

THE EFFECTIVENESS OF DIBELS ORAL READING FLUENCY  
AS A PREDICTOR OF READING COMPREHENSION  
FOR HIGH- AND LOW-INCOME STUDENTS

Except where reference is made to the work of others, the work described in this dissertation is my own or was done in collaboration with my advisory committee. This dissertation does not include proprietary or classified information.

---

Timon M. Paleologos

Certificate of Approval:

---

Susan K. Villaume  
Professor  
Curriculum and Teaching

---

Edna G. Brabham, Chair  
Associate Professor  
Curriculum and Teaching

---

Bruce A. Murray  
Associate Professor  
Curriculum and Teaching

---

Margaret E. Ross  
Associate Professor  
Educational Foundations,  
Leadership, and Technology

---

Stephen L. McFarland  
Dean  
Graduate School

THE EFFECTIVENESS OF DIBELS ORAL READING FLUENCY  
AS A PREDICTOR OF READING COMPREHENSION  
FOR HIGH- AND LOW-INCOME STUDENTS

Timon M. Paleologos

A Dissertation

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirement for the

Degree of

Doctor of Philosophy

Auburn, Alabama  
August 8, 2005

THE EFFECTIVENESS OF DIBELS ORAL READING FLUENCY  
AS A PREDICTOR OF READING COMPREHENSION  
FOR HIGH- AND LOW-INCOME STUDENTS

Timon Mark Paleologos

Permission is granted to Auburn University to make copies of this dissertation at its discretion, upon request of individuals or institutions and at their expense. The author reserves all publication rights.

---

Timon M. Paleologos

---

August 8, 2005

## VITA

Timon M. Paleologos, son of Nicholas C. Paleologos and Jacqueline M. Paleologos, was born April 16, 1968, in Saratoga Springs, New York. He graduated from Lassiter High School in 1986. He attended Auburn University in Auburn, Alabama, for four years and graduated with a Bachelor of Science degree in Business Management in December, 1990. He earned the degree of Master of Education in Special Education/Learning Disabilities from Auburn University in 1997 and certification as a reading specialist from Auburn University in 2004. He served as a teacher and reading coordinator at Loachapoka Elementary School and as a graduate teaching/research assistant at Auburn University. He is married to Tracy D. Paleologos, and he has one son, Calen Seth Paleologos.

DISSERTATION ABSTRACT  
THE EFFECTIVENESS OF DIBELS ORAL READING FLUENCY  
AS A PREDICTOR OF READING COMPREHENSION  
FOR HIGH- AND LOW-INCOME STUDENTS

Timon Mark Paleologos

Doctor of Education, August 8, 2005  
(M.Ed., Auburn University, 1997)  
(B.S., Auburn University, 1990)

85 Typed Pages

Directed by Edna G. Brabham

The purpose of this two-part, non-experimental study was to examine the predictive nature of reading fluency scores in relation to reading comprehension scores and determine if the associations between these scores were similar for high- and low-income children. In part one of this study, the researcher attempted to verify if statistically equivalent positive correlations existed between the Dynamic Indicators of Basic Early Literacy Skills – Oral Reading Fluency (DORF) and the Stanford Achievement Test – Tenth Edition (SAT-10), and between the DORF and Degrees of Reading Power (DRP), a secondary measure of reading comprehension. In part two of

this study, the researcher sought to determine whether students with proficient reading fluency skill but different economic backgrounds had statistically equivalent comprehension achievement. In addition, the study was designed to investigate whether or not proficient fluency skill is as strongly associated with low-income students' comprehension achievement as it is for their wealthier peers.

Participants in part one of this study consisted of 129 third-grade students who took the DORF, SAT-10, and DRP assessments during the 2003-2004 academic school year. Data analyses revealed statistically significant positive correlations between DORF and the SAT-10, between the DORF and DRP, and between the Sat-10 and DRP. Furthermore, correlational-comparison procedures showed these relationships did not differ at a statistically significant level between average- to high-income students and students receiving free or reduced lunch.

Participants in part two of this study consisted of 215 third-grade students (112 with proficient fluency skill, 103 without proficient fluency skill) who participated in the DORF and SAT-10 assessments during the 2003-2004 academic school year. Correlational, correlational-comparison, and sequential regression analyses indicated the following for students with proficient reading fluency skill: 1) Reading fluency did not equally predict reading comprehension for higher- and lower-income students; 2) statistically significant differences in reading comprehension, reading fluency, and reading vocabulary achievement existed between high- and low-income students; 3) reading vocabulary equally predicted reading comprehension for students of differing economic backgrounds; and 4) reading fluency did not predict reading comprehension for

low-income students proficient in reading fluency skill beyond what was accounted for by reading vocabulary.

## ACKNOWLEDGEMENTS

The author wishes to thank Dr. Edna G. Brabham for her guidance, flexibility, and enduring dependability. Thanks to Dr. Margaret E. Ross for helping me to think as a statistician and sacrificing her time whenever beckoned. Thanks to Dr. Bruce A. Murray and Dr. Susan K. Villaume for providing multiple perspectives on reading research and especially, for making my first post-secondary teaching experience possible. Thanks to Dr. Craig Darch for providing me with the content knowledge and instructional strategies that not only enabled me to be a successful teacher, but also to develop as a professional educator. Thanks are also due to Dr. Katherine Mitchell, Assistant State Superintendent of Education for Reading, for her interest and support of this project. Thanks, too, go to Shannon Henderson, Connie Buskist, Barbie Norvell Hall, and Jordan Barkley who helped keep my sanity intact throughout the doctoral process and to Scott D. Lubic and Lori Hardee for having the courage to proofread this project.

The writer would especially like to extend his deepest gratitude to his family. To my mother, who is responsible for teaching me right from wrong and who always has my best interests in mind. To my father, who taught me to appreciate people of all backgrounds and refused to let me fail. To my son, Calen, who taught me how to tune out distractions. To my wife, Tracy, for her understanding and support throughout our relationship – I love you and our son with all my heart.



Style Manual or journal used: Publication Manual of the American Psychological Association, Fifth Edition.

Computer software used: Microsoft Office Excel; Microsoft Word 2003; Statistical Package for Social Sciences 11.5 (SPSS 11.5).

TABLE OF CONTENTS

LIST OF TABLES . . . . . xii

I. INTRODUCTION . . . . . 1

    Introduction . . . . . 1

    Statement of Purpose . . . . . 3

    Overview of Related Literature . . . . . 3

    Research Questions . . . . . 5

    Significance of Study . . . . . 5

    Scope of Limitations . . . . . 7

    Assumptions . . . . . 8

    Definition of Terms . . . . . 8

    Organizational Overview . . . . . 10

II. REVIEW OF LITERATURE . . . . . 11

    Introduction . . . . . 11

    Reading Legislation . . . . . 11

    Reading Fluency and Comprehension . . . . . 12

    Vocabulary Knowledge . . . . . 13

    The Simple View of Reading . . . . . 15

    Socioeconomic Status and Reading . . . . . 16

    Reading Assessment . . . . . 21

        High-Stakes Testing . . . . . 21

        Curriculum-Based Measurement . . . . . 22

        DIBELS and the SAT-10 . . . . . 23

    Summary . . . . . 27

III. METHODS . . . . . 30

    Purpose of the Study . . . . . 30

    Research Design . . . . . 32

    Instrumentation . . . . . 33

        DIBELS – Oral Reading Fluency . . . . . 34

        The SAT-10 . . . . . 35

        Degrees of Reading Power . . . . . 37

Participants . . . . .	38
Part 1 . . . . .	38
Part 2 . . . . .	39
Data Collection and Procedures . . . . .	41
Statistical Method . . . . .	42
Part 1 . . . . .	43
Part 2 . . . . .	43
Summary . . . . .	45
 IV. DATA ANALYSIS AND RESULTS . . . . .	 46
Review of Research Questions . . . . .	46
Data Analysis and Results . . . . .	47
Part 1 . . . . .	47
Part 2 . . . . .	50
 V. SUMMARY, DISCUSSION AND RECOMMENDATIONS . . . . .	 57
Summary of the Study . . . . .	57
Discussion of Findings . . . . .	64
Limitations of the Study . . . . .	65
Need for Further Research . . . . .	67
Educational Implications and Recommendations . . . . .	68
 REFERENCES . . . . .	 71
 APPENDIX A . . . . .	 81
 APPENDIX B . . . . .	 84

## LIST OF TABLES

1.	Participants by Location, Gender, Race, and SES Status . . . . .	39
2.	Participants with Proficient Fluency Skill by Income level, Gender, and Race . . . . .	41
3.	Means and Standard Deviations for DORF, DRP, and SAT-10 Scores . . . . .	48
4.	Intercorrelations Among Variables for All Participants . . . . .	49
5.	Intercorrelations Among Variables . . . . .	50
6.	Means and Standard Deviations for Fluency, Comprehension, and Vocabulary Scores of Students with Proficient Fluency Skill . . . . .	51
7.	Intercorrelations Among Fluency and Comprehension for Students with Proficient Fluency Skill . . . . .	52
8.	Intercorrelations Among Vocabulary and Comprehension for Students with Proficient Fluency Skill . . . . .	55

## I. INTRODUCTION

### Introduction

Over the past decade, educational systems at the federal, state, and local levels have increasingly emphasized the importance of early literacy. Federal organizations, including the National Institute for Literacy (NIFL), have produced documents such as *Put Reading First* (2001) in attempts to work with researchers, educators, parents, and policy makers to ascertain the critical components of beginning reading and determine best practices in reading education. Equally important to teaching beginning reading is the accurate assessment of both prerequisite literacy skills and reading comprehension. This study examined relationships between state-mandated assessments that are widely used throughout the United States. Additionally, the relationships among reading fluency, reading vocabulary, and reading comprehension for students of wealthier and poorer backgrounds were investigated.

Most, if not all, state departments of education in the United States require public schools to administer standardized tests of students' progress in reading. Currently the Alabama State Department of Education (ALSDE) evaluates the effectiveness of schools' reading instruction at the third- through eighth-grade levels with the Stanford Achievement Test – Tenth Edition (SAT-10), a high-stakes achievement test that determines a schools' probational status. Additionally, beginning in the 2003-2004 school year, schools were required to utilize the Dynamic Indicators of Basic Early Literacy

Skills (DIBELS) assessment for students in kindergarten through third grade to identify children whose reading achievement was insufficient.

DIBELS measures include tests of pseudoword reading and oral reading speed and accuracy because these components of the reading process are indicators of children's understanding of the alphabetic principle and ability to decode words quickly and correctly, skills that have been identified as necessary components in reading for comprehension (Good, Simmons, & Kame'enui, 2001). The ALSDE chose to publish individual school system's 2003-2004 DIBELS scores, ranking them in order of year-end performance and cross-year improvement. These scores were made public, but not used to identify a specific school's academic status. However, publishing the scores makes DIBELS performance not only an assessment and instructional tool, but also a high-stakes test.

Considering the amount of time and money invested in purchasing and administering the DIBELS assessment in the kindergarten and primary grades, measures need to be taken to determine if DIBELS is a reliable predictor of reading achievement, which in Alabama and many other states is measured by student scores on the SAT-10 starting in the third grade. Research is also needed to see if there are any differences in DIBELS' predictability of SAT-10 reading achievement scores for students of differing socioeconomic backgrounds. Furthermore, studies designed to examine the various components of reading that may be factors related to any differences in the value of DIBELS scores in the primary grades as reliable predictors of reading comprehension on the SAT-10 for students with higher- and lower-socioeconomic status would be of educational value.

## Statement of Purpose

This two-part, non-experimental study examined the predictive nature of oral reading fluency (operationalized by DIBELS Oral Reading Fluency scores) in relation to reading comprehension (operationalized by SAT-10 and Degrees of Reading Power comprehension scores) and the associations between these scores for higher- and lower-income children (determined by free or reduced lunch status). In part one of the study, the researcher attempted to verify if statistically equivalent positive correlations existed between scores on DIBELS Oral Reading Fluency (DORF) tests and reading comprehension scores on the SAT-10, and between DORF and the Degrees of Reading Power (DRP), a secondary measure of reading comprehension. If these correlations are statistically equivalent, then one can assume that DORF is a valid predictor of reading comprehension.

In part two of the study, the researcher sought to determine whether students with proficient reading fluency skill, but different economic backgrounds have statistically equivalent comprehension achievement. If not, then an investigation into the relationships between reading fluency, reading vocabulary (operationalized by SAT-10 vocabulary scores), and reading comprehension for wealthier and poorer students was deemed necessary to examine the role that vocabulary knowledge may play in the reading comprehension exhibited by lower-income children who are fluent readers but not proficient comprehenders.

## Overview of Related Literature

Teachers and administrators in Alabama realize that the “bottom-line” for their third- through eighth-grade schools is the achievement scores students produce on the

SAT-10. Moreover, the ability of students to perform adequately on the SAT-10 is greatly influenced by a child's ability to derive meaning from written text. This ability is quantified on the SAT-10 by a series of questions developed to evaluate reading comprehension. The ALSDE has proactively sought to identify reading difficulties in the kindergarten through third grades by mandating that students be evaluated a minimum of three times per year with DIBELS. The ultimate goal is for all students to achieve 110 correct words per minute (cwpm) on the DORF measures by the Spring of third grade.

Although a previous study by Moscovitch conducted in 2004 showed positive correlations existed between reading fluency as measured by DORF and reading comprehension as measured by the SAT-10, no additional measure of comprehension was utilized to validate the SAT-10's assessment of comprehension achievement and the value of DORF as a reliable predictor of reading comprehension. Additionally, Moscovitch found the correlations between reading fluency and reading comprehension to be stronger for students of higher economic standing than those of lower economic standing. The DORF measures were mandated with the expectation that assessments showing specific reading skill deficiencies can be used to identify students in need of remedial reading instruction that will, in turn, improve reading comprehension scores on the SAT-10. This assumption and the widespread assessment procedures based on it warranted an investigation to examine the predictive nature of DORF scores on SAT-10 and DRP scores, see if the SAT-10's assessment of reading comprehension is corroborated by a secondary valid measure of reading comprehension, determine whether these relationships are similar for students of varying socioeconomic backgrounds, and explore the possibility that background vocabulary knowledge hinders the ability of



students from poor backgrounds to comprehend written material. In essence, this study was designed to determine whether or not proficient fluency skill is as strongly associated with low-income students' comprehension achievement as it is for their wealthier peers.

### Research Questions

The following questions directed part one of this study:

1. Do the DORF measures of reading fluency equally predict SAT-10 and DRP reading comprehension scores, and is there a positive correlation between the SAT-10 and DRP?

2. Do the DORF measures of reading fluency equally predict SAT-10 and DRP reading comprehension scores for students of differing economic backgrounds?

The following questions directed part two of this study:

3. Does reading fluency predict reading comprehension equally for higher- and lower-income students who have proficient fluency skill?

4. Do statistically significant differences in reading comprehension, reading fluency, and reading vocabulary exist between high- and low-income students with proficient reading fluency skill?

5. Does reading vocabulary equally predict reading comprehension for students with proficient reading fluency skill but differing economic backgrounds?

6. Does reading fluency predict reading comprehension beyond what can be accounted for by vocabulary for low-income students with proficient fluency skill?

### Significance of the Study

The current study was intended to determine the utility of current reading assessments administered to third-grade students, whether differences in reading

comprehension achievement between higher- and lower-income students with proficient fluency skill exist, and whether these discrepancies were associated with differing levels of background vocabulary knowledge. Specific purposes of this study are as follows:

- Evaluate DORF's effectiveness at predicting reading achievement outcomes on the SAT-10 and DRP and examine the concurrent validity between SAT-10 and DRP scores.
- Determine if the predictive nature of DORF in relation to the SAT-10 and DRP assessments of reading comprehension are statistically equivalent for students of differing economic backgrounds.
- Ascertain DORF's ability to identify students who have reading difficulties and, in turn, may require monitoring for future behavior problems, special education and Title 1 services, and dropout.
- Determine if socioeconomic status affects DORF'S ability to predict reading achievement scores.
- Determine whether students with proficient fluency skill, but differing economic backgrounds, perform equally on standardized measures of reading comprehension and reading vocabulary.
- Determine if the relationship between reading vocabulary and reading comprehension for fluent readers is similar or different for students of higher- and lower-income backgrounds.
- Examine whether or not proficient reading fluency skill equally predicts high-income and low-income students' reading comprehension achievement.

## Scope and Limitations

Although the current study was limited to eight participating schools located in rural and urban Alabama, the results may have implications for assessment and education anywhere DIBELS is administered kindergarten through third grade to identify students who are at risk of difficulties in reading and comprehending grade-level texts. These research questions and the results of this study may also provide information to guide additional educational research, administration at the state and local levels, and grade-level assessment and instruction. This study attempted to determine the relationships among several early literacy assessments as indicators of student reading outcomes and, ultimately, their ability to identify students at risk of reading failure, and to examine socioeconomic status (SES) as a factor influencing standardized scores of reading achievement. Schools participating in this study represented extremes on each end of a continuum for SES. Four schools were located in one of the most affluent areas in the state while the other four schools served low-income communities. Results obtained will enable the ALSDE, as well as other states currently administering DIBELS and high-stakes achievement tests, to better determine whether DIBELS accurately predicts future reading achievement on the SAT-10 and whether background vocabulary knowledge influences the performance of lower-income students on such assessments.

Due to the nature of the proposed study, several limitations require consideration. First, in view of the fact that the researcher utilized pre-existing data obtained from the ALSDE and local schools, absolute measurement of reading fluency, reading

vocabulary, and reading comprehension associated with the measures involved in the study cannot be assumed. For instance, the DORF assessment measures reading rate and accuracy, but does not account for other features of reading fluency such as expression and sensitivity to syntax. Also, the possibility of inconsistencies in the administration and scoring of all assessments, as well as errors in students recording of answers on the SAT-10 and DRP may exist. Second, correlational studies can produce spurious data that can be influenced by undetected independent variables and cannot be assumed as causal relationships. Third, specific reasons for any discovered relationship between variables cannot be assumed. Fourth, due to the time differences in the administration of the assessments, history and maturation may have influenced student scores. Fifth, results of the study are only generalizable to populations of similar children and not specific to any one child. Sixth, student characterizations of low socioeconomic status are based on free and reduced lunch data, which in some cases may be inaccurate. Finally, the results from the study's sample group are only generalizable to states with similar demographics.

#### Assumptions

The study is grounded in the following assumptions:

- The administration, scoring, and reporting of all student achievement scores as measured by DORF, the SAT-10, and the DRP were accurate.
- All school and student demographic data were reported correctly by both the individual schools and the Alabama State Department of Education.
- The data obtained represent each student's best effort on the employed measures of reading achievement.

## Definition of Terms

*Background vocabulary knowledge* – a person’s in-depth understanding of word meanings.

*Curriculum-based measurement* – easily used diagnostic measurement tools with multiple forms that allow for the on-going monitoring of progress.

*Decoding skill* – the ability to translate print into language.

*DORF Benchmark (Spring of third grade)* – the correct words read per minute on the Dynamic Indicators of Basic Early Literacy Skills – Test of Oral Reading Fluency (DORF) determined as a proficient level of reading fluency.

Benchmark:  $DORF \geq 110$

Below Benchmark:  $DORF < 110$

*High-stakes testing* – tests for which teachers, administrators, or schools are rewarded or reprimanded for student test performance.

*Listening comprehension* – the ability to draw on prior experiences and vocabulary to understand spoken language.

*Low socioeconomic status* – categorization of families whose children receive free or reduced public school lunches based on economic eligibility.

*Phonemic awareness* – the ability to identify and manipulate the individual sounds or vocal gestures in spoken words.

*Predictive validity* – the ability of an assessment of academic achievement to accurately forecast future performance on another assessment of academic achievement.

*Reading achievement* – operationally defined in this study as individual student scores on the Dynamic Indicators of Basic Early Literacy Skills – Oral Reading Fluency (DORF),

the Stanford Achievement Test – Tenth edition (SAT-10) vocabulary and reading comprehension subtests, and the Degrees of Reading Power (DRP) reading comprehension assessment.

*Reading assessment* – the measurement of individual reading skills.

*Reading comprehension* – the ability to utilize decoding skills, reading fluency, prior experiences, and vocabulary to understand written text.

*Reading fluency* – the ability to read text accurately and quickly while having the capacity to read with expression, divide text into meaningful chunks, and use emphasis and tone.

*Standardized assessment* - a test of achievement or knowledge in a specific academic domain whose scores are interpreted by reference to the scores of a norm group.

#### Organizational Overview

This chapter introduced the current study's purpose, research questions, significance, scope and limitations, assumptions, and definitions of terms. Additionally, a brief overview of related literature was provided. The subsequent chapters are organized as follows: Chapter Two considers the research literature and other publications related to this study; Chapter Three details the methodology utilized in this study and includes the investigation's purpose, design, instrumentation, participants, data collection procedures, and statistical analyses; Chapter Four describes the results of the study; and Chapter Five provides a summary, discussion of findings, conclusions, implications, and recommendations for further research.

## II. REVIEW OF LITERATURE

### Introduction

This chapter examines the literature associated with this study. The topics covered in this chapter include reading legislation, reading fluency and comprehension, vocabulary knowledge, a discussion of the simple view of reading, socioeconomic status as it relates to reading achievement, and reading assessment.

### Reading Legislation

The critical importance of early literacy in American schools has been repeatedly addressed at the highest levels of our country's administration to ensure a quality public education for all children. Federal organizations such as the National Institute of Child Health and Human Development (NICHD) and National Institute for Literacy (NIFL) have studied reading research and provided those responsible for our children's education information pertaining to literacy development. In 2001, the NIFL released *Put Reading First*, a synopsis of current early literacy research, to familiarize educators and parents with summaries of current reading education research. Also in 2001, the No Child Left Behind Act (NCLB) charged teachers to utilize scientifically valid reading practices when addressing the beginning reading needs of their students. Moreover, the NCLB Act addresses school accountability by mandating annual testing of all third- through eighth-grade students, and it requires that federal school funds be tied to the progress schools make on these standardized assessments. Before addressing reading assessment, a review

of the component skills necessary for proficient reading and the influence of wealth on children's reading achievement is necessary.

### Reading Fluency and Comprehension

In *Put Reading First* (Armbruster, Lehr, & Osborn, 2001), reading fluency is defined as “the ability to read text accurately and quickly” (p. 22) while having the capacity to read with expression, divide text into meaningful chunks, and use emphasis and tone. Unfortunately, the development of reading fluency has been regarded as the most neglected goal in our country's elementary schools (Allington, 1983). This belief was strengthened when Pinnell et al. (1995) discovered that 44% of fourth grade students in their sample population were reading grade-level texts with less than sufficient reading fluency. These findings have led researchers to question the effect that this deficiency has on children's ability to gain meaning from connected text.

Reading comprehension, often regarded as the “essence of reading” (Durkin, 1989, p.16), is the ability to draw on decoding skills, prior experiences, and vocabulary to understand written text. Proficient readers think actively when reading by having clear reading goals, previewing text, making predictions, integrating prior knowledge, and monitoring their understanding of written material (Duke & Pearson, 2002). Duke and Pearson (2002) discuss how good readers approach various texts differently, paying attention to characters and settings in narrative texts and developing summaries when reading expository texts. Furthermore, skillful readers continue to think and consider texts even when they are not reading.

There appears to be a strong connection between reading fluency and reading comprehension. The 1992 National Assessment of Educational Progress (NAEP) found a



relationship between the ability to read fluently and overall comprehension (Pinnell et al., 1995; Rasinski, 2000). This relationship exists because children with automatic and accurate word recognition are better able to gain meaning from written text than those who struggle to decode individual words (Armbruster et al., 2001).

### Vocabulary Knowledge

The assumption that vocabulary knowledge, the in-depth understanding of specific words, is critical to the ability to comprehend oral and written language seems to be apparent. A meta-analysis conducted by Stahl and Fairbanks (1986) demonstrated the significance of vocabulary knowledge and instruction in relation to text comprehension. Similarly, Anderson and Freebody (1981) determined that the number of difficult words in a text is the strongest indicator of a student's ability to comprehend written material. Interestingly in 2000, the National Reading Panel (NRP) reported that no experimental studies that met their rigorous criteria could be cited to justify a causal link between increased vocabulary knowledge and improved reading comprehension. The Panel reported that this lack of evidence was due to the complexity of defining and measuring vocabulary knowledge (National Institute of Child Health and Human Development [NICHD], 2000).

Merriam-Webster (2005) defines vocabulary as "a sum or stock of words employed by a language, group, individual, or work or in a field of knowledge." The NRP reported that the complexity of defining vocabulary knowledge goes much deeper. The Panel suggests that vocabulary can be subdivided into the following areas, some of which overlap: 1) receptive vs. oral; 2) oral vs. written; and 3) reading vs. writing. This suggests that vocabulary knowledge is multilayered and the various areas of vocabulary

knowledge sometimes contain similar characteristics. In addition, the NRP reported that the measurement of vocabulary knowledge with standardized tests is difficult not only because of the complexity of defining vocabulary, but also because standardized test items “can only ask a learner for a relatively small number of words” on any given administration (p. 4-16). Although test makers attempt to compensate for this problem by “selecting words that differ significantly in their familiarity,” it is simply not possible to accurately measure the size of an individual student’s vocabulary knowledge (p. 4-16). Considering the difficulties with defining and measuring vocabulary knowledge, the NRP acknowledged that “the importance of vocabulary in reading achievement has been recognized for more than half a century” (NICHD, 2000, p. 4-16).

The maturity of vocabulary knowledge evolves through oral communication, and this development is critical to making the transformation from verbal to written forms and to understanding written text (NICHD, 2000). Although there is little experimental evidence to support a causal link between vocabulary and reading comprehension, intensive vocabulary instruction designed to promote deep word knowledge has been associated with improved reading comprehension (McKeown, Beck, Omanson, & Perfetti, 1983). However, the direct instruction of vocabulary only accounts for a small proportion of a student’s vocabulary knowledge (Durkin, 1979; Jenkins & Dixon, 1983). Many researchers have concluded that most vocabulary words are acquired through incidental learning (Nagy, Anderson, & Herman, 1987; Nagy, Herman, & Anderson, 1985). This conclusion is supported by findings demonstrating that reading aloud to students increases the students’ vocabularies (Robbins and Ehri (1994). Moreover, students with stronger vocabularies have been found to be more motivated and to have

greater positive effects from incidental vocabulary learning than their peers with weaker vocabularies (Anderson, Hiebert, Scott, & Wilkinson, 1985; Nicholson & Whyte, 1992; Robbins & Ehri, 1994).

### The Simple View of Reading

In 1986, Gough and Tunmer introduced the following formula:  $R = D \times C$ .

In this simple view of reading, reading comprehension (R) is considered the product of decoding skill (D) and listening comprehension (C). This means that in order to read with understanding, one must be able to translate print into language and understand the message being conveyed. If either decoding skill or listening comprehension is inadequate, then the comprehension of written text is not possible.

Catts, Hogan, Adolf, and Barth (2003) conducted research that supported this view in a two-part longitudinal study. They first examined the variance in reading comprehension accounted for by word recognition and listening comprehension in students monitored in the second, fourth, and eighth grades, and then the changes in word recognition and listening comprehension abilities of poor readers in the second, fourth, and eighth grades. Part one of the study found that word recognition and listening comprehension accounted for most of the variance in reading comprehension across grades, but that the influence of these factors differed over time. Word recognition explained larger proportions of variance in reading comprehension in the second grade, less in the fourth grade, and only modest amounts in the eighth grade. In contrast, listening comprehension accounted for more variance in reading comprehension in the eighth grades as opposed to the earlier grades. Similar trends for poor readers at each

grade level were found in part two of the study, with the exception that poor readers generally exhibited listening comprehension deficits.

Although the study by Catts et al. (2003) exemplified the importance of both word recognition and listening comprehension for the skilled reader, it also showed that the amount of unique contribution of each varies over time. When considering students characterized as proficient readers, reading comprehension appears to be more dependent on word recognition in the early years; whereas, listening comprehension appears to be more influential in the later years. While this holds true for most children, poor readers continue to show insufficient listening comprehension skills in the later years. This difference may negatively affect these students' reading comprehension scores due to the increasing demands on vocabulary knowledge encountered beyond third grade on standardized examinations (Becker, 1977).

#### Socioeconomic Status and Reading Achievement

Educational policy and legislation such as the 2001 NCLB Act are attempts to equalize the inequalities seen between higher- and lower-income students. Kozol (1991) claimed that our educational system does not provide children from low-income families the same educational opportunities as their middle- and upper-income counterparts. To make matters worse, data support teachers' tendency to have lower learning expectations for at-risk students (Winfield, 1986). Although teachers' perceptions and school systems may view struggling and lower-income students in a different light than their wealthier peers, one must consider how the environment from which they come affects early literacy acquisition.

Drawing from the Bible's Gospel according to Matthew, Stanovich (1986) coined the term Matthew Effects to describe how slight differences between children's literacy development gradually develop into much larger differences in reading achievement throughout the elementary school years. Essentially, he states that "the rich get richer" (p.381) while "the poor get poorer" (p.382). Stanovich's research produced evidence showing that this is due to the progressive acquisition of the skills required to eventually become a successful reader. For good readers, environmental factors and instruction affect the development of phonemic awareness, which in turn advances decoding and word recognition skills. The ability to read words motivates developing readers to read more, which promotes better reading fluency, vocabulary, and text comprehension. Ultimately, the maturity of these reading skills leads to increased intelligence and a continued desire to learn. Stanovich described how the opposite of this is also true: Poor phonemic awareness hinders the progress of word recognition skills, inhibiting a child's desire to read, which leads to little reading practice and growth of reading fluency, vocabulary, and text comprehension. Considering this evidence, an investigation into the early literacy of lower socioeconomic children is warranted.

Research generalizations have shown that the early reading development of many children living in lower-income families is not nurtured to the same extent as it is for their middle- and upper-income peers (Desimone, 2001; Haycock, 2003; McCormick, 2003). Students living in lower-income homes were found to be exposed to fewer literary experiences, verbal and reading-related interactions with their parents, and opportunities to read. Academic progress of low-income students has also been found to be hindered by poor communication between their parents and the schools they attend (Desimone, 2001).

Furthermore, Senechal, LeFevre, Thomas, and Daley (1998) reported that children of parents who were read to in the home showed superior vocabulary and listening comprehension skills than those who were not. These differences, along with less game playing, which helps to develop critical thinking skills, tend to place lower-income children at a disadvantage for literacy development (Heath, 1991).

Haycock (2003) investigated the level of access children of low-income families had to literature versus the available access of middle- and upper-income children. His study included two low and two middle-to-high income neighborhoods, and he discovered significant differences. The higher income students were found to have approximately 4000 times the book titles available within their neighborhoods, significantly more available library time, better trained librarians, and more places appropriate for reading. These findings support concerns about the Matthew Effects in reading, but on a brighter note, there is evidence that lower SES children's motivation to read is similar to their wealthier peers in spite of having less exposure to books and support for reading (Baker & Scher, 2002).

Baker and Scher (2002) examined sixty-five six-year-old first graders' motivation for reading in relation to parental economic backgrounds. They utilized the Motivations for Reading Scale to rate four separate components of reading: enjoyment of reading, perceived value of reading, perceived competence in reading, and interest in library-related activities. Students responded to various questions by choosing which of two stuffed animals they were most alike, one with a positive face and one with a negative face. For example, "Regal thinks books are good places to find answers to questions, but Cha Cha doesn't think books are a good place to find answers to questions. Who are you

more like?” The researchers found SES background did not affect children’s motivation about reading, regardless of their familiarity with storybook reading or library visits. Essentially, students of both higher- and lower-income parents perceived reading with generally positive attitudes. The results of this study led the researchers to conclude that, as with any child, it is poorer students’ lack of print concepts and initial struggle with early literacy skills, such as the ability to recognize phonemes in spoken words and acquire decoding skills that, as Stanovich (1986) indicated, eventually decreases their motivation to read.

The English alphabetic writing system links letters and letter combinations to the phonemes heard in spoken language. By definition, phonemes are “the smallest units into which speech can be divided, and that make a difference to the meaning of a word” (Scarborough & Brady, 2002, p. 303). Furthermore, Liberman & Liberman (1992) describe phonemes as the basic vocal gestures found in oral language. Thus phonemic awareness is the “ability to notice, think about, and work with the individual sounds in spoken words” (NIFL, 2005, Glossary section, ¶1). Share, Jorm, Maclean, & Matthews (1984) determined that phonemic awareness was among the best predictors of a pre-literate child’s future reading success. Logically the predictive nature of phonemic awareness exists because the knowledge that spoken sounds are related to the mapping of spellings in written words provides insight into the alphabetic writing system (Stanovich, 1986). Fortunately, when students are taught to identify phonemes within words, they not only gain phonemic awareness, but they also “gain insight into the alphabetic principle and apply their insights in the early word identification” (Murray, 1998, p. 461).

Numerous studies have associated children of low socioeconomic status (SES) with low reading achievement (Au, 2000; Chall, Jacobs, & Baldwin; Desimone, 2001; Guthrie & Greaney; Molfese, Modglin, & Molfese, 2003). Further, this relationship is evident regardless of race or culture (Ratekin, 1978). Bowey (1995) sought to determine the impact of phonemic awareness skill on reading achievement of children of varying economic backgrounds. The longitudinal study included 148 English-speaking children from six preschools in Brisbane, Australia. Beginning in preschool and continuing through the completion of first grade, children were given various measures of oral language development, phonemic awareness, word-identification skill, vocabulary knowledge, and intelligence. Bowey found preschoolers' belonging to lower-income families were significantly less phonemically aware than those belonging to wealthier families and that "these differences remained robust even with performance IQ and verbal ability effects statistically controlled" (p.482). This evidence led the researcher to conclude that pre-existing phonemic awareness differences are at least partly responsible for lower SES children's early word-level reading achievement deficiencies.

In another study, Duncan and Seymour (2000) found that low socioeconomic status for children between the ages of four and eight was highly correlated with weaknesses on letter identification, phonemic awareness, and word identification tasks. Although the low SES students were approximately 17 months behind their wealthier peers in reading achievement, their socioeconomic status did not affect the speed at which they acquired letter and alphabetic knowledge. These findings led researchers to the conclusion that the lack of foundational literacy skills when entering school for children from poorer families impacted future reading achievement more than income



level itself. It appears the lack of materials, reading experiences, and literary activities in lower-income students' environment, rather than income, have the greatest negative impact on their reading achievement.

## Reading Assessment

### *High-Stakes Testing*

The recent trend towards educator accountability in our schools has become a passionate topic in today's society. Certainly there is a need to ensure the public that teachers and administrators are providing children with instruction that will allow them success in the professional and working worlds. In order to accomplish this objective, numerous states have decided the most effective strategy is to place great pressure on schools to improve standardized test scores. In these days of high-stakes testing, students' test performance determines which schools or teachers are rewarded or reprimanded.

Although the State of Alabama does not utilize standardized test scores to determine monetary rewards or employment status of its teachers, test scores are utilized to determine whether schools are sufficiently educating their children and whether or not a state-takeover is warranted. This movement towards increased school accountability requires reliable assessment measures that inform instruction and that also predict student outcomes as early in a child's development as possible (Carnine, 2000). The increased attention to ensuring effective early reading education seems necessary based on a study by Juel (1988). She found that students who were poor readers at the end of first grade had an 88% probability of remaining poor readers at the end of fourth grade. Considering that reading is a foundational skill fundamental to academic success, students who are not

proficient grade-level readers by the end of first grade are likely to experience future academic difficulties.

Due to the increased demands for improving reading achievement, the Alabama State Department of Education requires its schools to monitor the kindergarten through third grades' reading progress with the Dynamic Indicators of Basic Early Literacy Skills and the Stanford Achievement Test – Tenth Edition. The DIBELS assessments are Curriculum-Based Measures (CBM) used for the early identification of reading difficulties. These measures assess phonemic awareness skills, decoding skills, and (beginning in the first grade) reading fluency. Reading fluency is assessed using DIBELS Oral Reading Fluency (DORF) measures.

DORF is a standardized set of passages intended to identify children who may need additional instructional support. Student performance is measured by the total number of words read correctly while reading a passage aloud for one minute. Test passages have undergone numerous readability estimates to ensure that text difficulty is appropriate for each grade level (Good & Kaminski, 2002). Furthermore, progressive academic benchmarks are provided for each grade that establish the minimum levels of fluency proficiency. End-of-the-year benchmarks for the first, second, and third grades are 40, 90, and 110 correct words per minute, respectively (DIBELS, 2000-2003).

#### *Curriculum-Based Measurement*

Although most standardized assessments of reading are simple to administer and score, they are often expensive and time consuming, not well-suited for the consistent monitoring of student progress, have few alternate forms, and provide little relevant instructional information (Fuchs, Fuchs, & Maxwell, 1988; Warrington, 2003). In

contrast, Curriculum-Based Measures are brief, easily measured diagnostic tools with multiple forms that allow for on-going student progress monitoring in order to drive instruction. For these reasons, there is growing support for the use of CBM when assessing reading achievement (Knutson & Shin, 1991). The DIBELS measures are progressive Curriculum-Based Measures designed to assess early literacy skills for the purpose of identifying and remediating beginning reading deficiencies. They have been found not only to be predictive of each other but also correlated with teacher perceptions of kindergartners' academic readiness (Ritchey, 2004).

#### *DIBELS and the SAT-10*

If comprehension, the ultimate goal of reading, is to be measured through standardized achievement tests, and reading fluency is necessary for comprehension, then the assessment of younger students' fluency ability appears justified. A large body of correlational research has shown that DIBELS measures are valid and reliable predictors of student reading achievement on norm-referenced achievement tests. Researchers have found statistically significant positive correlations between kindergarten DIBELS scores and various measures of reading ability (Elliot, Lee, & Tollefson, 2001; Hintze, Ryan, & Stoner, 2002; Kaminski & Good, 1996; Speece, Mills, Ritchey, & Hillman, 2003). Positive correlations have also been noted between first, second, and third-grade DIBELS scores and the results of grade-level achievement measures (Cook, 2003; Kaminski & Good, 1996; Moscovitch, 2004).

In addition to the DIBELS measures, Alabama's third-grade students take the Stanford Achievement Test – Tenth Edition (SAT-10). Multiple studies have established the SAT-10 as a valid and reliable measure of reading achievement (Berk, 1998;

Harcourt, 2003). As a standardized examination, the SAT-10 is effective at determining a student's reading achievement relative to his or her peers, but it does not identify specific skill deficiencies. The Alabama State Department of Education (ALSDE) has determined the SAT-10 to be a valid and reliable measure a child's overall reading achievement, measured by word study, reading vocabulary, and reading comprehension scores. The ALSDE uses the SAT-10 to assess individual student reading achievement, and to determine a school's instructional effectiveness and probational status. In addition, the ALSDE publishes DIBELS and SAT-10 scores, making them open to public investigation and scrutiny.

Several studies have been conducted showing the predictive validity of third-grade DIBELS Oral Reading Fluency (DORF) and standardized high-stakes achievement tests (Barger, 2003; Buck & Torgesen, 2003; McKenna, 2003; Shaw & Shaw, 2002). Barger (2003) found a correlation of  $r = .73$  between third-grade students who met the 110 correct words per minute (cwpm) benchmark on the Spring DORF passages and those who achieved scores at or above grade-level on the North Carolina End of Grade Reading Assessment. Similarly, McKenna (2003) discovered a correlation of  $r = .66$  for students who obtained satisfactory scores on the DORF passages and the reading portion of the Oregon Statewide Assessment.

Shaw and Shaw (2002) extended this area of research by correlating Fall, Winter, and Spring DORF third-grade fluency levels with reading outcomes on the Colorado State Assessment Program (CSAP). They found Fall and Winter DORF scores produced correlations of  $r = .73$  and the Spring DORF scores correlated at a rate of  $r = .80$  with the

CSAP. Worth noting are the correlations the researchers found between the Fall and Winter DORF ( $r = .91$ ) and the Winter and Spring DORF ( $r = .93$ ) passages.

Buck and Torgesen (2003) found similar results when comparing third-grade DORF scores to the Florida Comprehensive Assessment Test (FCAT), both for the Sunshine State Standards Exam (FCAT-SSS) and the Norm Referenced Test (FCAT-NRT). Correlations of  $r = .73$ ,  $r = .74$ ,  $r = .53$  were found between the DORF and the FCAT-SSS, the FCAT-NRT, and the math portion of the FCAT-SSS, respectively. Additionally, the researchers disaggregated the data, including correlations for ethnicity and socioeconomic status. Regarding ethnicity (white,  $r = .70$ ; African-American,  $r = .62$ ; Hispanic,  $r = .78$ ), correlations were high, with Hispanics' DORF scores being most predictive of FCAT-SSS results. Students who received free or reduced lunch ( $r = .70$ ) showed minimal differences from students who paid for lunch ( $r = .69$ ). While this study supports DIBELS' ability to predict outcomes on high-stakes tests, it did not find that socioeconomic background nor ethnicity had statistically significant effects on DIBELS' predictive nature on standardized examinations of overall reading achievement.

Moscovitch (2004) examined the relationships between 2002-2003 third grade DORF scores and reading outcomes on the SAT-10. The following results were reported: DORF performance correlated with overall school quality; total proficiency didn't change after first grade for most students; poverty and minority status were associated with lower oral reading fluency scores; third grade DORF scores were predictive of SAT-10 reading achievement; and third grade DORF scores were less predictive of SAT-10 reading achievement for minority students and students from lower SES backgrounds. Unlike the

results reported by Buck and Torgesen (2003), these findings do suggest differences in DIBELS' predictive value depending on ethnicity and SES.

As with any research, however, there are some justifiable concerns about Moscovitch's methodology. SAT-10 and DORF scores were analyzed using percentages of total proficiency towards a desired goal, calculating the percentage of total number of items correct. DORF scores were examined by calculating the percentage of words read correctly per minute in relation to the established benchmark of 110 wcpm. For example, a DORF raw score of 110 revealed a total proficiency rating of 100%; whereas, a DORF raw score of 55 revealed a total proficiency rating of 50%. Moscovitch then compared DORF and SAT-10 total proficiency ratings, measured in percentages, to evaluate the existing relationships between DORF and SAT-10 achievement. While these methodologies appear to be a logical approach to calculating scores for statistical comparisons, there are reasons to interpret these results with caution.

First, DORF scores for students who read a minimum of 110 wcpm were all categorized as having a total proficiency rating of 100%. By doing this, there is no accounting for the differences between a student who read 110 wcpm and another who may have read 220 wcpm. This means that reading fluency scores used for subsequent correlations with SAT-10 reading achievement scores for students classified as having equal proficiency ratings could be the same for students reading at significantly different fluency levels. Second, although numerous correlational research studies have been conducted utilizing percentage scores, Huck (2004), states that true correlational research should employ the statistical analysis of continuous, equal interval scores. It appears that most of the studies conducted have found DORF scores to be predictive of standardized

reading achievement test scores, but research evidence has not yet provided a clear picture of the possible impact that socioeconomic status may have on correlations between DORF and standardized tests of reading comprehension.

### Summary

There is no doubt that all levels of our government consider the academic achievement of our country's children of the utmost importance. Each day the public is continually reminded of this through newspaper, radio, and television reports describing the adequacy or inadequacy of educational policy and practices. Certainly, reading achievement is among the most important of educational goals for this nation's children. Understanding that the foundation of a child's future reading success is often determined before leaving the third grade (Juel, 1988), the continual assessment of component literacy skills beginning when children enter the kindergarten has become increasingly popular and prevalent in today's public schools.

Although literacy is a complex and sometimes difficult skill to acquire, it can be viewed in relatively simple terms. Gough and Tunmer (1986) described the ability to gain meaning from the connected text as the product of decoding skill and listening comprehension. In such simplicity, this notion makes perfect sense – students who are able to decode text and understand the meaning that the language conveys can comprehend written material. In contrast, if one cannot read the words or does not understand the message being conveyed, meaning cannot be taken from a written passage. The level of reading proficiency is then determined by the strength of each component, ease of decoding skill, and ability to understand the message.

Even when considering reading in this simplistic manner, the potential influence of reading fluency and importance of background vocabulary knowledge for listening and reading comprehension cannot be discounted. When children are able to read quickly and accurately while also being able to read with expression, emphasis, and tone, they are said to be reading fluently (Armbruster, Lehr, & Osborn, 2001). Regrettably, evidence exists that many of our children do not read at a level of fluency deemed appropriate for their age (Allington, 1983; Pinnell et al., 1995). Additionally, the understanding of word meanings conveyed in language is essential for comprehension of written text. Decoding skill, albeit imperative, is worthless if one does not understand the vocabulary essential to a particular language.

Stanovich (1986) described the Matthew effects in reading to explain how children with strong early literacy skills enjoy progressively increased reading achievement over their peers with weaker early literacy skills. In essence, factors and skills leading to early reading success or failure contribute to the future practice and development of vocabulary knowledge and achievement. With success comes exponential development in reading skill, but with failure brings continued frustration and compounded deficiencies in reading skill development. Confounding this relationship is the influence of poverty, which has been associated with lower teacher expectations (Winfield, 1986), a lack of early reading experiences (Desimone, 2001; Haycock, 2003; McCormick, 2003), and, ultimately, poor reading development (Au, 2000; Duncan & Seymour, 2000; Molfese, Modglin, & Molfese, 2003).

The continual assessment of early literacy skills has become increasingly prevalent as teachers and schools are being held accountable to improve students'



performance on high-stakes standardized examinations of reading achievement. The State of Alabama currently uses the DIBELS assessment for the early identification reading problems in the areas of letter identification, phonemic awareness, decoding skill, and reading fluency throughout the first four years of school. These assessment procedures have been put into action to identify literacy deficiencies as early as possible so slowly developing readers can be remediated before students take the reading component of the SAT-10 at the end of third grade. Numerous studies have shown the DIBELS measures to be not only predictive of one another, but also predictive of reading achievement as measured by standardized examinations (Cook, 2003; Kaminski & Good, 1996; McKenna, 2003; Moscovitch, 2004; Richey, 2004). However, these correlational studies have only been undertaken with one measure of reading comprehension. This study was conducted to compare the predictive nature of DORF on two measures of reading comprehension and to determine whether or not proficient reading fluency skill is as strongly associated with low-income students' comprehension achievement as it is for their wealthier peers.

### III. METHODS

This chapter details this research project's methodological design. It is organized in the following six sections: purpose of the study, research design, description of the setting and sample, description of the instruments, data collection, and data analysis procedures.

#### Purpose of the Study

Previous research has suggested that third grade reading fluency as measured by DIBELS – Oral Reading Fluency (DORF) is predictive of third grade reading comprehension as measured by various standardized high-stakes examinations. Because previous research studies have raised questions about the value of DORF for predicting comprehension for high- as well as low-income children, the main research questions for this study focused on SES differences rather than ethnic differences. Part one of this project aims to reinforce these findings by not only evaluating the predictive value of DORF for scores on the Stanford Achievement Test – Tenth Edition (SAT-10) reading comprehension performance, but also on a secondary measure of reading comprehension, the Degrees of Reading Power (DRP). In addition, correlations between SAT-10 and DRP scores were examined. If these correlations are statistically equal, then one can assume that both the SAT-10 and DRP are measuring the same skill (reading comprehension) and that DORF is a valid and reliable predictor of reading

comprehension. In part two of the study, the researcher sought to determine whether students with proficient reading fluency skill but different economic backgrounds had equivalent comprehension achievement. If not, then an investigation into the relationships between reading fluency, reading vocabulary, and reading comprehension for wealthier and poorer students was deemed necessary to determine if vocabulary knowledge of lower-income children with proficient fluency skill may be a factor to address as efforts are made to improve their reading comprehension. A two-part, non-experimental correlational design was used to analyze the relationships between variables.

The following research questions directed part one of the study:

1. Do the DORF measures of reading fluency equally predict SAT-10 and DRP reading comprehension scores, and is there a positive correlation between the SAT-10 and DRP?
2. Do the DORF measures of reading fluency equally predict SAT-10 and DRP reading comprehension scores for students of differing economic backgrounds?

The following research questions directed part two of the study:

3. Does reading fluency predict reading comprehension equally for higher- and lower-income students who have proficient fluency skill?
4. Do statistically significant differences in reading comprehension, reading fluency, and reading vocabulary exist between high- and low-income students with proficient reading fluency skill?
5. Does reading vocabulary equally predict reading comprehension for students with proficient reading fluency skill but differing economic backgrounds?

6. Does reading fluency predict reading comprehension beyond what can be accounted for by vocabulary for students from low-income backgrounds with proficient fluency skill?

### Research Design

All testing procedures were administered independently of the researcher, and no treatment of any type was administered. The inability to control for extraneous variables was a limitation of this study. Consequently, cause and effect relationships cannot be inferred from the results of this correlational study as can be in true experimental designs (Keppel & Zedeck, 1989). A combination of correlational comparisons, descriptive statistics, and sequential regressions were employed for data analysis.

The correlational coefficient, known as Pearson's product-moment correlation ( $r$ ), was the principal statistical measure used in this study. This coefficient was used to determine the strength and direction of the relationships between independent and dependent variables. The  $R^2$  was used to measure the precise amount of change in the dependent variable that was accounted for by the independent variable (Huck, 2004).

Correlational-comparison analyses to compare the relationships between bivariate correlations were utilized. In part one of this study, bivariate correlations between the DORF and the SAT-10 and between the DORF and DRP were specific to the same group of participants. Steiger's  $Z$ -test was utilized to determine if statistical differences between two single sample correlations existed. Pearson's  $r$  coefficients were transformed to  $z$  scores and a normal curve  $Z$ -test determined the significance of correlational differences. According to Steiger (1980) and Meng, Rosenthal, and Rubin (1992), Steiger's  $Z$ -test has been established as a more valid and reliable method than Hotelling's  $t$ -test for

comparing single sample correlations. In part two of this study, bivariate correlations between reading fluency, reading vocabulary, and reading comprehension were conducted for high- and low-income participants. To compare these independent sample correlations, Fisher's *Z* transformation test was used.

Sequential regression was employed to determine the amount of variance accounted for by the independent variable on the dependent variable after controlling for a second independent variable. Furthermore, effect sizes were calculated to find the strength of association between variables. In accordance with Cohen (1988), effect sizes (*d*) of .20, .50, .80 indicated small, medium, and large differences between the sample means being compared (Huck, 2004). Additionally, descriptive statistics reporting group means and standard deviations were used to assist in the identification of group similarities and differences.

#### Instrumentation

Three instruments were utilized to measure the participants' reading achievement during the course of this study: The Stanford Achievement Test – Tenth Edition (SAT-10), the Dynamic Indicators of Basic Early Literacy Skills – Oral Reading Fluency (DORF), and the Degrees of Reading Power (DRP). The SAT-10 and DORF were chosen because the Alabama State Department of Education (ALSDE) had previously deemed them adequate measures of third-grade reading achievement. Furthermore, by virtue of the fact that most of Alabama's third-grade students are required to take the SAT-10 and DORF, they served as instruments of convenience. DRP scores were only available from schools that chose to use the DRP and were willing to participate in this study. These scores served as a secondary measure of student reading comprehension. In spite of the

fact that each measure has endured exhaustive reliability and validity studies, certain limitations do exist.

*Dynamic Indicators of Basic Early Literacy Skills – Oral Reading Fluency*

The Dynamic Indicators of Basic Early Literacy Skills – Oral Reading Fluency (DORF) is a set of fluency-based, individually administered standardized indicators designed to evaluate beginning reading skills (DIBELS, 2000-2003). The DORF is an example of Curriculum-Based Measurement, which are brief, easily measured diagnostic tools with multiple forms for on-going student progress monitoring that can be used to inform and alter instructional practices (Knutson & Shin, 1991; Warrington, 2003). Researchers have shown test-retest reliability for most Curriculum-Based Reading measures ranging from .92 to .97, alternate form reliability of different reading passages ranging from .89 to .94 (Tindal, Marston & Deno, 1983), and criterion-related validity ranging from .52 to .91 (Good & Jefferson, 1998).

Armbruster et al. (2001) defined reading fluency as “the ability to read text accurately and quickly” while having the capacity to read with expression, divide text into meaningful chunks, and use emphasis and tone (p. 22). According to the current accepted definition of fluency, the DORF assessment cannot be considered rich enough to measure all aspects of fluency, but it is an efficient measure of reading speed and accuracy, outcomes that are calculated from the number of words read correctly in one minute.

The DORF measures have undergone numerous reliability and validity studies containing heterogeneous samples of elementary students (Buck & Torgesen, 2003; Kaminski & Good, 1996; Shaw & Shaw, 2002). Additionally, the DORF fluency-based

measures are highly correlated to reading outcomes, aligned with best practices in early literacy reading instruction, and considered valid measures of reading fluency (Good, Simmons, & Smith, 1998; Hintze et al., 2002). However, an important limitation of DORF is the fact that the reliability of scores is dependent on the ability and integrity of the test administrator.

For the purpose of this study, DORF measures from the Spring assessments for third-grade students were utilized. The DORF measures contained three passages, each of which was an individually administered one-minute reading (Good, Simmons, & Kame'enui, 2001). After students read each passage, the total number of words read correctly per minute was calculated. Hesitations of more than three seconds, omissions, and substituted words were considered errors and did not count toward total words read. Words that were read incorrectly but then self-corrected within three seconds were counted as correct. Once the student read all passages, the median score of the three passages was recorded. The Spring benchmark, the minimum score which students can obtain to be considered achieving adequate grade-level reading fluency, is established at 110 correct words per minute.

#### *The Stanford Achievement Test – Tenth Edition*

The SAT-10 served as an indicator of third-grade students' reading vocabulary and reading comprehension achievement. SAT-10 is the most recent version of a series of assessments that originated in 1923 (Berk, 1998). The publishers have systematically updated the test to accommodate changes in school curriculum, update norms and materials, and include assessment of all current major academic domains. The SAT-10 is group administered and presents all items in a multiple-choice format. The total reading

achievement score for the SAT-10 is a composite measure based on reading vocabulary and reading comprehension subtest scores, each of which have independent norms and percentile rankings. The third-grade Spring SAT-10 test of reading comprehension yielded a Kuder-Richardson reliability coefficient of .93 (Harcourt, 2003).

Evidence supporting the validity of SAT-10 scores was based on content appropriateness and an item try-out program followed by item analyses. Additionally, standardization samples provided data to equate test levels as well as equate the SAT-10 to the SAT-9 scores. The test developers report that comprehensive reviews and analyses of major textbooks in all subject areas, state and district curricula and instructional standards, and “important educational trends and directions as expressed by the national professional organizations” (Harcourt, 2003, p.8) initially guided item content decisions. Based on these sources of information, a test blueprint was mapped out and reviewed by content area specialists and test construction professionals. Then the blueprints were revised based on feedback received. Finally, content area specialists and trained writers developed item pools, and items were screened by content experts, measurement experts, and editorial specialists. Items were designed to assess both basic understanding and thinking skills. For the purpose of this study, only the measures of reading vocabulary, reading comprehension, and total reading, all reported as raw scores, were considered.

Though the SAT-10 has undergone extensive reliability and validity testing, it is subject to the same criticisms associated with other standardized tests. In addition to the inability to satisfactorily monitor student progress (Fuchs et al., 1988), standardized tests are criticized for being technically inadequate (Galagan, 1985; Marston, 1989), providing little instructionally relevant information (Fuchs et al., 1988), being insensitive to small



changes in student performance (Marston, 1989), and often inadequately measuring the curriculum being taught (Fuchs et al., 1988; Fuchs, Fuchs, & Hamlett, 1989; Marston, 1989). Also worth noting is the failure of the SAT-10 and many other standardized multiple-choice tests to provide procedures factoring in correct or incorrect responses that are the result of guessing.

### *Degrees of Reading Power*

The Degrees of Reading Power (DRP) employs a modified cloze format that requires students to construct meaning across an entire paragraph or passage to assess third-grade students' ability to comprehend written material (Green, 2001). Students work independently under no time constraint to identify the appropriate answer that best supports the meaning of passages containing missing words. Kuder-Richardson (K-R-20) internal reliability coefficients of .92 with a standard deviation of 2.7 were found for the third-grade DRP (Green, 2001; Margolis, 2001). Additionally, correlations of .77 with the California Achievement Test and reported gains after 5 months of instruction, suggest that the DRP is sensitive to growth and support the validity of the DRP as a measure of reading comprehension (Margolis, 2001).

Although the DRP has withstood extensive tests of reliability and validity (Touchstone, 1995), several limitations are worth citing. First, the publisher's recommendation to administer the test only to students who have mastered decoding skills is ambiguous and leaves room for interpretation. Second, due to the DRP's multiple-choice format, guessing can inflate the number of correct answers. Finally,

group administration does not allow for clarification of procedures, meaning that a poor score may be due to misunderstanding the directions (Margolis, 2001).

## Participants

### *Part One*

Third-grade children were chosen as participants because third grade is the earliest grade for which the State of Alabama mandates the assessment of both reading fluency and reading comprehension. The researcher included data from both rural and urban schools to provide a representative cross-section of Alabama's student population. Only schools that administered the DRP and that were willing to provide the researcher with the necessary data were included. Most of the schools in the sample were Alabama Reading First Initiative (ARFI) schools (those with lower-income populations) and may not be truly representative of Alabama's third-grade population.

Part one of the study's sample consisted of 129 third-grade students who participated in the DORF, SAT-10, and DRP assessments during the 2003-2004 academic school year. Thirty-one participants were not included in analyses investigating students by income-level due to the fact their SES data were not available. The subjects represented three schools located in the state of Alabama, two rural and one urban. Characteristics such as age, race, gender, and income status were obtained from student assessment outputs. Subjects' ages ranged from 7.2 years to 11.0 years with a mean of 9.4 years and a standard deviation of .54 years. The sample included 56 Caucasian (43.4%), 61 African-American (47.3%), 9 Hispanic (7.0%), and 3 Asian (2.3%) students. Fourteen percent of the sample were reported as being recipients of special education

services. Additionally, 48.1% of the sample was reported to be lower SES as determined by free or reduced lunch (FRL) status. Data for all 129 participants were examined to address the first research question.

To address the second research question, the total sample was reduced to represent only those participants who could be identified as either of high- or low-income status. Thirty-six high-income participants and 62 low-income participants were included in analyses that investigated students by income-level. Thirty-one participants were not included because of insufficient FRL data. Permission to obtain all academic and demographic data was provided by both the Alabama State Department of Education and the subject’s principal or superintendent. Table 1 indicates the distribution of rural and urban participants by gender and SES (based on FRL status).

Table 1

*Participants by Location, Gender, and SES Status (Part One)*

	Rural		Urban		Total n
Gender	Male	Female	Male	Female	
	10	16	54	49	129
SES Status	Avg/High	Low	Avg/High	Low	
	0	26	36 <sup>a</sup>	36 <sup>b</sup>	98

<sup>a</sup>Sample included 11 additional participants with unknown SES status. <sup>b</sup>Sample included 20 additional participants with unknown SES status.

*Part Two*

Part two of the study’s sample consisted of 112 third-grade students with proficient fluency skill who participated in the DORF and SAT-10 assessments during

the 2003-2004 academic school year. In addition, follow-up analyses included 103 low-income third-grade participants without proficient fluency skill. Due to an insufficient sample size, follow-up analyses did not investigate high-income third-grade participants. The subjects represented eight schools located in the state of Alabama, two rural and six urban. Participants categorized as high-income did not receive free or reduced lunch and belonged to a community where the median family income was above \$100,000. Participants categorized as low-income either belonged to schools where 100% of the student population received free or reduced lunch or they received free or reduced lunch within schools in which at least 90% of the student population received free or reduced lunch. Median family income data was unavailable for low-income participants.

Information concerning age, race, gender, and SES status were obtained from student test data. Subjects' ages ranged from 8.5 years to 11.2 years with a mean of 9.5 years and a standard deviation of .55 years. The sample, including participants with and without proficient fluency skill, consisted of 86 Caucasian (40.0%), 121 African-American (56.3%), 6 Hispanic (3.0%), and 2 Asian (1%) students, and 109 male and 106 female students. Permission to obtain all academic and demographic data was provided by either the Alabama State Department of Education or the subject's principal or superintendent. Because initial data collection yielded a much larger sample of high-income (n=294) than low-income (n=56) students with proficient fluency skill, a random sampling utilizing a random number generator was employed to obtain equal sample sizes of 56 participants for high- and low-income participants. Table 2 indicates the distribution of participants with proficient fluency skill by income level and gender.

Table 2

*Participants with Proficient Fluency Skill by Income Level and Gender (Part 2)*

	High-Income		Low-Income	
Gender	Male	Female	Male	Female
	28	28	24	32

Data Collection Procedures

All data were collected ex-post facto with the permission of the Alabama State Department of Education and cooperating principals and superintendents. Schools that were sites where DORF, the SAT-10, and the DRP (required in part one but not part two of the study) were administered during the 2003-2004 school year and that were willing to provide the researcher with all necessary data were included in the study. The Alabama State Department of Education, cooperating school systems, and building-based administrators were provided signed confidentiality statements from the researcher ensuring that no school or student identifiers would be utilized during analysis and reporting (see Appendix A). Additionally, the researcher obtained signed data release statements from administrators indicating their willingness to support this study (see Appendix B).

DORF and the SAT-10 were administered during the months of April and May of the 2003-2004 school year. The DRP was conducted during the first week of May 2004. DORF was administered individually requiring students to read several grade-level

passages orally to the assessor. In some schools, the classroom teacher conducted the assessment, while in other schools an assessment team was organized to test all third-grade students. The SAT-10 and DRP assessments were both multiple-choice formats and group administered. The classroom teacher conducted the administration of the SAT-10 with the assistance of a proctor, whereas either the classroom teacher or reading specialist administered the DRP.

As previously cited, reliability and validity of all measures have been established by the test publishers. Although test administrators for all three assessments were required to attend training, the researcher was unable to address inter-rater reliability due to the number of participating schools and security measures during test administration.

DORF and SAT-10 data were provided by the Alabama State Department of Education and participating schools. Hard copies of DRP data were provided to the researcher by participating schools. All student data, including assessment scores and demographic information, were compiled by the researcher using Microsoft Excel. All data were collected, compiled, and coded so that all school and student identifying information was removed before leaving school supervision. All statistical analyses were conducted using SPSS.

#### Statistical Method

Part one of the study determined whether two measures of reading comprehension, the SAT-10 and the DRP, were equally predicted by DORF and whether SAT-10 and DRP scores were positively correlated. Part two of the study examined whether students with proficient reading fluency skill but different economic backgrounds had statistically equivalent comprehension achievement. The research

questions listed at the beginning of this chapter guided this study. The following procedures were used for data analysis.

### *Part One*

To examine the relationships between DORF and the SAT-10 as well as DORF and the DRP, bivariate correlations were used. DORF served as the independent variable while the SAT-10 and DRP served as dependent variables. Bivariate correlational procedures were also employed to examine the relationship between the SAT-10 and DRP. The SAT-10 served as the independent variable while the DRP served as the dependent variable. Pearson's product moment correlation coefficients were transformed to  $z$  scores to compare the predictive nature of DORF, operationalized by DORF reading fluency raw scores, on the SAT-10 and DRP, operationalized by SAT-10 and DRP comprehension raw scores. Steiger's  $Z$ -test was utilized to determine if statistical differences existed between  $r$  comparisons of DORF and the SAT-10, and DORF and the DRP. This analysis was utilized to evaluate these relationships for all participants, students receiving free or reduced lunch, and students not receiving free or reduced lunch.

### *Part Two*

To examine the relationships between reading fluency and reading comprehension for higher- and lower-income students with proficient fluency skill, bivariate correlations were used. Reading fluency served as the independent variable while the reading comprehension served as the dependent variable. Pearson's product coefficients were transformed to  $z$  scores to compare the predictive nature of reading fluency (operationalized by DORF raw scores) on the reading comprehension (operationalized by

SAT-10 comprehension raw scores). Fisher's  $Z$  transformation test was utilized to determine if statistical differences existed between  $r$  comparisons of reading fluency and reading comprehension. This analysis was utilized to evaluate these relationships for two groups of participants, students from high-income backgrounds and students from low-income backgrounds.

To determine if significant differences in reading comprehension, reading fluency, and reading vocabulary existed between high- and low-income students with proficient reading fluency skill, one-way analyses of variance were utilized. Income level served as the independent variable while reading comprehension, reading fluency, and reading vocabulary served as dependent variables.

To ascertain whether or not reading vocabulary equally predicted reading comprehension for students with proficient reading fluency skill, but differing economic backgrounds, bivariate correlations were used. Reading vocabulary served as the independent variable while reading comprehension served as the dependent variable. Pearson's product coefficients were transformed to  $z$  scores to compare the predictive nature of reading vocabulary (operationalized by SAT-10 vocabulary raw scores) on the reading comprehension (operationalized by SAT-10 comprehension raw scores). Fisher's  $Z$  transformation test was utilized to determine if statistical differences existed between  $r$  comparisons of reading vocabulary and reading comprehension for two groups of participants, students from high-income backgrounds and students from low-income backgrounds.

Finally, sequential regression was utilized to determine if reading fluency predicted reading comprehension for students from low-income backgrounds with



proficient fluency skill, after statistically controlling for reading vocabulary. Reading fluency served as the independent variable while reading comprehension served as the dependent variable.

### Summary

Third-grade reading achievement data were collected from participating schools. This study examined the predictive nature of the DORF measures of reading fluency in relation to the SAT-10 and DRP measures of reading comprehension and the correlation between the SAT-10 and DRP. It also investigated whether or not proficient reading fluency skill is as strongly associated with low-income students' comprehension achievement as it is for their wealthier peers. Additional analyses were used to identify any significant differences between high- and low-income students' reading comprehension, reading fluency, and reading vocabulary achievement scores and to determine if statistical differences existed between high- and low-income students' reading vocabulary and reading comprehension. Finally, sequential regression was employed to investigate the value of reading fluency as a predictor of reading comprehension for students with proficient reading fluency skill from low-income backgrounds after statistically controlling for vocabulary.

#### IV. DATA ANALYSIS AND RESULTS

The purpose of this two-part study was to confirm the predictive nature of DIBELS Oral Reading Fluency (DORF) in relation to Stanford Achievement Test – Tenth Edition (SAT-10) reading comprehension performance, as well as on a secondary measure of reading comprehension, the Degrees of Reading Power (DRP) and examine the relationship between the SAT-10 and DRP. In addition, the study aimed to ascertain whether students with proficient reading fluency skill but different economic backgrounds have similar vocabulary and comprehension achievement and to examine the relative values of fluency and vocabulary as predictors of reading comprehension.

##### Review of Research Questions

The following questions directed part one of this study:

1. Do the DORF measures of reading fluency equally predict SAT-10 and DRP reading comprehension scores, and is there a positive correlation between the SAT-10 and DRP?
2. Do the DORF measures of reading fluency equally predict SAT-10 and DRP reading comprehension scores for students of differing economic backgrounds?

The following questions directed part two of this study:

3. Does reading fluency equally predict reading comprehension for higher- and lower-income students who have proficient reading fluency skill?

4. Do statistically significant differences in reading comprehension, reading fluency, and reading vocabulary exist between high- and low-income students with proficient reading fluency skill?
5. Does reading vocabulary equally predict reading comprehension for students with proficient reading fluency skill but differing economic backgrounds?
6. Does reading fluency predict reading comprehension beyond what can be accounted for by vocabulary for students from low-income backgrounds with proficient fluency skill?

## Data Analysis and Results

### *Part One*

In part one of this study, student assessment scores for 129 student participants were entered into an SPSS data file (11.5 version for Windows) for analyses. Analyses included all participants, as well as those categorized receiving free or reduced lunch ( $n=62$ ) and not receiving free or reduced lunch ( $n=36$ ). Free or reduced lunch status was unavailable for 31 participants, who were not included in analyses investigating students by income level. The research questions posed in part one of the study were designed to establish if there were statistically significant indicators that reading fluency scores predict comprehension scores for students of differing economic backgrounds. Reading fluency, operationalized by reading rate and accuracy scores, was measured by DORF while reading comprehension, operationalized by comprehension achievement scores, was measured by the comprehension subtest of the SAT-10 and the DRP. All analyses utilized raw scores provided by the Alabama State Department of Education and cooperating schools. Means and standard deviations are presented in Table 3.

Table 3

*Means and Standard Deviations for Student DORF, SAT-10, and DRP Scores (Part One)*

	DORF		SAT-10		DRP	
	Mean	SD	Mean	SD	Mean	SD
All Participants	102.34	38.64	29.72	11.08	23.60	9.30
Avg/High SES	117.25	32.40	35.61	9.94	28.67	10.26
Low SES	96.87	40.90	26.27	10.77	21.05	8.47

The following research hypotheses addressed the research questions presented in part one of this study:

*Hypothesis 1.* The DORF measures of reading fluency equally predict SAT-10 and DRP reading comprehension scores, and significant positive correlations exist between the SAT-10 and DRP. DORF was determined to be the independent variable because the study aimed to examine DORF's predictability of the SAT-10 and DRP. Also, students are tested on the DORF prior to being tested on either of the measurements of reading comprehension. To test this hypothesis, Pearson product moment correlation coefficients between DORF, SAT-10, and DRP scores for all participating students ( $n=129$ ) were obtained. Zero-order correlations for the DORF, SAT-10, and DRP are presented in Table 4. Statistically significant positive correlations between DORF and the SAT-10, DORF and the DRP, and SAT-10 and DRP were found. To determine if these correlations differed at a statistically significant level, Steiger's Z-test to compare single sample correlations was utilized. Comparisons showed that the DORF measures'

predictability of SAT-10 and DRP reading comprehension scores did not differ at a statistically significant level ( $Z=1.00, p>.05$ ).

Table 4

*Intercorrelations Among Variables for All Participants (Part One)*

	DORF	SAT-10	DRP
DORF	1.00		
SAT-10	.722**	1.00	
DRP	.683**	.781**	1.00

\*\* $p<.001$

*Hypothesis 2.* The DORF measures of reading fluency equally predict SAT-10 and DRP reading comprehension scores for students of differing economic backgrounds. To test this hypothesis, Pearson product moment correlation coefficients between DORF, SAT-10, and DRP scores for students not receiving free or reduced lunch ( $n=36$ ) and students receiving free or reduced lunch ( $n=62$ ) were obtained. Zero-order correlations for the DORF, SAT-10, and DRP are presented in Table 5. Statistically significant positive correlations between DORF and the SAT-10, DORF and the DRP, and the SAT-10 and DRP were found. To determine if these correlations differed at a statistically significant level, Dunn and Clark's (1969) Fisher's  $Z$  transformation test was used to compare single sample correlations. Comparisons showed that the DORF measures' predictability of SAT-10 and DRP reading comprehension scores did not differ at a statistically significant level for average- to high- SES students ( $Z=0.16, p>.05$ ) or low SES students ( $Z=0.00, p>.05$ ).

Table 5  
*Intercorrelations Among Variables (Part One)*

	Avg/High SES			Low SES		
	DORF	SAT-10	DRP	DORF	SAT-10	DRP
DORF	1.00			1.00		
SAT-10	.736**	1.00		.717**	1.00	
DRP	.748**	.847**	1.00	.717**	.737**	1.00

\*\*p<.001

*Part 2*

Part two of this study examined students with proficient reading fluency skill but differing economic backgrounds. The relationships between reading fluency, reading vocabulary, and reading comprehension for these two groups were investigated to determine whether reading fluency predicts reading comprehension to the same extent for high- and low-income students. Student assessment scores for 215 student participants were entered into an SPSS data file (11.5 version for Windows) for analyses. In order to conduct correlation comparisons containing equal sample sizes, analyses included randomly selected participants categorized as high-income with proficient fluency skill ( $n=56$ ) or low-income with proficient fluency skill ( $n=56$ ). Additionally, follow-up analyses included low-income students with insufficient fluency skill ( $n=103$ ).

The research questions posed in part two of this study focused on whether or not proficient reading fluency skill is as sufficient a predictor of reading comprehension for students of low-income families as it is for their wealthier peers. Reading fluency was measured by DORF while reading comprehension and reading vocabulary were measured

by the comprehension and vocabulary subtests of the SAT-10. Means and standard deviations are presented in Table 6.

Table 6

*Means and Standard Deviations for Fluency, Comprehension, and Vocabulary Scores of Students with Proficient Fluency Skill (Part Two)*

	Fluency		Comprehension		Vocabulary	
	Mean	SD	Mean	SD	Mean	SD
High-Income	150.14	27.97	45.66	5.69	27.13	2.62
Low-Income	134.66	24.02	37.20	8.11	23.25	4.75

The following research hypotheses addressed the research questions presented in part two of this study:

*Hypothesis 3.* Reading fluency does not equally predict reading comprehension for higher- and lower-income students who have proficient reading fluency skill. To test this hypothesis, Pearson product moment correlation coefficients between fluency and comprehension scores for high-income ( $n=56$ ) and low-income ( $n=56$ ) students were obtained. Zero-order correlations of reading fluency and comprehension scores for students with proficient fluency skill are presented in Table 7. Statistically significant positive correlations between fluency and comprehension were found for high-income students but not low-income students. To determine if these correlations differed at a statistically significant level, Fisher's  $Z$  transformation test to compare independent sample correlations was utilized. Comparisons showed that reading fluency scores' predictability of reading comprehension scores differed at a statistically significant level

for high-income and low-income students ( $Z=2.363, p<.05$ ). For high-income students with proficient fluency skill, analyses indicated a statistically significant positive correlation between reading fluency and reading comprehension. For low-income students with proficient fluency skill, the correlation between reading fluency and reading comprehension was positive, but not statistically significant.

Table 7  
*Intercorrelations Among Fluency and Comprehension by Fluency Skill (Part Two)*

Fluency Skill		High-Income		Low-Income	
		Fluency	Comp	Fluency	Comp
Proficient	Fluency	1.00		1.00	
	Comp	.600**	1.00	.229	1.00
Not Proficient	Fluency	—		1.00	
	Comp	—	—	.652**	1.00

\*\* $p<.001$

Since reading fluency was not found to be a statistically significant predictor of reading comprehension for low-income students with proficient reading fluency skill, a finding inconsistent with previous correlational results between reading fluency and reading comprehension scores, part two of the study also included a follow-up investigation into the relationship between reading fluency and reading comprehension for low-income students with insufficient fluency skill. To determine if statistically insignificant results using these variables was also found for non-fluent, low-income students, Pearson product moment correlation coefficients between fluency and



comprehension scores for low-income ( $n=103$ ) students without proficient fluency skill were obtained. Zero-order correlations of reading fluency and comprehension scores for low-income students with insufficient fluency skill yielded statistically significant positive correlations ( $r=.652, p<.001$ ). Due to an insufficient sample size, analyses were not conducted for non-fluent high-income students.

*Hypothesis 4.* Statistically significant differences in reading comprehension, reading fluency, and reading vocabulary exist between high- and low-income students with proficient reading fluency skill. To test this hypothesis, three one-way analyses of variance were utilized to compare reading comprehension, reading fluency, and reading vocabulary means scores for high- and low-income children. A statistically significant difference was found between wealthy and poor students' comprehensions scores ( $F=40.90, p=.001$ ), with the high-income mean ( $M= 45.66, SD=5.69$ ) higher than the low-income mean ( $M= 37.20, SD=8.11$ ). According to Cohen (1988), the calculated effect size ( $d=1.21$ ) showed income level to have a large effect on reading comprehension scores. A statistically significant difference was also found between wealthy and poorer students' fluency scores ( $F=9.88, p=.045$ ), with the mean for high-income students ( $M= 150.14, SD=27.97$ ) significantly higher than the mean for low-income students ( $M= 134.66, SD=24.02$ ). The calculated effect size ( $d=.59$ ) showed income level to have a medium effect on reading fluency scores. Finally, a statistically significant difference was also found between wealthy and poorer students' vocabulary scores ( $F=28.57, p=.000$ ), with the mean for high-income students ( $M= 27.13, SD=2.62$ ) significantly higher than the mean for low-income students ( $M= 23.25, SD=4.75$ ). The calculated effect size ( $d=1.01$ ) showed income level to have a large effect on reading vocabulary scores.

*Hypothesis 5.* Reading vocabulary equally predicts reading comprehension for students with proficient reading fluency skill but differing economic backgrounds. To test this hypothesis, Pearson product moment correlation coefficients between vocabulary and comprehension scores for high-income ( $n=56$ ) and low-income ( $n=56$ ) students were obtained. Means and standard deviations were provided previously in Table 6. Zero-order correlations of reading vocabulary and reading comprehension scores for students with proficient fluency skill are presented in Table 8. Statistically significant positive correlations between vocabulary and comprehension were found for high-income and low-income students. To determine if these correlations differed at a statistically significant level, Fisher's  $Z$  transformation test to compare independent sample correlations was utilized. Comparisons showed no significant differences in reading vocabulary scores' predictability of reading comprehension scores between high-income and low-income students ( $Z=.619, p>.05$ ).

Table 8

*Intercorrelations Among Vocabulary and Comprehension for Students with Proficient Fluency Skill (Part 2)*

	High-Income		Low-Income	
	Vocabulary	Comprehension	Vocabulary	Comprehension
Fluency	1.00		1.00	
Comprehension	.606**	1.00	.677**	1.00

\*\*p<.001

*Hypothesis 6.* Reading fluency does not predict reading comprehension beyond what can be accounted for by vocabulary for students from low-income backgrounds with proficient fluency skill. Results from part one of this study implied that the relationship between reading fluency and reading comprehension is statistically equivalent for both high- and low-income students, meaning that as reading fluency scores increase so do reading comprehension scores. However, results from part two of this study suggested that reading fluency is a statistically significantly better predictor of reading comprehension for high-income students than low-income students with proficient fluency skill. Furthermore, mean comprehension and vocabulary scores were statistically significantly higher for high-income students than low-income students with proficient fluency skill. To further test the stated hypothesis, sequential regression was utilized to determine if reading fluency predicts reading comprehension for children with proficient fluency skill from low-income backgrounds ( $n=56$ ) beyond what was accounted for by reading vocabulary. Reading fluency was not found to predict increases in reading

comprehension scores for low-income students proficient in fluency skill after controlling for reading vocabulary ( $R^2=.465$ ,  $R^2_{change}=.007$ ,  $p>.05$ ).

A follow-up sequential regression analysis was used to determine if reading fluency predicted improved reading comprehension for students with proficient fluency skill from high-income ( $n=56$ ) backgrounds after controlling for reading vocabulary. Unlike the findings for low-income students, results indicate that reading fluency was found to predict increases in reading comprehension scores of high-income students proficient in fluency skill after controlling for reading vocabulary ( $R^2=.462$ ,  $R^2_{change}=.095$ ,  $p<.05$ ).

## V. SUMMARY, DISCUSSION, AND RECOMMENDATIONS

### Summary of the Study

Part one of this study examined whether or not the DORF measures of reading fluency equally predicted SAT-10 and DRP reading comprehension scores and the relationship between SAT-10 and DRP scores. Bivariate correlations of 129 student participants found the DORF measures to be strong predictors of both SAT-10 and DRP scores and that the SAT-10 and DRP were positively correlated at statistically significant levels. Additionally, correlation comparisons showed no statistical differences in DORF's ability to predict SAT-10 and DRP scores. Moreover, similar findings were found between these assessments for both high-income students not receiving free or reduced lunch (n=36) and low-income students receiving free or reduced lunch (n=62).

Part two of this study explored the relationships between reading fluency, reading vocabulary, and reading comprehension for students with proficient reading fluency skill but differing economic backgrounds. Specifically, this study sought to determine whether or not proficient fluency skill is as strongly associated with low-income students' comprehension achievement as it is for their wealthier peers. Bivariate correlations among 56 high-income and 56 low-income student participants found reading fluency to be a strong predictor of reading comprehension for high-income but not low-income fluent readers. Additionally, correlation comparisons showed that these differences in the

value of using reading fluency for predicting high- and low-income students' reading comprehension scores were statistically significant.

Three one-way analyses of variance were used to compare reading comprehension, reading fluency, and reading vocabulary mean scores for high- and low-income children with proficient reading fluency skill. Statistically significant differences were found between high- and low-income students' reading comprehension, reading fluency, and reading vocabulary scores. Also, income-level produced large effect sizes for reading vocabulary and reading comprehension and medium effect sizes for reading fluency.

Bivariate correlations of high- and low-income student participants found reading vocabulary to be a strong predictor of reading comprehension for high- and low-income fluent readers. Additionally, correlation comparisons between high- and low-income students showed no statistical differences in the value of reading vocabulary as a predictor of reading comprehension scores. Lastly, sequential regression analyses found reading fluency to be a more powerful and statistically significant predictor of reading comprehension scores after controlling for reading vocabulary for fluent high-income students than for fluent low-income students. A review of findings associated with each research question is presented next.

The following research questions directed part one of this study:

#### *First Research Question*

The first research question explored whether or not the DORF measures of reading fluency equally predict SAT-10 and DRP reading comprehension scores and the relationship between SAT-10 and DRP scores. Data analyses established the DORF

measures to be statistically significant predictors of SAT-10 ( $r=.722, p<.001$ ) and DRP ( $r=.683, p<.001$ ) scores. These findings were consistent with previous research evaluating DORF's ability to predict test scores on various standardized measures of comprehension (Barger, 2003; Buck & Torgesen, 2003; McKenna, 2003; Moscovitch, 2004; Shaw & Shaw, 2002). Additionally, no significant differences were discovered between  $r$  comparisons of DORF and the SAT-10, and DORF and the DRP ( $Z=1.00, p>.05$ ). Results from this study's sample, along with previous research findings, provide evidence for DORF's capacity to reliably predict student achievement on multiple tests of reading comprehension achievement. In addition, the statistically significant positive correlation between the SAT-10 and DRP ( $r=.781, p<.001$ ) supports the concurrent validity of these two different measures of reading comprehension.

#### *Second Research Question*

The second research question asked whether or not the DORF measures of reading fluency equally predict SAT-10 and DRP reading comprehension scores for children of differing economic backgrounds. Analyses found the DORF measures to be statistically significant predictors of SAT-10 scores ( $r=.736, p<.001$ ) and DRP scores ( $r=.748, p<.001$ ) for students not receiving free or reduced lunch. No significant differences were discovered between  $r$  comparisons of DORF and the SAT-10, and DORF and the DRP ( $Z=.16, p>.05$ ). In addition, a statistically significant positive correlation between the SAT-10 and DRP ( $r=.847, p<.001$ ) was found. When considering students receiving free or reduced lunch, analyses determined the DORF measures to be statistically significant predictors of SAT-10 scores ( $r=.717, p<.001$ ) and DRP scores ( $r=.717, p<.001$ ). As with students not receiving free or reduced lunch, no significant

differences were realized between  $r$  comparisons of DORF and the SAT-10, and DORF and the DRP ( $Z=0.00, p>.05$ ) for students receiving free or reduced lunch. Similar to their wealthier peers, a statistically significant positive correlation between the SAT-10 and DRP ( $r=.737, p<.001$ ) was found for students receiving free or reduced lunch.

The DORF measures of reading fluency are utilized in the State of Alabama to identify students who may be at risk for learning to read and comprehend texts as early as possible. Although Moscovitch (2004) found DORF to be predictive of SAT-10 reading comprehension scores, no other measure of reading comprehension was employed to validate that this relationship was not specific to the SAT-10. By utilizing the DRP, the results of this study suggest that the predictive nature of DORF does indeed extend beyond the SAT-10. Although this study only examined third-grade students, these findings substantiate the use of DORF to predict reading comprehension test scores for students of varying backgrounds.

The following research questions directed part two of this study:

### *Third Research Question*

The third research question examined whether or not reading fluency equally predicts reading comprehension for high- and low-income students who have attained proficient levels of reading fluency. Data yielded statistically significant positive correlations between reading fluency and reading comprehension for high-income students but not for low-income students. Additionally, comparisons of these independent sample correlations showed reading fluency scores' predictability of reading comprehension scores differed at a statistically significant level between high-income and low-income students with proficient fluency skill ( $Z=2.363, p<.05$ ). On the other hand, a



follow-up investigation found reading fluency statistically significantly predicts reading comprehension for low-income students who are not fluency proficient ( $r=.652, p<.001$ ).

The positive and predictive relationships between reading fluency and comprehension established in several studies, including this one, are based on the notion that fluency provides a link between speed and accuracy of word recognition in the connected text and comprehension (Armbruster et al., 2001). Results from this study, however, appear to support this claim for high-income students but not for low-income students with proficient levels of reading fluency. Pinnell et al. (1995) found elements of reading fluency (such as phrasing, adherence to syntax, and expressiveness) that extend beyond reading rate and accuracy are also associated with higher scores on reading comprehension measures. Results from this study support these findings and suggest that there may be a ceiling effect for the value of using reading rate and accuracy ability as predictors of comprehension for low-income students.

#### *Fourth Research Question*

The fourth research question investigated whether or not achievement differences in reading comprehension, reading fluency, and reading vocabulary exist between high- and low-income students with proficient reading fluency skill. High-income students' achievement was higher for each subtest at statistically significant levels with large effects sizes obtained for reading vocabulary and reading comprehension. A medium rather than large effect size for reading fluency may have been due to the attempt in this study to control for fluency skill by only including participants assessed as fluent readers. Similar to previous research findings (Au, 2000; Desimone, 2001; Guthrie et al., 2003), the results of this study show that students of wealthier backgrounds appear to have

higher levels of achievement on measures of reading comprehension than students from poorer backgrounds. Furthermore, it does not appear that reading with proficient speed and accuracy, as measured by DORF, is a reliable predictor of low-income students' reading comprehension achievement.

The simple view of reading suggests reading comprehension is the product of decoding skill and listening comprehension ( $R = D \times C$ ). Utilizing this model, this study's findings suggest that when decoding skill is proficient, the higher performance on measures of reading comprehension may be dependent on listening comprehension skill and in-depth vocabulary knowledge. Therefore, other skills such as vocabulary acquisition that extend beyond reading rate and accuracy may be associated with low-income students' inability to comprehend text equally to their high-income peers.

#### *Fifth Research Question*

The fifth research question evaluated whether or not reading vocabulary equally predicts reading comprehension for students with proficient reading fluency skill but differing economic backgrounds. Analyses discovered reading vocabulary to be a statistically significant predictor of reading comprehension for both high- ( $r=.606$ ,  $p<.001$ ) and low-income ( $r=.677$ ,  $p<.001$ ) students. Additionally, comparisons of these independent sample correlations showed reading vocabulary's predictive nature in relation to reading comprehension did not differ at a statistically significant level between high-income and low-income students with proficient levels of fluency ( $Z=.619$ ,  $p>.05$ ). These findings imply that regardless of income level, background vocabulary knowledge is related to reading comprehension and, in turn, support the premise that when students

are given vocabulary instruction designed to support deep understanding of word meanings, their text comprehension may improve (McKeown et al., 1985).

### *Sixth Research Question*

The sixth research question addressed whether or not reading fluency predicts reading comprehension beyond what can be accounted for by vocabulary for students of low-income backgrounds and for their high-income peers. Sequential regression analyses found reading fluency scores did not predict reading comprehension scores for low-income students proficient in reading fluency skill after controlling for reading vocabulary ( $R^2=.465$ ,  $p>.05$ ), and that only seven percent of the variance in comprehension was accounted for by fluency above and beyond the variance predicted by vocabulary ( $R^2_{change}=.007$ ,  $p>.05$ ). Conversely, the analyses in this study indicated that reading fluency had increased predictability for reading comprehension scores of high-income students proficient in fluency skill after controlling for reading vocabulary ( $R^2=.462$ ,  $p<.05$ ), and that ninety-five percent of the variance in comprehension was accounted for by fluency above and beyond the variance predicted by vocabulary ( $R^2_{change}=.095$ ,  $p>.05$ ). Additionally, the data analyses for students with proficient fluency skill indicated the following: 1) Reading fluency achievement was a better predictor of reading comprehension for high-income than for low-income students at a statistically significant level; 2) ceiling effects existed for reading rate and accuracy as predictors of reading comprehension for low-income students; 3) differences in mean comprehension and vocabulary scores were statistically significant and greater for high-income students than low-income students; and 4) statistically equivalent positive

relationships existed between reading vocabulary and reading comprehension for both high- and low-income students.

Data suggest that reading fluency, at least as measured by reading rate and accuracy, does not predict reading comprehension as well for students of low-income backgrounds as it does for their high-income peers. Furthermore, there appears to be a point at which rapid and automatic word recognition becomes limited as a predictor of reading comprehension. These results, like those of many other studies, indicate that students must be able to read quickly and accurately, but at some point, they must also have deep understandings of word meanings to fully comprehend written text. This study investigated students within schools on opposite ends of the economic continuum, those from affluent communities and those from low-income communities, in order to control for oral vocabulary that is acquired during encounters with other students. Considering this study's sample of high- and low-income fluency proficient students, it seems that the relative weakness of low-income students' background vocabulary, along with other unidentified factors, may have a more powerful influence on their inability to perform as well as their wealthier peers on tests of reading comprehension than the levels of reading fluency they attain.

#### Discussion of Findings

To depict what reading involves, it appears that a more complex equation than reading equals the product of decoding skill and listening comprehension (Gough & Tunmer, 1986) is required. Certainly, background vocabulary knowledge and the understanding of oral language structures associated with listening comprehension are critical to reading; however, there may be much more to processing meaning for known

vocabulary in reading printed language than listening to and understanding those same words in the speech stream. Reading comprehension is probably more complex and challenging than listening comprehension. For example, readers must be able to discern appropriate stress patterns or pronunciations and the meanings for words such as *con-tent*/*con*'-tent, *pro-duce*/*prod*'-uce, and *re-cord*/*rec*'-ord. Reading and comprehending these words in written language requires the construction of meaning and pronunciation across printed phrases, sentences, or paragraphs, increasing the complexity of processing and comprehending beyond what is required when those words are spoken and heard.

The effortless and automatic recognition of words is essential, but so are other aspects of reading fluency, including (a) knowledge and applications of syntax; (b) expression; and (c) attention to the prosodic features such as intonation, stress, and pausing (Pinnell et al., 1995). A more complex view of reading would also include knowledge of text structures and the awareness of the language seen in expository and narrative texts. Moreover, proficient readers think actively when reading by having clear reading goals, previewing text, making predictions, integrating prior knowledge, and monitoring their understanding of text (Duke & Pearson, 2002). A more complex view of reading is needed to account for not only all features of reading fluency and the knowledge of text structures and language, but also for the metacognitive and comprehension strategies that skillful readers possess and actively use.

#### *Limitations of the Study*

1. The absolute validity of pre-existing data collected from school systems cannot be assured. Inconsistencies in the administration and scoring of all assessments, as well as errors in students' recording of answers on the SAT-10 and DRP may exist.

2. The utilized measure of reading fluency only accounts for reading rate and accuracy but not for all features of reading fluency.

3. The relationships measured in correlational research may be affected by unidentified variables. For this reason, no claim of causation between variables can be assumed.

4. This study only analyzed test data from third-graders living in the state of Alabama; therefore, the identified relationships are not generalizable to the world's student population as a whole.

5. Correlations among DORF, SAT-10, and DRP scores were established only for these measures when administered to third-grade students, so the value of DORF scores in the first and second grades for predicting reading comprehension on the SAT-10 and DRP cannot be assumed.

6. All student scores are subject to a phenomenon called regression towards the mean. Due to the fact that participants were grouped by reading fluency scores on a test, retest scores can be expected to be closer to the mean than they were on the original test.

7. Levene's test of homogeneity found one-way analyses of variance for reading comprehension, reading vocabulary, and reading fluency between high- and low-income participants to be statistically significant. Although caution should be used interpreting results when the assumption of variance is violated, analyses of variance results were robust to this assumption and *F* values (estimates of a population variance based on the information in two or more random samples) were not inaccurate enough to warrant concern. However, the results should be used acknowledging the lack of homogeneity of variance as a limitation.

### *Need for Further Research*

1. In this study, the DORF measures of reading fluency were found to be statistically equivalent predictors of SAT-10 and DRP comprehension achievement scores. Due to the fact that DORF only measures reading rate and accuracy, research utilizing a more complete measure of reading fluency is needed to determine whether or not other aspects of reading fluency predict comprehension achievement beyond what is accounted for by rate and accuracy.

2. The purpose of utilizing the DORF measures of reading fluency is to identify students with insufficient fluency so that intervention and instruction may be provided to prevent difficulties developing reading comprehension. However, it is possible that DORF data does not provide information beyond what teachers already know about their students. An investigation to determine the accuracy of teacher perceptions of student fluency achievement and the relationship of those perceptions to DORF fluency, and SAT-10 and DRP comprehension scores would help answer this question.

3. Investigations that help researchers, investigators, and teachers construct a comprehensive view of reading that provides a more detailed map of the processes underlying reading comprehension are required. Further research addressing the impact of reading fluency – including phrasing, expressiveness, and adherence to syntax, above and beyond reading rate and accuracy in relation to reading comprehension – is needed. Moreover, more thorough studies that create a better understanding of the extent to which reading fluency, reading comprehension, and reading vocabulary influence each other also will be valuable.

4. Studies developed to determine how and to what degree high-income students utilize metacognitive and comprehension strategy skills to a greater extent than their low-income peers is necessary. Also, research investigating instructional strategies that develop these skills is justified.

5. This study presented correlational evidence that reading fluency skill alone is not sufficient as a predictor of reading comprehension and should not be the only factor considered in efforts to help low-income students close the reading achievement gap. For this reason, experimental research analyzing the impact of vocabulary instruction on low-income but fluent readers' comprehension is warranted.

6. Research with a larger and more representative sample of the population and data for high- and low-income students as they move from the first through third grades is needed to test and extend this study's findings.

#### *Educational Implications and Recommendations*

1. Results of this study suggest the DORF measures of reading fluency are adequate predictors of comprehension for many students, regardless of ethnicity or economic background. This evidence supports the use of DORF to identify students who have reading difficulties and, in turn, may require future academic progress monitoring. Teachers should be adequately trained in the administration, scoring, and interpretation of the DORF measures.

2. Reading rate and accuracy, as measured by DORF, were found to predict reading comprehension for high-income but not low-income students with proficient reading fluency skill. This may be due to a multitude of factors, including achievement differences in the other features of reading fluency, reading vocabulary, metacognition,



and comprehension strategy use. Teachers need to realize the complexity of how these skills interrelate and be careful not to overemphasize any one aspect of literacy.

3. High-income students were found to have greater reading fluency, reading vocabulary, and reading comprehension achievement than their low-income peers even though both groups were considered to have achieved benchmarks for proficient fluency skill. Also, the relationships between reading vocabulary and reading comprehension were statistically equivalent for both wealthy and poor students. Evidence suggests that early vocabulary instruction that helps children build deep understanding of word meanings may benefit students of poorer backgrounds and help them achieve higher levels of reading comprehension as much or more than, and certainly along with, fluency instruction to increase speed and accuracy.

4. Teachers need to communicate to parents research-proven strategies and practices that encourage the development of reading comprehension. Parents should be made aware of the importance of decontextualized conversations with language that includes specific and rich vocabulary instead of general referents that are tied to immediate contexts. Parents, especially those of lower incomes, also need greater access to children's books and opportunities to learn how to use those books to expose children to higher-order, more abstract vocabulary and language structures than is typical of abbreviated, rudimentary, and highly contextualized everyday communications.

5. School administrators and policy-makers must realize the scope and limitations of early literacy assessments. They must understand which skills such assessments intend to measure and the importance of those skills to beginning reading. Furthermore, administrators must communicate that the end-goal of reading is not any

one specific skill, but a culmination of skills and strategies that allow proficient readers to gain meaning from written text.

## REFERENCES

- Allington, R.L. (1983). Fluency: The most neglected goal in reading instruction. *The Reading Teacher*, 36, 556-561.
- Anderson, R.C., & Freebody, P. (1981). Vocabulary knowledge. In Arena Press (2001). *Reading research anthology: The why of reading instruction*. Navato, California.
- Anderson, R.C., Hiebert, E.H., Scott, J.A., & Wilkinson, I.A.G. (1985). *Becoming a nation of readers*. Washington, DC: National Institute of Education.
- Armbruster, B. B., Lehr, F., & Osborn, J. (2001). *Put reading first: The research building blocks for teaching children to read*. Washington DC: Center for Improvement of Early Reading Achievement.
- Au, K.H. (2000). A multi-cultural perspective on policies for improving literacy achievement: Equity and excellence. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson & R. Barr (Eds.), *Handbook of reading research: Vol. 4* (pp. 835-851). Mahwah, NJ: Erlbaum.
- Baker, L., & Sher, D. (2002). Beginning readers' motivation for reading in relation to parental beliefs and home reading experiences. *Reading Psychology*, 23, 239-268.
- Barger, J. (2003). *Comparing the DIBELS Oral Reading Fluency indicator and the North Carolina end of grade reading assessment*. (Technical Report). Asheville, NC: North Carolina Teacher Academy.

- Becker, W.C. (1977). Teaching reading and language to the disadvantaged: What we have learned from field research. *Harvard Educational Review*, 47, 518-543.
- Berk, R.A. (1998). Review of the Stanford Achievement Test, Ninth Edition. *Mental Measurements Yearbook*, 13. Retrieved October 13, 2003, from Mental Measurements Yearbook database.
- Bowey, J.A. (1995). Socioeconomic status differences in preschool phonological sensitivity and first grade reading achievement. *Journal of Educational Psychology*, 87(3), 476-487.
- Bradley, L. & Bryant, P.E. (1983). Categorizing sounds and learning to read: A causal connection. *Nature*, 301, 419-421.
- Buck, J. & Torgesen, J. (2003). *The relationship between performance on a measure of oral reading fluency and performance on the Florida comprehensive assessment test*. (FCRR Technical Report #1) Tallahassee, FL: Florida Center for Reading Research.
- Carnine, D.W. (2000). *A consortium for evidence in education (CEE)*. Unpublished manuscript.
- Catts, H.W., Hogan, T.P., Adolf, S.M., & Barth, A.E. (2003). *The simple view of reading: Changes over time* (Language and Reading Disorders Laboratory at the University of Kansas). Retrieved March 3, 2005, from <http://www.ku.edu/~splh/Catts/poster7.pdf>
- Chall, J.S. (1967). Learning to read: The great debate. In Marilyn Adams (Ed.), *Beginning to read: Thinking and learning about print* (p. 36). Chicago: University of Illinois, Center for the Study of Reading.

- Chall, J.S., Jacobs, V.A., & Baldwin, L.E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Cook, R.G. (2003). *The utility of DIBELS as a curriculum based measurement in relation to reading proficiency on high stakes tests*. Unpublished master's thesis, Marshall University, West Virginia.
- Desimone, L. (2001). Linking parent involvement with student achievement: Do race and income matter? *The Journal of Educational Research*, 93(1), 11-30.
- Duncan, L.G., & Seymour, P.H. (2000). Socio-economic differences in foundation-level literacy. *British Journal of Psychology*, 91, 145-166.
- Duke N.K., & Pearson, P.D. (2002). Effective practices for developing reading Comprehension. In T.S. Stewart, Early literacy instruction in the climate of No child left behind, *The Reading Teacher*, 2004, 57(4).
- Durkin, D. (1979). What classroom observations reveal about reading comprehension Instruction. *Reading Research Quarterly*, 14, 481-533.
- Durkin, D. (1989). *Teaching them to read* (5<sup>th</sup> ed.). Boston, MA: Allyn and Bacon.
- Dynamic Indicators of Basic Early Literacy Skills* (2000-2003). Retrieved October 1, 2003, from <http://dibels.uoregon.edu/>
- Elliott, J., Lee, S., & Tollefson, N. (2001). A reliability and validity study of the Dynamic Indicators of Basic Early Literacy Skills – modified. *School Psychology Review*, 30, 33-49.
- Fuchs, L.S., Fuchs, D. & Hamlett, C.M. (1989). Effects of instrumental use of

- curriculum based measurement to enhance instructional programs. *Remedial and Special Education, 10*(2), 43-52.
- Fuchs, L.S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9*, 20-28.
- Galagan, J.E. (1985). Psycho educational testing; Turn out the lights, the party's over. *Exceptional Children, 52*, 288-289.
- Good, R. H. & Kaminski, R. A. (2002). *DIBELS Oral Reading Fluency Passages for First through Third Grades (Technical Report No. 10)*. Eugene, OR: University of Oregon.
- Good, R.H. & Jefferson, G. (1998). *Contemporary perspectives on Curriculum-based measurement validity*. New York: Guilford.
- Good, R. H., Simmons, D., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*, 257-288.
- Good, R. H., Simmons, D., & Smith, S. (1998). Effective academic interventions in the United States: Evaluating and enhancing the acquisition of early reading skills. *School Psychology Review, 27*(1), 45-56.
- Green, F.J. (2001). Review of the degrees of reading power (Revised). *Mental Measurements Yearbook, 14*. Retrieved September 22, 2003, from Mental Measurements Yearbook database.
- Gough, J.B., & Tunmer, W.E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6-10.

- Guthrie, J.T., & Greany, V. (1991). Literacy acts. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research: Vol. 3* (pp. 835-851). Mahwah, NJ: Erlbaum.
- Harcourt Educational Measurement. (2003). *Stanford Achievement Test – Tenth Edition Spring Technical Data Report*. US: Author.
- Haycock, K. (2003). Closing the Disparity Gap. *Teacher Librarian*, 30(4), 38.
- Heath, S.B. (1991). The sense of being literate: Historical and cross-cultural features. In McCormick, S. *Instructing students who have literacy problems* (p.43). New Jersey: Pearson Education.
- Hintze, J.M., Ryan, A.L., & Stoner, G. (2002). *Concurrent validity and diagnostic accuracy of the Dynamic Indicators of Basic Early Literacy Skills and the Comprehensive Test of Phonological Processing*. Amherst, MA: University of Massachusetts.
- Huck, S.W. (2004). *Reading statistics and research*. Boston, MA: Pearson Education, Inc.
- Jenkins, J.R., & Dixon, R. (1983). Vocabulary learning. *Contemporary Educational Psychology* 8, 237-260.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80, 437-447.
- Kaminski, R. A., Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review*, 25, 215-227.
- Keppel, G., & Zedeck, S. (1989). *Data analysis for research designs: Analysis of variance and multiple regression/correlation approaches*. New York: W.H.

Freeman & Company.

- Knutson, N., & Shinn, M. (1991). Curriculum-based measurement: Conceptual underpinnings and integration into problem solving assessment. *Journal of School Psychology, 29*, 371-393.
- Kozol, J. (1991). *Savage inequalities: Children in America's schools*. New York: Crown.
- Liberman, I.Y., & Liberman, A.M. (1992). Whole language versus code emphasis: Underlying assumptions and their implications for reading instruction. In B.A. Murray (Ed.) *Gaining alphabetic insight: Is phoneme manipulation skill or identity knowledge causal?* pp. 461-475.
- Margolis, H. (2001). Review of the Degrees of Reading Power (Revised). *Mental Measurements Yearbook, 14*. Retrieved September 22, 2003, from Mental Measurements Yearbook database.
- Marston, D.B. (1989). A curriculum-based measurement approach to assessing academic performance: What is it and why do it? In M.R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18-78). New York: The Guildford Press.
- McCormick, S. (2003). *Instructing students who have literacy problems* (p.43). New Jersey: Pearson Education.
- McKenna, M. K. (2003). *Assessing reading comprehension: The relation between DIBELS Oral Reading Fluency, DIBELS Retell Fluency and Oregon State Assessment scores*. Unpublished master's thesis, University of Oregon.
- McKeown, M/G., Beck, I.L., Omanson, R.C., & Perfetti C.A. (1983). The effects of long-



- term vocabulary instruction on reading comprehension: A replication. *Journal of Reading Behavior*, 15(1), 3-18.
- Meng, X., Rosenthal, R., & Rubin, D.B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, 3(1), 172-175.
- Merriam-Webster Online* (2005). Retrieved March 5, 2005, from <http://www.m-w.com/>
- Molfese, V.J., Modglin, A., & Molfese, D.L. (2003). The role of environment in the development of reading skills: A longitudinal study of preschool and school-age measures. *Journal of Learning Disabilities*, 36(1), 59-67.
- Moscovitch, E. (2004). *Evaluation of the Alabama Reading Initiative*. Cape Ann Economics.
- Murray, B.A. (1998). Gaining alphabetic insight: Is phoneme manipulation skill or identity knowledge causal? *Journal of Educational Psychology*, 90(3), 461-475.
- Nagy, W.E., Anderson, R.C., & Herman, P.A. (1987). Learning words from context during normal reading. *American Educational Research Journal*, 24, 237-270.
- Nagy, W.E., Herman, P.A., & Anderson, R.C. (1987). Learning words from context. *Reading Research Quarterly*, 20, 233-253.
- National Institute for Literacy* (2005). Retrieved March 2, 2005, from <http://www.nifl.gov/partnershipforreading/glossary/glossary.html>
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

- Pinnell, G.S., Pikulski, J.J., Wixson, K.K., Campbell, J.R., Gough, P.B., & Beatty, A.S. (1995). *Listening to children read aloud*. Washington DC: Office of Educational Research and Improvement, U.S. Department of Education.
- Rasinski, T.V. (2000). Speed does matter in reading. *The Reading Teacher*, 52(2), 146-151.
- Ratekin, N. (1978). A comparison of reading achievement among three racial groups using standard reading materials. In S. McCormick (Ed.), *Instructing students who have literacy problems* (p. 44). Upper Saddle River, NJ: Pearson Education.
- Ritchey, K.D. (2004). From letter names to word reading: The development of reading in kindergarten. *Reading Research Quarterly*, 39(4), 374-376.
- Robbins, C., & Ehri, L.C. (1994). Reading storybooks to kindergartners helps them learn new vocabulary words. *Journal of Educational Psychology*, 86, 54-64.
- Scarborough, H.S., & Brady, S.A. (2002). Toward a common terminology for talking about speech and reading: A glossary of the “phon” words and some related terms. *Journal of Literary Research*, 34(3), 299-336.
- Senechal, M., Fevre, J., Thomas, E.M., & Daley, K.E. (1998). Differential effects of home literacy experiences on the development of oral and written language. *Reading Research Quarterly*, 33, 96-116.
- Share, D., Jorm, A., Maclean, R., Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Educational Psychology*, 76, 1309-1324.
- Shaw, D., Shaw, R. (2002). *DIBELS Oral Reading Fluency-Based Indicators of Third Grade Reading Skills for Colorado State Assessment Program (CSAP)*. (Technical Report). Eugene, OR: University of Oregon.

- Speece, D.L., Mills, C., Ritchey, K.D., & Hamilman, E. (2003). Initial evidence that letter fluency tasks are valid indicators of early reading skill. *The Journal of Special Education, 36*, 223-233.
- Stahl, S.A., & Fairbanks, M.M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*(1), 72-110.
- Stanovich, K.E. (1986). Matthew effects on reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407.
- Steiger, J.H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*, 245-251.
- Tindal, G., Marston, D., & Deno, S.L. (1983). *The reliability of direct and repeated measurement* (Research Rep. No. 109). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.
- Touchstone Applied Science Associates, Inc. (1995). *DRP Handbook: G & H Test Forms*. Brewster, NY: Author.
- Warrington, R. (2003). *The effects of revised directions for Dynamic Indicators of Basic Early Literacy Skills Oral Reading Fluency (DORF): Oral Retell Fluency*, Unpublished master's thesis, University of Oregon.
- Winfield, L. (1996). Teachers beliefs toward academically at-risk students in inner-city Schools. In McCormick, S. *Instructing students who have literacy problems* (p. 43). New Jersey: Pearson Education.

## APPENDICES

APPENDIX A  
CONFIDENTIALITY STATEMENT

Date

Name

Position

School System

Address

Name,

This correspondence concerns a request for your assistance in acquiring School System third grade STAR and DRP Reading data for the 2003-2004 school year. For my dissertation requirements, I am conducting a correlational study between third grade DIBELS Oral Reading Fluency, SAT-10 Reading, DRP Reading, and STAR Reading scores. Originally, I planned to utilize DRP comprehension scores provided by the State Department of Education, but they have informed me that only a few schools administered the DRP. For this reason, I have decided to use STAR Reading scores as a secondary measure of reading achievement in addition to any available DRP data. I have already been allowed access to an unnamed city school system's STAR Reading data and was hoping to gain the same access from a rural or city school system such as yours. I have been granted permission from the Alabama State Department of Education to obtain third-graders' 2003/2004 DIBELS Oral Reading Fluency (DORF) scores, and reading vocabulary and reading comprehension scores on the Stanford Achievement Test – Tenth Edition (SAT-10).

There exists a body of research that suggests that the third grade DORF assessment is a relatively strong predictor of outcomes on the SAT-10 for most populations of children. On the other hand, several studies have shown that this relationship is not as powerful for children of lower SES or minority backgrounds. These relationships often lead teachers to believe that students are reading well, but not comprehending. My belief is that the SAT-10 measures of vocabulary and comprehension are not as sensitive to comprehension gains for these populations of children. By utilizing STAR Reading and DRP data, which have been shown to be more sensitive to students of minority and poorer backgrounds, I hope to show that other measures of comprehension may be more indicative of certain children's reading achievement than the SAT-10. I will also be considering other characteristics such as gender and special education services.

Important to note is that Dr. Katherine Mitchell has made it clear that all assessment data is to be collected, compiled, and then re-coded so that no identifying student information is known during analysis and reporting. This procedure will be utilized to secure the privacy of sensitive student data. I believe that demographic information such as race, gender, and special education services will be provided through my access to the State's DIBELS database. My request of you is to provide the 2003/2004 third grade STAR Reading and DRP reports (if available).

If you have any questions, please contact me. If not, I appreciate your consideration

Sincerely,

Timon Paleologos/Edna Brabham  
(334) 745-2151 / 750-8600  
paleotm@auburn.edu

APPENDIX B  
DATA RELEASE STATEMENT



September 28, 2004

M E M O R A N D U M

TO: Whom it May Concern

FROM: Name  
Title  
School System

RE: Correlational Study to be conducted by Timon Paleologos

The following letter is in reference to the correlational study to be conducted by Timon Paleologos using secured test data. The XYZ School System will grant Mr. Paleologos access to 2003/2004 third grade STAR and DRP Reading data by means of Floppy Disk, CD-ROM, or printed hard copy. Mr. Paleologos will only have access to data containing student identifiers until all data is re-coded in anonymous fashion. Mr. Paleologos will only be granted access to the previously mentioned assessment scores after he has signed a waiver of confidentiality ensuring all data will be recorded and reported as anonymous with no identifying information.

Additionally, Mr. Paleologos will not be provided copies of any assessments as they are secured examinations that cannot be distributed.

Sincerely,

Name  
Title  
School System