

THE UTILITY OF EARLY VOCABULARY SCREENING FOR AT-RISK READERS

by

JONATHAN BENNETT POTTER

A DISSERTATION

Presented to the Department of Special Education
and Clinical Sciences
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2008

University of Oregon Graduate School

Confirmation of Approval and Acceptance of Dissertation prepared by:

Jonathan Potter

Title:

"The Utility of Early Vocabulary Screening for At-Risk Readers"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the degree in the Department of Special Education and Clinical Sciences by:

Roland Good, Chairperson, Special Education and Clinical Sciences
Kenneth Merrell, Member, Special Education and Clinical Sciences
Elizabeth Harn, Member, Special Education and Clinical Sciences
Helen Neville, Outside Member, Psychology

and Richard Linton, Vice President for Research and Graduate Studies/Dean of the Graduate School for the University of Oregon.

June 14, 2008

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

An Abstract of the Dissertation of
Jonathan Bennett Potter for the degree of Doctor of Philosophy
in the Department of Special Education and Clinical Sciences
to be taken June 2008

Title: THE UTILITY OF EARLY VOCABULARY SCREENING FOR AT-RISK
READERS

Approved: _____
Dr. Roland H. Good, III

Early literacy instruction provided in kindergarten through third grade plays an important role in the development of literacy skills (Adams, 1990; National Reading Panel, 2000). Children who do not develop early literacy skills within this time frame do not have a high likelihood of catching up to their normally achieving peers (Good, Simmons, & Smith, 1998). Schoolwide early literacy assessment data helps to inform this instruction, which is most effective when it is driven by the five core components of reading. These core components include phonemic awareness, phonics, accuracy and fluency with connected text, vocabulary, and comprehension (National Reading Panel, 2000). Current schoolwide literacy screening systems have been effective in identifying those students who may struggle in the areas of phonemic awareness, phonics, and

accuracy and fluency with connected text. However, despite evidence that early vocabulary skills are predictive of later reading comprehension outcomes in grades three and above, early vocabulary assessments are not widely utilized in schools (Scarborough, 1998).

The present study examined the utility of adding a brief measure of vocabulary, the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Word Use Fluency (WUF) measure, to the other commonly used DIBELS screening measures of phonics, phonemic awareness, and oral reading fluency. Using a longitudinal sample of 29 current third grade students, data analysis examined the predictive utility of kindergarten and first grade WUF scores with third grade reading comprehension and vocabulary outcomes. The relation between WUF scores and the DIBELS comprehension measure, the Retell Fluency (RTF), was also explored to help examine the relation between reading comprehension and vocabulary over time.

Results indicated that kindergarten WUF scores were predictive of third grade reading comprehension and vocabulary. First grade WUF scores were not predictive of third grade reading comprehension and vocabulary. These results are discussed with the limited sample size and specific data collection procedures in mind. An examination of the relation between vocabulary (WUF) and comprehension (RTF) revealed an increasingly stronger relation over time. Limitations of the study, in addition to implications for practice and future research are discussed.

CURRICULUM VITAE

NAME OF AUTHOR: Jonathan Bennett Potter

PLACE OF BIRTH: Eugene, Oregon

DATE OF BIRTH: May 19, 1979

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
University of Denver

DEGREES AWARDED:

Doctor of Philosophy, 2008, University of Oregon
Master of Science, 2006, University of Oregon
Bachelor of Arts, 2001, University of Denver

AREAS OF SPECIAL INTEREST:

Early Literacy Instruction and Assessment
Response to Intervention
Teacher Training and Instructional Consultation

PROFESSIONAL EXPERIENCE:

School Psychologist, Heartland AEA 11, Johnston, Iowa, August, 2007 to present

Graduate Teaching Fellow, University of Oregon, Eugene, Oregon, September
2005 to June 2006

GRANTS, AWARDS AND HONORS:

DIBELS Student Support Award Recipient, 2006

Project INTEGRATE Training Grant, University of Oregon, 2006-2007

VANGUARD Training Grant, University of Oregon, 2004-2005

ADHD Training Grant, University of Oregon, 2003-2004

ACKNOWLEDGMENTS

I would like to gratefully acknowledge the commitment and willingness to participate of the administrators, teachers, and students in Wyoming. I would like to thank Dr. David Chard for his assistance in reviewing the dissertation proposal and providing feedback that was utilized in the preparation of the final manuscript. I would also like to thank Dr. Roland Good for his assistance in developing this research project and providing ongoing support in the development of the final manuscript. I want to express my gratitude to all of the committee members for their time and support in the development and review of the final project. I would also like to give special thanks to all of the members of my cohort, whom I learned so much from in my studies and in life. In particular, I would like to thank Rohanna Buchanan and Amy Kauffman for being the last remaining cohort members whom I could share the graduate experience with for the full five years. I would have never been able to complete my graduate studies without their support and friendship. Finally, I would like to thank all my family and friends for their love and support. This investigation was supported by funding provided through the DIBELS Student Support Award in the School Psychology program at the University of Oregon.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
Development of Reading Skills	3
Comprehending Written Text	7
The Changing Relation Between Oral Language and Reading Comprehension	9
Vocabulary	13
Unanswered Questions	15
Purpose of the Study	16
II. LITERATURE REVIEW	18
Importance of Early Identification and Intervention	18
Oral Language and Reading	21
Early Oral Language and Early Reading	22
Early Oral Language and Later Oral Language	26
Early Oral Language and Later Reading	28
Later Oral Language and Later Reading	36
The Importance of Vocabulary	39
Voabulary Instruction/Intervention	42
Purpose of the Current Study	47
III. METHODOLOGY	48
Participants and Setting	49
Measures	49
DIBELS Measures	50
DIBELS Phoneme Segmentation Fluency (PSF)	50
DIBELS Nonsense Word Fluency (NWF)	50
DIBELS Oral Reading Fluency (DORF)	51
DIBELS Retell Fluency (RTF)	52

Chapter	Page
DIBELS Word Use Fluency (WUF)	52
Gates-MacGinitie Reading Test, Fourth Edition (GMRT-4)	53
Procedures	54
Human Subjects Review	54
Participant Selection/Consent	54
Data Collection	55
Data Analysis	57
Research Question 1	57
Research Question 2	57
Research Question 3	58
Research Question 4	59
Summary	60
IV. RESULTS	62
Descriptive Statistics	63
DIBELS Measures	63
Gates-MacGinitie Reading Test-4 (GMRT-4)	66
Longitudinal Continuity	67
Predictive Validity	72
Vocabulary and Reading Comprehension	80
Summary of Results	87
V. DISCUSSION	89
Interpretation of Results	90
WUF as a Measure of Vocabulary Skills Over Time	90
Additional Benefit of the WUF Measure	91
The Relation Between Vocabulary and Comprehension	93
Implications for Practice	94

Chapter	Page
Limitations	97
Technical Adequacy of the WUF	98
Design and Implementation	99
Criterion Measures	100
Directions for Future Research	101
Conclusions	105
APPENDICES	107
A. K AND 1ST GRADE WUF DISTRIBUTIONS	107
B. SCATTERPLOTS	108
REFERENCES	112

LIST OF FIGURES

Figure	Page
1. The changing relation between measurements of oral language, phonological processing, and reading comprehension in kindergarten through third grade.	11

LIST OF TABLES

Table	Page
1. Timeline for Data Collection	56
2. Descriptive Statistics for DIBELS Measures	65
3. Descriptive Statistics for GMRT-4	67
4. Within-Grade WUF Benchmark Correlation Coefficients	68
5. Between-Grade WUF Composite Score Correlation Coefficients	71
6. Predictive, Criterion-related Validity Coefficients for K and First Grade WUF with End-of-Third Grade Criterion Measures	73
7. Hierarchical Multiple Regression Between End-of-K Early Literacy and Language and End-of-Third Grade Reading Outcomes	77
8. Hierarchical Multiple Regression Between Middle-of-First Grade Early Literacy and Language and End-of-Third Grade Reading Outcomes	79
9. Correlation Coefficients Between Concurrent RTF and WUF Scores	81
10. Hierarchical Multiple Regression Between First Grade ORF, WUF and RTF	83
11. Hierarchical Multiple Regression Between Second Grade ORF, WUF, and RTF	84
12. Hierarchical Multiple Regression Between Third Grade ORF, WUF, and RTF	85

CHAPTER I

INTRODUCTION

Educators are charged with providing quality literacy instruction for all children with a very limited amount of resources. Regrettably, current evidence paints a grim picture of student reading outcomes. More than one in six children in grades one through three experience reading difficulty (Kame'enui, 1996). Approximately 36% of fourth grade students and 27% of eighth grade students in the United States do not read well enough to complete grade level work adequately (National Center for Education Statistics, 2005). Additionally, the longer we wait to address these reading problems, the more resistant to change they become as children who start off as poor readers have a high likelihood of remaining poor readers (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Good, Simmons, & Kame'enui, 2001; Good, Simmons, & Smith, 1998; Juel, 1988; National Reading Panel, 2000).

Early research by Juel (1988) found that the probability of a child who is a poor reader in the first grade remaining a poor reader in the fourth grade is .88. Good, Simmons, and Smith (1998) examined the trajectories of reading growth from first through fifth grade. The difference in reading trajectories between poor readers and good readers becomes apparent towards the end of first grade and the gap tends to grow wider with each grade. These differences in literacy skills are apparent even earlier on measures of early literacy skills, before children can even read connected text (Adams, 1990). In

addition to the important early literacy skills of phonological awareness and phonics, deficits in early oral language put children at a significant disadvantage, as these deficits grow larger over time (Hart & Risley, 1995). Children with early oral language and vocabulary deficits are at a higher risk for developing reading difficulties later in their academic careers (Bishop & Adams, 1990; Catts, Fey, Tomblin, & Zhang, 2002). Those children who do not develop initial competency in early literacy skills, in addition to oral language skills, are likely to fall further and further behind their normally achieving peers unless powerful instructional programs are provided to them. This “Matthew Effect”, as it was termed by Stanovich (1986), can be explained by a lack of competence in basic skills impeding the development of more complex reading skills. In addition, children who experience early reading difficulties often develop a strong dislike for reading due to the difficulty of the task. Their dislike can lead to decreased motivation to read, thus exposing them to a significantly smaller amount of text than encountered by good readers (Cunningham & Stanovich, 1998; Juel, 1988).

For children who lack initial early literacy skills, time is a precious resource. We simply do not have time to waste in setting all children on the path to literacy. If children without the necessary skills progress through school unnoticed, without appropriate intervention, it will become more resource intensive and more difficult to improve their reading skills. Given the limited amount of resources provided to schools and our current awareness of the disparities in outcomes between children who are afforded a healthy start and those lacking fundamental initial skills, the need for early intervention is clear. A recent emphasis on all children becoming competent readers has allowed for the

development of systems of schoolwide reading support (Simmons et al., 2002). These systems provide early literacy screening to all students in order to determine which students may need additional reading support. Though these systems of early screening have been effective in identifying many children who are at risk for reading difficulties, many systems have focused on the skills that help children decode written text, such as phonemic awareness, phonics, and accuracy and fluency with connected text. As a result, instructional programs for young children have emphasized these skills, sometimes at the expense of other important skills such as vocabulary acquisition (Beck, McKeown, & Kucan, 2002; Biemiller, 2005). Effective reading instruction should include all the big ideas of reading: phonemic awareness, phonics, accuracy and fluency with connected text, vocabulary, and comprehension (NRP, 2000). Research is needed to expand on these early screening systems and improve their ability to identify children who may struggle in the non-phonological aspects of reading such as vocabulary, oral language, and reading comprehension. The proposed study aims to examine the utility of a non-phonological screening measure. Specifically, this study will examine the utility of adding a measure of vocabulary to currently used early literacy screening systems.

Development of Reading Skills

Reading is a complex construct made up of several important skills, including phonological awareness, alphabetic principle, accuracy and fluency with connected text, vocabulary, and comprehension (Adams, 1990; National Reading Panel, 2000). Students need to master all of these skills in order to fully develop as competent readers. In the

development of reading, children need to acquire certain skills that lay the foundation for more advanced skills. For example, phonological awareness and mastery of the alphabetic principle must be obtained before they are able to fluently read complex texts, which will then impact their ability to comprehend what they are reading (Ehri, 2005; Vellutino, 1991).

The interaction and development of these reading components are described in Chall's (1983) stages of reading development. Within these stages of reading development, Stages 0 through 4 address the initial development of literacy skills during the elementary years. Stage 0, the Prereading Stage, occurs from birth to around age 5 or 6 when children enter school. During this stage, children become competent using and understanding language and begin to develop print awareness. It is during this stage that children's phonological skills and oral vocabularies develop. This stage lays the foundation for the later integration of oral language and decoding into competent reading, as children have not yet begun to fully interact with written text. Next is Stage 1, the Decoding Stage, which occurs during grades 1 and 2. During this stage, children begin to master the alphabetic principle, gaining an understanding of the phonetic structure of written text. Children continue to develop oral vocabulary, although their immature level of decoding proficiency limits the ability to increase their vocabularies through reading. Vocabulary development must then continue through interactions involving oral language, allowing children to initially understand thousands of words more than what they can identify or decode in written text. Estimates of first grade children's oral vocabularies have been around 5,000 words, with word meanings being learned mostly

through oral interactions with adults (Lorge & Chall, 1963). Phonics instruction at this stage plays an important role in helping children's decoding skills catch up with their knowledge of word meanings. Stage 2 is the Ungluing from Print Stage, and occurs in grades 2 and 3. During this stage, children continue to develop their decoding skills, becoming more automatic with the task. They start to recognize patterns of letters and words, allowing them to read more efficiently and begin to focus on the meaning of text. Stage 3, the Reading to Learn Stage, takes place from grades 4 through 8 and denotes an important shift in the purpose of reading. It is at this stage that children begin to focus on reading for comprehension, which becomes a more difficult task due to the increased complexity of the language. It is also at this stage when children begin primarily to learn new word meanings through written text (Nagy & Anderson, 1984). Word recognition and decoding skills typically have been fully developed by this point, allowing children access to unfamiliar words not encountered in verbal interactions. Within these early stages of reading development, both decoding and knowledge of word meanings are important early components of overall reading proficiency. These two components of reading have been described as the medium and message of reading, respectively (Chall, 1983).

In the area of decoding, explicit instruction in the phonological structure of written language has a direct positive impact on future reading achievement, particularly in the early grades (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Scanlon, Vellutino, Small, Fanuele, & Sweeney, 2005; Torgesen et al., 1999). Without the ability to utilize phonic decoding to analyze written text, children almost invariably will not

develop competent reading skills later in school, as they have not mastered the medium of reading (Juel, 1988; Scanlon et al., 2005; Torgesen, 1998). But though mastery of the medium is essential to success in reading, it does not necessarily ensure an understanding of the message contained within those words. The relation between effective phonics-based instruction and a subsequent increase on measures of reading comprehension is not as powerful as the relation between phonics-based instruction and word reading skills (Torgesen, 2002). This relation between phonics-based instruction and reading comprehension also changes over time. The National Reading Panel (2000) reported that the impact of phonics-based instruction on reading comprehension is significant in the earlier grades, particularly kindergarten and first grade. But in grades 2 through 6, there was not a significant relation between phonics-based instruction and reading comprehension. Decoding and word recognition skills directly impact comprehension early on, but as the level of text complexity becomes more advanced around Stage 3, decoding skills are necessary but not sufficient to enable effective comprehension of written material.

The impact of decoding on reading comprehension can be seen in the decreasing relation between concurrent decoding and comprehension and the increasing relation between concurrent listening comprehension and reading comprehension as children get older (Catts, Hogan, & Adlof, 2005; Gough, Hoover, & Peterson, 1996; Yovanoff, Duesbery, Alonzo, & Tindal, 2005). As word reading skills become more established in the middle elementary years, general language comprehension skills play an increasingly larger role in reading comprehension. However, extremely low word reading skills would

still be indicative of low comprehension skills. But in general, we must consider the importance of understanding word meanings in addition to recognizing the words. Early oral language skills have been found to be predictive of future reading performance, independent of decoding and word recognition (Catts, Fey, Zhang, & Tomblin, 1999; Senechal, Ouellette, & Rodney, 2005; Storch & Whitehurst, 2002). Therefore, it is important to examine the overall process of reading comprehension and the individual components that are necessary for readers to effectively extract meaning from written text.

Comprehending Written Text

In order for children, and adults, to comprehend what they read, a number of complex processes must be in place. Perfetti (1999) has outlined a model of written language comprehension that discusses three general processes that contribute to successful comprehension. These three general processes include (1) a visual processing system, (2) a system that encodes the visual input into a linguistic form, and (3) a language comprehension system that operates on the encoded linguistic information. The visual processing system is initially at work as the reader actually looks at the written text and is able to visually discern the shapes of the letters on the page. After receiving the visual input, the second system, or the linguistic encoding system, helps convert that visual input into linguistic information. This process allows for the reader to know that certain shapes represent certain letters, which in turn represent certain auditory sounds. Finally, the language comprehension system attempts to extract meaning from the newly

encoded linguistic information. This is where the reader begins to understand what they read. Breakdowns in any of these three systems can lead to comprehension failure. For children who are just beginning to read, we are most commonly able to identify breakdowns in the second process, the encoding of the visual input into linguistic, or phonological, information (Adams, 1990; Snow, Burns, & Griffin, 1998). If children do not know what the words on the page say, they will not be able to know what the words mean. However knowledge of what the words *say* is not sufficient for knowledge of what the words *mean*. A breakdown in the ability to extract meaning from the phonological information can be just as detrimental to overall comprehension.

Gough and Tunmer's (1986) Simple View of Reading provides another way to conceptualize written language comprehension that is similar to Perfetti's general model. The Simple View of Reading states that there is a multiplicative relation between decoding and linguistic comprehension that results in reading comprehension. This can be represented as $D \times C = R$, where D is decoding, C is linguistic comprehension, and R is reading comprehension. In parallel to Perfetti's model, decoding is the system that encodes the visual input into linguistic information, and linguistic comprehension is the system that acts upon the linguistic information in order to extract meaning. This model also presupposes that the visual processing system is sufficient to enable the reader to effectively receive the visual input. Within this model, if either component fails, or is significantly deficient, reading comprehension will suffer.

The recent development of educational policy has put a premium on scientific-based reading programs, and a focus on the basic skills of reading including phonemic

awareness and phonics (No Child Left Behind, 2001). This focus is an important one as students must grasp the basic skills in order to lay the foundation for higher order reading processes. However within the present models of reading comprehension, the role of general language skills is of equal importance to the ability to decode written text. Within these models, accounting for early oral language skills has important implications for the appropriate development of reading comprehension, as these skills have demonstrated the ability to help predict reading comprehension outcomes (Catts, et al., 1999; Scarborough, 1998; Storch & Whitehurst, 2002). Though phonological measures tapping the decoding aspect of reading have historically been demonstrated to be one of the most powerful predictors of future reading achievement, the addition of early oral language measures may increase the power to predict future outcomes. It is then important to examine the particular relation between oral language and reading comprehension.

The Changing Relation of Oral Language and Reading Comprehension

Within the discussion of oral language and reading comprehension, there are two types of relations to examine. One is the concurrent relation, which tells us the immediate contribution of oral language to the reading process. This is accomplished through comparisons of oral language and reading comprehension measures administered at the same time. The second relation is the predictive relation, which helps us identify students who may be at risk for future reading difficulties. This is accomplished by examining the ability of initial measures of oral language in preschool, kindergarten, and first grade to predict future reading outcomes after first grade.

An examination of the concurrent relation between oral language and reading comprehension reveals a changing pattern over time. In the early grades of kindergarten and first grade, oral language skills and reading comprehension are not highly related (Catts, Hogan, & Adlof, 2005; Gough, Hoover, & Peterson, 1996; Hoover & Tunmer, 1993). During this time, phonological measures and measures of decoding skill account for much of the variability in children's ability to comprehend written text. This is not surprising, as the level of text complexity in early texts does not place a high demand on a student's language comprehension skills. If a student cannot decode in the earliest stages of reading, they have no access to the written material. So the relation between decoding and comprehension is fairly strong. However, as students get older and the level of text difficulty becomes more advanced, students must understand increasingly more complex vocabulary contained within the text. Simply decoding the words on the page may not be enough to fully comprehend the information contained within those words.

Although early oral language skills do not have a powerful concurrent relation with reading comprehension, the predictive relation of these skills to later reading comprehension is compelling. Early oral language measures administered in kindergarten and first grade have demonstrated utility in predicting reading comprehension outcomes around grades 2 and above (Catts, et al., 1999; Roth, Speece, & Cooper, 2002; Senechal, Ouellette, & Rodney, 2005; Storch & Whitehurst, 2002). Early oral language skills, although not immediately related to reading comprehension, still play an extremely important role in the reading process. The development of phonological skills in the early grades is of foremost importance in providing children with the tools to access written

texts. However, the neglect of early oral language skills could have a significant delayed effect on reading comprehension. The hypothesized changing relation between oral language, phonological processing and reading comprehension is depicted in Figure 1.

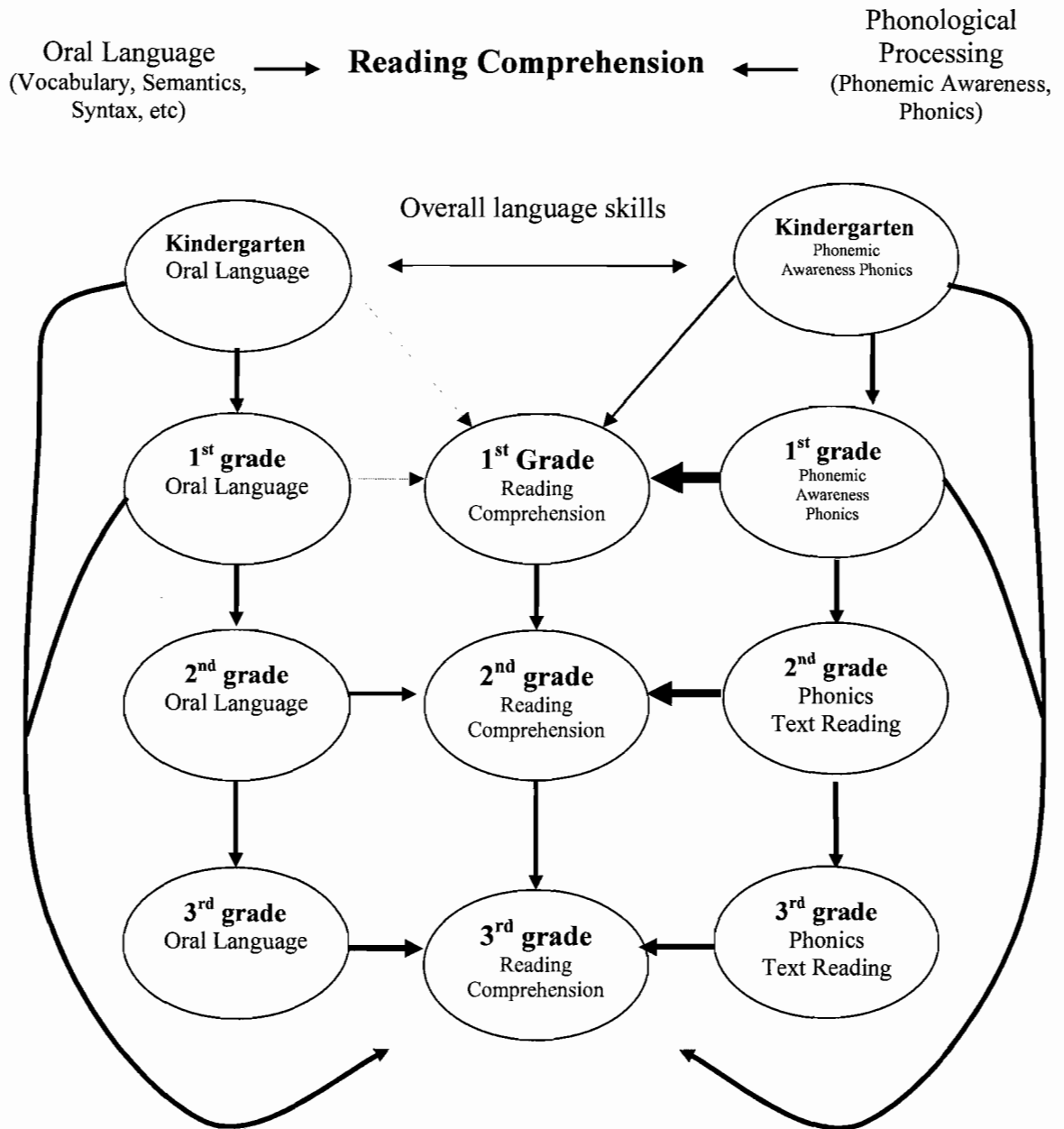


Figure 1. The changing relation between measurements of oral language, phonological processing, and reading comprehension in kindergarten through third grade.

This model was developed with Perfetti's (1999) and Gough and Tunmer's (1986) models of reading comprehension in mind. Within these models, recall that there are two main components that contribute to reading comprehension. These two components are the phonological encoding system that recognizes and organizes the text into linguistic information, and the language comprehension system that extracts meaning from that linguistic information. Due to the complexity of oral and written language, the language comprehension system is affected by a wide array of contributing factors such as syntax, semantics, and general background knowledge. But for the purposes of this discussion, the two main components of phonological encoding and language comprehension are represented as the phonological processing domain and the oral language domain, respectively. The entire model in Figure 1 shows how the relation between measurement of these two main components and reading comprehension changes over time in grades K through 3.

The concurrent relation between oral language and reading comprehension is weak at the beginning of students' academic careers. This relation then grows increasingly stronger as they progress through school, as indicated by the thicker arrows linking oral language to reading comprehension in grades 2 and 3 of Figure 1. Oral language begins to play a more important direct role in grades three and above, when the demands of understanding the language in the text become higher. This coincides with the transition into the Reading to Learn stage in Chall's model of reading development (Chall, 1983). Similarly, the relation between phonological processing and reading comprehension is strong initially, but loses strength as students progress through school,

as indicated by the decreasing thickness of the arrows linking phonological processing skills to reading comprehension from first through third grade of Figure 1. This is due to the increasing contribution of oral language skills needed to deal with the increasing level of text complexity, as well as the typical establishment of phonological skills early on. Within this particular model, the direct impact of oral language on reading comprehension is not initially apparent in the early grades. However, though oral language skills measured in kindergarten and first grade are not highly related to reading comprehension in those grades, they do establish the foundation for later language skills and thus help predict future reading comprehension in third grade and beyond (Senechal, Ouellette, & Rodney, 2005; Storch & Whitehurst, 2002). The predictive nature of early oral language skills in kindergarten and first grade make it an important target of opportunity for early assessment and intervention. The neglect of these early oral language skills may have a hidden effect on reading comprehension that, though not immediately observed, will become apparent in subsequent academic years when children transition from “learning to read” to “reading to learn”.

Vocabulary

Within the broader framework of oral language, vocabulary knowledge has been positively linked to reading comprehension (Anderson & Freebody, 1981; Kamhi, 1989). Early vocabulary measures, used with other early literacy measures, have demonstrated the ability to predict future reading achievement (Catts, et al., 1999; Roth, Speece, & Cooper, 2002; Senechal, Ouellette, & Rodney, 2005; Scarborough, 1998; Storch &

Whitehurst, 2002). Understanding the meaning of the individual words within a text allows the reader to comprehend the text on a higher level. Similar to how basic word identification is a necessary prerequisite for understanding the meaning of a written word, understanding of the meaning of the words in a text is necessary for the application of more advanced reading comprehension strategies. Anderson and Freebody (1981) have proposed that vocabulary knowledge reflects a broader level of background knowledge that allows the reader to understand a text. Readers with larger vocabularies have more background knowledge to draw upon when comprehending a text. Vocabulary difficulty also affects a very basic understanding of text, limiting the ability to recall information and answer explicit comprehension questions (Stahl, Jacobson, Davis, & Davis, 2006). In general, it is difficult to understand the meaning of a text at its most basic level if the reader cannot understand the individual meaning of each word (Kamil, 2004).

Early and frequent measurement of vocabulary provides us with an effective means of identifying students who may be at risk for reading difficulties. In measuring vocabulary, two particular types of measures are most frequently utilized. These types of measures include those tapping receptive vocabulary and expressive vocabulary (Scarborough, 1998). Receptive vocabulary tasks involve identifying objects that names are provided for. Expressive vocabulary tasks involve providing names, definitions, or descriptions of objects or words. Though both tasks provide adequate measurements of vocabulary, there are some indications that expressive tasks may be a better index of children's actual vocabulary (Roth, Speece, & Cooper, 2002; Scarborough, 1998; Senechal, Ouellette, & Rodney, 2005), as the expressive tasks tend to examine the depth

of vocabulary knowledge. Receptive tasks focus on the breadth of a child's vocabulary. Early vocabulary assessments that tap expressive vocabulary skills may be more predictive of future reading comprehension, and thus more useful to educators.

Unanswered Questions

Current early literacy screening systems are in place in many schools, providing measures of letter knowledge, phonemic awareness, phonics, and oral reading fluency. These systems, focusing on the skills necessary to decode written text, are helpful to educators in identifying those children who are at risk for future reading failure and may require additional support. But given available models of reading comprehension (Gough & Tunmer, 1986; Perfetti, 1999) and the current evidence of the contributions of early oral language skills to later reading, it is important to account for oral language skills in predicting future reading outcomes. The hypothesized model of reading comprehension in K-3 places unique importance on the early identification of oral language skills, and particularly vocabulary skills.

Research examining the utility of early vocabulary has shown that these skills are highly predictive of future vocabulary and reading comprehension performance in grades 3 and above (Catts, et al., 1999; Roth, Speece, & Cooper, 2002; Senechal, Ouellette, & Rodney, 2005; Scarborough, 1998; Storch & Whitehurst, 2002). Even after controlling for the predictive nature of measures of phonological skills, letter identification and print awareness, vocabulary measures explain a significant additional amount of unique variance in future reading performance. However, the vocabulary measures typically used

within previous studies are more comprehensive and time consuming, and thus may not be feasible to implement as a screening tool for all children. Using these measures, with other measures of phonological awareness and print identification skills, creates a strain on a school's already limited resources. There is then a need to examine the use of brief and efficient measures of vocabulary that can be administered to all children as part of a comprehensive screening battery.

Purpose of the Study

The purpose of the proposed study is to examine the utility of a brief screening measure that examines children's vocabulary skills in kindergarten and first grade. Specifically, this study will examine whether a brief measure of expressive vocabulary administered in kindergarten and first grade explains performance on reading comprehension outcomes in third grade. This study will also examine the additional benefit of the vocabulary measure in a comprehensive early literacy screening system. Finally, this study will examine the relation between the vocabulary measure and a brief measure of reading comprehension. The proposed study will then address the following research questions:

- (1) What is the longitudinal continuity of a measure of students' skills in defining or using vocabulary words in a sentence from kindergarten through third grade?

- (2) What is the predictive validity of a measure of student's skills in defining or using vocabulary words in a sentence, given in kindergarten and first grade, with a measure of reading comprehension given at the end of third grade?
- (3) Does the vocabulary screening measure add to the predictive validity of commonly utilized early literacy measures such as phonological awareness, phonics, and oral reading fluency in predicting outcomes on an end-of-third grade measure of reading achievement?
- (4) What is the relation between the vocabulary screening measure and a reading comprehension measure in which students retell what they orally read?

CHAPTER II

LITERATURE REVIEW

Importance of Early Identification and Prevention

Given the cost and current prognosis for intervening in the later grades with struggling students, early intervention and prevention of academic and behavioral difficulties has become a recent focus in education. This recent focus is the result of research indicating the effectiveness of these early intervention and prevention efforts in comparison to later intervention after problems have already developed, particularly in the area of literacy (Snow, Burns, & Griffin, 1998). Within this early intervention and prevention framework, systems of early identification have been developed and implemented to identify those students at-risk for future difficulties (Simmons et al., 2002; Snow, Burns, & Griffin, 1998). In the particular area of reading, these systems of early identification focus on the basic early literacy skills that predict future reading achievement such as letter identification, print awareness, phonological awareness, and phonics. Although some single measures, such as phonological awareness ($r = .46$) and letter identification ($r = .52$) have been reported as strong correlates of future reading achievement, a combination of several measures of early literacy and language significantly increases the correlations with future reading outcomes ($r = .75$) (Scarborough, 1998).

As one example of a more comprehensive early literacy screening system, the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 2002) is a useful system for identifying students at-risk for reading difficulties. Current research has demonstrated that the DIBELS measures align with important literacy outcomes in kindergarten through third grade, ensuring a continuum of progress in reading skills (Good, Simmons, & Kame'enui, 2001; Good, Simmons, Kame'enui, Kaminski, & Wallin, 2002). For example, it was reported by Good and colleagues (2002) that 87% of children who met the beginning-of-kindergarten benchmarks for a letter identification and a phonological awareness task met benchmarks in oral reading fluency at the end of first grade. 92% of children who met the end-of-first grade oral reading fluency benchmark met the end-of-second grade oral reading fluency benchmark. 89% of children who met the end-of-second grade oral reading fluency benchmark met the third grade oral reading fluency benchmark. Finally, children who met the end-of-third grade oral reading fluency benchmarks also performed highly on statewide, standardized tests of reading competence. In separate examinations of various state tests, students who reached benchmark on oral reading fluency in third grade were over 85% likely to meet or exceed expectations on statewide tests of reading competence in North Carolina, Florida, Oregon, Colorado, and Arizona (Barger, 2003; Buck & Torgesen, 2003; Good, Simmons, & Kame'enui, 2001; Shaw & Shaw, 2002; Wilson, 2005).

Many of these early screening assessment systems have focused on phonological skills that are related to word-level decoding. These phonological skills have been found to be highly predictive of reading outcomes. In addition, direct teaching of these skills

has been shown to result in better reading outcomes (Scanlon, Vellutino, Small, Fanuele, & Sweeney, 2005; Torgesen et al., 1999; Vandervelden & Siegel, 1997). Though word level decoding skills are critical in the reading process, the development of effective decoding and word recognition skills may not be enough to enable effective reading comprehension, which is the goal of reading (Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003; Nation & Snowling, 1998; Riedel, 2007; Stothard & Hulme, 1992, 1995). This is especially true for children with a history of speech-language impairments (Snowling et al., 2000; Storch & Whitehurst, 2002). For example, Jenkins and colleagues (2003) found that reading words in context, as measured by oral reading fluency, was a much better predictor of reading comprehension than reading decontextualized word lists. It was reported that after accounting for oral reading fluency, word list reading explained only an additional 1% of the variance in reading comprehension. Additionally, they found that after accounting for word list reading skill, comprehension explained an additional 27% of the variance in oral reading fluency. These results suggest that simply being able to decode words quickly is not enough to fully comprehend written text. Some level of language and text comprehension is necessary to achieve adequate fluency.

Studies examining the specific profiles of children with comprehension deficits also highlight the importance of non-phonological skills in reading comprehension (Cain, Oakhill, & Bryant, 2000; Catts, Hogan & Adloff, 2005; Nation, Marshall, & Snowling, 2001; Nation & Snowling, 1998; Stothard & Hulme, 1992, 1995). These studies have identified a group of children with normal phonological skills, but poor comprehension skills. These poor comprehenders exhibited deficits in listening comprehension (Stothard

& Hulme, 1992) and picture naming speed and accuracy (Nation, Marshall, & Snowling, 2001), and identified fewer irregular and low frequency words (Nation & Snowling, 1998) than children with normal phonological and normal comprehension skills. These deficits were apparent despite normal phonological skills in these children, suggesting the importance of the role of non-phonological skills in reading comprehension.

Given the contribution of non-phonological skills to the reading comprehension process, there is a need to examine the specific utility for schools in screening these additional non-phonological skills. Is the addition of extra screening measures to already useful screening batteries, such as DIBELS, worthwhile to schools? Within the framework of the Simple View of Reading (Gough & Tunmer, 1986) and Perfetti's (1999) model of reading comprehension, the particular addition of a brief measure that addresses oral language comprehension holds promise. Within these models, screening for oral language comprehension, in addition to phonological skills, would provide a more comprehensive way of predicting reading comprehension outcomes. This comprehensive screening system would then be even more effective than current screening batteries in identifying additional students who are at-risk for reading difficulties.

Oral Language and Reading

Though the ability to decipher written language is a much more complex language task than understanding oral language, linguistic processes lay the foundation for successful decoding and comprehension of written text (Kamhi & Catts, 1999; Liberman,

Shankweiler, & Liberman, 1989). Models of reading comprehension also highlight the importance of oral language in addition to decoding skills (Gough & Tunmer, 1986; Perfetti, 1999). Given the language basis of literacy, and the importance of oral language in the comprehension process, it is important to examine the particular contribution of oral language to reading. Research examining the impact of a variety of early language and literacy skills has demonstrated important linkages between language and reading across a range of age and grade levels. However, the particular relation of reading and language changes over time. This discussion will focus on the relation of (1) early oral language and early reading, (2) early oral language and later oral language, (3) early oral language and later reading, and (4) later oral language and later reading. Early oral language and reading will be defined within the range of preschool through first grade, which falls within Chall's (1983) Pre-reading and Decoding stages. Later language and reading will be defined within the general range of grade 2 and above, with an emphasis on the transition to the Reading To Learn stage in grades 3 and above (Chall, 1983).

Early Oral Language and Early Reading

Several studies have examined the relation of early oral language and literacy measures at the earliest stages of formal schooling. Within the Pre-reading and Decoding stages, the focus of reading is on the development of initial print awareness, phonological awareness, and an understanding of the alphabetic principle. Children who successfully master the phonetic structure of written language are highly successful in comprehending written text (Catts, Hogan, & Adlof, 2005; Gough, Hoover, & Peterson, 1996). This is

especially true in the early grades, as studies examining the relation between kindergarten and first grade measures of oral language with reading comprehension have generally demonstrated a stronger relation between phonological skills and reading comprehension than between oral language skills and reading comprehension.

Storch and Whitehurst (2002) examined the contribution of oral language and code-related skills to reading outcomes from preschool through fourth grade. Code-related skills were identified as those skills that served as the precursors to formal reading such as print awareness and phonological awareness. These skills were assessed in preschool and kindergarten. Oral language skills measured in preschool and kindergarten included receptive and expressive vocabulary, narrative recall, conceptual knowledge of language, and syntax. The longitudinal design of the study followed a sample of 626 children from families who qualified for Head Start programs. In addition to the early assessment of code-related and oral language skills in preschool and kindergarten, other measures of receptive vocabulary, reading comprehension, and word recognition were also administered in first through fourth grades.

Results indicated that early oral language skills and code-related precursors to reading were related during the Pre-reading stage in preschool and kindergarten. In preschool, oral language skills accounted for 48% of the variance in code-related skills. In kindergarten, this relation was significantly weaker, although still significant, with oral language skills accounting for only 10% of the variance in code-related skills. This finding suggests an interrelation of these two skill domains within the more general domain of overall language skills. However, as children began to transition into the

Decoding stage in first grade, there was no significant relation found between oral language skills and word reading ability. Prior code-related skills were the primary mechanisms of influence on reading ability at this grade levels. The direct relation between early oral language and reading comprehension in first grade was not reported in this study. The results of this study are somewhat limited by the restricted sample of participants. All participants in this study came from low-income family backgrounds. Results should then be cautiously applied to children from middle- and upper-income families.

Senechal, Ouellette, and Rodney (2005) analyzed longitudinal data examining the contribution of kindergarten receptive vocabulary in predicting outcomes on a variety of language and literacy tasks across grade levels. Their sample included 84 children. Regression analyses indicated that kindergarten receptive vocabulary was related, although minimally, to kindergarten phonological awareness, explaining a unique 4% of the variance in the phonological measure in kindergarten and first grade. These results are similar to the findings of Storch and Whitehurst (2002), suggesting an overall language domain that impacts both early oral language and phonological processing. However, receptive vocabulary did not predict any unique additional variance in first grade reading comprehension, after controlling for parent literacy level and kindergarten measures of early literacy skills, listening comprehension, and phonological awareness. Kindergarten early literacy skills included measures of letter naming and simple decoding and spelling.

Share and Leikin (2004) examined the contribution of early oral language and literacy skills in kindergarten to reading in first grade in a sample of 454 children from

varying SES backgrounds. A semantic-syntactic oral language composite was created combining a measure of vocabulary and a measure of syntax in which children matched a spoken sentence to a picture. These measures were administered in kindergarten with a measure of phonemic awareness. In first grade, measures of decontextualized word list reading, reading accuracy with connected text, and reading comprehension were administered.

Regression models were used to examine the contributions of kindergarten phonemic awareness and the oral language composite to the first grade reading measures. In the model predicting decontextualized word list reading, phonemic awareness predicted the most unique variance (12.8%) after controlling for age, gender, IQ and SES. The semantic-syntactic oral language composite did not explain a significant amount of unique variance at less than 1%. Results were similar for reading accuracy with connected text as phonemic awareness predicted a unique 9% of the variance and the semantic-syntactic oral language composite predicted a unique 2%. The pattern of results was different for reading comprehension. Within the model predicting reading comprehension, semantic-syntactic oral language predicted a unique 6%, with phonemic awareness predicting a unique 6.3% of the variance.

A longitudinal study of oral language and reading by Roth, Speece, and Cooper (2002) found similar results. In this study, the researchers administered several oral language measures, print awareness measures, and phonological measures in kindergarten to 39 children and examined the relation of these measures to measures of word reading and reading comprehension administered in first grade. In the regression

model predicting first grade reading comprehension, kindergarten expressive oral vocabulary explained only a small but significant 3% of the variance, after controlling for family literacy level and other early literacy variables such as print awareness and the ability to provide narrative descriptions of a familiar story. Oral vocabulary did not predict performance in word reading. Measures of phonological awareness and phonic decoding were the primary predictors of word reading at this grade level.

Results from these studies indicate several findings about the relation between early oral language and reading skills in kindergarten and first grade. First, there exists a weak relation between early oral language skills and phonological skills in kindergarten, with kindergarten oral language explaining around 4-10% unique variance in kindergarten phonological skills (Senechal, Ouellette, & Rodney, 2005; Storch & Whitehurst, 2002). This suggests the possibility of an overall language domain that impacts both processes. Second, even with this relation, phonological skills are the primary determinant of word reading in first grade, with oral language skills having very little impact. Finally, though oral language skills are not related to word reading, they are related to reading comprehension in first grade, explaining around 3-6% unique variance depending on which early literacy skills are controlled for in the analysis.

Early Language and Later Oral Language

Oral language skills, like literacy skills, develop rather continuously. This is particularly true in the area of vocabulary development, with children's vocabularies becoming larger with each new word meaning they acquire. Children who have early

deficits in oral language have a high likelihood of continuing difficulties with oral language (Hart & Risley, 1995). Some longitudinal data from studies of elementary-age children have highlighted this continuity. The longitudinal analysis conducted by Storch and Whitehurst (2002) indicated a high level of continuity across the early elementary grades for oral language skills, which included vocabulary and expressive language skills. It was reported that preschool oral language skills explained 90% of the variance in kindergarten oral language skills, kindergarten oral language skills explained 96% of the variance in first and second grade oral language skills, and first and second grade oral language skills explained 88% of the variance in third and fourth grade oral language skills. Similar longitudinal data reported by Dickinson and Tabors (2001) indicated high correlations between a kindergarten measure of receptive vocabulary and the same measure of receptive vocabulary in fourth ($r = .76$) and seventh ($r = .63$) grades.

In the area of vocabulary, Hart and Risley (1995) studied the early language environments of 42 children from 42 families of varying socioeconomic backgrounds. They examined a variety of factors including the vocabulary use and vocabulary growth of these children from the time they started using language to three years old. Within this time frame, they noticed differences in children's vocabulary usage and growth related to the amount of language and parenting practices used in the home. Children who were exposed to more rich language experiences in the home had larger and more developed vocabularies. Researchers were also able to follow up with 29 of these children at age 9-10 years. Children's early vocabulary use at age 3, as measured by the number of words these children used per hour, correlated very well with other vocabulary and oral

language measures at age 9-10 years, such as the Peabody Picture Vocabulary Test-Revised (PPVT-R) ($r = .57$) and the Test of Language Development (TOLD) ($r = .72$).

Early oral language skills lay the foundation for later oral language skills. Children who experience early oral language deficits are at a higher risk for later oral language deficits and reading difficulties (Catts et al., 2002; Hart & Risley, 1995). However, as Beck and colleagues (2002) have pointed out, this could be due to the historical lack of effective vocabulary instruction in schools. Placing a greater emphasis on the development of early oral language skills may allow educators to change the trajectories of those children who start out with oral language deficits in school. This is especially important given the nature of the relation between oral language and literacy.

Early Oral Language and Later Reading

The primary benefit of early literacy and language screening systems lies in their ability to predict future performance on educationally important outcomes. The main purpose of these systems is to provide educators with a way to identify students who may be at risk for reading difficulties. In this way, educators can identify these students early and provide intervention services aimed at preventing later, more severe reading difficulties that can develop if left unaddressed. The strong positive relation of several measures of early literacy skills, such as letter identification and phonological awareness, to later reading outcomes has been well documented (Scarborough, 1998; Torgesen, Wagner, & Rashotte, 1994). Early oral language is also positively related to later reading

outcomes. The strength of this relation depends on the grade level of reading outcomes measured.

In their study of 39 children, Roth and colleagues (2002) extended their analysis of the predictive utility of early language and literacy measures to explain second grade outcomes. Results indicated that kindergarten oral vocabulary explained a unique 9% of the variance in second grade reading comprehension, after controlling for kindergarten print awareness. A kindergarten expressive naming task also accounted for an additional 6% of the variance. This percent of variance explained at second grade was higher than the 3% reported for the prediction of first grade reading comprehension. An additional finding in this study was that a measure of kindergarten oral vocabulary predicted second grade reading comprehension better than kindergarten phonological awareness. Though phonological skills in kindergarten predicted word reading in first and second grade, vocabulary measures provided a better indicator of comprehension outcomes in second grade. However, this study is somewhat limited by its small sample size.

Catts and colleagues (1999) examined a much larger sample of 604 children, and followed them from kindergarten through second grade. This sample consisted of 328 children who were identified in kindergarten as having a language impairment or nonverbal impairment, and 276 typically developing children. Due to the high number of children with identified language impairments, a weighting procedure was used to ensure the results were representative of the normal population. This group of 604 children was divided into two groups based on performance on a second grade reading comprehension measure. The poor readers were those readers who performed 1 *SD* below the entire

group's mean performance on the comprehension measure. The good readers were identified as those children who performed above this standard. Final group sizes included 421 good readers and 183 poor readers. All students in the sample were provided a battery of early oral language and literacy measures in kindergarten, and measures of word recognition and reading comprehension in second grade. Attrition in this study was minimal leaving 593 students with full data available for analysis at both kindergarten and second grade.

Several key findings were discussed by the authors. First, children who were identified as poor readers in second grade, exhibited early oral language deficits as commonly as phonological deficits in kindergarten. The percentages of poor readers in the second grade who exhibited expressive language, receptive language, and phonological deficits in kindergarten were 50.3%, 57.4%, and 56.0% respectively. A closer analysis of the particular deficits of the poor readers in second grade revealed that a greater percentage of these children only exhibited oral language deficits (21.9%) than children who only exhibited phonological deficits (14.3%). Regression analyses conducted on all children in the sample indicated that even after controlling for kindergarten phonological skills and rapid naming, an oral language composite that combined expressive and receptive language measures still explained a unique 13.8% of the variance in second grade reading comprehension scores. This percentage was even more than the unique variance (5.4%) explained by phonological awareness. Though this study controlled for phonological awareness and rapid naming, it did not include a measure of print awareness or letter identification, which have also been identified as

highly predictive of later reading outcomes. These results are similar to the findings of Roth and colleagues (2002), describing the emerging relation of early oral language measures to reading comprehension performance in grade 2.

Another longitudinal study conducted by Catts and colleagues (2002) examined the reading outcomes of a group of children with early language impairments in kindergarten. These children were identified as having early language impairments based on deficits in two out of five domains of language including vocabulary, grammar, narration, and expressive and receptive language. These procedures identified a sample of 208 children who met the criteria for early language impairment. In examining these children in later grades, the authors reported that 52.9% were classified as poor readers in second grade and 48.1% were classified as poor readers in fourth grade. The classification of a poor reader was determined by poor performance on a measure of reading comprehension. Within this sample, roughly half of children identified with early language impairment developed a reading difficulty, which is much higher than the percentage of the overall population of children that are identified as having reading difficulties.

The particular contributions of each language domain to later reading outcomes was also reported, with the kindergarten grammar composite being the most related to second grade ($r = .67$) and fourth grade ($r = .67$) reading comprehension. The kindergarten vocabulary composite was also significantly related to second ($r = .50$) and fourth ($r = .55$) grade reading comprehension. Regression analyses indicated that after controlling for letter identification, nonverbal IQ, and rapid naming in kindergarten, the

kindergarten grammar composite explained an additional 6.5% of the variance in second grade. No oral language measure explained significant additional variance in fourth grade comprehension. The main limitation of this study is the utilization of a sample of children already identified as having language impairment. Due to the use of a sample of already impaired children, the authors have already controlled for oral language skills in some capacity. These findings should be cautiously applied to children without language impairment.

In order to determine the relation of oral language skills over time, the researchers also examined the reading comprehension outcomes of students whose language impairments improved from kindergarten to second and fourth grade. It was reported that children who no longer met the criteria for language impairment in second grade had significantly better reading outcomes than those children whose impairments had not improved. This finding suggests that improvements in language skills may result in improved reading outcomes.

Scarborough (1998) conducted a meta-analysis examining the predictive ability of various kindergarten literacy and language measures to later reading outcome. Studies included in the meta-analysis administered some combination of early literacy and language measures in or around kindergarten with reading outcome measures one to three years later. Mean correlations for the various language and literacy measures were reported. Receptive vocabulary tasks, which required kindergarten children to select a picture that corresponds to a word provided for them, correlated with later reading outcomes at a mean of .33 over 20 studies reviewed. In other words, these receptive

vocabulary tasks explained about 10% of the variance in later reading outcomes.

Expressive vocabulary tasks, which place a greater demand on vocabulary skills by asking students to verbally produce words to name objects, correlated with later reading measures at a mean of .45, or explained about 20% of the variance in later reading outcomes, over 5 studies reviewed. Kindergarten phonological awareness and letter identification have been identified as two of the best early predictors of reading achievement. As a comparison in this meta-analysis, phonological awareness correlated with later reading outcomes at a mean of .42 over 27 studies and letter identification correlated at a mean of .53 over 24 studies.

Within this meta-analysis, it was also reported that studies using multiple measures to predict reading outcomes produced greater correlations ($r = .75$) than any single measure. This finding is interesting in light of the fact that few of the multiple measure batteries reported in the meta-analysis included measures of receptive or expressive vocabulary. The significant predictive power of these vocabulary measures suggests that the average correlation of the multiple skill batteries may have actually been higher had these important measures been utilized.

As part of the Home-School Study, Dickinson and Tabors (2001) examined the correlations of early language and literacy measures administered in kindergarten with later comprehension and vocabulary measures administered in fourth and seventh grade. Due to the attrition of some subjects throughout the multi-year study, sample sizes for each correlation varied depending on which outcome measure was used and when it was administered. The range of sample size was between 51 to 56 students. Significant

correlations were reported between kindergarten measures of narrative production ($r = .47$), formal definitions ($r = .55$), emergent literacy as measured by print awareness, letter naming, and identifying sounds in words ($r = .62$), and receptive vocabulary ($r = .60$) with a fourth grade measure of reading comprehension. A similar pattern was observed when seventh grade reading comprehension was used as the outcome measure. Of particular interest was the finding that kindergarten receptive vocabulary was an even stronger predictor of reading comprehension in seventh grade than it had been in fourth, as demonstrated by a correlation of $.71$. These correlations are impressive given the length of time between criterion measures. Results of this study should be interpreted with the small sample size in mind.

Senechal, Ouellette, and Rodney (2005) analyzed longitudinal data examining the contribution of kindergarten vocabulary in predicting outcomes on a variety of language and literacy tasks. Kindergarten vocabulary was assessed using the PPVT-R, which measures receptive vocabulary. In their examination of the data, kindergarten receptive vocabulary did not predict any unique additional variance in first grade reading comprehension, after controlling for parent literacy level and kindergarten measures of early literacy skills, listening comprehension, and phonological awareness. Kindergarten early literacy skills included measures of letter naming and simple decoding and spelling. Alternatively, after controlling for those same literacy variables, in addition to first grade reading comprehension, kindergarten receptive vocabulary did predict a small but statistically significant additional 4% of the variance in third grade reading comprehension. It should also be noted that only the kindergarten vocabulary measure

and the kindergarten listening comprehension measure predicted any significant amount of variance in third grade reading comprehension, after controlling for reading comprehension in first grade. Even more compelling is the finding that kindergarten vocabulary predicted an additional 15% of the variance in fourth grade reading comprehension, after controlling for parent literacy level, reading fluency in fourth grade, reading scores in first grade, and the early literacy and phonological awareness measures in kindergarten.

The results from these studies provide evidence for the predictive power of early oral language skills as students begin formal schooling. These early oral language skills, though not strongly related to early reading performance, have an impact on reading comprehension outcomes that starts to become apparent in second grade and becomes increasingly stronger in subsequent grades. In general, kindergarten oral language measures explain an additional 10-15% of the variance in reading comprehension outcomes in and around second through fourth grades, when children are transitioning into the reading to learn stage. This variance explained is in addition to the predictive power of commonly used measures of print awareness, letter recognition, and phonological awareness. These findings suggest a particular utility of oral language measures in identifying children who may experience difficulty transitioning into the more difficult and advanced texts characteristic of the upper elementary grades.

Later Oral Language and Later Reading

Once students have mastered the phonological system of reading and writing, they gain access to the written material. However, within Gough and Tunmer's (1986) Simple View of Reading and Perfetti's (1999) model of reading comprehension, the reader must also be able to process the written language system and apply meaning to it. Effective oral language skills become especially important as children are transitioning into the Reading to Learn stage around third grade. Given the predictive nature of early oral language skills, it is also important to discuss the concurrent relation between oral language skills and reading in the later grades.

Lombardino, Riccio, Hynd, and Pinheiro (1997) examined the relation of various phonological and oral language skills to reading comprehension and decoding in a group of 80, 9-year-old children, which is during the critical transition into the Reading to Learn stage. This sample of 80 children included 32 children identified as reading disabled, 34 children identified as ADHD, and 14 normal functioning children. All language and literacy measures were administered at the same time to examine the concurrent relation of these skills at this age, which is typically around a third grade level. Results indicated that the best predictors of reading comprehension in this sample were expressive and receptive language measures combined with phonemic awareness. Expressive language alone was the most significant predictor, explaining a unique 49% of the variance in reading comprehension at this age. Conversely, phonemic awareness was the best predictor of decoding, explaining a unique 49% of the variance. This pattern suggests that, at this particular age, phonological skills allow children to read text and oral

language skills ensure that they can comprehend text. However, results of this study should be cautiously applied to normal populations of children, as the sample included in the study had a high rate of students with disabilities.

Within their longitudinal analysis, Storch and Whitehurst (2002) also examined the concurrent relation of oral language skills to reading in third and fourth grade. Within their statistical model, significant predictors of reading comprehension at these grade levels included prior reading achievement, current reading accuracy, and current oral language skills, accounting for 18%, 16%, and 7% of the variance respectively. It should also be noted that oral language skills at these grade levels were not significantly related to word reading accuracy.

A study by Ouellette (2006) examined vocabulary and reading comprehension in a group of 60 typically developing fourth grade students. All students were assessed using the Woodcock-Johnson Passage Comprehension subtest to measure reading comprehension. These students also received measure of decoding and sight word recognition in addition to four measures of vocabulary. The vocabulary measures consisted of an expressive and receptive vocabulary task and two measures of vocabulary depth, requiring students to provide definitions and synonyms of vocabulary words. Results indicated that after controlling for measures of nonverbal IQ, decoding, and sight word recognition, depth of vocabulary knowledge accounted for a statistically significant, additional 12% of the variance in reading comprehension.

Catts, Hogan, and Adlof (2005) examined a sample of children that took part in another longitudinal study of literacy and language previously discussed (see Catts et al.,

1999). In this analysis, the children were examined in second, fourth, and eighth grades to determine the relative contributions of word recognition and listening comprehension to reading comprehension at each of these grade levels. The listening comprehension composite included measures of vocabulary, listening to stories and answering comprehension questions, and a syntactic measure of oral language. The word recognition composite included the Word Attack and Word Identification subtests of the Woodcock Reading Mastery Test-Revised. It was observed that concurrent word recognition skill explained a decreasing amount of unique variance in reading comprehension at 27%, 13%, and 2% at each grade level, respectively. The opposite trend was observed for the concurrent listening comprehension composite, explaining 9%, 21%, and 36% at each respective grade level. Listening comprehension, which measured several oral language skills, was a better predictor of reading comprehension as students got older.

Similarly, Gough, Hoover, and Peterson (1996) used meta-analytic techniques to evaluate the relation between decoding and listening comprehension to reading comprehension at various grade levels. Results were included from 10 studies using 17 different samples of students. For analysis purposes, grade levels were grouped into first and second (Group 1), third and fourth (Group 2), fifth and sixth (Group 3), and seventh and eighth (Group 4). It was reported that decoding measures correlated with reading comprehension in a decreasing trend, with correlations of .61, .53, .48, and .39 for groups 1, 2, 3, and 4, respectively. Listening comprehension showed an increasing trend in correlating with reading comprehension at .41, .50, .72, and .68 for groups 1, 2, 3, and 4,

respectively. These results are similar to the sample reviewed by Catts, Hogan, and Adlof (2005).

These findings are compelling in light of the prior discussion of a high level of continuity in oral language skills over time. These results suggest that though early oral language skills may not directly contribute to early reading outcomes in kindergarten and first grade, the impact of these oral language skills will be felt later as reading comprehension tasks become more difficult around grades two and above. It is then important to account for early oral language skills in predicting reading comprehension outcomes in the upper grades.

The Importance of Vocabulary

Within the larger domain of oral language, measures of vocabulary knowledge have been identified as some of the most powerful predictors of reading comprehension outcomes (Dickinson & Tabors, 2001; Scarborough, 1998; Senechal, Ouellette, and Rodney, 2005; Storch & Whitehurst, 2002). These vocabulary measures can serve as indicators of the much broader domain of oral language (Hart & Risley, 1995). Students who have a large and developed vocabulary have better oral language skills.

Riedel (2007) has suggested that vocabulary may serve as a critical link between oral reading fluency and reading comprehension. In his study of 1,518 first graders, it was reported that roughly 15% of students who met or exceeded oral reading fluency benchmarks at the end of first grade had poor reading comprehension, as measured by performance on a standardized reading comprehension test, the Group Reading

Assessment and Diagnostic Evaluation (GRA+DE). These students with low reading comprehension, despite their satisfactory oral reading fluency, had much lower vocabulary skills than those students who met oral reading fluency benchmarks and had satisfactory reading comprehension. Alternatively, those students who had low oral reading fluency scores and satisfactory reading comprehension had significantly better vocabulary skills than those students who had low oral reading fluency scores and low reading comprehension scores. Results suggest that vocabulary plays a key role, above and beyond oral reading fluency, in ensuring adequate reading comprehension.

In order to further understand more specifically how vocabulary affects reading comprehension, Tannenbaum, Torgesen, and Wagner (2006) examined the contributions of different types of vocabulary assessments. The researchers investigated the role of two factors, vocabulary breadth and depth/fluency, to reading comprehension in a sample of 203 third grade students. The PPVT-III and the vocabulary subtest of the WISC-III were used to assess vocabulary breadth. The Multiple Meanings subtest of the Language Processing Test-Revised (LPT-R) was used to assess depth. The DIBELS WUF measure and an experimental measure requiring students to name items from a given category were used to assess fluency. Reading comprehension was assessed using the Florida Comprehensive Assessment Test–Sunshine State Standards (FCAT-SSS) and the Stanford Achievement Test (SAT-9) reading comprehension subtest. Results indicated that the two vocabulary factors accounted for 50% of the total variance in reading comprehension. Breadth of vocabulary contributed 19% unique variance, depth/fluency contributed 2% unique variance, and 29% of the variance was shared between the two

factors. This study did not control for decoding or any other word reading skills. Results suggests that knowing more vocabulary words has a stronger effect on reading comprehension than knowing fewer words in more depth and being able to retrieve those meanings more fluently. However, the authors point out that the zero-order correlations between both the breadth ($r = .70$) and depth/fluency ($r = .56$) indicate that both factors are significantly related to reading comprehension. Although measures of vocabulary breadth are more related to reading comprehension, measures of depth/fluency may be more appropriate and feasible for use in educational settings.

The simplest explanation for the effect of vocabulary on reading comprehension is that one cannot understand written text, unless they understand the meanings of each of the individual words on the page (Anderson & Freebody, 1981; Kamil, 2004). In this way, vocabulary serves as an indicator of general language comprehension skills. A vocabulary rich in breadth and depth allows for quick retrieval of meaning from written text, freeing up resources for the use of more advanced comprehension strategies. Similarly, Stanovich (1986) described the benefit of vocabulary on reading through a relation of reciprocal causation. Through this process, children who have a well developed vocabulary have an easier time reading, which then exposes them to more text and more vocabulary within that text. So the more children read, the larger their vocabularies grow. Both the size and depth of a child's vocabulary and the amount they read affect each other, through an interactive process. However, it should also be noted that the simple act of reading may not be sufficient to build vocabulary in all students. A meta-analysis conducted by Swanborn and de Glopper (1999) found that on average

children only adequately learn about 15% of unknown words they encounter while reading, highlighting the importance of effective vocabulary instruction for all children.

Vocabulary Instruction/Intervention

In addition to studies of the relation between measures of oral language/vocabulary to measures of reading comprehension, studies of vocabulary instruction have examined the causal relation between teaching vocabulary and improved vocabulary and reading comprehension. A meta-analysis conducted by Stahl and Fairbanks (1986) demonstrated that teaching vocabulary words contained within a story improves children's comprehension of that story. Within this meta-analysis, the authors examined the findings of 52 studies of vocabulary instruction. It was reported that vocabulary instruction had an effect size of .97 on the reading comprehension of passages containing taught vocabulary words, and an effect size of .30 on the reading comprehension of more general passages that did not contain directly taught vocabulary words. Instructional methods that included both definitional and contextual information for the taught vocabulary words appeared to be the most effective in improving reading comprehension, with a mean effect size of 1.05 reported. However, it should be noted that the studies included in this meta-analysis included only grades two and above, and did not examine earlier vocabulary instruction.

Research with younger children has pointed to the effectiveness of using teacher-directed storybook reading activities for increasing vocabulary (Beck & McKeown, 2007; Justice, Meier, & Walpole, 2005; Robbins & Ehri, 1994; Senechal, 1997). The

improvements in children's vocabulary obtained from simply having children listen to storybooks with unfamiliar words may not be significant to educators. However, the addition of a brief elaborated description of the unfamiliar words in the story as well as active student participation can impact even larger effects.

Beck and McKeown (2007) examined the effectiveness of a vocabulary instructional activity known as Text Talk with a group of 98 kindergartners and first graders. A group of 52 students received the Text Talk activity, which involved the use of read-aloud activities combined with more in-depth discussion of target vocabulary words after the story reading. A comparison group of 46 students only engaged in read-aloud activities without the additional discussions of the vocabulary words. Results indicated that children who participated in the Text Talk activities learned an average of 3-4 words more than the students who did not receive the Text Talk activity.

Justice and colleagues (2005) examined the use of a storybook reading strategy to enhance the vocabulary of kindergarten students. This strategy involved providing more elaborate descriptions of several vocabulary words within story books read by the teacher. The more elaborate descriptions consisted of providing a definition of the word after it was encountered in the story, and then using the word in a supportive context. Results of this study indicated an effect size of 1.22 on elaborated words learned for students who engaged in the elaborated storybook reading activity. Even children with low initial vocabularies made significant gains on elaborated words learned as compared to their no-treatment peers, learning approximately 5 more words.

Senechal (1997) indicated the power of repeated exposure to words in increasing vocabulary acquisition. In this study, three groups of preschool children were compared on number of target vocabulary words learned. The single-reading group received only a one-time reading of the story. The repeated-reading group and the questioning group received three readings of the story, on two successive days. In the questioning group, the adult reading the story asked the children “what” or “where” questions that elicited the students to use the target vocabulary words in the story when they occurred. Results indicated that both the repeated-reading group and the questioning group learned significantly more of the target vocabulary words as measured by a receptive vocabulary task, with a reported effect size of 1.06. The questioning group also learned significantly more words than the repeated-reading group as measured by the receptive vocabulary task, with a reported effect size of .54. A similar pattern was observed for the expressive vocabulary task. Results from this study indicate that repeated reading and actively engaging children in the text by asking them questions about the vocabulary words improves vocabulary more than simply reading the stories to the children. The findings of the National Reading Panel (2000) have supported the benefit of repetition and multiple exposures to vocabulary words in vocabulary instruction.

Beck and colleagues (2002) have highlighted the need for effective vocabulary instruction in schools. They have identified a framework for prioritizing which vocabulary words to teach, in order to affect important vocabulary and comprehension growth in young children. This framework uses three tiers of instructional words. Tier One words consist of basic words that young children will encounter frequently in text,

such as *happy*, *baby*, *run*, and *ball*. These are words that children should be familiar with already through oral language interactions, and should require minimal explicit instruction. Tier Two words are more advanced words that children will encounter in texts but will hear with less regularity in oral language interactions, such as *devour*, *reluctant*, and *desperate*. These words should be the focus of explicit vocabulary instruction, and will have the most impact on improving reading comprehension in young children. Tier Three words are those words that are domain specific and used rather infrequently in everyday texts and conversation. Examples of these words include *photosynthesis*, *judiciary*, and *ulcer*. Focusing on Tier Two words, teachers can help explain and expand these words to build a conceptual understanding in their students and improve overall reading comprehension.

Results of vocabulary instruction studies indicate that direct vocabulary instruction does impact vocabulary growth in children. There are research-based instructional activities that have been shown to improve oral language skills and expand the vocabulary size of children. There is also some support for the effect of vocabulary instruction on reading comprehension outcomes. Specifically, vocabulary instruction that involves multiple exposures to words, involves contextual information in addition to definitions of words, and directs children to engage in a deeper level of understanding with the words has the greatest effect on reading comprehension outcomes (Baumann, Kame'enui, & Ash, 2003; Stahl & Fairbanks, 1986). However, the research linking vocabulary instruction to improved reading comprehension is not extensive. More studies are needed to examine the utility of various instructional strategies, and to determine

which particular vocabulary words should be taught in order to produce the greatest impact on reading comprehension (Baumann, et al., 2003). Although important, direct vocabulary instruction is not sufficient to improve reading comprehension outcomes. This finding could be due to the sheer breadth of vocabulary children are exposed to in text (Beck, McKeown, & Kucan, 2002). The amount of vocabulary that children need to acquire in order to understand the wide range of texts they encounter is extensive. For vocabulary instruction to be effective in increasing reading comprehension, it needs to take place over an extended time period, and include opportunities for children to gain vocabulary knowledge through reading a wide range of texts. This idea highlights the need to provide effective vocabulary support as early as possible.

Overall, improving the vocabulary skills of young children represents an instructionally important and feasible outcome in the area of reading (NRP, 2000). We currently have effective instructional methods for teaching vocabulary. Assessment of vocabulary skills provides educators a way to identify students who may need extra vocabulary support and helps to determine the effectiveness of their instructional efforts in this area. In order for school-based assessments to be useful, they need to be efficient and linked to specific instructional outcomes. Vocabulary also provides a useful indicator of a student's broader oral language skills. This indicator can provide educators with information about children's general language comprehension skills before they are even able to read. Brief and efficient assessments that provide information on the vocabulary skills of their students can help teachers identify which students may need additional oral language support in order to reach educationally meaningful literacy outcomes.

Purpose of the Current Study

Early intervention and prevention efforts rely on the ability of effective early screening systems that can identify those students who are at risk for reading difficulties. Current screening systems exist that are tied to important literacy outcomes (Good, Kame'enui, & Simmons, 2001). These batteries of early screening primarily have consisted of measures of print awareness, letter knowledge, phonemic awareness, phonics, and even oral reading fluency at later stages of development. Though these systems are highly effective in identifying at-risk readers, they are not perfect, as some children who acquire appropriate early literacy skills still experience reading difficulties (Al-Otaiba & Fuchs, 2006; Riedel, 2007).

Currently, there is a need to investigate the additional benefit of screening other early literacy and language skills. Measures of oral language and more specifically, vocabulary, have been shown to play an important role in predicting reading comprehension outcomes in grades two and above (Catts, Fey, Zhang, & Tomblin, 1999; Scarborough, 1998; Senechal, Ouellette, and Rodney, 2005; Storch & Whitehurst, 2002). Vocabulary instruction has also been linked to increases in reading comprehension (Stahl & Fairbanks, 1986). Vocabulary measures provide a unique target of opportunity for early screening systems in identifying children who are at-risk for reading comprehension difficulties. The proposed research aims to examine the utility of a brief early measure of vocabulary, using vocabulary words in a verbal utterance. More specifically, this study will examine the utility of the DIBELS WUF measure in combination with existing early screening systems to identify students who may be at risk for reading difficulties.

CHAPTER III

METHODOLOGY

A correlational research design was used to examine the technical adequacy and utility of a measure of early vocabulary screening, the DIBELS Word Use Fluency (WUF) measure. Specifically this study examined (a) the longitudinal continuity of a measure of students' skills in defining or using vocabulary words in a sentence from kindergarten through third grade, (b) the predictive validity of a measure of student's skills in defining or using vocabulary words in a sentence, given in kindergarten and first grade, with measures of reading comprehension and vocabulary given at the end of third grade, (c) whether scores on the vocabulary screening measure significantly add to the predictive validity of measures of phonemic awareness, phonics, and oral reading fluency in predicting end-of-third grade outcomes on standardized measures of reading comprehension and vocabulary, and (d) the relation between a vocabulary screening measure and a reading comprehension measure in which students retell what they orally read, the DIBELS Retell Fluency (RTF) measure. The purpose of this chapter is to provide a description of the participants and setting, measures, procedures, and data analysis methods used in this study.

Participants and Setting

The sample in this study consisted of 29 third grade students from a single elementary school in rural Wyoming. This school has collected their own schoolwide DIBELS data for the past 4 academic years. As part of their data collection, they have administered the WUF and RTF measures, in addition to all other DIBELS measures. Out of the final total sample of 29 students, 19 of these students had complete data from kindergarten through third grade, 23 students had complete data from first grade through third grade, and 27 students had complete data from second grade through third grade. Written consent to participate in the study was obtained from the principal, and parents of current third grade students received information and consent letters inviting the participation of their child. Data from all 29 students was included in the study.

Measures

Several measures of literacy were used in this study. Four DIBELS measures were utilized as independent variables: the DIBELS Phoneme Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), Oral Reading Fluency (DORF), and Word Use Fluency (WUF) measures. The dependent measures consisted of four end-of-third grade criterion measures, including the Gates-MacGinitie Reading Test-Fourth Edition (GMRT-4) (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) Comprehension and Vocabulary subtests and the DIBELS ORF and RTF measures. Additionally, the RTF measure in first through third grade was also used as a dependent measure.

DIBELS Measures

DIBELS Phoneme Segmentation Fluency (PSF) is an individually administered, standardized measure of a student's skills in segmenting orally provided words into individual phonemes (Good & Kaminski, 2002). For this measure, students are orally provided with a three or four phoneme stimulus word and are instructed to orally segment the word into individual phonemes. The student is then presented with subsequent words until a time period of one minute is reached. The unit of measure is the number of phonemes correctly produced in one minute. For example, if the examiner provides the student with the word *sat*, the student must say the individual phonemes, /s/ /a/ /t/, to receive the full possible three points for that particular word. Partial credit is awarded to words segmented into sound-parts, but not completely into individual phonemes. The measure has a two-week, alternate-form reliability of .88 in kindergarten (Kaminski & Good, 1996). It has a concurrent criterion-related validity of .54 with the Woodcock-Johnson Psychoeducational Battery Readiness score, and a predictive validity of .62 with the DIBELS NWF score in the middle of first grade and .62 with Curriculum-Based Measurement-Reading in the middle of first grade (Good et al., 2003).

DIBELS Nonsense Word Fluency (NWF) is an individually administered, standardized measure of a student's skills in producing letter sounds and blending letter sounds into words (Good & Kaminski, 2002). For this measure, students are visually presented with a list of one-syllable, three-letter or two-letter nonsense words and are instructed that they can either pronounce the sound for each individual letter or say the whole word. They are also instructed to read the words the best they can. The unit of

measure is the number of letter sounds produced correctly by the student in one minute. The entire measure takes about two to three minutes to administer for each child. The technical adequacy of this measure is well documented as it has demonstrated an alternate-form reliability of .83 in January of first grade (Good et al., 2004). The predictive validity of the NWF measure in January with Curriculum Based Measurement-Reading (R-CBM) is .81 for the end of first grade and .68 for May of second grade (Good et al., 2004).

DIBELS Oral Reading Fluency (DORF) is an individually administered, standardized measure of a student's skills in accurately and fluently reading grade level connected text (Good & Kaminski, 2002). For this measure, students are presented with a grade level passage and asked to read it aloud for one minute. Words pronounced incorrectly, omitted, or hesitated upon for three seconds are counted as incorrect. The unit of measure is the number of words read correctly in one minute. The DORF measure has an estimated alternate-form reliability of .96 when three probes are given (Roberts, Good, & Corcoran, 2003). The DORF measure is based on CBM-R procedures, which have criterion-related validity of anywhere from .60 to .90 depending on the criterion (Good & Jefferson, 1998; Marston, 1989). Good, Simmons, and Kame'enui (2001) have also shown that students achieving DORF benchmark goals have a high probability of meeting future reading goals and becoming competent readers. DORF has also been shown to be a good indicator of a student's overall literacy skills, making it a good measure for use in this study (Fuchs, Fuchs, Hosp, & Jenkins, 2001).

DIBELS Retell Fluency (RTF) is an individually administered, standardized measure of a student's skills in understanding written text that they have read. The RTF measure is given in conjunction with the DORF. It is intended to provide an additional measure of reading comprehension to ensure that students understand what they orally read during the DORF measure administration. After students complete the DORF one-minute reading, they are asked to orally retell what they have just read. The unit of measure is the number of words that are relevant to the actual text read that the student orally provides. Alternate-form reliabilities of the RTF measure, used in conjunction with the DORF measure, range from .57 to .90 depending on the grade and the number of probes provided (Roberts et al., 2003). Typical DORF/RTF administration provides three probes to students, which has an estimated reliability of at least .80 (Roberts et al., 2003). The RTF measure alone has demonstrated good criterion-related validity, ranging from .27 to .72, depending on the criterion measure. Combining the DORF and RTF scores improves the criterion-related validity, as demonstrated by their correlation of .73 with the Oregon State Assessment (OSAT) (McKenna, 2003) and .81 with the Woodcock Johnson Broad Reading Cluster (Roberts et al., 2003).

DIBELS Word Use Fluency (WUF) is an individually administered, standardized measure of a student's skills in correctly using orally presented vocabulary words in verbal utterances (Good & Kaminski, 2002). For this measure, students are orally provided with a stimulus word and asked to use it in a sentence, phrase, or utterance. The examiner continues to present words until a one-minute time limit is reached. If the student hesitates on any word for 5 seconds, they are provided the next word. The unit of

measure is the total number of words spoken in correct verbal utterances provided by the student. Utterances are counted as correct if the stimulus word is used correctly in the utterance. Correct usage is defined as an utterance that conveys the correct meaning of the stimulus word or provides a definition or synonym for the stimulus word. Initial research has demonstrated an alternate-form reliability of .71 in kindergarten, .65 in first grade, and .66 in second grade (Kaminski et al., 2004). Inter-rater reliability data is not currently available from the WUF technical report. The WUF has a demonstrated concurrent criterion related validity of .44 to .55 with the Test of Language Development, .44 to .47 with a comprehensive language sample, .34 to .42 with the DORF, and .35 with the Oregon State Assessment at the end of third grade.

Gates-MacGinitie Reading Test, Fourth Edition (GMRT-4)

The GMRT-4 is a group-administered, standardized, norm-referenced measure of reading achievement. It provides a variety of scores, including scaled scores and percentile ranks and has separate forms for kindergarten through twelfth grade, in addition to a form for adults. For the purpose of this study, Level 3 was administered. This level offers two subtests, Vocabulary (GMRT-V) and Reading Comprehension (GMRT-C). For the Vocabulary subtest, test words are visually presented in the test booklet in a brief context intended to suggest part of speech but not to provide clues to meaning. Students select the word or phrase that means most nearly the same as the test word. For the Comprehension subtest, students read a series of reading passages and answer questions about the passages. Passages represent a variety of writing types and

content areas. Both the vocabulary and comprehension subtests, as well as the total reading test score, have demonstrated internal consistency reliabilities at or above .90 (MacGinitie, MacGinitie, Maria, & Dreyer, 2000). The test was developed through a rigorous process involving pilot studies, field testing, and the use of input from education experts, improving content validity. The GMRT-4 also is highly correlated with the GMRT-3, which has demonstrated high correlations with other measures of reading comprehension such as the Stanford-Achievement Test.

Procedures

Human Subjects Review

Procedures and logistics of this study were reviewed and approved by the University of Oregon Human Subjects Review Committee. The committee reviewed a detailed description of the purposes of the project, participant selection and recruitment procedures, parental consent, confidentiality of participant identities, and all activities involving participants. This process ensured that the rights and wellbeing of all participants were protected.

Participant Selection/Consent

Once participating schools were identified and district and school approval was secured, parents of all third grade students received information letters providing a description of their child's participation in the study. These information letters indicated for parents to contact their child's school if they did not wish for their child to participate.

There were no parents who refused their child's participation in the study. In addition, all students were provided with an oral description of the study and were given the option to not participate. All students indicated they wished to participate.

Data Collection

Data used in this study was collected from multiple sources. Existing DIBELS data collected by the participating school was used in the analysis. The participating school collected this data as part of their typical DIBELS benchmarking procedures over the past several years. However, the researcher and two trained graduate student data collectors collected the kindergarten DIBELS data for this sample in the spring of 2004. This kindergarten data collection was done as a service to the school district. The district was beginning to administer DIBELS to all their students in the fall of 2004. The graduate student data collectors helped to provide a model to the school in how to collect school-wide DIBELS data, as part of their initial DIBELS training. Graduate student data collectors utilized standardized DIBELS administration and scoring procedures.

The GMRT-4 was used as a criterion measure and administered by the researcher to all participating third grade students in the spring of 2007 as a group-administered assessment. Test protocols were given to each student, standardized directions were read aloud, and the students were given the time allotted in the administration guide to complete each subtest. The protocols were scored at a later time, and rescored by another graduate student to ensure accuracy of scores. The entire data collection schedule for this study is outlined in Table 1. Data collected by the researcher is underlined in the table.

Table 1.

Timeline for data collection

Year	Grade	Fall	Winter	Spring
1 (2003-2004)	Kindergarten			<u>WUF</u> <u>PSF</u> <u>NWF</u>
2 (2004-2005)	1 st Grade	WUF PSF NWF	WUF PSF NWF DORF RTF	WUF PSF NWF DORF RTF
3 (2005-2006)	2 nd Grade	WUF DORF RTF	WUF DORF RTF	WUF DORF RTF
4 (2006-2007)	3 rd Grade	WUF DORF RTF	WUF DORF RTF	WUF DORF RTF <u>GMRT-4</u>

Note. Underlined measures were administered by the researcher. Non-underlined measures were administered by school personnel.

Data Analysis

The following analyses were conducted to answer each of the research questions.

Research Question 1

What is the “longitudinal continuity” of a measure of students’ skills in defining or using vocabulary words in a sentence from kindergarten through third grade?

“Longitudinal continuity” refers to the stability of a student’s skills over time. To answer this question, a WUF composite score for each grade level was created, taking the mean of the fall, winter, and spring WUF scores at each grade level. Creating a grade level composite score helped to improve the within-grade stability of the scores. A Pearson correlation coefficient was calculated between each grade’s WUF composite scores to determine the relative continuity and stability between grade levels of scores on the WUF. In addition to between-grade correlations, within-grade correlations were also calculated between each concurrent benchmark period’s WUF scores for each grade level.

Research Question 2

What is the predictive validity of a measure of student’s skills in defining or using vocabulary words in a sentence, given in kindergarten and first grade, with reading comprehension and vocabulary measures given at the end of third grade?

To answer this question, a Pearson correlation coefficient was calculated between WUF scores for the end-of-kindergarten and all first grade benchmark periods (fall,

winter, and spring) with the end-of-third grade GMRT-4 reading comprehension and vocabulary, and DORF and RTF scores. In addition, a correlation coefficient was calculated between the first grade WUF composite score with the end-of-third grade GMRT-4, DORF, and RTF scores. The ability of the WUF measure administered in kindergarten and first grade to predict later reading achievement and more specifically, comprehension and vocabulary performance, is critical in developing a framework for the specific utility of this measures.

Research Question 3

Does the vocabulary screening measure, WUF, add to the predictive validity of commonly utilized early literacy measures such as phonological awareness (PSF), phonics (NWF), and oral reading fluency (DORF) in predicting outcomes on end-of-third grade measures of reading achievement?

Research has indicated that phonological awareness measures provide some of the most powerful prediction of future reading achievement (Torgesen et al., 1994). Additionally, it has been demonstrated that ORF is a powerful indicator of overall reading proficiency (Fuchs, Hosp, Hosp, & Jenkins, 2001). A series of hierarchical multiple regression analyses were conducted for each benchmark period to determine if WUF explains a significant amount of additional variance after accounting for commonly used DIBELS measures. In kindergarten, PSF or NWF scores were entered into the model first. In first grade, NWF or ORF scores were entered into the model first. WUF scores were the final variable entered into the model to examine the additional variance

predicted after accounting for the current DIBELS measures. Tabachnick and Fidell (2001) have indicated that in conducting hierarchical regression analyses, a minimum requirement of sample size should be 4 to 5 times greater than the number of IV's. The current study meets this requirement by using no more than 2 IV's in each hierarchical regression model, due to the sample size between 20 and 30. These analyses will help determine if the addition of the WUF is beneficial to educators who already administer the other DIBELS measures.

Research Question 4

What is the relation between a vocabulary screening measure and a reading comprehension measure in which students retell what they orally read?

The DIBELS WUF and RTF measures were designed to provide additional indicators of reading competence in the areas of vocabulary and comprehension (Good & Kaminski, 2002). Vocabulary and reading comprehension are related, in that basic vocabulary knowledge has been described as a necessary prerequisite to understanding connected text (Anderson & Freebody, 1981; Kamil, 2004). One must understand the meanings of individual words before they can understand the meaning of connected text. The relation between vocabulary and reading comprehension also seems to grow larger as students get older (Storch & Whitehurst, 2002).

An investigation of the relation between WUF and RTF will help examine whether these measures demonstrate the increasing relation between vocabulary and reading comprehension found in previous research. This investigation will also help

examine whether the WUF measure is useful to educators in predicting how children will do on reading comprehension measures, and possibly help identify vocabulary as an area for additional support for children who do not have sufficient reading comprehension skills despite good decoding skills. In order to examine this relation, a Pearson correlation coefficient was calculated between the student's WUF and RTF scores for each individual benchmark period. Additionally, a hierarchical multiple regression analysis was conducted using RTF as the dependent variable, and DORF and WUF as the independent variables entered into the regression model. This analysis was conducted for each benchmark period for which RTF, DORF, and WUF data was available and helped to determine if WUF explains any additional significant variance in RTF scores after accounting for DORF scores.

Summary

The participants in this study included 29 current third grade students from a rural school in Wyoming. A longitudinal DIBELS data set was available for these students. This data set included DIBELS measures of phonological awareness (PSF), phonics (NWF), oral reading fluency (DORF), reading comprehension (RTF), and vocabulary (WUF) in kindergarten through third grade. In addition to the existing DIBELS data, a standardized measure of reading comprehension (GMRT-C) and vocabulary (GMRT-V) was administered at the end of third grade as criterion outcome measures. All of this data was analyzed to examine (a) the "longitudinal continuity" of the DIBELS vocabulary measure (WUF) over time, (b) the predictive validity of the WUF measure, given in

kindergarten and first grade, with end-of-third grade reading comprehension and vocabulary outcomes, (c) whether scores on the DIBELS WUF measure significantly add to the predictive validity of measures of phonological awareness, phonics, and oral reading fluency in predicting end-of-third grade reading comprehension and vocabulary outcomes, and (d) the relation between the DIBELS WUF measure and the DIBELS reading comprehension measure (RTF) over time.

CHAPTER IV

RESULTS

The purpose of this chapter is to present the results of this study. Results will be examined to better understand the utility of the DIBELS WUF measure as an early indicator of oral language skills. First, descriptive statistics of all DIBELS measures, including the WUF, and the criterion outcome measures will be described. Next, the examination of the results will focus on the four main research questions. The first research question will examine the continuity of vocabulary skills, as measured by the WUF, over time. Scores on the WUF measure will be compared at each successive benchmark to examine their general continuity. The second research question will examine the predictive validity of the WUF measure, given in kindergarten and first grade, on end-of-third grade reading comprehension and vocabulary outcomes. Three separate standardized measures of reading comprehension, administered at the end of third grade, will be used as the criterion measures for this research question. The third research question will examine the added contribution of the WUF measure in kindergarten and third grade in predicting end-of-third grade reading comprehension outcomes, after taking into account other measures of phonemic awareness, phonics, and oral reading fluency. The final research question will examine the nature of the relation between the WUF and RTF measures, as measures of vocabulary and reading

comprehension, respectively. Concurrent scores on both of these measures throughout the grade levels will be examined to answer this question.

Descriptive Statistics

This section will present the descriptive statistics for all measures used in the study. First, the means and standard deviations for all DIBELS measures will be discussed. The DIBELS measures were administered to a single cohort of students starting in their kindergarten year in 2003-2004 through their third grade year in 2006-2007. This discussion will include the DIBELS ORF and RTF measures, used as dependent variables in this study. Second, the means and standard deviations of the GMRT-4 vocabulary and comprehension subtests will be discussed. These scores were collected at the end of participating students' third grade year, in the spring of 2007.

DIBELS Measures

Means and standard deviations for all DIBELS measures are presented in Table 2. An increasing trend for all measures is expected due to the improving student mastery of early literacy and oral language skills throughout the elementary years. A general increasing pattern of performance was observed on each individual DIBELS measure from kindergarten through third grade. Mean scores reported on the measures fell within the average range as compared to national DIBELS norms (Good, Wallin, Simmons, Kame'enui, & Kaminski, 2002).

A ceiling effect was observed for the WUF measure. Scores on this measure showed an increasing trend in the early grades, as demonstrated by a large jump in scores from the beginning to the middle of first grade. This large increase was followed by a leveling off in second and third grade, as scores remained fairly stable with little growth observed. This pattern of rapid early growth followed by slower progress in second and third grade is similar to the results found by Kaminski and colleagues (2004), and may be indicative of a ceiling effect for the WUF measure, in general.

DIBELS Oral Reading Fluency (DORF) scores showed an increasing trend across benchmark periods. Mean scores within this sample fell within the average range as compared to national norms (Good et al., 2002), indicating a representative sample.

Table 2.

Descriptive Statistics for DIBELS Measures

Measure	Fall		Winter		Spring	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
2003 – 2004 Kindergarten (<i>n</i> = 19)						
PSF					37.26	16.41
NWF					22.37	12.38
WUF					6.68	10.50
2004 – 2005 First Grade (<i>n</i> = 23 ^a)						
PSF	34.05	9.40	50.13	10.67	56.22	9.84
NWF	27.00	15.51	53.65	20.28	69.65	24.66
WUF	12.09	13.93	44.83	15.49	50.17	18.99
DORF			27.87	19.04	62.17	26.63
RTF			11.70	10.90	25.04	10.92
2005 – 2006 Second Grade (<i>n</i> = 27 ^a)						
WUF	42.23	14.74	47.78	11.03	52.81	13.86
DORF	46.58	21.40	81.00	27.09	98.00	26.27
RTF	19.85	12.52	32.52	11.33	41.15	13.32
2006 – 2007 Third Grade (<i>n</i> = 29 ^a)						
WUF	43.57	8.60	46.59	12.66	40.76	12.06
DORF	77.29	25.49	99.90	25.57	116.79	28.71
RTF	38.04	15.39	44.76	18.79	45.79	16.91

^a*n* – 1 in Fall.

Of particular interest is the finding that the initial mean WUF scores in the present study were much lower than those reported by Kaminski and her colleagues. In that study, the mean WUF score at the end of kindergarten was much higher at 30, as compared to 6.68 in the current study. One possible explanation for this substantial difference in scores could be due to a practice effect. Kaminski and colleagues initially administered the measure at the beginning of kindergarten. In the present study, the initial administration of the WUF measure was at the end of kindergarten. In both studies, mean WUF scores at the first administration time were low at 6.68 and 10.40 respectively. Also in both studies, student WUF scores experienced a substantial jump from the second to third administration of the measure. This could explain why initial scores in the present study were much lower than those reported by Kaminski and her colleagues (2004). By the end of first grade, mean WUF scores in both studies are similar through the end of third grade.

Gates-MacGinitie Reading Test-4 (GMRT-4)

The Gates-MacGinitie Reading Test-Fourth Edition (GMRT-4) was used as a criterion measure in this study. It was administered to all students in the sample in the spring of their third grade year. Means and standard deviations for the GMRT-4 are presented in Table 3.

The GMRT-4 subtests of vocabulary and comprehension have maximum raw scores of 45 and 48, respectively. Mean scores of 30.79 and 29.28 were observed for the vocabulary and comprehension subtests, respectively. These mean scores fell at the 51st

and 45th percentiles, as compared to national norms, suggesting a fairly representative study sample. Three students in the sample did not wish to complete the comprehension subtest. Their comprehension scores were thus omitted from the final analyses, but their scores on the vocabulary test were used.

Table 3.

Descriptive Statistics for GMRT-4.

Measure	<i>M</i>	<i>SD</i>	<i>N</i>
Vocabulary raw score (GMRT-V)	30.79	9.24	28
Comprehension raw score (GMRT-C)	29.28	11.01	25

Longitudinal Continuity

Research Question 1: What is the “longitudinal continuity” of a measure of students’ skills in defining or using vocabulary words in a sentence from kindergarten through third grade?

The term “longitudinal continuity” refers to the relative stability of the WUF as a measure of vocabulary over time. It can also be conceptualized as the reliability of the measure, although correlations between successive benchmarks were taken approximately three to four months apart. Although this long delay between assessment periods is not ideal for examining test-retest reliability, the correlations still provide some

limited information about the measure's reliability. Moderate to high correlations between successive benchmark WUF scores would indicate relative continuity in the WUF measure's ability to assess vocabulary skills in developing readers, or a high test-retest reliability over the three to four month time period. Alternatively, low correlations would indicate a lack of continuity in the WUF measure's ability to assess vocabulary skills in developing readers, or a low test-retest reliability. Correlation coefficients between successive within-grade benchmark WUF scores are presented in Table 4.

Table 4.

Within-Grade WUF Benchmark Correlation Coefficients

Grade	Beginning to middle		Middle to end		End to beginning	
	<i>r</i>	<i>n</i>	<i>r</i>	<i>n</i>	<i>r</i>	<i>n</i>
K					-.09	19
1	.51*	22	.48*	23	.63*	23
2	.09	26	-.04	27	.12	27
3	.33	28	.44*	29		

Note. End to beginning correlations indicate the relation between end of the year and beginning of the subsequent academic year WUF scores.

* $p < .05$.

The highest correlations were observed in the earlier grades. In particular, moderate correlations ranging from .48 to .63 were observed from the beginning of first

grade through the beginning of second grade, with all of these correlations being significantly different than zero ($p < .05$). Correlations between benchmark periods in second grade were not significant. In third grade, correlations were generally higher, with a statistically significant correlation of $r = .44$ from the middle to the end of third grade.

Of particular interest is the low, negative correlation between the end of kindergarten and the beginning of first grade, $r = -.09$. This low correlation was unexpected, and may be explained by the particular data collection procedures utilized for each of the benchmark periods. For the end of kindergarten data collection period, highly trained graduate students collected all DIBELS data. These data collectors had an extensive background in educational assessment and received several trainings on the DIBELS measures in particular. They also had administered the DIBELS measures several times to students in various settings. Starting at the beginning of first grade, all subsequent data collection was conducted by resident school staff. These school staff members, though trained in DIBELS, did not share the background in educational assessment held by the graduate student data collectors. The higher difficulty with administering the WUF, combined with the relative inexperience of school staff in administering the DIBELS measures, could have contributed to this low correlation. This finding would then suggest the need for more extensive and effective training on all DIBELS measures in order to ensure reliable and valid administration of the measures from the very beginning.

Alternatively, the scores collected by the resident school staff may have been the more accurate representation of student vocabulary skills. It is possible that students

performed better when tested in a more natural testing setting by familiar teachers and school staff, in comparison to the graduate student data collectors, whom the students were not familiar with. It is also possible that as students became more familiar with the measure their scores became a more accurate representation of their skills over time.

The pattern of low or negative correlations between successive benchmark WUF scores in second grade was also unexpected. This sudden drop in continuity, or test-retest reliability, is even more surprising given the moderate, and significant, correlations of the measure in first and third grade. This finding could be the result of several different factors. First, the lower reliability of the WUF measure, in relation to the other DIBELS measures, could influence the correlations. The reliability of the WUF measure decreases slightly from kindergarten through second grade, making it a more optimal measure in the early grades. This decreasing reliability may suggest difficulty in the WUF's ability to accurately assess vocabulary skills after first grade. As children achieve a higher level of linguistic complexity in higher elementary grades, the measure becomes more difficult to administer. However, higher and significant reliabilities were observed in third grade. These higher reliabilities in third grade could have been due to the ceiling effect of the measure, creating greater stability in scores once students have reached the ceiling around that time period. A lack of growth in scores during that grade would inflate correlations and increase the stability of the measure. It is also very possible that due to the small sample size, all correlations were greatly influenced by a few particular scores. This small sample size could explain the lack of a consistent pattern in within-grade WUF correlations.

In addition to analyzing the relations of within-grade benchmarks, relations between grade levels were also analyzed. In order to do this, the average of the fall, winter, and spring benchmark at each grade level was computed and used for the correlations presented in Table 5. These correlations provide a clearer picture of between-grade continuity in the measure's ability to assess vocabulary skills. The WUF measure has reported reliability scores lower than the other DIBELS measures, in the .6 to .7 range. These reported reliabilities are appropriate for screening purposes, which is the intent of the measure. However, increasing the reliability of the measure, by increasing the number of probes the score is taken from, provides more confidence in reported scores used for data analysis. It should be noted that the kindergarten score used in this analysis was based only on the spring benchmark score, as this was the first and only time the measure was administered in kindergarten. Results should be interpreted with this in mind.

Table 5.

Between-Grade WUF Composite Score Correlation Coefficients

Benchmark period	<i>r</i>	<i>n</i>
K to 1 st grade	.08	19
1 st to 2 nd grade	.58**	23
2 nd to 3 rd grade	.28	27

** $p < .01$.

Results indicate positive correlations between grade levels. The correlation between first and second grade WUF scores was the strongest at $r = .58$. The difference in test administration conditions previously noted may explain the low correlation between kindergarten and first grade scores. The lower correlation between second and third grade could be due to the ceiling effect observed at this time. Still, results suggest a moderate level of continuity of the WUF measure when administered consistently across time, from first to second grade.

Overall results indicate the WUF measure to be a more stable and reliable measure of vocabulary in first to the beginning of second grade. This result is promising given the focus of the DIBELS measures in providing an early indicator of risk status. Higher utility of the DIBELS measures at earlier grades provides teachers with an early indication of which students may require additional support in early literacy and oral language.

Predictive Validity

Research Question 2: What is the predictive validity of a measure of student's skills in defining or using vocabulary words in a sentence, given in kindergarten and first grade, with the GMRT-4 at the end of third grade?

The DIBELS measures were created to provide an early indicator of students' level of risk with later literacy and oral language. In order to examine the particular ability of the WUF to predict reading outcomes, a series of correlation coefficients were

calculated between early WUF scores in kindergarten and first grade with end of third grade reading outcomes. Criterion outcome measures included the Gates-MacGinitie Vocabulary (GMRT-V) and Comprehension (GMRT-C) subtests, and the end-of-third grade DIBELS ORF and RTF scores. Table 6 presents the predictive, criterion-related validity coefficients of the WUF measure in kindergarten and first grade.

Table 6.

Predictive, Criterion-related Validity Coefficients for K and First Grade WUF with End-of-Third Grade Criterion Measures

WUF Benchmark	GMRT-V		GMRT-C		DORF		RTF	
	<i>r</i>	<i>n</i>	<i>r</i>	<i>n</i>	<i>r</i>	<i>n</i>	<i>r</i>	<i>n</i>
End of K	.50*	18	.61*	16	.52*	19	.42	19
Beg of 1 st	.22	21	.23	18	-.09	22	.16	22
Mid of 1 st	.35	22	.21	19	.12	23	.41	23
End of 1 st	.19	22	-.09	19	.02	23	.10	23
1 st Composite	.35	21	.15	18	.08	22	.30	22

Note. A 1st Grade composite score was created by taking the average of all three first grade benchmark scores.

* $p < .05$.

Results indicate a moderate relation between end of kindergarten WUF scores and end-of-third grade reading outcomes. Correlation coefficients for the end-of-kindergarten

WUF scores with each of the end-of-third grade criterion measures were all significant, with the exception of the RTF measure, which approached significance. Correlation coefficients ranged from $r = .42$ to $r = .61$. These correlations are impressive, given the time period between the two measures from kindergarten to third grade. No other significant correlations were observed. Correlations ranged from $r = -.09$ to $r = .23$ for the beginning of first grade, from $r = .12$ to $r = .41$ for the middle of first grade, and from $r = -.09$ to $r = .19$ for the end of first grade. Correlations of similar magnitude have been found in previous research and may be educationally meaningful (Scarborough, 1998). However, due to the small sample size and the lack of power, the similar first grade correlations obtained in this study were not significant. The lack of significant first grade correlations does not support the relation between first grade WUF and third grade reading outcomes. Alternatively, the modest power makes it difficult to firmly conclude that the relation does not exist. Thus, results should be interpreted with caution. Appendix C provides the scatterplots for the correlations between kindergarten and first grade WUF measures with the criterion outcome measures at the end of third grade.

These findings are not surprising, given the nature of the low correlations between the end-of-kindergarten WUF scores and the first grade WUF scores. The lack of consistency in data collection procedures from kindergarten to first grade may help to explain why kindergarten correlations are significant and first grade correlations are not significant. It is possible that the WUF measure, when administered appropriately, is powerful in predicting outcomes but lacks predictive power when the administration is less standardized. Alternatively, it is also possible that the significant correlations

observed in kindergarten are due to the small sample size and not representative of the WUF's true predictive utility.

An examination of the distribution of kindergarten WUF scores in Appendix B reveals a high number of zero scores. This floor effect in the spring of kindergarten may have had a large effect on the correlations in kindergarten. Due to the small sample size, a few scores could have unduly influenced the correlations. However, an examination of the scatterplots in Appendix C reveals that although some students who scored zero on K WUF were able to perform well on the third grade outcome measures, no student who achieved a score greater than zero performed poorly on the third grade outcome measures. This finding would indicate that although initial zero scores might be difficult to interpret, higher scores on the WUF measure might allow for more confidence that students are on track for reading success. All results should be interpreted with the limited sample size in mind.

Research Question 3: Does the vocabulary screening measure, WUF, add to the predictive validity of measures of phonological awareness (PSF), phonics (NWF), and oral reading fluency (DORF) in predicting outcomes on end-of-third grade measures of reading achievement?

Examining the degree to which WUF predicts comprehension and vocabulary outcomes is important. However, the WUF measure was never intended to individually predict outcomes. It was designed as a part of a comprehensive battery of DIBELS

assessments that target all of the core components of reading. Currently, WUF is one of the most under-utilized measures. Educators have typically administered the DIBELS measures that assess phonemic awareness (PSF), phonics (NWF), oral reading fluency (DORF), and to a lesser extent, reading comprehension (DORF and RTF). These measures, when given in combination, provide a powerful indicator of future reading outcomes. Their ability to predict outcomes with a high level of accuracy is even more impressive, given that they are all one-minute measures. It is then important to examine the ability of the WUF measure to predict outcomes above and beyond that of the currently utilized DIBELS measures.

A series of hierarchical multiple regression analyses were conducted in order to examine how much additional variance in reading outcomes was predicted by early WUF scores. End-of-kindergarten WUF scores and first grade WUF scores for each benchmark period were used in the regression models to examine how much additional variance in end-of-third grade reading outcomes they would explain after accounting for the other DIBELS measures.

Results indicated that only the end-of-kindergarten WUF data yielded a significant amount of additional variance explained in reading outcomes. Table 7 shows the results of the regression analyses comparing the end-of-kindergarten DIBELS data with the end-of-third grade GMRT-Comprehension, GMRT-Vocabulary, and DORF scores.

Table 7.

Hierarchical Multiple Regression Between End-of-K Early Literacy and Language and End-of-Third Grade Reading Outcomes

Variable	R^2	R^2 change	Partial correlations	Zero-order correlations
GMRT-Comprehension ($n = 16$)				
NWF	.51		.61*	.71**
WUF	.60	.10	.45	.61*
GMRT-Vocabulary ($n = 18$)				
PSF	.46		.71**	.68**
WUF	.63	.17*	.56*	.50*
Oral Reading Fluency (DORF) ($n = 19$)				
NWF	.27		.37	.52*
WUF	.37	.10	.37	.52*

** $p < .01$. * $p < .05$.

Due to a small sample size, only two independent variables were entered into each regression model. The second variable entered in each model was end-of-kindergarten WUF scores. The first variable entered into the model was determined based on an examination of the correlation coefficients between the kindergarten measures of NWF and PSF with each of the three dependent variables used in the regression models. The early literacy measure that demonstrated the highest correlation with the particular criterion measure was used in the first step of each final regression model. This was done

to ensure the highest percentage of variance explained by the initially entered variable and to avoid overestimation of the added contribution of the WUF scores.

Each regression model explained a significant amount of variance in end-of-third grade reading outcomes. The addition of WUF scores explained a minimum of 10% additional variance after accounting for the basic early literacy skills of phonemic awareness or phonics. However, only the model using the GMRT-Vocabulary test as the dependent variable yielded an additional amount of variance explained by the WUF, 17%, that was significant.

Although first grade WUF scores for each benchmark period did not explain a significant amount of additional variance in reading outcomes, middle-of-first grade results approached significance. Middle-of-first grade WUF scores explained an additional 10% of the variance in the third grade GMRT-V scores, after accounting for middle-of-first grade DORF scores. Table 8 presents the results for the middle-of-first grade regression models. Results for all other remaining first grade WUF benchmark scores explained less than 8% additional variance after accounting for the other DIBELS measures, and were non-significant.

Table 8.

Hierarchical Multiple Regression Between Middle-of-First Grade Early Literacy and Language and End-of-Third Grade Reading Outcomes

Variable	R^2	R^2 change	Partial correlations	Zero-order correlations
GMRT-Comprehension ($n = 19$)				
DORF	.31		.56*	.56*
WUF	.34	.03	.21	.21
GMRT-Vocabulary ($n = 22$)				
DORF	.31		.57**	.56**
WUF	.41	.10	.38	.35
Oral Reading Fluency (DORF) ($n = 23$)				
DORF	.33		.57**	.57**
WUF	.34	.01	.10	.12

** $p < .01$. * $p < .05$.

These results suggest that the WUF measure, when administered in kindergarten with currently utilized DIBELS measures of early literacy, may explain an additional 10-17% of the variance in end-of-third grade reading outcomes. In educationally meaningful terms, this could mean that a few more kindergarten children that are at risk for reading difficulties in the area of comprehension and vocabulary would be identified if the WUF measure were used as part of the DIBELS battery. Without the use of the WUF measure at this particular time, some children who may require additional early literacy and

language support may not be identified and would thus not receive that additional support. The percent of unique variance explained by the WUF measure in first grade was similar to results of previous studies examining the added contribution of oral language measures to predicting later reading outcomes (Roth, Speece, & Cooper, 2002; Senechal, Ouellette, & Rodney, 2005). Results of these previous studies may be educationally meaningful in identifying at-risk readers. However, due to the small sample size and the lack of power, the similar first grade results obtained in this study were not significant, and make it difficult to firmly conclude whether a relation between first grade WUF and later reading outcomes does or does not exist. All results should be interpreted cautiously due to the limited sample size and the non-significant findings in first grade. A few particular scores in kindergarten or first grade could have overly influenced the percent of additional variance explained.

Vocabulary and Reading Comprehension

Research Question 4: What is the relation between a vocabulary screening measure (WUF) and a reading comprehension measure in which students retell what they orally read (RTF)?

Both the WUF and RTF measures were designed to provide additional indicators of reading competence in the areas of vocabulary and comprehension (Good & Kaminski, 2002). Basic vocabulary knowledge is a necessary prerequisite to understanding connected text (Anderson & Freebody, 1981; Kamil, 2004). One must understand the

meanings of individual words before they can understand the meaning of connected text. There is some support for the hypothesis that vocabulary and general oral language skills play an increasingly larger role in reading comprehension as students progress through school (Senechal, Ouellette, and Rodney, 2005; Storch & Whitehurst, 2002).

In order to examine the ongoing relation between WUF, a measure of vocabulary, and RTF, a measure of reading comprehension, a Pearson correlation coefficient was calculated between these two measures at each benchmark period from the middle of first grade through the end of third grade. The middle of first grade is the first time both the RTF and WUF measures are typically given. Table 9 presents the correlation coefficients.

Table 9.

Correlation Coefficients Between Concurrent RTF and WUF Scores

Grade	Beginning		Middle		End	
	<i>r</i>	<i>n</i>	<i>r</i>	<i>N</i>	<i>r</i>	<i>n</i>
1			.23	23	.04	23
2	.13	26	.26	27	.38*	27
3	.41*	28	.23	29	.14	29

* $p < .05$.

Results indicate a non-significant relation between the two variables for most benchmark periods. A moderate correlation was found for the two measures during the end of second grade and the beginning of third grade. These two correlation coefficients

were $r = .38$ and $r = .41$, respectively. Both of these correlations were significantly different from zero ($p < .05$). A pattern of increasing correlation coefficients was observed from the end of first grade through the beginning of third grade, although many of the correlations were not significant. The strength of the relation then fades towards the end of third grade. In first grade, correlations between WUF and RTF ranged from .04 to .23. In second grade, the correlations ranged from .13 to .38. In third grade, the correlations ranged from .14 to .41. Although many of the correlations were not found to be significant, these findings do show some support for the hypothesis that vocabulary and oral language skills play a larger role in reading comprehension as students progress through school. In general, the data reflects a weak relation between vocabulary and reading comprehension that exists as students begin to read connected text in first grade. This weak relation then becomes stronger as the level of text difficulty becomes more advanced and a more extensive knowledge of word meanings is necessary in second and third grades. However, it is unclear whether a larger sample size would have yielded significant correlations in the same pattern as the correlations observed in this study.

Prior research has demonstrated the utility of oral reading fluency as a global measure of overall reading competence, with oral reading fluency explaining much of the variance in measures of reading comprehension (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Jenkins et al., 2003). Readers who have good comprehension skills generally read more fluently, and vice versa. As a result, many educators use oral reading fluency as a general indicator of reading comprehension. A hierarchical multiple regression analysis was conducted in order to examine whether vocabulary plays an additional role in reading

comprehension, above and beyond the impact of fluent reading. For this analysis, RTF was entered as the dependent variable into the regression model. DORF was entered as the initial independent variable followed by WUF. This analysis was conducted for each benchmark period for which RTF, DORF, and WUF data is available. This allowed for an examination of whether the contribution of vocabulary skills to reading comprehension changes over time, from first through third grade. Results of the regression models for first, second, and third grades are represented in Tables 10, 11, and 12, respectively.

Table 10.

Hierarchical Multiple Regression Between First Grade DORF, WUF and RTF

Variable	R^2	R^2 change	Partial correlation	Zero-order correlation
Middle ($n = 23$)				
DORF	.61		.79**	.78**
WUF	.64	.03	.28	.23
End ($n = 23$)				
DORF	.23		.48*	.48*
WUF	.24	.01	.08	.04

** $p < .01$. * $p < .05$.

Table 11.

Hierarchical Multiple Regression Between Second Grade DORF, WUF, and RTF

Variable	R^2	R^2 change	Partial correlations	Zero-order correlations
Beginning ($n = 26$)				
DORF	.42		.65**	.65**
WUF	.44	.02	.17	.13
Middle ($n = 27$)				
DORF	.31		.56**	.55**
WUF	.36	.05	.22	.26
End ($n = 27$)				
DORF	.22		.41*	.47*
WUF	.29	.07	.29	.38*

** $p < .01$. * $p < .05$.

Table 12.

Hierarchical Multiple Regression Between Third Grade DORF, WUF, and RTF

Variable	R^2	R^2 change	Partial correlations	Zero-order correlations
Beginning ($n = 28$)				
DORF	.23		.48*	.48*
WUF	.37	.14*	.43*	.41*
Middle ($n = 29$)				
DORF	.35		.61**	.59**
WUF	.40	.05	.28	.23
End ($n = 29$)				
DORF	.14		.39*	.37*
WUF	.16	.02	.17	.14

** $p < .01$. * $p < .05$.

Results indicate that vocabulary skills, as measured by WUF, do not generally contribute a significant amount of additional variance in reading comprehension, as measured by RTF, after controlling for oral reading fluency. Only in the beginning of third grade did vocabulary explain a significant additional amount of variance in reading comprehension at 14%. There are several possible explanations for this finding.

One possible explanation for this finding is the role that vocabulary plays in fluent reading. An extensive vocabulary knowledge base allows for more accurate and fluent reading. As a student reads a familiar vocabulary word, a process of semantic activation

occurs, priming the student for all of the known words related to the vocabulary word they just encountered (Posner & Snyder, 1975). As a result, they can apply top-down processes to match words in the sentence they are currently reading to those words activated in their semantic network. In this way, students who have good semantic networks, or well developed vocabularies, may be able to read more fluently than those students who share the same phonological skills but possess less extensive oral vocabularies. So controlling for oral reading fluency in this analysis may have already accounted for the effect of vocabulary, allowing little variance remaining to be explained by the WUF measure.

A second explanation is that the interrelatedness of the DORF and RTF measures may make it difficult for WUF to explain any additional variance. The DORF and RTF scores are based off of the same reading passages. Students receive the DORF score from orally reading a passage. They receive the complimentary RTF score for retelling everything they recall from that same passage. The WUF measure tests students on vocabulary words, which may be unrelated to the particular passages the DORF and RTF scores are based on. Even if a student has a good knowledge of the specific vocabulary words on the WUF probe, it may not carry over to the DORF and RTF scores if the material in those passages is unrelated to the material on the WUF probes.

Despite the lack of results indicating a significant amount of additional variance in RTF explained by the WUF scores, the overall pattern of results fits well with previous research examining the relation between vocabulary and comprehension through the elementary years. Overall, the amount of additional variance explained by WUF

increased from first through third grade. In first grade, the WUF scores explained an additional 1-3% of the variance in RTF scores. In second grade, WUF scores explained an additional 2-7% of the variance. In third grade, WUF scores explained an additional 2-14% of the variance. Even after controlling for oral reading fluency skills, vocabulary continues to play an increasingly larger role in reading comprehension, as students get older. However, these results were not significant, and should be interpreted cautiously. The lack of significant results in the findings of this particular study could have been due to a number of factors. Small sample size or other methodological limitations, may have limited the magnitude of the relations. Alternatively, the WUF and RTF measures may not have had sufficient validity as measures of vocabulary and comprehension to produce significant results.

Summary of Results

An analysis of the descriptive statistics indicated that despite lower initial scores on the WUF measure than previously seen in the research, student scores on all measure fell within the average range and followed an increasing trend over time. The WUF measure appeared to be a fairly stable measure of oral language skills in first to the beginning of second grade, but low correlations in second and early third grade indicated difficulty of the WUF measure in adequately demonstrating the stability of oral language skills at that particular time. Across grade levels, the WUF was most stable from first to second grade, but demonstrated lower positive correlations between all grade levels. An examination of *predictive validity* revealed several findings. First, moderate correlations

between end-of-kindergarten WUF scores and end-of-third-grade criterion measure scores indicated good predictive validity of the WUF measure in predicting later reading comprehension and vocabulary outcomes. Even after accounting for the high predictive ability of concurrent phonological skills, early oral language skills in kindergarten still accounted for around 10-17% additional variance explained in end-of-third-grade reading outcomes. But despite these promising findings for the WUF measure administered at the end of kindergarten, first grade WUF scores did not predict end-of-third-grade reading outcomes. This may have been due to the different data collection procedures utilized at kindergarten and first grade. Additionally, small variations in a select number of scores could have greatly influenced the analysis, due to the small sample size. Finally, an examination of the relation between vocabulary and comprehension revealed an increasingly stronger relation from first through third grade. Although analyses were generally not significant, the pattern of correlations and regression results suggested an increasing relation between vocabulary and comprehension as students progressed to the higher grade levels.

CHAPTER V

DISCUSSION

The DIBELS Word Use Fluency (WUF) measure was originally designed to be a measure of early oral language skills. More specifically, it was designed to be an indicator of vocabulary and expressive language skills, one of the five core components of reading identified by the National Reading Panel report (2000). Given the intent of the measure, the incorporation of the WUF into the full battery of DIBELS measures should provide educators with a more complete picture of students' early literacy and language skills. A comprehensive screening of all students would allow educators to identify those particular students with early deficits, and thus provide them with preventative interventions aimed at ensuring they become readers by the end of third grade. The goals of this study were to a) examine the ability of the WUF measure to provide a stable assessment of vocabulary over time, b) examine the predictive utility of the WUF measure in identifying those students who would later struggle in reading, and c) examine the relation between WUF and RTF as measures of vocabulary and reading comprehension, respectively. The results of the study suggest that a) the WUF measure does provide some stability in measuring early vocabulary across some grade levels with mixed results within grades, b) WUF scores in kindergarten do explain some additional variance in later reading comprehension and vocabulary, even after controlling for phonological measures, and c) the WUF and RTF measures are more related in third

grade than in first and second grades. The following section will discuss the implications of the findings, limitations of the study, and directions for future research.

Interpretation of Results

This study examined the utility of the DIBELS WUF measure in helping to predict children who may struggle in later reading comprehension and vocabulary. The following section will discuss the WUF measure's technical adequacy as it relates to predicting reading comprehension and vocabulary outcomes.

WUF as a Measure of Vocabulary Skills Over Time

The DIBELS WUF measure was originally intended to be an early indicator of students' vocabulary skills. As an early indicator, it was never intended to provide a comprehensive assessment of all vocabulary and oral language skills. It provides a way to screen all students to determine if they may be at-risk for reading difficulties as a result of limited oral language skills and a way to monitor progress in vocabulary and expressive language. One of the goals of this study was to examine whether the WUF measure provides a stable indicator of vocabulary skills over time. Stability in measurement is desirable to allow educators to track student progress over time.

Results of this study were mixed, depending on grade level. In examining within-grade correlations, the WUF provided a more stable measure of vocabulary skills in first grade, with a low level of stability in second and the beginning of third grade. In examining between-grade correlations, WUF scores were most stable between grades one

and two, and between grades two and three. Kindergarten and first grade scores were not related. These findings were somewhat surprising. One possible explanation for the lack of stability in the WUF scores at the various grade levels is the variability in test administration procedures over time. Kindergarten WUF scores were collected by graduate student data collectors. School testers began giving the WUF measure to this cohort of students in the fall of their first grade year. This change in testers and testing conditions may have influenced the kindergarten to first grade correlation. As school testers became more familiar with the measure over time, tester drift may have also influenced test scores. A further discussion of this limitation is provided in a following section. Results suggest that the WUF measure, when administered under consistent conditions, may allow for a stable measure of vocabulary skills as demonstrated by moderate, significant correlations in first grade.

Additional Benefit of the WUF Measure

In general, educators currently utilizing the DIBELS assessments administer the following DIBELS measures as part of their comprehensive assessment system: Letter Naming Fluency (LNF), Initial Sound Fluency (ISF), Phoneme Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), Oral Reading Fluency (ORF), and Retell Fluency (RTF). These measures provide research-based indicators of early literacy skills in four of the five big ideas identified by the National Reading Panel, with the exception of vocabulary. Currently, these measures possess good predictive validity in identifying future reading performance. For example, students who meet or exceed the research-

based benchmarks on these particular assessments are about 90% likely to meet future reading benchmarks, putting them on track to become successful readers by the end of third grade. But despite these research-based benchmarks, not all students meeting these benchmarks become successful readers (Riedel, 2007). Decoding skills and even reading fluency have been described as a necessary, but not sufficient, prerequisite for reading comprehension (Perfetti, 2007). This could be due to deficiencies in the instructional environment, student disabilities, oral language deficits, or any other number of additional factors impacting reading performance.

The primary purpose of this study was to examine whether adding an indicator of early vocabulary skills to commonly utilized early literacy measures would improve the ability to predict reading outcomes. Results were mixed, depending on the grade level. In kindergarten, results indicated that WUF scores were able to predict 10-17% additional variance in end-of-third-grade reading comprehension and vocabulary outcomes. In educational terms, this could mean the early identification of one or two additional students per class that will experience reading comprehension difficulties. These students are those children that may not be otherwise identified without including the WUF measure as part of the DIBELS battery of assessments. Kindergarten results should be interpreted with caution due to the high number of zero scores and positively skewed distribution. These scatterplots can be reviewed in Appendix B. First grade WUF scores predicted 1-10% additional variance in end-of-third grade reading outcomes. These results were not significant and should also be interpreted with caution due to the limited sample size and modest power of the analyses.

These results are most likely due to the lack of consistent test administration procedures over time. The same testers did not administer the measures at the two separate grade levels. Graduate students administered the kindergarten measures and teachers administered the first grade measures. In first grade, teachers were just learning to administer the measure, which may have influenced the reliability and validity of those scores. Alternatively, it is also possible that first grade scores may have been a more accurate representation of students' vocabulary skills than their kindergarten scores. Test administration by familiar teachers in first grade could have produced more accurate results from these students. This particular explanation would indicate a more limited utility of the WUF measure in helping to predict future reading outcomes. Previous research has indicated that reading vocabulary measures correlate more highly with measures of reading comprehension than do oral vocabulary measures (Tannenbaum, Torgesen, & Wagner, 2006). Reading vocabulary and reading comprehension measures both tap into word-reading abilities, whereas oral vocabulary measures do not measure any sort of reading ability. Future research is needed to clarify the nature the WUF measure as it contributes to predicting future reading outcomes.

The Relation Between Vocabulary and Comprehension

In examining the relation between the WUF and RTF measures, results indicated an increasing relation between these measures as students progressed in school. Though many of the results were non-significant, the patterns in the data are similar to previous research that has indicated an increasing relation between vocabulary and comprehension

as students get older and the level of text difficulty gets harder (Catts, Hogan, & Adlof, 2005; Storch & Whitehurst, 2002; Yovanoff, Duesbery, Alonzo, & Tindal, 2005). The results of the present study indicate that WUF and RTF may serve as useful indicators of vocabulary and reading comprehension. However, due to the previously discussed inconsistencies in the stability of WUF scores over time in this particular sample, results should be interpreted with caution.

Implications for Practice

One interesting finding of this study is the possible utility of WUF scores in predicting later reading comprehension outcomes. Significant bi-variate correlations of .50 to .61 were observed between end-of-kindergarten WUF scores and end-of-third-grade reading comprehension scores. Even after controlling for measures of phonological awareness and phonics, vocabulary did play a role in predicting reading comprehension and vocabulary outcomes two to three years later, explaining an additional 10-17% of the variance. First grade WUF predicted an additional 1-10% additional variance in third grade reading outcomes after controlling for phonic decoding. However, these results were not significant. Use of the WUF measure with other DIBELS measures could result in the identification of a small number of additional students who may be at risk for difficulties in the area of reading vocabulary that would not otherwise be identified. These implications should be received cautiously due to the small sample size, the high number of zero scores in kindergarten, and the positively skewed distributions. These

statistical limitations limit the ability to make firm conclusions. More research is still needed to further explore the relation between WUF and future reading outcomes.

Although kindergarten findings were promising, the first grade WUF scores were not helpful in predicting later reading comprehension or vocabulary outcomes. Within the proposed theoretical model of vocabulary and reading comprehension, correlations between first grade vocabulary and later reading comprehension should be much higher. Once again, the lack of consistency in test administration procedures over time may have accounted for this surprising finding. Highly trained graduate student data collectors administered all kindergarten DIBELS measures, including the WUF measure. In addition to their extensive training in DIBELS standardized administration, these data collectors also had a background in more general standardized testing procedures and foundational knowledge in educational testing and measurement. In contrast, resident school staff administered the first grade DIBELS measures. Though the staff did receive appropriate training in DIBELS test administration procedures, they did not have the foundational knowledge and extensive practice that the graduate student data collectors had.

Improvements in DIBELS training for school staff may be a useful way to improve accuracy of scores, thus increasing the utility of the measure. The WUF measure is one of the more difficult measures to administer, making it less reliable. Testers must record the number of words in verbal utterances provided by students who may be speaking quietly or quickly, and with unclear enunciation. This makes it difficult to accurately hear the number of words spoken by the student, much less determine if the

utterance is correct or not. Although data collectors can be trained for standardized administration, it is sometimes difficult to react to actual testing situations when children provide unusual answers or exhibit challenging behavior. When this occurs, a background in educational testing and measurement can help to ensure that the most reliable and valid scores are obtained.

In improving the reliability of test administration, there are several suggestions. First, trainings should include multiple opportunities for testers to administer the measure to a variety of children of various grade levels. This will help testers develop an understanding of the range of responses children will typically provide, and allow them to develop fluency with the scoring procedures before collecting data that will be used for educational decision-making. Second, whereas many schools only train the basics of actual test administration, training that includes a foundation of why they are collecting the data and what the data actually means should help to improve test administration and the quality of the data. Schools are then encouraged to include this foundational knowledge into their trainings, as the extra time taken during training may result in more reliable and valid data.

There are also currently efforts to improve the reliability of the measure through test development. For example, current research is examining the utility of clarifying standardized administration directions and providing different pools of words used in probes. Research is also examining the relation of the WUF to other oral language measures. These changes to the WUF measure should help to improve its technical adequacy, addressing one of the biggest concerns with the current measure.

The changing relation between the WUF and RTF measures also has interesting implications for educators. Measures of oral reading fluency (ORF) provide a good indication of overall reading comprehension (Fuchs, Fuchs, Hosp & Jenkins, 2001). The addition of retell fluency (RTF) measures to ORF provides an additional comprehension check that can ensure students who are reading fluently also understand what they are reading (Roberts, Good, & Corcoran, 2005). Results of the current study indicate that once students are reading connected text in the middle of first grade, ORF and RTF provide a better indication of text comprehension than WUF. After controlling for the effects of ORF and RTF, WUF does not explain meaningful additional variance in reading comprehension outcomes, with the exception being in the beginning of third grade. However, WUF may be useful to educators as an earlier comprehension check before students begin to read connected text. Kindergarten WUF scores were predictive of later reading comprehension and vocabulary outcomes. This finding suggests the WUF measure may serve as a downward extension of reading comprehension indicators by providing a more general language comprehension check. Educators would then be able to have a better idea of which students are likely to struggle with reading comprehension, despite mastering the prerequisite skills of phonological awareness, phonics, and fluency and accuracy with connected text.

Limitations

There are several limitations of this study including the technical adequacy of the WUF measure, the design and implementation of the study, and the particular criterion

measures used in this study. These limitations pose threats to internal and external validity and should be considered in the interpretation of the final results.

Technical Adequacy of the WUF

The WUF measure is currently one of the least utilized DIBELS measures by educators. This low usage is partially due to the lower reliability and validity of the measure, and a lack of clear benchmarks for student achievement. The reported reliability of the WUF, ranging from .65 to .71 (Kaminski et al., 2004), makes it one of the less reliable of the DIBELS measures. This lower reliability is most likely due to the increased difficulty of administration and scoring, as compared to the other DIBELS measures. In this study, a high number of zero scores were observed on the kindergarten WUF measure. Appendix B provides the distribution of kindergarten WUF scores. This high number of zero scores may have been due to difficulty understanding the directions and confusion with the task for some children. As a result, zero scores on the kindergarten WUF measure may not have necessarily been indicative of each student's actual oral language skills. Results indicated that several children with zero scores in kindergarten actually performed highly on future measures of reading comprehension and vocabulary. The positive skew in kindergarten WUF scores, in addition to the small sample size makes it difficult to have full confidence in the technical adequacy of these scores and the correlation coefficients reported in this study. Though the WUF measure increases in reliability through administration of multiple probes, participant scores in this study were based on single probes. The full benefit of the WUF measure in

educational decision-making would be the utilization of scores based on an aggregate of multiple probes. Results should be interpreted with this mind.

Design and Implementation

This study was intended to provide an examination of the utility of the WUF measure as predictive of later reading comprehension outcomes. In order to achieve this, a longitudinal data set containing four years of a single cohort's data was needed. As a result, the study used an existing data set that had been collected over the past four years in an actual school. Utilization of an existing data set for this study provided unique strengths and limitations. On one hand, the data does provide a view of how the WUF measure functions in actual school settings, with real teachers and school staff collecting the data. However, the data collection procedures make it difficult to have full confidence in the technical adequacy of the data. The researcher was unable to control for the training received by the data collectors and the specific data collection procedures utilized in first through third grades in this sample. Alternatively, the researcher collected the kindergarten DIBELS data and the GMRT-4 criterion outcome data. This lack of consistent data collection procedures over time makes it difficult to examine the continuity of the WUF measure in kindergarten and first grade, as well as the predictive validity of data collected under different conditions.

In addition to the problem of data collection procedures, the small sample size in this particular study limits the interpretation of the results. Due to the strict participation requirements of the study, there was difficulty in locating schools that met eligibility

criteria. Participating schools had to have collected comprehensive DIBELS data, including the WUF measure, for the past four years. Whereas the WUF measure is a relatively newer measure, lacking the extensive research base of the other DIBELS measures, most schools do not regularly collect WUF data as part of their school-wide DIBELS administration. This resulted in a very limited sample, both in size and diversity. With a sample size this small, it is possible that a few scores could have overly influenced the results of the analyses. All results should be interpreted with this limited sample in mind.

Finally, the correlational design of this study makes it difficult to determine if certain variables have a causal effect on other variables. For example, despite moderate correlations between the WUF measure and future measures of reading comprehension and reading vocabulary, this study makes it difficult to determine whether improving vocabulary actually improves reading outcomes. There is the possibility that a third variable, such as socio-economic status (SES), is mediating the relation between the early vocabulary skills and later reading outcomes. Thus, there is a need for prospective experimental research that manipulates vocabulary skills in kindergarten and examines the impact on third grade reading comprehension outcomes.

Criterion Measures

Due to limited resources, criterion measures were selected to be efficient and cost-effective. The criterion reading comprehension measures used in this study included DORF, RTF, and the GMRT-4 measures of reading comprehension and vocabulary. The

GMRT-4 measure was chosen based on its relatively good technical adequacy and because it is a standardized, group-administered assessment. Though this allowed for efficient criterion measure data collection procedures, an individually administered and more specific comprehension and vocabulary measure may have provided different results. This may have also allowed for more accurate representations of participants' vocabulary and reading comprehension skills.

The DORF and RTF measures were also used as comprehension measures in this study. Oral reading fluency has been found to be a good indicator of overall reading competence and can subsequently be used as an indicator of reading comprehension (Fuchs, Fuchs, Hosp, & Jenkins, 2001). Research has also indicated the utility of the RTF measure as a reading comprehension indicator (Roberts, Good, & Corcoran, 2005). These measures were included as reading comprehension measures due to their availability to the researcher. The participating school had already collected this data as a part of their normal DIBELS administrations. However, reading comprehension is a complex process involving many sub-skills and processes. The ORF and RTF measures were never intended to provide a comprehensive view of student skills in this area. Rather, they were designed to provide a brief indicator of students' overall understanding of written text. Results should be interpreted with this measurement limitation in mind.

Directions for Future Research

There are several implications for future research as a result of this study. The first, and most important implication is the need for a replication of this study with a

larger sample size. The restricted nature of the sample, both in size and diversity, limits the ability to discuss the results as they apply to theoretical and practical implications. Generalizations to the larger population are limited as a result of the small sample. Though initial study findings are interesting, a much larger sample size is necessary to examine whether the same patterns of data emerge and would provide a much clearer picture of the magnitude of the relations of that data. A larger sample size would help clarify a) whether WUF is a stronger predictor in kindergarten than first grade, b) if there really is no relation between kindergarten and first grade WUF scores, and c) if WUF and RTF are more related in later elementary grades than in the earlier grades. Within the national DIBELS dataset, a larger sample of schools utilizing the WUF measure is available. Access to that dataset would provide a better examination of the research questions.

Second, the different patterns between the researcher-collected data and the school-collected data point to a need to collect a data set under stricter, more consistent administration standards. A longitudinal study with WUF data collected under stricter administration conditions across the early elementary years, would allow for a more accurate representation of the true relation between the DIBELS measures and comprehension and vocabulary outcomes at the end of third grade. A study comparing researcher collected data with school collected data for the same students would clarify the utility of the WUF measure as it is used by schools.

Third, data collection procedures that allow for the collection of multiple probes at each benchmark period from kindergarten through third grade would be helpful.

Multiple probes would help examine whether an aggregate WUF score would improve the magnitude of the predictive validity coefficients with future reading comprehension and vocabulary outcomes. The improvements in these validity coefficients could be weighed against the additional time and resources needed to collect the additional probes at each benchmark period.

Fourth, the use of alternative comprehension and vocabulary assessments would provide additional support for the technical adequacy and predictive utility of the WUF measure. The GMRT-4 provides one efficient measure of reading comprehension. Additional criterion outcome assessments, such as high stakes state testing results, would help to further clarify the WUF measure's predictive utility. This study also relied on the use of RTF and ORF measures as approximations of reading comprehension in first through third grade. Additional comprehension assessments provided in these grades would also help to clarify the concurrent relation between vocabulary, decoding, and reading comprehension and how that relation changes over time.

Future research could also clarify the question of whether a reading vocabulary measure would be more predictive of reading comprehension than the WUF, which is an oral vocabulary measure. Administering the WUF measure by having the students read the stimulus words and provide oral definitions may yield different results than having the stimulus words orally provided to the students through standard WUF administration procedures. The addition of the reading task within the WUF measure may increase the predictive utility for reading comprehension outcomes.

In order to further understand the relation between early vocabulary and later reading outcomes, experimental manipulation of the independent variable would also be helpful. Future research could include intervention studies that use the WUF measure to assess the progress of students receiving oral language instruction over time, as compared to students who are not receiving that instruction. Significant increases in WUF scores as related to specific oral language instruction would provide much more powerful evidence that vocabulary is important for future reading outcomes.

Another particularly important implication for future research is the extension of these research questions to populations of English-language learners (ELL). Many ELL's do not have the same exposure to English language vocabulary as compared to native English-speaking students, resulting in gaps in vocabulary between ELL's and non-ELL's (August, Carlo, Dressler & Snow, 2005). Similar to native English speaking students, research has demonstrated the high linkage of vocabulary to reading comprehension in ELL's, above and beyond the contribution of decoding skills (Proctor, Carlo, August, & Snow, 2005). The WUF measure may have more utility in identifying ELL's who will struggle with reading comprehension despite developing sufficient decoding skills in English. Many ELL's who learn to decode in English may appear on track to becoming successful readers. But their lack of English vocabulary knowledge can prevent them from fully comprehending what they read. It is possible that the results of this study may have looked very different with a population of ELL's. For example, kindergarten and first grade WUF scores may have been much more predictive of future

reading comprehension than the results obtained in this sample of native English speaking students.

Finally, future research should continue to enhance the technical adequacy of the WUF measure. Though current reliabilities of the measure are appropriate for screening purposes, improving the technical adequacy of the measure will help to enhance its utility in making better educational decisions. In its current state, an aggregate of four to six WUF probes would provide a reliability of over .90 (Kaminski et al., 2004). Improving the reliability of the measure would allow educators to make better educational decisions based on fewer probes. Improving the technical adequacy of the measure would also serve to improve the measure's ability to help predict reading comprehension outcomes.

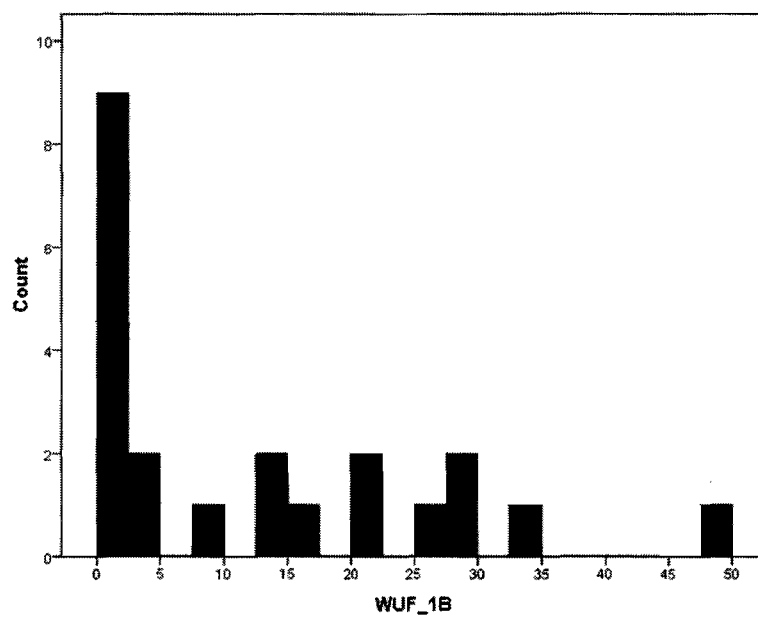
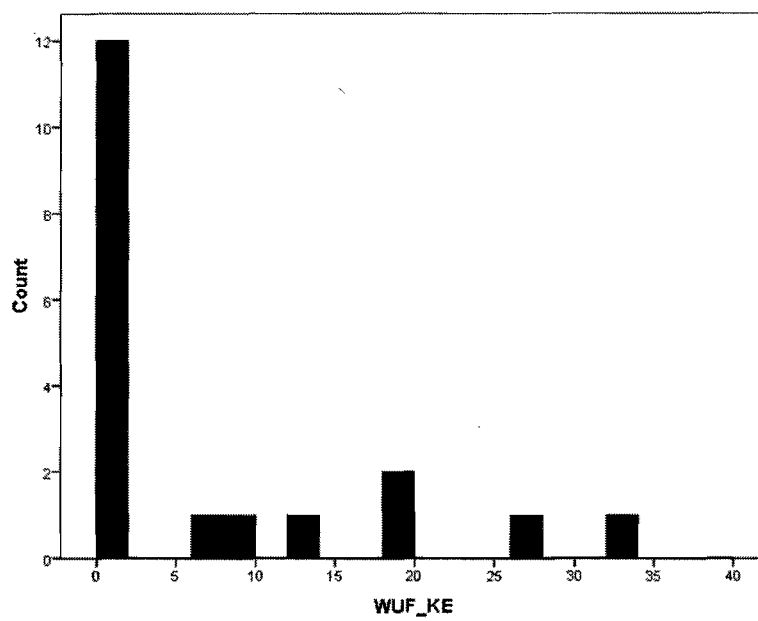
Conclusions

The early identification of students who are at-risk for reading difficulties serves as the foundation for preventing reading failure. The earlier we can identify those children who may require additional support, the higher likelihood we have of ensuring those students become successful readers. There is simply no time to waste in setting all children on the path to literacy, with prevention efforts being much more effective than later interventions after problems have already developed.

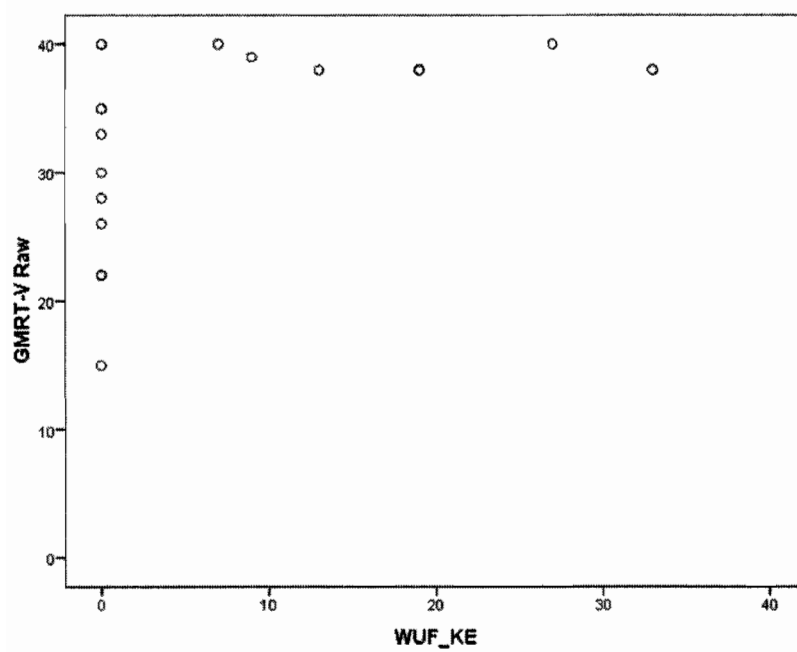
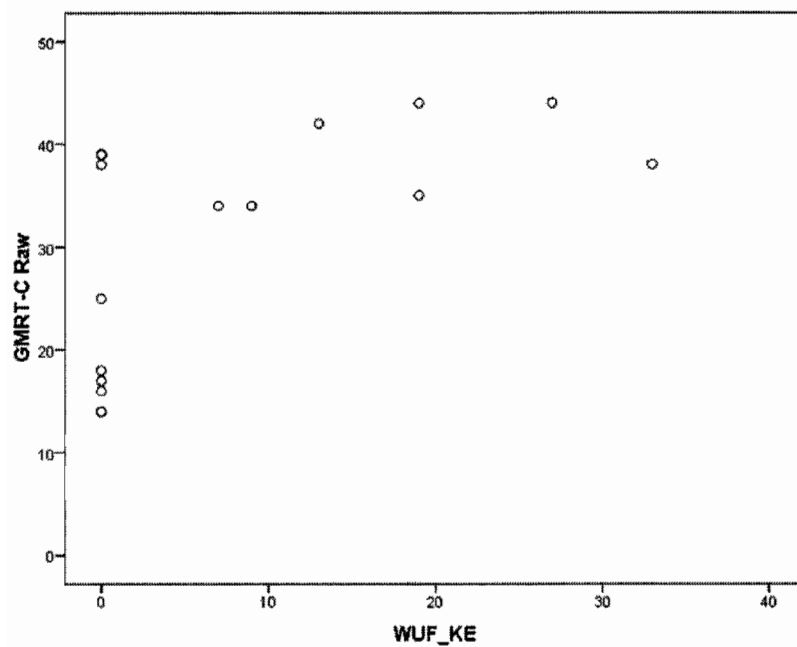
Current research-based screening batteries have improved the ability of educators to accurately and efficiently identify those students who may require additional early support. The most commonly utilized screening batteries contain measures of phonological awareness, phonics, and accuracy and fluency with connected text.

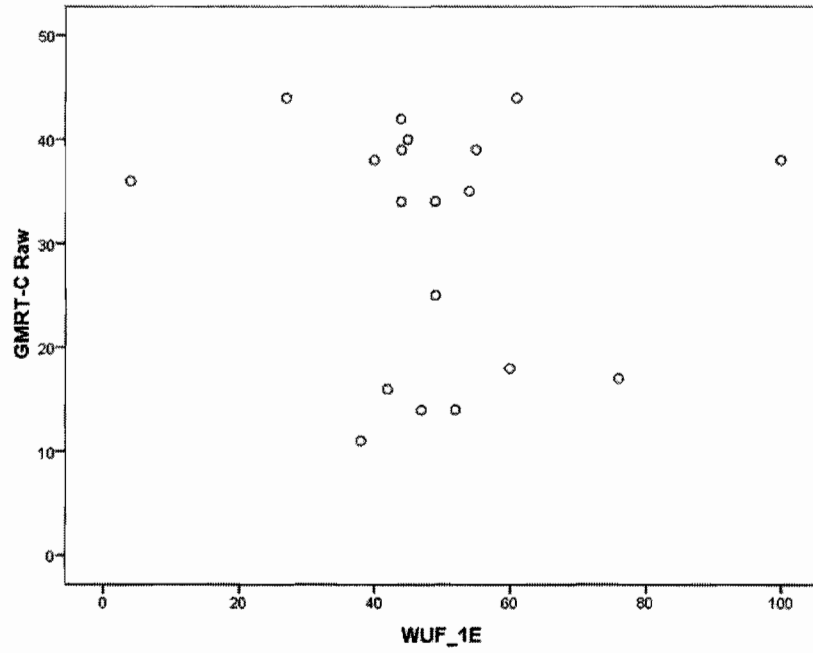
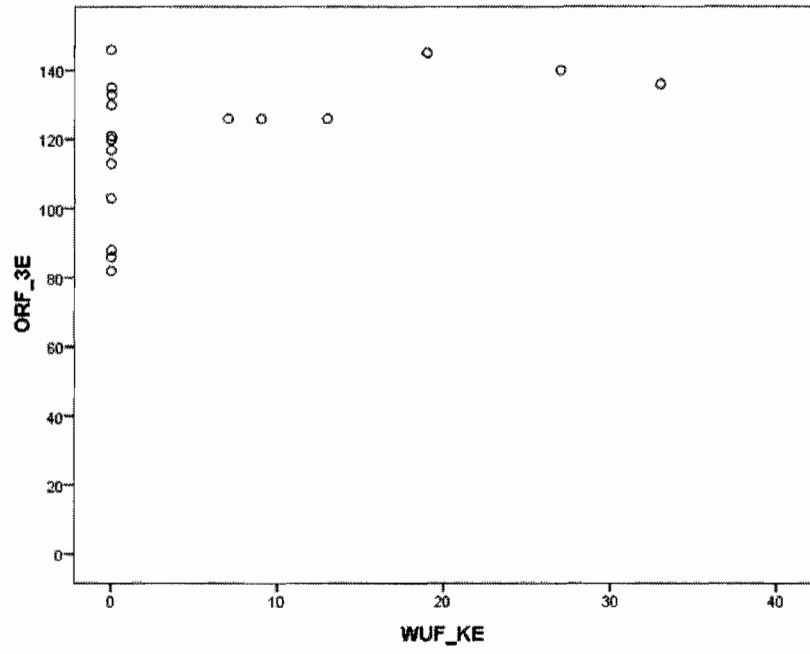
Research has demonstrated these skills to be some of the most reliable predictors of future reading achievement. However, research has also pointed towards the utility of measures of oral language and vocabulary in helping to predict reading comprehension outcomes (Scarborough, 2000). But the nature of most early vocabulary assessments makes them too time intensive to provide for all children as screening measures. The resources needed to implement these assessments on a large-scale basis outweigh the benefit of early identification. Results of the present study suggest that the addition of a brief, early vocabulary measure may provide additional benefit in identifying children who could later struggle in reading comprehension and vocabulary. Results also highlight the promising nature of the WUF measure as an adequate indicator of overall early vocabulary and oral language skills. However, inconsistencies in the results across grade levels indicate the need for further research to clarify the specific relation between the WUF measure and measures of reading comprehension over time, and to improve the ability of the measure to help predict reading comprehension outcomes.

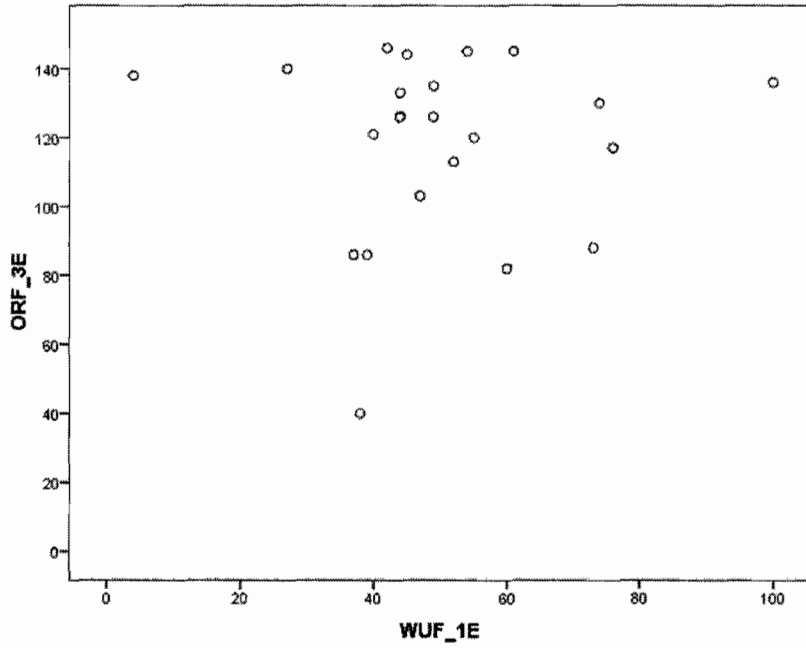
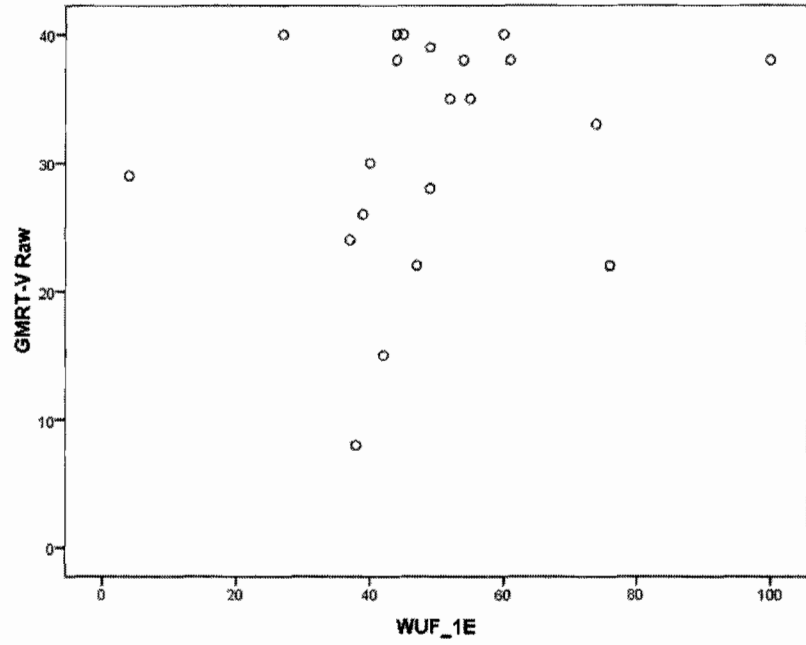
APPENDIX A

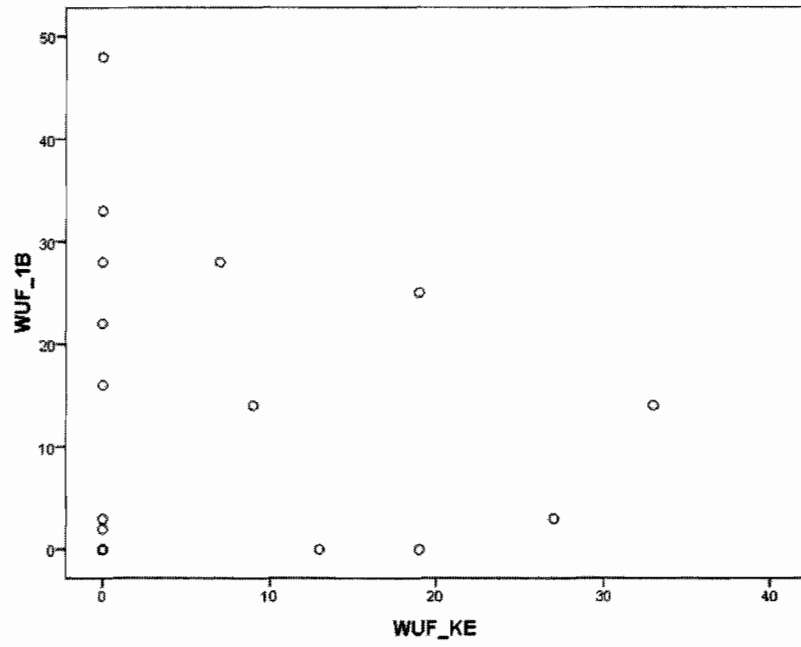
K AND 1ST GRADE WUF DISTRIBUTIONS

APPENDIX B
SCATTERPLOTS









REFERENCES

- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-116). Newark, DE: International Reading Association.
- August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20*(1), 50-57.
- 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.
- Baumann, J. F., Kame'enui, E. J., & Ash, G. E. (2003). Research on vocabulary instruction: Voltaire redux. In D. Lapp, J. Squire, & J. Jensen (Eds.), *Handbook of research on teaching the English language arts* (2nd ed., pp. 752-785). Mahwah, NJ: Erlbaum
- Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal, 107*, 251-271.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York: Guilford Press.
- Biemiller, A. (2005). Vocabulary development and instruction: A prerequisite for school learning. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research: Volume 2* (pp. 41-51). New York: Guilford Press.
- Bishop, D. V. M., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry, 21*, 1027-1050.

- 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.
- Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research, 45*, 1142-1157.
- Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence for a longitudinal investigation. *Scientific Studies of Reading, 3*(4), 331-361.
- Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In H.W. Catts & A.G. Kamhi (Eds.), *Language and reading disabilities*. (pp. 25-40). Boston: Pearson/Allyn and Bacon.
- Chall, J. S. (1983). *Stages of reading development*. New York: McGraw-Hill.
- Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. *American Educator, 22*, 8-15.
- Dickinson, D. K., Cote, L., & Smith, M. W. (1993). Learning vocabulary in preschool: Social and discourse contexts affecting vocabulary growth. *New Directions for Child Development, 61*, 67-78.
- Dickinson, D. K., & Snow, C. E. (1987). Interrelationships among prereading and oral language skills in kindergartners from two social classes. *Early Childhood Research Quarterly, 2*, 1-25.
- Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language*. Baltimore: Brookes Publishing Co.
- Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.). *Word recognition in beginning literacy*. (pp. 3-40). Mahwah, NJ: Lawrence Erlbaum Associates.
- Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*, 167-188.
- Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*, 37-55.

- Francis, D. J., Shaywitz, S. E., Steubing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*, 3-17.
- Fuchs, L. S., Fuchs, D., Hosp, M., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*, 239-256.
- Good, R. H., Gruba, J., & Kaminski, R. A. (2002). Best practice in using Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in an outcomes-driven model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 699-720). Bethesda, MD: National Association of School Psychologists.
- Good, R. H., & Jefferson, G. (1998). *Contemporary perspectives on Curriculum-based measurement validity* (pp. 61-88). New York: The Guilford Press.
- Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic Indicators of Basic Early Literacy Skills (6th ed.)*. Eugene, OR: Institute for the Development of Education Achievement. Available: <http://dibels.uoregon.edu/>
- Good, R. H., Kaminski, R. A., Shinn, M., Bratten, J., Shinn, M., Laimon, L., Smith, S., & Flindt, N. (2004). *Technical adequacy and decision making utility of DIBELS* (Technical Report No. 7). Eugene, OR: University of Oregon.
- 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. S., Kame'enui, E. J., Kaminski, R. A., & Wallin, J. (2002). *Summary of decision rules for intensive, strategic, and benchmark instructional recommendations in kindergarten through third grade* (Technical Report No. 11). Eugene, OR: University of Oregon.
- Good, R. H., Simmons, D. C., & Smith, S. B. (1998). Effective academic interventions in the United States: Evaluating and enhancing the acquisition of early reading skills. *School Psychology Review, 27*, 45-56.
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7(1)*, 6-10.

- Gough, P. B., Hoover, W. A., & Peterson, C. L. (1996). Some observations on a simple view of reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 1-13). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in everyday experience of young American children*. Baltimore: Brookes Co.
- Hoover, W. A., & Tunmer, W. E. (1993). The components of reading. In G. G. Thompson, W. E. Tunmer, & T. Nicholson (Eds.), *Reading acquisition processes* (pp. 1-19). Philadelphia: Multilingual Matters.
- Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology, 95*(4), 719-729.
- 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.
- Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners. *Language, Speech, and Hearing Services in Schools, 36*, 17-32.
- Kame'enui, E. J. (Ed.). (1996). *Learning to read/reading to learn: Helping children with learning disabilities to succeed information kit*. (Technical Report). Eugene, OR: National Center to Improve the Tools of Educators.
- Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25*, 215-227.
- Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1-33). Ann Arbor, MI: U. of Michigan Press.
- Lombardino, L. J., Riccio, C. A., Hynd, G. W., & Pinheiro, S. B. (1997). Linguistic deficits in children with reading disabilities. *American Journal of Speech-Language Pathology, 6*, 71-78.
- Lorge, I., & Chall, J. S. (1963). Estimating the size of vocabularies of children and adults: An analysis of methodological issues. *The Journal of Experimental Education, 32*, 147-157.

- MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie Reading Tests (4th ed.)*. Rolling Meadows, IL: Riverside Publishing.
- 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.
- Nagy, W., & Anderson, R. (1984). The number of words printed in school English. *Reading Research Quarterly, 20*, 233-253.
- National Center for Education Statistics. (2005). *National Assessment of Educational Progress: Reading 2004 Major Results*. Washington, DC: Institute of Education Sciences, U.S. Dept. of Education. Retrieved November 10, 2006 from <http://nces.ed.gov/nationsreportcard/>
- Nation, K., Marshall, C., & Snowling, M. J. (2001). Phonological and semantic contributions to children's picture naming skills: Evidence from children with developmental reading disorders. *Language and Cognitive Processes, 16*, 241-259.
- Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39*, 85-101.
- National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: Author.
- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*, 554-566.
- Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 167-208). Oxford University Press.
- Posner, M. I., & Snyder, C. R. R. (1975) Attention and cognitive control. In R. L. Solso (Ed.), *Information Processing and Cognition: The Loyola Symposium*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology, 97*(2), 246-256.

- Riedel, B. W. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade students. *Reading Research Quarterly, 42*, 546-567.
- Robbins, C., & Ehri, L. C. (1994). Reading storybooks to kindergartners helps them learn new vocabulary words. *Journal of Educational Psychology, 86*, 54-64.
- Roberts, G., Good, R. H., & Corcoran, S. (2005). Story retell: A fluency-based indicator of reading comprehension. *School Psychology Quarterly, 20*, 304-317.
- Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *The Journal of Educational Research, 95*(5), 259-272.
- Scanlon, D. M., Vellutino, F. R., Small, S. G., Fanuele, D. P., & Sweeney, J. M. (2005). Severe reading difficulties – Can they be prevented? A comparison of prevention and intervention approaches. *Exceptionality, 13*(4), 209-227.
- Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75-119). Timonium, MD: York Press.
- Senechal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language, 24*, 123-138.
- Senechal, M., Ouellette, G., & Rodney, D. (2005). The misunderstood giant: On the predictive role of early vocabulary to future reading. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research: Volume 2* (pp. 173-182). New York: Guilford Press.
- Share, D. L., & Leikin, M. (2004). Language impairment at school entry and later reading disability: Connections at lexical versus supralexical levels of reading. *Scientific Studies of Reading, 8*(1), 87-110.
- Shaw, R. & Shaw, D. (2002). *DIBELS Oral Reading Fluency-Based Indicators of Third Grade Reading Skills for Colorado State Assessment Program (CSAP)*. (Technical Report) Eugene, OR: University of Oregon.
- Snow, C. E. (1991). The theoretical basis for the relationship between language and literacy in development. *Journal of Research in Childhood Education, 6*, 5-10.

- Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*, 72-110.
- Stahl, S. A., Jacobson, M. G., Davis, C. E., & Davis, R. L. (2006). Prior knowledge and difficult vocabulary in the comprehension of unfamiliar text. In K. A. D. Stahl & M. C. McKenna (Eds.), *Reading research at work: Foundations of effective practice* (pp. 284-302). New York: Guilford Press.
- Stanovich, K. E. (1986). Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407.
- Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*(6), 934-947.
- Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children: The role of language comprehension and working memory skills. *Reading and Writing, 4*, 245-256.
- Stothard, S. E., & Hulme, C. (1995). A comparison of phonological skills in children with reading comprehension difficulties and children with decoding difficulties. *Journal of Child Psychology and Psychiatry, 36*, 399-408.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Allyn & Bacon.
- Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading, 10*, 381-398
- Torgesen, J. K. (1998). Catch them before they fall: Identification and assessment to prevent reading failure in young children. *American Educator, 22*, 32-39.
- Torgesen, J. K. (2002). Lessons learned from intervention research in reading: A way to go before we rest. *Learning and Teaching Reading, 89-103*.
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities, 27*, 276-286.

- Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., & Garvin, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91*, 579-593.
- Vandervelden, M. C. & Siegel, L. S. (1997). Teaching Phonological processing skills in early literacy: A developmental approach. *Learning Disabilities Quarterly, 20*, 63-81.
- Wilson, J. (2005). *The relationship of Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Oral Reading Fluency to performance on Arizona Instrument to Measure Standards (AIMS)*. Tempe, AZ: Tempe School District No. 3.
- Yovanoff, P., Duesbery, L., Alonzo, J., & Tindal, G. (2005). Grade-level invariance of a theoretical causal structure predicting reading comprehension with vocabulary and oral reading fluency. *Educational Measurement Issues and Practices, 4-12*.