promote academic achievement. The participating school district used a partial response to intervention (RTI) approach to support students in general education and identify potential learning disabilities. This method aims to ensure that students receive appropriate instruction and interventions in the general education setting before being considered eligible for requiring special education services. Using this method of intervening and monitoring progress to determine intervention efficacy may prevent student academic failure and referral for special education evaluation (Reschly, 2008). In addition,

the early intervention method employed when using the RTI approach may "prevent" some students from requiring SLD services. While only 2.9% of the sample had been identified as having SLD, a national average of 5.9% of children aged 6-21 were identified as part of this group in 2003-2004 (U.S. Department of Education, National Center for Education Statistics, 2006). Access to school-based academic interventions may have led to a lower incidence of SLD as well as overall higher achievement observed in the current sample.

In addition to school-based academic support, children participating in the study also likely benefited from more family-based support to promote academic achievement. Community and school data showed that the participating students did not qualify as "low income" based on family income and free/reduced lunch status. Previous research has demonstrated that students with higher income levels had more exposure to enriching out-of-school activities, benefiting their academic skills (Miller, 2007). School data also indicated that parental involvement was high in the school district from which the sample was drawn (Illinois Interactive Report Card, 2008). This level of participation was high compared to other children in the state, and presumably as well as general population. This combination of academic support from both parents and school programs may have led to higher achievement. Furthermore, pre-test scores showed that the majority of the sample (68%) of the students assessed had reading scores that met benchmark expectations for their grade. In other words, more than two-thirds of the school showed a baseline level of average or better academic achievement in terms of reading fluency.

In addition to the use of a convenience sample, the timing of data collection was another factor that impacted the study's results and generalizability to other populations. When looking at summer reading regression, data collection would ideally have taken place as close to the

academic break as possible. Such a scenario would have prevented any interim academic instruction from impacting students' pre and post-test scores. Due to school events and benchmarking procedures, the current study's data were collected two weeks prior to the summer break. In addition, due to similar circumstances, post-test data were collected almost two weeks after the first day of school. Thus, some instructional time was included in the gap between pre and post-test data, which ideally would have had no undocumented reading instruction taking place. Knowing that as a group, students who regressed improved their reading performance four weeks after the post-test, having some instruction during that time could have compromised the level of regression observed in the sample.

One factor missing in the current study's analysis regards the amount of reading instruction or practice students had over the summer. Unfortunately, information regarding students' independent reading or program attendance over the summer was unavailable. The researcher attempted to send home a parent survey asking about students' reading practice during the summer. The survey asked a) approximately how many hours per week the participating children spent reading, and b) whether the students attended an academically focused reading program during the summer. Unfortunately, these data could not be obtained due to lack of parent participation in the surveys. Thus, the only information about students' summer reading practice or participation in intervention programs came from attendance and registration records provided by the school. Not having these data led to a lack of information that may have explained skill improvement or maturation in students given access to practice or instruction in reading during the break.

Implications for Research

Given the current study's limitations, further study and follow-up research is recommended. For example, using a more representative sample, selecting a sample with students who have SLD, having repeated measures closer to the beginning and end of the school year, using alternate scoring procedures, and using data about exposure to reading instruction or practice during the summer may have led to different results. Thus, implementing some of these changes may assist in more thoroughly and accurately answering the question the study addressed. A more representative sample would be warranted. The finding that most students improved slightly or did not change in ORF from spring to fall may be atypical to the general population. Assessing students from different school districts, with different family income levels, and living in different types of areas (urban, rural, etc.) may help determine whether the current study's results describe the population of interest. The inclusion of students with low-income or free/reduced lunch status would provide another important factor that the current study lacked. Finding a more diverse sample would provide a further comparison to determine factors that may lead some groups of students to regress.

In addition to finding a more representative sample, future research could supplement and improve on the current study by assessing more students with SLD. While the current study had some students with low achievement, the small number of students with SLD made it difficult to compare the two groups. In addition, having so few students with SLD made it difficult to compare students with SLD to those in general education. Longitudinally looking back at the performance of students identified later may find more students with SLD in the current or other samples. Another direction for further research would be to study only students with SLD from a variety of schools or settings in a larger sample.

Next, follow-up research that assesses students' ORF closer to the end of the school year and beginning of the following school year may have shown more regression than that found in the current study. When combining time elapsed between assessments, the current study had about 3-4 weeks of school in addition to the summer break between the pre-test and post-test. Thus, some instructional time was included during this time. Given that students who regressed regained their lost ORF skills within four weeks, it is possible that others regained lost skills during the first two weeks of school. Further research should be done with assessment dates as close as possible to the beginning and end of the school year. For example, doing an assessment the day after students return to school, followed by repeated measures two and four weeks later would aid in determining the extent to which instruction helps students recoup ORF skills.

In addition, it is recommended that further study considering regression and recoupment skills use identical probes for pre-test, post-test, and follow-up measures. Current benchmarking practices often repeat administrations of identical probes when collecting benchmarking data in the fall, winter, and spring. These measurements occur about three months apart, as did the pre-test and post-test measures. This suggests that this length of time is appropriate for conducting identical-probe measurements. Using different probes for pre-test and post-test measurements as students advance in grade level also leads to a difference in difficulty levels between measurements. Using identical probes prevents possible distortions in data arising from differences in the measurement tools. Thus, it is recommended that further researchers use the same grade-level probes in a repeated-measures design as used in the current study.

Finally, further research should aim to determine the amount of reading exposure, intervention, or practice students have over the summer. Collecting parent survey data, as attempted in the current study, may

explain differences between students' ORF change over the summer. Also, documenting which students received summer intervention in the form of summer school, enrichment programs, etc. likely would yield valuable information regarding reading exposure during students' time out of school. For example, asking parents about how much reading practice students had during the summer, access to books, and other academically enriching experiences likely would provide valuable information regarding strategies to promote achievement. Replicating studies such as Kim's (2006) analysis of the effects of giving low-income children access to books over the summer using CBM measures also would assess the benefits of that practice. Finally, using pre-test and post-test measures to evaluate the academic effects of summer camps or programs also would provide useful information about their indirect benefits (as well as the general nature of regression).

Overall, further study of regression and recoupment would supplement the current study's findings. Implications include more precise information about how much individual and groups of students with disabilities regress, aiding in ESY determinations.

Implications for the Practice of School Psychology

The current study's results have several implications for the practice of school psychology. Given the focus on data-based decision making and accountability and participation on school teams determining need for extended school year (ESY) services, school psychologists may consider using R-CBM repeated measures as a tool to aid in determining eligibility for such programs. Given the finding that some students in general education regressed, collecting regression and recoupment data for at-risk students may assist in determining whether reading interventions are recommended for those students. School psychologists working in school districts that already use R-CBM as a progress-monitoring or benchmarking

tool might recommend collecting data for individual or small groups of students to assist in such decision making.

Regarding ESY, school psychologists might recommend pre-test and post-test assessments before and after summer (and other breaks from school) for all students with disability to aid in making ESY decisions. Such data also might help schools pinpoint which students might benefit from having access to books or reading programs over the summer. Gathering further normative data in order to compare students with disabilities to a local norm via benchmarking also likely would benefit individual schools.

According to the National Association of School Psychologists (1997, p. 7), data-based decision-making and accountability should "permeate every aspect of the practice" of school psychology. Collecting regression and recoupment data for all students may provide benefits to the profession beyond knowing more about ESY eligibility for students with disabilities. For example, collecting and using R-CBM data also has the potential to inform school psychologists about normative student performance. Knowing more about "normative" performance of students without disabilities helps the profession to put disabilities into perspective. Findings that indicate similarities between students with SLDs and those in the general population may allow school psychologists as a profession to more accurately aid students with SLDs and determine which other students may benefit from similar intervention.

In addition to using data to understand individual students and those with disabilities, school psychologists also have expertise in evaluating programs for large groups of students. Having access to data about the general nature of regression and recoupment among general education students would allow school psychologists to evaluate summer programs. Such information would help schools to assess the benefits of such programs, and whether they should continue. Collecting

regression/recoupment data might also inform decisions such as whether schools should engage in "year round" schooling. Given some earlier studies' findings that differences between income groups strongly impact academic skill regression, shorter breaks may benefit some (or all) students. Having shorter breaks might reduce the need for review to recoup losses seen in students who regress. Similarly, given that "year round" schools often have the same number of days in school with more evenly distributed breaks or "inter-sessions", such breaks may serve as opportunities for further academic intervention.

Knowing about the nature of regression and recoupment for general education students would benefit the practice of school psychology. First, such students provide a normative comparison for students with disabilities. Second, such data may inform larger scale policies such as the adoption of a year-round calendar or specific summer interventions programs. These data would contribute to the field of school psychology, and by extension, the field of education as a whole.

Summary

The current study aimed to increase the understanding of ORF regression and recoupment for students in general education as well as those with SLDs. The use of repeated measures using identical probes was unprecedented, leading to lack of concurrence with norms using increasingly difficult probes. In addition, the current study's use of R-CBM with a group of general and special education students distinguished it from earlier studies of summer regression. Instead of showing regression, as some previous studies have found, the current study showed that students showed only slight differences in ORF from pre-test to post-test.

Few factors differentiated among students' ORF changes from pre-test to post-test. For example, different grade levels, sexes, students who

attended a SRP, and those with initial low achievement were not found to differ significantly in terms of regression. Too few students with SLDs participated to analyze the extent to which they experienced a different level of regression than their peers.

The convenience sample used in the current study had too few students with SLDs to make a comparison with general education peers. In addition, participants in the sample were not receiving free or reduced lunch, instead coming from "middle income" families. Given that income levels have been found to be important in terms of reading practice out of school, having a more economically diverse sample may have led to differences in regression and recoupment. Additionally, knowing how much summer reading practice the students in the sample had would have helped to determine differences among ORF regression based on summer reading exposure. Although few students regressed significantly, data indicated that those who did recouped their lost skills within a month.

The current study yielded information about regression and recoupment for a convenience sample. Due to the lack of representativeness in the sample, the results of the current study may not be generalized to other samples. In addition, the finding that few students regressed in ORF may have arisen due to characteristics of the school or students' participation in summer reading activities. Thus, it is recommended that further studies use a more economically diverse sample. In addition, soliciting information regarding the degree of summer reading practice students have would help to determine other factors leading to regression. Finally, including more students with SLDs in the sample would have led to a stronger analysis to determine whether this group differs significantly from the general population.

Following the aforementioned recommendations for further study would aid school psychologists in making sound data-based decisions. These data

could be used for students in special education, both as individuals and as a normative comparison. Collecting ORF data for regression and recoupment also would aid in evaluating programs and assessing need for a large-scale changes such as a longer school year for all students. Overall, collecting similar data would help school psychologists and the field of education by providing quality data to assist in educational decision-making for individuals and groups of students.

References

- Ahearn, E.M. (2000). *Extended school year- state regulations and policies. Quick turn around project forum*. Alexandria, VA: National Association of State Directors of Special Education.
- AIMSweb (2002). *Reading curriculum-based measurement (R-CBM) standard oral reading fluency assessment passages*. Retrieved June 2007 from Edformation, Inc., <https://AIMSweb.edformation.com>.
- AIMSweb (2008). *R-CBM Oral Reading Fluency Aggregate Norm Table*. Retrieved July 2008 from Edformation, Inc., <https://AIMSweb.edformation.com>.
- Alamo Heights ISD v. Texas State Board of Education 790 F.2d 1153, 1158 (5th Cir. 1986).
- Alexander, K.L., Entwisle, D.R., & Olson, L.S. (2001). Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation and Policy Analysis, 23*(2), 171-91.
- Allinder, R.M., & Eicher, D.D. (1994). Bouncing back: Regression and recoupment among students with mild disabilities following summer break. *Special Services in the Schools, 8*(2), 129-142.
- Allington, R.L. (2006). *What really matters for struggling readers: Designing research-based programs* (2nd ed.) Boston: Allyn & Bacon.
- Anderson, R.L., Wilson, P., & Fielding, L. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly, 23*, 285-303.
- Arizona State Department of Education (1993). *Extended school year. AZ-TAS themes & issues: A series of topical papers on special education*. Phoenix Division of Special Education: Author.
- Armstrong v. Kline, 476 F. Supp 583 (E.D. Pa 1979).
- Battle v. Commonwealth of Pennsylvania, 629 F.2d 269 (3rd Cir., 1980).
- Board of Education v. Rowley, 458 U.S. 176 (2nd Cir. 1982).
- Chaplin, D., & Capizzano, J. (2006). *Impacts of a summer learning program: A random assignment study of building educated leaders for life (BELL)*. Greensboro, NC: Smith Richardson Foundation, Inc.
- Christ, T.J., & Silberglitt, B. (2007). Estimates of the standard error of measurement for curriculum-based measures of oral reading fluency. *School Psychology Review, 36*(1), 130-146.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Erlbaum.

- Cook, J.J., & Schwarz, R.H. (1969). A look at some assumptions in special education. *Educational Training of the Mentally Retarded*, 4(4) 149-57.
- Cooper, H., Nye, B., & Charlton, K. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, 66(3), 227-68.
- Cordrey v. Eukert, 917 F.2d. 1460 (6th Cir. 1990).
- Cornelius, P.L., & Semmel, M.I. (1982). Effects of summer instruction on reading achievement regression of learning disabled students. *Journal of Learning Disabilities*, 15(7), 409-413.
- Crawford v. Pittman, 708 F.2d. 1028, 1037 (5th Cir. 1983).
- Dale, E., & Chall, J.S. (1948). A formula for predicting readability. *Educational Research Bulletin*, 27, 37-53.
- Deno, S. L. (1985). Curriculum-Based Measurement: The emerging alternative. *Exceptional Children*, 52, 219-232.
- Deno, S. L., & Fuchs, L. S. (1987). Developing curriculum-based measurement systems for data-based special education problem solving. *Focus on Exceptional Children*, 19(8), 1-15.
- Edgar, E., Spence, W.M., & Kenowitz, L.A. (1977). Extended school year for the handicapped: Is it working? *Journal of Special Education*, 11(4), 441-447.
- Education for All Handicapped Children Act of 1975 (P.L. 94-142), 20 U.S.C. sec. 1401.
- Espin, C.A., & Tindal, G. (1998). Curriculum-Based Measurement for secondary students. In M.R. Shinn (Ed.), *Advanced applications of Curriculum- Based Measurement* (pp. 214-253). New York: Guilford.
- Fewster, S., & MacMillan, P.D. (2002). School-based evidence for the Validity of curriculum-based measurement of reading and writing. *Remedial and Special Education*, 23(3), 149-156.
- Flannery, K., Liederman, J., Daly, L., & Schultz, J. (2000). Male prevalence for reading disability is found in large sample of black and white children free from ascertainment bias. *Journal of the International Neuropsychological Society*, 6, 433-442.
- Fletcher, J.M., Shaywitz, S.E., Shankwater, D.P., Katz, L., Liberman, I.Y., Stuebing, K.K., Francis, D.J., Fowler, A.E., & Shaywitz, B.A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement profiles. *Journal of Educational Psychology*, 86(1), 6-23.
- Florida Department of Education (2002). *Determining an Individual Student's Need for Extended School Year (ESY) Services: Technical Assistance Paper*. Jacksonville, FL: Author.

- Freeman, C.E. (2005). *Trends in educational equity of girls & women: 2004*. Washington DC: U.S. Department of Education, National Center for Education Statistics.
- Fry, E. (1968). A readability formula that saves time. *Journal of Reading*, 11(7), 265-71.
- Fuchs, L.S., & Fuchs, D. (1992). Identifying a measure for monitoring student reading progress. *School Psychology Review*, 21(1), 45-58.
- Fuchs, D., Fuchs, L.S., Mathes, P.G., Lipsey, M.W., Roberts, P.H. (2001) *Is "learning disabilities" just a fancy term for low achievement? A meta-analysis of reading differences between low achievers with and without the label: Executive summary*. Paper presented at the Learning Disabilities Summit: Building a Foundation for the Future.
- Fuchs, L.S, Fuchs, D., & Deno, S.L. (1983). *The nature of inaccuracy among readability formulas: Technical report*. Washington, DC: Office of Special Education and Rehabilitative Services.
- Georgia Association for Retarded Citizens v. McDaniel, 716 F. 2d 1565 (11th Cir. 1983).
- Ginsburg, A. (1981). *Summer Learning and the Effects of Schooling: A Replication of Heyns*. Los Angeles, CA: Paper presented at the 65th Annual Meeting of the American Educational Research Association.
- Good, R.H. III, & Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. In M.R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp.61-88). New York: Guilford Press.
- Gottlieb, J., Alter, M., Gottlieb, B.W., & Wishner, P. (1994). Special education in urban America: It's not justifiable for many. *The Journal of Special Education*, 27(4), 453-465.
- Hayes, D.P., & Grether, J. (1983). The school year and vacations: When do students learn? *Cornell Journal of Social Relations*, 17, 56-71.
- Hewett, F.M., & Forness, S.R. (1974). *Education of Exceptional Learners*. Boston: Allyn & Bacon.
- Heyns, B. (1978). *Summer Learning and the Effects of Schooling*. New York: Academic Press.
- Ho, A. (2004). To be labelled or not to be labeled that is the question. *British Journal of Learning Disabilities*, 32(2), 86-92.
- Hoskyn, M., & Swanson, H.L. (2000). Cognitive processing of low achievers and children with reading disabilities: A selective meta-analytic review of the published literature. *School Psychology Review*, 29(1), 102-119.
- Howe, K.B., & Shin, M.M. (2002). *Standard reading assessment passages (RAPs) for use in general outcome measurement: A manual for describing development and technical features*. Retrieved September 2008 from <https://AIMSweb.edformation.com>.

- Illinois Detailed Profile* (2008). Retrieved from <http://www.city-data.com/html> on September 29, 2008.
- Illinois Interactive Report Card* (2008). School data retrieved from <http://iirc.niu.edu/> on April 25, 2008.
- Individuals with Disabilities Education Act. (1990). 20 United States Code Annotated, Section 1400. Washington, DC: U.S. Congress.
- Individuals with Disabilities Education Improvement Act (2004). Retrieved February 13, 2008 from www.ed.gov/about/offices/list/osep/index.html
- Katsiyannis, A. (1991). Extended school year policies: An established necessity. *Remedial and Special Education, 12*, 24-28.
- Kim, J.S. (2006). Effects of a voluntary summer reading intervention on reading achievement: Results from a randomized field trial. *Educational Evaluation and Policy Analysis, 28*(4), 335-355.
- Lauer, P.A., Akiba, M., Wilkerson, S.B., Apthorp, H.S., Snow, D., Martin-Glenn, M.L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research, 76*(2) 275-313.
- Lichtenstein, R. (2008). Best practices in identification of children with learning disabilities. *Best Practices in School Psychology V*. Bethesda, MD: National Association of School Psychologists.
- Lichtenstein, R., & Klotz, M.B. (2007). Deciphering the federal regulations on identifying children with specific learning disabilities. *NASP Communiqué, 36*(3), 13-16.
- Liederman, J., Kantrowitz, L., Flannery, K. (2005). Male vulnerability to reading disability is not likely to be a myth: A call for new data. *Journal of Learning Disabilities, 38*(2), 109-129.
- MacMillan, D.L., Gresham, F.M., & Bocian, K.M. (1998). Discrepancy between definitions of learning disabilities and school practices: An empirical investigation. *Journal of Learning Disabilities, 31*(4), 314-326.
- Marston, D.B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why to do it. In M.R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18-78). New York: Guilford Press.
- McCoach, D.B., O'Connell, A.A., Reis, S.M., & Levitt, H.A. (2006). Growing readers: A hierarchical linear model of children's reading growth during the first 2 years of school. *Journal of Educational Psychology, 98*(1), 14-28.
- McLeskey, J., & Waldron, N.L. (1991). Identifying students with learning disabilities: The effect of implementing statewide guidelines. *Journal of Learning Disabilities, 24*(8), 501-506.

- Miller, B.M. (2007). *The Learning Season: The Untapped Power of Summer to Advance Student Achievement*. Quincy, MA: The Nellie Mae Education Foundation.
- Microsoft Corporation (2009). *Test Your Document's Readability*. Retrieved January 2009 from <http://office.microsoft.com/>.
- Mirkin, P. (1982). *Direct and repeated measurement of academic skills: An alternative to traditional screening, referral, and identification of learning disabled students*. Washington, DC: Office of Special Education and Rehabilitative Services.
- Mraz, M., & Rasinski, T.V. (2007). Summer reading loss. *Reading Teacher*, 60(8) 784-789.
- National Association of School Psychologists (1997). *School Psychology: A Blueprint for Training and Practice II*. Bethesda, MD: Author.
- National Association of School Psychologists (2006). *School Psychology: A Blueprint for Training and Practice III*. Bethesda, MD: Author.
- National Center for Education Statistics (2007). *Dropout Rates in the United States: 2005*.
- National Center for Learning Disabilities (2008). *Talking about LD*. Retrieved from <http://www.nclld.org/content/view/256/> on 2-27-08.
- NCS Pearson, Inc. (2008). *AIMSweb® Growth Table: Reading Curriculum Based Measurement Multi-Year Aggregate*. Retrieved on November 14, 2008 from <https://AIMSweb.edformation.com>.
- Oakland, T., & Lane, H. B. (2004). Language, reading, and readability formulas: Implications for developing and adapting. *International Journal of Testing*, 4(3), 239-252.
- Olmi, D.J., Walker, D.W., & Ruthven, A.J., (1995). Extended school year services: Prediction, description, and impact of judicial precedence. *Journal of Special Education*, 29(1), 72-83.
- O'Malley, K.J., Francis, D.J., Foorman, B.R., Fletcher, J.M., & Swank, P.R. (2002). Growth in precursor and reading-related skills: Do low-achieving and IQ-discrepant readers develop differently? *Learning Disabilities: Research & Practice*, 17(1), 19-34.
- Prasse, D.P. (2002). Best practices in legal issues. In A. Thomas & J. Grimes (Eds.). *Best practices in school psychology IV*. Washington DC: National Association of School Psychologists.
- Publication manual of the American Psychological Association (5th ed.)*. (2001). Washington, DC: American Psychological Association.
- Raffaele Mendez, L.M., Mihalic, S.T., & Hardesty, R. (2006). Gender differences in academic development and performance. In G. Bear & K. Minke (Eds.). *Children's Needs III: Development, Prevention, and Intervention*. Bethesda, MD: National Association of School Psychologists.

- Readability Calculations* (1999). Micro Power & Light Co. Dallas: TX.
- Reschly, D.J. (2008). School psychology paradigm shift and beyond. *Best Practices in School Psychology V*. Bethesda, MD: National Association Of School Psychologists.
- Reusch v. Fountain, 872 F. Supp. 1421 (D.Md. 1994).
- Ross, P.A. (1974). Regression study- now the long hot summer. *Reading Teacher*, 28(1), 28-30.
- Shaywitz, B.A., Fletcher, J.M., Holahan, J.M., & Shaywitz, S.E. (1992). Discrepancy compared to low achievement definitions of reading disability: Results from the Connecticut longitudinal study. *Journal of Learning Disabilities*, 25(10) 639-648.
- Shaywitz, S.E., Shaywitz, B.A., Fletcher, J.M., & Escobar, M. D. (1990). Prevalence of reading disability in boys and girls. Results of the Connecticut Longitudinal Study. *Journal of the American Medical Association*, 264, 998-1002.
- Shepard, L.A., & Smith, M.L. An evaluation of the identification of learning disabled students in Colorado. *Learning Disability Quarterly*, 6(2) 115-127.
- Shinn, M.M., & Shinn, M.R. (2002) *AIMSweb® Training Workbook: Administration and Scoring of Reading Curriculum-Based Measurement (R-CBM) for Use in General Outcome Measurement*. Retrieved from <http://www.AIMSweb.com/support-training/training/training-materials/> on September 29, 2008.
- Shinn, M.R. (2002). *AIMSweb™ Training Workbook: Organizing and implementing a benchmark assessment program*. Eden Prairie, MN: Edformation. Retrieved from www.edformation.com on 7-18-06.
- Siegel, S., & Sleeter, C.E. (1991). Transformingt: Next stages for the school-to-work transition movement. *Career Development for Exceptional Individuals*, 14(1), 27-41.
- Spache, G. (1953). A new readability formula for primary-grade reading materials. *The Elementary School Journal*, 55, 410-413.
- Spear-Swerling, L., & Sternberg, R.J. (1998). Curing our "epidemic" of learning disabilities. *Phi Delta Kappan*, 79(5), 397-401.
- Tressoldi, P.E., Lorusso, M.L., & Brenbati, F. (2008). Fluency remediation in dyslexic children: Does age make a difference? *Dyslexia*, 14(2), 142-152.
- Tschantz, J., & Markowitz, J. (2003). *Gender and special education: current state data collection. Quick turn around (QTA)*. Alexandria, VA: National Association of State Directors of Special Education.
- Thesis/dissertation manual* (4th ed.). Indiana, PA: Graduate School, Indiana University of Pennsylvania.
- U.S. Department of Education (1998). *Twentieth annual report to Congress*

- on the implementation of the Individuals with Disabilities Education Act.* Washington, DC: Author.
- U.S. Department of Education, Office of Special Education and Rehabilitative Services (2006). *Twenty-sixth annual report to Congress on the implementation of the Individuals with Disabilities Education Act.* Washington, DC: Author.
- U.S. Department of Education, Office of Special Education and Rehabilitative Services, Office of Special Education Programs (2007). *Twenty-seventh annual (2005) report to Congress on the implementation of the Individuals with Disabilities Education Act, Vol. 1.* Washington, D.C.: Author.
- U.S. Department of Education, National Center for Education Statistics (2006). *Digest of Education Statistics, 2005.* Washington, D.C.: Author.
- Whitehouse, E.M. (2006). Poverty. In G. Bear & K. Minke (Eds.). *Children's Needs III: Development, Prevention, and Intervention.* Bethesda, MD: National Association of School Psychologists.
- Wright, J. (2001). OKAPI Readability Analysis [On-line Software]. Available from <http://www.interventioncentral.org/htmdocs/tools/okapi/okapi.php>
- Woodcock, R. W., & Johnson, M. (1977). Woodcock-Johnson Psychoeducational Battery. Allen, TX: DLM.
- Yaris v. Special School District of St. Louis County, 728 F. 2d. 1055 (8th Cir., 1984).
- Ysseldyke, J., Algozzine, B., Richey, L., & Graden, J. (1982). Declaring students eligible for learning disability service: Why bother with the data? *Learning Disability Quarterly*, 5, 37-44.