Use of the Dynamic Indicators of Basic Early Literacy Skills in Predicting Future Reading Achievement for English-Learner and non-English Learner Students

By

Rachel K. McCarthy

A Thesis Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Education Specialist
School Psychology

At

The University of Wisconsin – Eau Claire

June, 2016
Graduate Studies

The members of the Committee approve the thesis of

Rachel K. McCarthy presented on June 21, 2016

Dr. Mary Beth Tusing, Chair

Dr. Michael Axelrod

Dr. Vicki Samelson

APPROVED: ____________________________
Dean of Graduate Studies
Use of the Dynamic Indicators of Basic Early Literacy Skills in Predicting Future Reading Achievement for English-Learner and non-English Learner Students

By

Rachel K. McCarthy

The University of Wisconsin-Eau Claire, 2016
Under the supervision of Dr. Mary Beth Tusing

The current study examined predictive validity of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) for future performance on the Wisconsin Knowledge and Concepts Examination (WKCE). The intent of the study was 1) to explore the overall validity for the use of DIBELS assessments with English-learner (EL) students and 2) to compare the predictive validity between EL and non-EL students. Spring of first grade performance on the DIBELS assessments of Nonsense Word Fluency (NWF), Phoneme Segmentation Fluency (PSF), and Oral Reading Fluency (ORF) was compared to Fall of third grade performance on the WKCE for 450 non-EL students and 58 EL students. The EL students were primarily of Hmong-speaking background. Comparisons were then made between the predictive validity of DIBELS assessments for EL and non-EL students to determine if there was a difference in how well these assessments correlated with future WKCE performance. NWF and ORF appear to have the strongest predictive validity for future WKCE performance for EL students, with ORF being the more significantly correlated assessment between the two. ORF and NWF were also the two most strongly correlated assessments for the non-EL
students. When reflecting on the predictive validity differences between EL and non-EL students, the correlation was much stronger for the non-EL students, suggesting that the DIBELS assessments are better predictors for WKCE performance for non-EL students overall. Findings are discussed in terms of the relationship between the DIBELS assessments' predictive validity for the WKCE for EL and non-EL students. Limitations and implications for future research are also discussed.
# TABLE OF CONTENTS

**LIST OF TABLES** ........................................................................................................ vi

**Chapter**

| I. INTRODUCTION | ........................................................................................................ 1 |
| Statement of the Problem | ........................................................................................................ 4 |
| Purpose of the Present Study | ........................................................................................................ 4 |
| Definition of Terms | ........................................................................................................ 5 |

| II. REVIEW OF THE LITERATURE | ........................................................................................................ 8 |
| Reading Development in Students | ........................................................................................................ 8 |
| Introduction to Curriculum-Based Measurements | ........................................................................................................ 10 |
| Evaluating CBMs for Use in RtI Practice | ........................................................................................................ 13 |
| DIBELS as a Specific CBM | ........................................................................................................ 14 |
| History and Present Use of DIBELS Assessments with EL and non-EL Students | ........................................................................................................ 16 |
| Oral Reading Fluency | ........................................................................................................ 17 |
| Predictive Validity of ORF on State-Wide Assessments | ........................................................................................................ 17 |
| Predictive Validity of ORF on Classroom Performance | ........................................................................................................ 21 |
| DIBELS Early Literacy Assessments | ........................................................................................................ 22 |
| Summary | ........................................................................................................ 25 |

| III. METHODOLOGY | ........................................................................................................ 27 |
| Setting | ........................................................................................................ 27 |
| Participants | ........................................................................................................ 27 |
| Procedure | ........................................................................................................ 28 |
| Measures | ........................................................................................................ 30 |
| Data Analysis | ........................................................................................................ 36 |

| IV. RESULTS | ........................................................................................................ 38 |
| Data Screening | ........................................................................................................ 38 |
| Question One | ........................................................................................................ 41 |
| Question Two | ........................................................................................................ 42 |

| V. DISCUSSION | ........................................................................................................ 44 |
| Predictive Ability of DIBELS Assessments | ........................................................................................................ 45 |
| Implications for School Psychology Practice | ........................................................................................................ 48 |
| Limitations | ........................................................................................................ 49 |
| Future Research | ........................................................................................................ 52 |
| Conclusion | ........................................................................................................ 53 |

**REFERENCES** ........................................................................................................ 55
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Student Demographics, Home Language, and Education Status</td>
<td>28</td>
</tr>
<tr>
<td>2. Descriptive Statistics for DIBELS Scores during Spring of First Grade</td>
<td>40</td>
</tr>
<tr>
<td>3. Correlations Between Spring DIBELS Assessments and WKCE</td>
<td>40</td>
</tr>
<tr>
<td>4. Regression Analysis for Spring DIBELS Assessments and WKCE for non-EL and EL Groups</td>
<td>43</td>
</tr>
</tbody>
</table>
CHAPTER I

Introduction

Over the past several decades, the demographics of the United States have changed significantly, resulting in a more diverse population year in and year out. Due to this shift within our communities, the overall student composition in public schools has been impacted. This demographic swing is being reflected in schools in a variety of ways, one way specifically being the array of native home languages that individual students are entering school with. When the native home language of a student is something other than English, the student may demonstrate a limited proficiency in the English language. Consequently, they may require support in the school setting as they increase their communication and academic fluency in the English language. Students who fit into this category are identified as English Learners, or EL students. Sandberg & Reschly (2011) found that between the years 2009-2013, the percentage of public school students in the United States whose language spoken at home was something other than English was about 21.8 percent of the student population, an estimated 11.8 million students. Of these 11.8 million students, 9.4 percent come from a home in which English is spoken “less than very well”. This means that at least 1.1 million students are in a setting in which their language spoken at home and language spoken at school are likely different. While the Spanish language represents the largest contingent of students who do not speak English as a first language, a staggering 79 percent, over 460 languages in all are spoken in school districts across the nation (Sandberg & Reschly, 2011).

This shift presents an interesting problem for public schools. If the English language is not fully mastered, basic skills that teachers assume all students have may be
absent from EL students’ toolbox, such as the early literacy skills of phonemic awareness, a strong vocabulary, and alphabetic principle (Haynes, 2009). Often, differences in language and cultural experience leave EL students at-risk for academic struggles, such as poor reading development, grade retention, and disproportionate rates of special education referrals (Sandberg & Reschly, 2011; Gyovai, Cartlidge, Kourea, Yurick, & Gibson, 2009). According to recent data, 31 percent of EL students do not graduate from high school compared to 10 percent of students who come from a monolingual, English-speaking (non-EL) home. Furthermore, 28 percent of EL students continue on to a postsecondary institution, while 37 percent of non-EL students complete that same venture (Sandberg & Reschly, 2011).

While many schools continue to develop systems to address the unique learning needs of EL students, the stakes continue to rise on a national level for all students to achieve proficiency in core academic areas. Educational reform efforts, such as the No Child Left Behind Act (NCLBA) (No Child Left Behind [NCLB], 2002), focus on measurable academic outcomes for students. In 2001, NCLBA stated that schools must provide reading instruction based on both scientific research and informed by data from valid and reliable assessments used to continuously monitor students’ reading progress (Paleologos & Brabham, 2011). Essentially, schools are evaluated by how well their students perform on standardized assessments, such as the Wisconsin Knowledge and Concepts Examination (WKCE) in the state of Wisconsin. Rather consistently, EL students perform poorly on these types of assessments. According to the WISEdash Public Portal, which is a data source that collects information to provide multi-year educational data about Wisconsin school districts, only 6.4 percent of fourth grade and
3.4 percent of eighth grade EL students scored in the proficient range on the WKCE reading assessment. In comparison, 38.7 percent of their fourth grade and 35.5 percent of their eighth grade non-EL counterparts achieved that same standard (WISEdash Public Portal, 2013). Findings such as these exemplify the need for early intervention and intensive supports so that EL students are able to achieve important educational outcomes.

Response to Intervention (RtI) practices hold promise for improving educational outcomes of EL students, primarily because RtI focuses on preventing and responding to early reading difficulties through the use of research-based practices in an early intervention model (Gunn et al., 2005; Riedel, 2007). The use of universal screening to promote early identification of students in need of additional academic support is an integral part of RtI. After early identification, struggling learners are provided with interventions at increasing levels of intensity to hopefully accelerate their rate of learning. Progress is closely monitored to assess both the learning rate and level of performance of students receiving additional support (RTI Action Network, n.d.). Essential to a strong RtI system are assessment tools and strategies that 1) accurately identify students in need of additional support and 2) accurately assess whether additional supports lead to progress on desired academic outcomes.

As schools develop practices for RtI, it is essential that assessment tools used for screening and progress monitoring are reliable and valid for RtI decision-making. For example, progress monitoring tools need to adequately reflect changes in important academic skills so that educators can determine whether students are benefitting from additional intervention. Similarly, screening tools need to be predictive of future
academic outcomes, such as reading achievement. Screening tools also need to accurately identify those students who are truly at risk for future academic difficulties and in need of additional support (Hughes & Dexter, 2011).

Statement of the Problem

English learners have unique educational needs. When considered in the broader context of RtI, educators need to ensure that the assessment practices used with all students are also valid for EL students. Although some research on the validity of RtI assessment practices when used with EL students has been completed (Haager, 2007), more evidence on the validity of screening tools for use with EL students is needed. As it stands, there is significantly more research concerning the use of screening tools with monolingual English-speaking (non-EL) students than with EL students (Keller-Margulis, Clemens, Im, Kwok, & Booth, 2012). Research that has been completed with non-EL students suggests that screening assessments consistent with RtI are associated with improved outcomes for students in their early literacy skills (e.g., Burns, Griffiths, Parson, Tilly, & VanDerHayden, 2007; Fletcher, Lyon, Fuchs, & Barnes, 2007; Haager, Klingner, & Vaughn, 2007). Because success has been demonstrated within the non-EL student population, it only intensifies the need to establish a valid screening tool for EL students as well.

Purpose of the Present Study

The Dynamic Indicators of Basic Early Literacy Skills, or DIBELS (Good & Kaminski, 2002) is a screening tool commonly used in RtI systems to identify students at risk for early reading difficulties, such as struggles with letter identification and letter-sound correspondence, decoding, and reading fluency. The objectives of the present
study are to 1) assess the predictive validity of DIBELS for English Learners by examining how well Spring of first grade screening assessments predict third grade performance on a broader assessment of reading achievement, and 2) examine whether differences exist in the degree to which DIBELS assessments predict third grade reading achievement for monolingual English-speaking (non-EL) students and English-learner (EL) students.

The following research questions are examined:

1) Which DIBELS indicators administered to EL students during the Spring of first grade are the most predictive of third grade students’ performance on the WKCE?

2) Is there a difference in predictive validity of DIBELS assessments when used with students who do not speak English as a first language as compared to students who do?

Definition of Terms

To ensure complete understanding, the following terms have been defined.

**Curriculum-Based Measurement (CBM):** CBM is a standardized procedure for assessing student progress in the core academic areas of reading, mathematics, written expression, and spelling as well as monitoring readiness skills (Patton, Reschly, & Appleton, 2014). CBM is widely used across the United States to monitor student progress, screen for students who are academically at-risk, and set individualized goals for students in general, remedial, and special education (Reschly, Busch, Betts, Deno, & Long, 2009).
**English Learner:** An English Learner is defined as a student who comes from a home where a language other than English is spoken as the primary language (Yesil-Dagli, 2011). The *No Child Left Behind Act* furthered defined EL students as students who may not have been born in the United States or speak a native language other than English. It also includes students who are Native American, Alaska Native, or native resident of outlying areas and comes from an environment where language other than English has had a significant impact in the individual’s level of English language proficiency (Muyskens, Betts, Lau, & Marston, 2009).

**Monolingual Student:** For the purpose of this thesis, monolingual student refers to a student who speaks English as their primary and only language, also referred to in this paper as “non-EL students”. Monolingual students do not experience a dichotomy in their home language and the primary language spoken at school. The English language has always been the prominent language that they have been exposed to in their home environment and in their education.

**Predictive Validity:** The extent to which performance on an initial assessment, such as a CBM or screener, is related to performance on a later assessment that the initial tool was designed to predict. Predictive validity is meant to answer the question, “Does this assessment measure what it is intended to measure and can the results be used to predict future performance of the participants?” Essentially, predictive validity addresses how well a specific tool predicts future behavior (Vogt, King, & King, 2004).

**Universal Screening Assessment:** Universal screening is the systematic assessment of all children within a given grade, school, or district, on academic and/or social-emotional factors that the school personnel and community
believe to be important (Ikeda, Neessen, & Witt, 2008). It is the means for identifying students who are struggling within the school setting even after being provided a scientific, evidence-based general education (Jenkins, Hudson, & Johnson, 2007). Universal screening is typically conducted three times per school year: Fall, Winter, and Spring.
CHAPTER II

Review of the Literature

This chapter reviews research relevant to the development of reading for students and the use of curriculum-based measurements with non-EL and EL students. How students acquire reading skills, the history of curriculum-based measurements, a description of their use with varied populations of students, and DIBELS as a specific CBM will be analyzed. Current research on DIBELS will be reviewed and research on the use of DIBELS assessments with EL students will be critiqued.

Reading Development in Students

Research has shown that knowledge of alphabetic principle and phonological awareness are strong predictors of successful literacy development in students. It is particularly important to recognize the role that phonological awareness plays as EL students learn to read, especially in the English language. Several models of reading achievement propose that students progress through a series of stages when they learn to read (Gough, Juel, & Griffith, 1992). Chiappe & Siegel (2006) state that most students initially begin reading through a logographic approach, in that they begin to form connections between distinctive features of a word and its meaning. For example, a student may recognize the word “mom” by remembering that it has the same symbol (letter) at the beginning and end of the word. The second stage is called “alphabetic” or “phonological” and is more analytical than the logographic stage. This stage is characterized by phonological decoding, in which students use their knowledge of letter-sound correspondence to read new words. The final stage is the orthographic stage. In this stage, students have a more comprehensive knowledge of morphemes and their
corresponding sounds. Students' knowledge of reading whole words enables them to use patterns of known words to decode new words. For example, a student may use their knowledge of the pronunciation of the word “could” to pronounce the word “should”.

Subsequently, students must enter the orthographic stage to become proficient readers. There is a general agreement that reading acquisition is a cooperative process that depends on students’ oral language skills in the progression through these stages (Chiappe & Siegel, 2006). Specifically, researchers have agreed upon the importance of phonological awareness, syntactic processing, and verbal working memory in the acquisition of reading skills, with phonological awareness being one of the most crucial skills to possess (McBride-Chang & Kail, 2002).

There is a growing body of research examining the role of phonological awareness in reading acquisition for EL students. While research suggests that phonological awareness may transfer from a student’s first to second language, it is unclear if phonological awareness follows the same developmental path and contributes to reading acquisition in the same way for students learning to read in a second language (Chiappe & Siegel, 2006). It is suggested that phonological awareness may develop later for EL students when instruction is provided in a second language. For example, Chiappe, Siegel, and Gottardo (2002) found that EL students performed more poorly on measures of phonological awareness in English than non-EL students when they were in kindergarten. However, these differences were resolved by the time these students completed the first grade (Chiappe, Siegel, & Wade-Woolley, 2002). Nevertheless, the development of phonological awareness in EL students needs to be better understood.
Because English learners may experience a disconnect in their phonological awareness skills due to acquiring two languages at once, EL students may demonstrate difficulties in reading acquisition and place them at-risk for future reading difficulties. Although poorer readers have been found to show a wide range of deficits on oral language tasks, it is less clear what role language deficits beyond the domain of phonology play in reading struggles (Chiappe & Siegel, 2006). The bottom line is that acquiring reading skills is a complex process and the complexity is only exacerbated for EL students as they navigate the reading world balancing between two languages. Solidifying reading skills at a young age is paramount for future academic success, which intensifies the need to ensure that students can be screened and identified as being at-risk for reading difficulties at an early age. Curriculum-based measurements (CBM) have shown promise in providing the screening abilities necessary to identify these at-risk students so that early intervention can be provided.

**Introduction to Curriculum-Based Measurements**

The development of CBMs for reading began in the late 1970s at the University of Minnesota Institute for Research of Learning Disabilities (Grima-Farrell, 2014). Researchers and educators identified a need for a tool that could provide quick, efficient, and accurate information regarding academic progress in the classroom. CBMs were subsequently designed to be given frequently to track student progress toward annual goals, monitor the effectiveness of interventions, and make instructional changes as needed throughout the year. As research began on CBMs, evidence emerged that suggested that the generic measurement principles of CBM can provide technically sound
and instructionally relevant data using materials that have been drawn from resources related to a school’s curriculum (Fuchs & Deno, 1994).

Christ and Vining (2006) suggested that CBM procedures and instrumentation can be devised as either a subskill mastery measure (SMM) or a general outcomes measure (GOM) (Fuchs & Deno, 1991). SMMs are used to assess a narrow range of skills that typically develop within a short period of time and in response to a single instructional unit. As such, SMMs typically assess a single skill, such as identifying individual letters of the alphabet. In contrast, GOMs are used to analyze a broad range of skills that evolve during the course of an academic school year and across several instructional units. GOMs typically assess multiple skills.

While CBMs are one tool used by classroom teachers to assess student performance and outcomes, there are other ways to gain useful information on student performance in the classroom as well. In contrast to CBMs, traditional standardized, state-wide assessments are usually administered once per school year. They provide a summative assessment of a student’s achievement of grade level learning expectations. They are lengthier assessments, sometimes taken over the course of two-three hours. While meaningful in assessing a student’s learning, statewide assessments are not intended for screening or progress monitoring purposes. They provide information from one point during the student’s school year and results are not typically immediately available (Wisconsin Department of Public Instruction, n.d.).

In contrast to the state-wide assessments, CBM allows teachers to quickly obtain information on the progress of their students several times within a school year (Fuchs & Fuchs, 2011). CBMs have become popular in the classroom in that they allow teachers to
compare each individual student’s data to data on other students in their classroom or to that student themselves. With CBMs, teachers are able to reflect on the performance of their own students and set their own goals and standards for the students to achieve based on their present level of performance soon after completing the assessment. These goals are typically long-term and are designed to be achieved within several months or sometimes within the entire school year (Miller, Bell, & McCallum, 2015).

Research has shown that the data gathered from CBMs can be used in several educational decisions (Deno, 2003). In an RtI context, CBMs have been used to help identify students in need of interventions, decide which level of intervention is most appropriate, and determine if an intervention is successful (Mellard & Johnson, 2008). They also can be used to determine whether a student is making adequate progress in the general education classroom, whether a student requires a more intensive level of intervention, and whether a student has responded successfully to an intervention (Mellard & Johnson, 2008).

Monitoring the progress achieved in the classroom and during intervention is done using CBM assessments called probes. Probes are standardized in the way they are administered, meaning they are administered the same way every time, using the same types of materials, the same directions and protocol, and the same set of expectations (Deno, 2003). Every probe is unique, but the probes assess the same skills at the same level of difficulty. CBMs are always timed, allowing for an assessment of both accuracy and automaticity with basic academic skills. The students’ scores can be graphed to determine whether or not the student is making adequate progress and whether changes/adaptations to instruction or intervention are needed (Fuchs & Fuchs, 2011).
When used properly within an RtI framework, reading CBMs have been found to be highly sensitive to change over the course of the school year and have been deemed valid in being good indicators of reading abilities (Hintze & Christ, 2004).

**Evaluating CBMs for Use in RtI Practice**

Due to increase in popularity of CBMs, in 2004, Fuchs reviewed the status of research on the uses of CBM for RtI purposes. Fuchs suggested that stages of CBM research were necessary to address the type of validity questions essential to RtI decision-making. The first stage consists of what Fuchs (2004) calls the “technical features of the static score”. Research at this stage examines student performance on CBMs at one point in time, which is important when validating CBMs and their use as screening tools. Stage two research examines the “technical features of slope”, which addresses important validity questions surrounding the use of CBMs for progress monitoring. Stage three research reflects “instructional utility”, which evaluates the degree to which use of CBMs lead to important instructional decisions that improve student achievement. At the time of the Fuchs (2004) review, a growing body of research had been established to ascertain the utility of reading CBMs at all three stages of use. Fuchs (2004) concluded that significant CBM research had been completed and that CBMs should be considered a validated, research-based tool that could be used in an RtI framework.

When CBMs are used for screening purposes, validity of individual scores is important. First, individual scores need to validly assess the broader construct they are meant to represent. If used to predict future academic performance, CBMs need to show a strong association with future, broader assessments of academic achievement. Second, if used to identify students at risk for future learning difficulties, CBMs need to
accurately predict risk status. Fuchs (2004) demonstrated that CBMs show an ability to accurately model learning over time and subsequently identify students’ achievement levels. In the area of reading, CBMs like oral reading fluency have consistently demonstrated strong criterion-related validity with high-stakes reading tests (McGlinchey & Hixson, 2004; Clarke & Shinn, 2004; Ardoín, Witt, Suldo, Connell, Koenig, Resetar, & Slider, 2004).

**DIBELS as a Specific CBM**

In 1997, Congress and the U.S. Department of Education formed the National Reading Panel (NRP). The intent of this panel was to review research on how students learn to read and to determine which methods of teaching reading are most effective. The panel included members from varied areas of expertise, including school administrators, teachers, and educational research scientists. The NRP’s final analysis stated that the best approaches to reading instruction focused on instruction in phonemic awareness, alphabetic principle, reading fluency, vocabulary, and reading comprehension (NICHD, 2000). Based on the NRP’s findings, the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) were developed to assess key early literacy skills important for reading acquisition (Good & Kaminski, 2002).

DIBELS extends CBM practices to younger grade levels and helps identify reading concerns earlier in a student’s education. DIBELS distinctly assesses three out of five critical early literacy skills: phonological awareness, alphabetic principal, and reading fluency (Good & Kaminski, 2002). Explicitly, these three areas are assessed with four different tasks: Initial Sound Fluency (ISF), Phoneme Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), and Oral Reading Fluency (ORF). Initial Sound
Fluency, Phoneme Segmentation Fluency, and Nonsense Word Fluency are administered in kindergarten and first grade in order to assess phonological awareness and the development of alphabetic principle. Oral Reading Fluency is added to the mix in the middle of first grade to assess for the application of those early reading skills to connected text (Riedel, 2007). Consistent with RtI practices, DIBELS was designed as a way to efficiently identify students at risk for early reading difficulties, monitor at-risk students receiving specialized instruction, and assess the overall effectiveness of instruction (VanDerHeyden, Witt, & Gilbertson, 2007).

Each DIBELS task serves a specific and unique purpose in assessing early reading skills. Initial Sound Fluency is a phonological awareness task that assesses a student's ability to recognize and produce the initial sound in an orally presented word. Phoneme Segmentation Fluency is a phonological awareness task that assesses a student's ability to segment three- and four-phoneme words into their individual phonemes fluently. Nonsense Word Fluency is an alphabetic principle task in which students are presented a page of vowel-consonant and consonant-vowel-consonant nonsense words and are expected to either verbally sound out each sound in the word or to blend the sounds together to read the word as a whole. Finally, Oral Reading Fluency measures a student’s ability to read connected text accurately and fluently (Good & Kaminski, 2002).

According to its authors, DIBELS is a solid match with federal standards, such as those outlined by NCLBA, is quick and easy to use, and matches the design principles of CBMs (Good & Kaminski, 2002).

DIBELS are unique from other CBMs in that, by design, they assess early literacy skills (Good & Kaminski, 2002). Similar to other CBMs which attempt to identify a
GOM or robust indicator of future academic performance, DIBELS does not assess all possible early literacy skills. Instead, DIBELS was designed to be an indicator of a student's progress toward the long-term early literacy outcomes. For example, instead of assessing all aspects of phonological awareness, Phoneme Segmentation Fluency (PSF) and Initial Sound Fluency (ISF) were designed to be indicators of broader phonological awareness skills. Other phonological awareness skills like rhyming and blending are not part of the measurement, because PSF and ISF are the most salient indicators of phonological awareness skills for children in kindergarten and first grade (Good & Kaminski, 2002). The authors emphasize that the notion of DIBELS as dynamic indicators is crucial for RtI purposes. The “dynamic” aspect is important because it suggests that the tool is sensitive to small changes. The “indicator” aspect is important because it suggests that the assessment could be used as a predictor of future reading performance (Kaminski, Cummings, Powell-Smith, & Good, 2008).

**History and Present Use of DIBELS Assessments with EL and non-EL Students**

Several studies have examined the predictive and concurrent validity of DIBELS assessments for use with EL and non-EL students. While research has been conducted on all DIBELS assessments, DIBELS ORF has been studied with the greatest frequency. For the purpose of this thesis, the literature has been reviewed in two specific categories: predictive validity of DIBELS to state-standardized tests and the predictive validity of DIBELS to classroom performance. A further delineation of the predictability of the DIBELS assessments with the specific subgroups of EL students and non-EL students is also examined.
**Oral reading fluency.** As mentioned previously, ORF is one of the more frequently researched DIBELS assessments. Goffreda & DiPerna (2010) completed a thorough and systematic review of DIBELS assessments, including ORF. This was done by carefully analyzing 26 research articles that specifically focused on DIBELS and DIBELS ORF. The articles included dissertations, peer-reviewed articles, and technical reports. Through their comprehensive review, reliability and criterion-related validity consistently were found to be strong across the multiple research papers. Goffreda & DiPerna concluded that overall, ORF had moderate to high concurrent validity with general assessments related to basic reading skills and reading comprehension. Findings in relation to this were strong across grade levels except for first grade. Goffreda & DiPerna found that there was less predictive validity in first grade due to the limited number of DIBELS assessments that first graders complete, and those few assessments did not seem to have strong predictive validity beyond their correlations with DIBELS ORF. The conclusion suggested that ORF is a strong predictor of future performance both on state-wide assessments and classroom performance.

**Predictive validity of ORF on state-wide assessments.** Presently, there is adequate research that suggests that ORF is a strong predictor of performance on future state tests. Shaw & Shaw (2002) found that ORF performance in the Spring of third grade was strongly associated with subsequent performance on the Colorado State Assessment Program. In 2005, Vander Meer et al. found similar results. While monitoring performance of third and fourth graders using DIBELS ORF, there was a .61-.65 correlation between performance on ORF and subsequent performance on the Ohio Proficiency Test (OPT) in reading, which is considered a strong correlation. Buck and
Torgesen (2003) found that Oral Reading Fluency performance was strongly correlated with subsequent performance on the Florida Comprehensive Assessment. Similarly, Good, Simmons, & Kame‘enui (2001) found a 0.67 strong correlation rate between ORF and performance on the Oregon Statewide Assessment. Furthermore, DIBELS ORF has demonstrated good diagnostic accuracy. Schilling, Carlisle, Scott, & Zeng (2007) reported that 72 percent of third graders who met the DIBELS ORF benchmark passed the fourth grade Reading Proficiency Test in Ohio. In the same study, it was reported that 90 percent of students who achieved the Spring benchmark goal for ORF also scored in the proficient or advanced levels on the Colorado State Assessment Program. Daane, Campbell, Grigg, Goodman, & Oranje (2005) examined the relationship between DIBELS ORF and performance on reading portions of the National Assessment of Educational Progress (NAEP). Strong correlations between ORF and NAEP reading scores for fourth grade students were reported. Findings further illustrate the robust relationship between ORF and assessments of reading comprehension.

While these findings are significant, it is also important to note that the majority of research evaluating the predictive validity of ORF has been completed with monolingual, English-speaking students. The populations of students participating in the studies were primarily of Caucasian, non-Hispanic origin. Because of this, it is difficult to generalize the current findings of the relationship between CBMs/DIBELS and performance on state tests to EL students (Stokes, 2010). Presently, only about a dozen studies were found that directly analyzed EL students and DIBELS assessments. With the research that has been completed, very few have investigated the relationship between EL students and performance on DIBELS ORF specifically. The EL studies that do exist
primarily involve Spanish speaking students. This thesis intends to add to the pool of literature regarding EL students and DIBELS performance, specifically adding data on Hmong-student performance.

Reflecting on EL-specific studies, Goffreda & DiPerna (2010) reviewed only one study that controlled for ethnicity when examining DIBELS performance. Riedel (2007) examined how predictive the DIBELS assessments of PSF, NWF, and ORF were of reading comprehension skills at the end of first and second grades, as measured by the TerraNova Reading test and the GRA+DE reading test. Of the 1500 students who participated in this study, 59 were classified as EL students, all of whom were of Spanish-speaking background, so comparisons between EL and non-EL students were able to be made. Riedel (2007) suggested that DIBELS ORF was able to predict students’ first- and second-grade reading comprehension status (satisfactory or not) with 80 percent and 71 percent accuracy, respectively. He also found that the predictive ability of ORF did not differ from studies that had primarily Caucasian, non-EL students.

Similarly, Wiley and Deno (2005) examined the relationship between both ORF and Maze passages, which assess reading comprehension, and the Minnesota Comprehensive Assessment (MCA) for EL students. They chose to include Maze passages in the study to formally evaluate the concern that ORF may not be associated with reading comprehension outcomes for EL students. Moderate to moderately high correlations between ORF and the MCA were reported. Further, no notable differences in the size of the associations existed for EL students when compared to students who only spoke English. For EL students specifically, oral reading fluency was a better predictor of overall reading ability than Maze measures, suggesting that ORF measures
alone could be used to predict performance on broader assessments of reading comprehension.

Muyskens, Betts, Lau, and Marston (2009) also examined the relationship between ORF and the MCA. Their study primarily focused on fifth-grade English-learners representing three different language groups: Spanish, Hmong, and Somali. Fall fifth grade ORF scores were compared to the reading scores achieved on the MCA. The researchers separated the overall MCA scores into two categories: students who scored at the “proficient” level or higher on the MCA and students who scored below the “proficient” level on the MCA. Strong correlations between Fall of fifth-grade ORF performance and Spring reading achievement scores on the MCA were reported for all three language groups. These results suggest that ORF has a strong ability to predict which students would pass or fail the MCA, regardless of which language group they fell into.

The results of the previously cited study were similar to the conclusions of Klein and Jimerson (2005), but there were some subtle differences in what they inferred. Klein and Jimerson examined the predictive validity of ORF when factors including language background, gender, socioeconomic status, and ethnicity were considered. ORF was administered to first through third grade students in the Fall and Spring, and scores were used to predict their performance on the reading portion of the Stanford Achievement Test-Ninth Edition (SAT-9). When ethnicity was controlled for, the difference in ORF’s ability to predict SAT-9 scores was found to be a function of socioeconomic status and home language spoken. Spanish-speaking students from lower socioeconomic backgrounds performed significantly lower on the ORF measure than Spanish-speaking
students from higher socioeconomic backgrounds. The authors concluded that a combination of factors (e.g., gender, home language, and socioeconomic status) are associated with differences in the relationship between ORF and future reading performance. This means that home language is an important component as to why the predictive ability of ORF with future reading performance can be different when assessing varied language backgrounds, but is not the only factor. Klein and Jimerson built on the research from Muyskens, et al. (2009) in that it provides another layer of analysis beyond only considering home language status to help educators understand their students’ performance on these measures. Overall, these studies appear to imply that the ORF measure of DIBELS is a valid predictor of future reading success.

**Predictive validity of ORF on classroom performance.** In 2005, Graves, Plasencia-Peñado, Deno, & Johnson evaluated the relationship between English language proficiency scores and reading growth on assessments of NWF and ORF for kindergarten and first grade students. Language proficiency scores from the beginning of kindergarten were not associated with end of first grade ORF scores or first-grade ORF gains, suggesting that initial English language skills for ELs were not significantly associated with later reading skills in English. Overall, the increases found among EL students on ORF from the beginning to the end of first grade were commensurate with findings from other studies examining the growth in oral reading of first-grade students who are non-ELs (Deno et al., 2001; Fuchs, Fuchs, Hosp, & Jenkins, 2001). The results of this study suggest that the reading progress achieved for EL students between kindergarten and first grade is similar to the reading progress achieved by non-EL students during the same time frame as measured by DIBELS ORF and NWF.
Ramírez and Shapiro (2006) collected Spanish and English ORF data from students in first through fifth grades to compare monolingual (non-EL) students’ growth in English reading to EL students’ growth in English and Spanish reading. Overall, ORF was found to be sensitive to reading growth for both EL and non-EL students. Monolingual students read more fluently and had higher growth rates than EL students across grades, but the gap in level of overall reading performance between the groups on ORF growth rates significantly diminished by fifth grade. A second study (Ramírez & Shapiro, 2007) found positive, significant correlations between ORF assessments in Spanish and English, suggesting that Spanish ORF can predict English ORF. The results of Ramírez and Shapiro’s studies suggest that ORF can be useful for progress monitoring in Spanish and English, may be informative for the purposes of educational decision making, and may serve as a helpful piece of information when trying to distinguish between delays due to limited language proficiency and learning disabilities.

**DIBELS early literacy assessments.** Previous research on the development of early literacy skills in EL students permits us to make some general assumptions about the nature of the relationship of CBMs and predictive validity for EL students. Existing literature supports that phonological awareness is an important predictor of later English-language word reading for EL students (Durgunoglu, Nagy, & Hancin-Bhatt, 1993; Quiroga, Lemos-Britton, Mostafapour, Abbott, & Berninger, 2002). Additionally, research has noted a significant relationship between Spanish and English basic early literacy skills (e.g., Cicero & Royer, 1995; Gottardo, 2002; Gottardo, Yan, Siegel, & Wade-Woolley, 2001), which might indicate that early literacy skills in both languages is important to performance on later English ORF. Associations have been found in
Spanish-speaking EL students between Spanish- and English-language phonological awareness tasks and word and pseudoword reading tasks (Durğunoglu et al., 1993; Quiroga et al., 2002). English letter knowledge in EL students is also recognized as a poignant foundational skill for later literacy outcomes (Chiappe, Siegel, & Gottardo, 2002).

Less research on the early literacy assessments of Nonsense Word Fluency (NWF) and Phoneme Segmentation Fluency exists. Goffreda & DiPerna (2010) found NWF to be “adequate” for screening purposes. However, it was mentioned that there were no reviewed studies that addressed interrater reliability evidence or convergent validity for NWF. Several peer-reviewed articles reported moderate to high concurrent validity coefficients across 11 different literacy measures, meaning similar results were found across the 11 different measures. Predictive validity evidence was also displayed across three studies, also ranging from moderate to high. Fien and colleagues (2008) evaluated the use of NWF in Kindergarten and first and second grade to screen EL students and non-EL students for potential reading difficulties. They used students from fourteen different school districts in which 34 percent of the sample was identified as EL, although the exact ethnicity and language of the EL sample was not specified in the study. Almost 12000 students participated in this study, split into five different cohorts. The concurrent and predictive validity of NWF for ORF and the Stanford Achievement Test (SAT) was examined. DIBELS NWF in the Winter of kindergarten was moderately to strongly correlated with DIBELS ORF scores in the first and second grades along with scores from the SAT from the spring of kindergarten through second grade. Strong intercorrelations were found between NWF and ORF, with an increase in the size of the
relationship when the measures were used in first and second grades. Strong within-grade correlations were also found between NWF and the SAT. Researchers concluded through this study that NWF can be an effective screener for EL students.

Vanderwood, Linklater, and Healy (2008) also studied the relationship between NWF and later reading performance. Their study specifically examined whether or not NWF measures in first grade were related to performance on three third-grade literacy outcome measures with EL students. A correlational analysis was completed to determine if NWF measures were related to the Stanford Achievement Test (SAT), AIMSweb reading CBM fluency probes (R-CBM) and Maze probes, and the California Achievement Test, Sixth Edition (CAT-6). First grade NWF scores for EL students, mostly from Spanish-speaking homes, were found to have moderate to strong correlations with their performance on third grade R-CBM and Maze probes, and a moderate relationship with performance on the CAT-6 and SAT. Predictive accuracy analyses revealed NWF measures to be 80 percent accurate in correctly identifying the students above the 25th percentile on the R-CBM, CAT-6, and Maze probes in third grade. When scores were below the 25th percentile on the outcome measures, NWF scores were not as accurate in predicting who would perform below the desired level. The researchers concluded that NWF measures should be used in conjunction with other methods when determining which students need intensive intervention services.

Regarding PSF, most of the research centering on these tasks to predict later reading skills utilizes monolingual English-speaking (non-EL) students or does not disaggregate the findings for EL students (Good et al., 2001; Rouse & Fantuzzo, 2006). Goffreda & DiPerna (2010) found that two peer-reviewed articles and one dissertation
reported reliability evidence for PSF. One study in particular reported test-retest reliability coefficients and two studies reported alternate-form reliability coefficients that suggested “adequate” strength for screening and individual decision-making purposes. Convergent validity was reported in one study, which yielded a “moderate” relationship, and several other articles reported concurrent validity, with coefficients across studies ranging from low to high. Predictive validity coefficients also ranged from low to moderate.

Summary

As we continue to adapt to our ever-changing school populations in attempts to foster and develop programs that will elicit the most positive outcomes for every student, research suggests that curriculum-based measurements, specifically DIBELS, can be useful tools in screening students in the classroom. DIBELS ORF appears to have the strongest predictive ability for both EL and non-EL students for both their performance in the classroom and on high-stakes standardized state tests. DIBELS NWF appears to also have a moderate relationship with predictive ability for EL and non-EL students, although the relationship isn’t as strong as it is with DIBELS ORF. Findings have been mixed in regards to the complete impact that home language/English language proficiency plays with DIBELS’ predictive validity. It appears that home language plays a role, but it is difficult to desegregate it out from other factors, such as socioeconomic status and overall language proficiency.

The current study set out to determine the predictive validity of DIBELS with the high-stakes standardized test, the WKCE. Student performance from both DIBELS and WKCE test scores were evaluated in order to establish 1) which DIBELS assessments
administered in the Spring of first grade are the most predictive of future third-grade performance on the WKCE for ELs, and 2) if there was a difference in predictive ability of DIBELS assessments when used with students who do not speak English as a first language (EL students) as compared to students who do (non-EL students). This study extends the research already established by focusing on the specific EL-subgroup of Hmong-speaking students.
CHAPTER III

Methodology

This study evaluated the predictive validity of DIBELS for subsequent performance on a comprehensive, standardized reading assessment. Specifically, the study examined EL students’ first grade performance during the Spring benchmarking period on DIBELS assessments of Phoneme Segmentation Fluency, Nonsense Word Fluency, and Oral Reading Fluency to determine how predictive of 3rd grade performance on the Wisconsin Knowledge and Concepts Examination each was. The study also examined if there was a difference in the strength of correlation and predictive validity between DIBELS assessments and the WKCE for English learner students when compared to English-speaking (non-EL) students. This study aligns closely with a stage one research study, as outlined by Fuchs (2004). The selection of participants, assessment tools, and procedures for data collection are described below, followed by a discussion of data analysis procedures.

Setting

This study was conducted using archival data from a Midwestern school district in Wisconsin. At the time of this study, almost 11,000 students were enrolled between 4K and twelfth grade. There were a total of twelve schools serving elementary-age students, but only five of these schools were included in the current study.

Participants

During the 2008-2009 school year, 9.5 percent of the total student population was Hispanic, 2.8 percent was Native American, 2.0 percent was Asian, 1.1 percent was African-American, and 84.4 percent was Caucasian. 508 students were included in the
final sample; 450 were non-EL students and 58 were EL students. EL status is
determined by performance on the W-APT, which stands for the WIDA-ACCESS
Placement Test. The ACCESS Placement Test is administered to identify English
language proficiency skills of individual students. A school committee then determines if
students need EL services, and if so, an EL service plan is provided, as outlined by the
state. Service provision can vary by school. The majority of students identified as
English Learners (EL) were Asian (96.2 percent) and spoke Hmong as their primary
home language (96.2 percent). Comparatively, in the non-EL student population, 92.4
percent of the students were of Caucasian ethnicity. Access to specific English
proficiency levels was not available for this study.

Table 1

Student Demographics. Home Language, and Education Status (N=508)

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Home Language</th>
<th>Non-EL (%)</th>
<th>EL (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>African-American</td>
<td>English</td>
<td>0.7</td>
<td>0.0</td>
</tr>
<tr>
<td>Hispanic</td>
<td>Spanish</td>
<td>1.5</td>
<td>1.9</td>
</tr>
<tr>
<td>Am. Indian</td>
<td>Hmong</td>
<td>0.7</td>
<td>0.0</td>
</tr>
<tr>
<td>Asian</td>
<td>Arabic</td>
<td>2.9</td>
<td>96.2</td>
</tr>
<tr>
<td>Caucasian</td>
<td></td>
<td>92.4</td>
<td>1.9</td>
</tr>
</tbody>
</table>

Procedure

Approval for this study was granted by the university’s Institutional Review
Board. After the approval was obtained, existing academic assessment data were
collected from the school district. Data representing the 2004-2005 and 2008-2009
school years were accessed. The district assigned individual and unique identification numbers to each student which were used to match the first grade DIBELS data to the appropriate third grade WKCE data. No other identifiers were used in order to protect student identity. Due to the fact that the archival data were owned and collected by the school district, there was no need for parent permission to complete this study. The school district received the results of this study for their own planning. No other incentives were offered or provided to the school district for the use and release of this student information.

In this study, the predictive relationship between the Spring benchmarking assessments used with DIBELS and subsequent performance on the WKCE was of primary interest. The Spring benchmarking period was chosen because it was thought that this timeframe would elicit the most accurate reading skills for the first grade students. Students at this point would have participated in two earlier benchmarking periods (Fall and Winter) and also would have a full year's exposure to the first grade curriculum. Extraneous variables such as unfamiliarity with expectations of the DIBELS assessments and lack of exposure to a general education curriculum would theoretically be less impactful during the Spring benchmarking period as compared to earlier benchmarking periods. Specifically, the data collected included student performance on DIBELS measures of Nonsense Word Fluency, Phoneme Segmentation Fluency, and Oral Reading Fluency. The school district reported that the data were collected by teachers and graduate students who were trained in DIBELS administration. The WKCE was completed in the Fall of third grade. The WKCE was also administered by school district employees who followed state guidelines for appropriate proctoring of the exam.
These data were then used to establish the predictive validity of DIBELS assessments to state mandated reading assessments. A further analysis was also completed to determine if there was a difference between the level of predictive validity for EL and non-EL students.

**Measures**

**Dynamic Indicators of Basic Early Literacy Skills (DIBELS), 6th Edition.**

DIBELS is the official name given to the work of Roland Good and Ruth Kaminski (2002) and their colleagues at the University of Oregon. DIBELS is a set of six individually administered, standardized assessments of early literacy skills which the authors assert reflect the key reading domains discussed in the National Reading Panel. DIBELS are designed to be quickly administered, as students have one minute to complete each assessment. DIBELS are intended to be administered three times per school year. Three of the six measures designed by DIBELS were included in the study and are discussed more in depth below.

**Nonsense Word Fluency (NWF).** NWF is a standardized, individually administered test of alphabetic principle. The tasks include identifying letter-sound correspondence and the ability to blend letters into words in which letters represent their most common sounds (Good, et al., 2001). NWF is typically administered between the Winter of Kindergarten through the Fall of second grade. The student is presented a sheet of paper with randomly ordered nonsense words (e.g., sig, ray, ov) and asked read the words out loud. Students can either read the individual sound of each letter or read the whole word. For example, if the stimulus word is "sim" the student could say /s/ /i/ /m/ or say the word /sim/ to score a total of three letter sounds correct. The student is
allowed 1 minute to read as many letter sounds as they can, and the final score is the
number of letter sounds produced correctly in that one minute. Because the measure is
fluency based, students typically receive a higher score if they read the whole word and
receive a lower score if they sound the word out letter by letter.

According to the instructional recommendations within the DIBELS administration and scoring manual (Good, Simmons, Kame’enui, & Wallin, 2002, as cited in Good & Kaminski, 2002), students producing 71 correct letter sounds per minute at the end of first grade would be considered at benchmark. Students scoring below 62 correct letter sounds during that same time frame would be considered at-risk and potentially in need of more intensive instructional support (Good & Kaminski, 2002).

Good and colleagues (2004) reported adequate median alternate-form reliability for NWF in first grade. NWF demonstrated small to moderate correlations with the Woodcock-Johnson Psycho-Educational Battery-Revised in January and February of first grade respectively, where the total reading cluster score was .66. According to Good, et al., 2004, correlations between NWF in January of first grade with future assessments of ORF were .82 in May of first grade and .60 in May of second grade.

**Phoneme Segmentation Fluency (PSF).** PSF is a standardized, individually administered screener of phonological awareness. PSF assesses a student's ability to fluently segment three- and four-phoneme words into their individual sounds. A phoneme is defined as any of the perceptually distinct units of sound in a specified language that distinguish one word from another (Good & Kaminski, 2002). According to Good & Kaminski (2002), PSF is a good predictor of later reading achievement. In the
DIBELS 6th edition, PSF is administered in the Winter and Spring of kindergarten as well as Fall and Winter of first grade (Good & Kaminski, 2002).

During administration, the examiner orally presents words of three to four phonemes and the student is to segment the words at the phoneme level. For example, the examiner says "sat" or "trick", the student is to reply with "/s/ /a/ /t/" or "/t/ /r/ /i/ /ck/" to receive all possible points for the word. If a student blends phonemes together instead of breaking them apart, they receive fewer points for the number of total phonemes provided (e.g., a reply of "/tr/ /i/ /ck/" would earn 3 points. After the student responds, the examiner presents the next word, and the number of correct phonemes produced in one minute determines the final score.

According to the DIBELS administration and scoring manual (Good, Simmons, Kame'enui, & Wallin, 2002, as cited in Good & Kaminski, 2002), students producing 35 correct phonemes per minute at the end of first grade would be considered at grade-level. Students scoring below that mark during that same time frame would be considered at-risk for needing more intensive instructional support (Good & Kaminski, 2002).

The two-week, alternate-form reliability for the PSF measure is .88, and the one-month, alternate-form reliability is .79. Concurrent validity of PSF with the Woodcock-Johnson Psycho-Educational Battery readiness cluster was .54 in Spring of kindergarten and .68 in Spring of first grade. The predictive validity of Spring of kindergarten PSF with other measures is as follows: Winter of first grade NWF is .62 and Spring of first grade ORF is .62 (Good, et al., 2001).

**Oral Reading Fluency (ORF).** ORF is also an individually administered standardized assessment of reading fluency. In the DIBELS system, ORF is administered
beginning in the Winter of first grade and can continue through the Spring of sixth grade. When administering ORF, students are presented with a grade level passage to read aloud. Students are then given the instructions, "Please read this out loud. If you get stuck, I will tell you the word so you can keep reading. When I say 'stop' I may ask you to tell me about what you read, so do your best reading. Start here. Begin" (Good & Kaminski, 2002). After one minute of reading, the number of words read correctly is tallied. The student is asked to read a total of three passages, from which the median number of words read correctly are used as the overall ORF score. Incorrectly read words include words that are omitted (skipped over), substituted (for example, the student reads the word “cat” instead of “car”), or instances when the student hesitates for more than three seconds on a word. When students pause for three seconds on a word, the examiner reads the word to the student.

Reliability coefficients for elementary students range from .89 to .94. Similarly, criterion-related validity with norm-referenced achievement tests range from .49 to .94. Test-retest reliability for elementary students range from .92 to .97; alternate form reliability ranges from .89 to .94, and criterion-related validity from eight separate studies reported coefficients ranging from .52 to .91 (Good & Kaminski, 2002).

**Wisconsin Knowledge and Concepts Examination (WKCE).** The WKCE is a standardized, state-wide assessment that evaluates a student’s skills in the areas of Reading, Language Arts, Social Studies, Mathematics, and Science. This assessment is aligned with state-curriculum expectations. Results of the WKCE are reported in two different ways: a scaled score and a performance level. The scaled score is a numerical representation of the number of questions answered correctly along with their value.
Each question is worth an assigned number of “points”, i.e., an assigned value. Some questions deemed as easier may be worth one point and some questions deemed as more difficult may be worth three points. Therefore, students who answer ten questions right may actually earn more than ten points, and this is reflected in their overall scaled score. The performance level is a criterion referenced score that categorizes student performance as advanced, proficient, basic, or minimal performance. An “Advanced” score means that the student demonstrates in-depth understanding of academic knowledge and skills tested for that grade level. The scaled score equivalent for an “Advanced” score is 466 or higher. “Proficient” means that the student is demonstrating competency in the academic knowledge and skills tested for that grade level. A “proficient” scaled score is a score between 430-465. A score of “Basic” means that the student demonstrates some academic knowledge and skills tested for that grade level. “Basic” scaled scores would be between 394-429. And finally, “Minimal Performance” suggests that the student is demonstrating very limited academic knowledge and skills tested for that grade level. “Minimal performance” scaled scores are 393 or lower (WKCE, 2014).

The WKCE Reading section assesses four skill domains: word meaning, understanding text, analyzing text, and evaluating and extending text. In third grade, about 25% of the test questions relate to word meaning, 30% to understanding text, 35% to analyzing the text, and 10% to evaluating and extending the text. Some specific tasks that students may be asked to do include: using context clues, knowledge of word structure, and/or reference materials to determine the meaning of words and phrases (such as knowledge of synonyms and antonyms, recognizing irregular verb tenses, and
using a dictionary), demonstrating knowledge of using literary and informative texts (such as identifying events in a story and story elements), analyzing informative and literary texts (such as making inferences and identifying implied relationships), and/or evaluating and extending information from literary and informational texts (such as making predictions and identifying point of view). In general, all of the tasks involve the ability to comprehend text presented and build upon that comprehension.

To complete the reading portion of the WKCE, students read six to eight short passages. Two of the passages are literary, two passages are informational, at least one passage is poetry, and at least one passage is considered an “everyday text”, which is classified as a chart, a graph, an advertisement, etc. Students answer multiple choice responses or respond to short answer questions via paper and pencil (Burmaster, 2005).

During test development, content validity was established by having a group of educators individually review and rate test items according to how close the item aligned with an objective or subskill consistent with curriculum standards. The WKCE was described as having “adequate” validity and reliability as a whole. According to CTB & McGraw-Hill (2003), student performance on the WKCE is comparable to performance on the TerraNova, which was used to develop the WKCE. Test developers ensured that no items were biased according to gender, ethnicity, or age (CTB/McGraw-Hill, 2003).

In order to ensure the least amount of bias possible, a group of educators individually reviewed the different test items and rated each of them based on how closely they believed they aligned with an objective or subskill that was laid out by the state of Wisconsin. This way, the committee could safeguard against questions that could potentially be biased towards or against a certain ethnicity, language background, etc.
The Madison Metropolitan School District compiled 2009-2012 3rd through 10th grade WKCE data for students who identified their primary language spoken at home as Hmong (McCready & Vaade, 2013). Students who came from a home in which they are ethnically Hmong but do not consider Hmong as their primary home language were not included. In this sample, Hmong students consistently scored lower than non-Hmong speaking peers, which included peers who did not speak Hmong or English as their home language, but it is to be noted that the level of language proficiency was not fully described in the study (McCready & Vaade, 2013). So while efforts were attempted to control for bias, this information suggests that the WKCE may still have some bias concerns regarding ethnicity and language spoken.

Data Analysis

Statistical calculations were completed using SPSS, Version 18.

A multiple linear regression was chosen over a simple linear regression or correlation analysis because multiple variables were analyzed. The performance on the WKCE was plotted as a function against the performance on the individual DIBELS assessments. Linear regression consists of finding the best-fitting straight line (regression line) through the plotted points. R and R-squared (R²) are statistical measures of how close the data are to the fitted regression line. Correlations between the Spring DIBELS assessments and the WKCE were also analyzed.

Question One. The degree to which DIBELS assessments administered during the Spring in first grade (PSF, NWF, ORF) were predictive of third grade EL students' performance on the WKCE was first analyzed. A linear regression analysis was completed to determine the predictive relationship between performance on first grade
DIBELS measures and third grade WKCE scores. English-proficiency was not controlled for.

**Question Two.** The second research question compared regression results for groups of EL and non-EL students. The analyses compared each of the indicators used (Phoneme Segmentation, Nonsense Word Fluency, and Oral Reading Fluency) during the Spring benchmarking period.
CHAPTER IV

Results

The predictive validity between performance on specific Dynamic Indicators of Basic Early Literacy Skills (DIBELS) assessments administered during the Spring of first grade and performance on the WKCE administered during the Fall of third grade is discussed in this chapter. Furthermore, a comparison of the predictive validity of these measures between English-learner (EL) students and non-English-learners (non-EL) students is also examined.

Data Screening

The data set obtained from the school district included ethnicity of the participants, gender of the students, whether or not the student was an English-learner, performance on the DIBELS assessments in the Spring of first grade, and performance on the reading portion of the Wisconsin Knowledge and Concepts Examination (WKCE) completed in the Fall of third grade. All data were included in the final analyses except data for special education students who took the Wisconsin Alternate Assessment for Students with Disabilities (WAA-SwD) instead of the WKCE. The WAA-SwD is administered to students with significant cognitive abilities in place of the WKCE.

Mean scores, standard deviations, skewness and kurtosis statistics (Table 2) were reviewed for the DIBELS and WKCE scores to examine normality of the sample data. For the non-EL population, each assessment analyzed approximated a normal distribution. For the EL population, PSF and ORF approximated a normal distribution. NWF demonstrated was positively skewed, as the sample skewness values exceeded 1.0. This suggests is that there was less variability in the NWF scores for EL students,
meaning the standard deviation was smaller than what would be expected. Conversely, all kurtosis scores were within the recommended range of +/- 2 standard deviations (Brown, 2016).

When compared to DIBELS benchmark goals for the Spring of first grade (Good & Kaminski, 2002), mean scores for the EL and non EL samples exceeded PSF benchmark goals, met (EL sample) or exceeded NWF goals, and met (EL sample) or exceeded ORF goals. Although mean scores for the sample were at or above benchmark standards, there were aspects of the students’ data that were not consistent with DIBELS technical data. In the DIBELS technical manual, the means and standard deviations for Spring of first grade measures are: PSF 50.68 (14.65), NWF 71.41 (34.58), and ORF 60.65 (37.99) (Good, et al, 2002). As seen in Table 2, The EL student sample scored lower than the overall mean outlined in the DIBELS manual for PSF, NWF, and ORF. The non-EL student sample scored lower than the overall mean outlined in the DIBELS manual for PSF and ORF.
Table 2

*Descriptive Statistics for DIBELS Scores during Spring of First Grade*

<table>
<thead>
<tr>
<th></th>
<th>Descriptive Statistics</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Skewness</td>
<td>Kurtosis</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td><strong>Non-EL Students</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSF</td>
<td>-0.259</td>
<td>0.357</td>
<td>48.16 (12.43)</td>
</tr>
<tr>
<td>NWF</td>
<td>0.685</td>
<td>-0.272</td>
<td>61.48 (32.62)</td>
</tr>
<tr>
<td>ORF</td>
<td>0.611</td>
<td>-0.305</td>
<td>60.36 (37.34)</td>
</tr>
<tr>
<td><strong>EL Students</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSF</td>
<td>-0.120</td>
<td>-0.208</td>
<td>40.48 (15.05)</td>
</tr>
<tr>
<td>NWF</td>
<td>1.491</td>
<td>1.951</td>
<td>50.34 (29.11)</td>
</tr>
<tr>
<td>ORF</td>
<td>0.679</td>
<td>-0.540</td>
<td>43.16 (28.70)</td>
</tr>
<tr>
<td>WKCE Reading</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaled Score</td>
<td>-1.09</td>
<td>4.46</td>
<td>457.55 (40.52)</td>
</tr>
</tbody>
</table>

Table 3

*Correlations Between Spring DIBELS Assessments and WKCE*

<table>
<thead>
<tr>
<th></th>
<th>EL Students</th>
<th>Non-EL Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSF</td>
<td>.09</td>
<td>.20</td>
</tr>
<tr>
<td>NWF</td>
<td>.35*</td>
<td>.41*</td>
</tr>
<tr>
<td>ORF</td>
<td>.53*</td>
<td>.64*</td>
</tr>
</tbody>
</table>

* Significant at p < .000.
Question One

Which Spring first grade DIBELS assessments are the most predictive of performance on the WKCE for EL students?

According to Cohen (1992), correlations up to .29 are considered small, correlations between .30 to .49 are considered medium, and correlations .50 and above are considered large. When examining the relationships between first grade DIBELS assessments administered in the Spring and performance on the WKCE, DIBELS ORF had the strongest correlation with WKCE performance for both EL and non-EL students, with correlations equal to $r(58) = .53$, $p < .000$ for the EL students and $r(450) = .64$, $p < .000$ for the non-EL students. PSF had the weakest correlation with the WKCE for both EL and non-EL groups. Correlations equaled to $r(58) = .09$, $p = .247$ for the EL students and $r(450) = .20$, $p < .000$ for the non-EL students. The correlations with NWF scores were lower than ORF correlations, but still significant for each group, with correlations equaling $r(58) = .35$, $p < .004$ for the EL students and $r(450) = .41$, $p < .000$ for the non-EL students.

To review the predictive relationship between DIBELS assessments and WKCE for EL students, data from the DIBELS benchmarking assessments administered during the Spring of first grade and subsequent performance on the Fall third grade WKCE assessment were compared. A multiple linear regression was conducted to evaluate how well Spring DIBELS assessments, Phoneme Segmentation Fluency, Nonsense Word Fluency, and Oral Reading Fluency, predicted third grade WKCE performance for EL students. A significant regression equation was found wherein all three DIBELS measures predicted third grade WKCE performance (adjusted $R^2 = .32$), $F (3, 53) = 8.56$, $p < .001$; partial $\eta^2 = .48$ (Table 4). What these results suggest is that DIBELS overall is
a valid predictor of future WKCE performance for EL students and that DIBELS ORF has the greatest predictive validity when looking at each individual assessment.

**Question Two**

**Is there a difference in predictive ability of DIBELS assessments for the WKCE when used with EL versus non-EL students?**

When reviewing the predictive validity of the DIBELS assessments for the WKCE for non-EL students, a multiple linear regression was also conducted to determine the relationship between first grade Spring DIBELS assessments and WKCE performance. A significant regression equation was found with all three DIBELS measures predicting third grade WKCE performance (adjusted $R^2 = .43$), $F(3, 446) = 110.54$, $p < .000$; partial $\eta^2 = .74$ (Table 4). These results are similar to the EL students’ performance in that DIBELS assessments were found to be significantly correlated with later WKCE performance. However, DIBELS assessments appear to be more strongly correlated with non-EL students compared to EL students, which answers the research question that was being analyzed. Reviewing the $R^2$ data, DIBELS assessments appear to account for 11 percent more of the variance for non-EL students compared to EL students.

When reviewing the beta weights for DIBELS measures for each group, the greatest predictor of WKCE performance was DIBELS ORF. However, the relationship between ORF and WKCE performance was stronger for non-EL students ($\beta = .667$) compared to EL students ($\beta = .478$). There was a negative relationship for NWF with non-EL students ($\beta = -.065$), although the relationship is not significant. For EL students, the NWF assessment demonstrated a positive relationship, with a beta weight of .227.
Finally, for PSF, the data suggested that for non-EL students, there was a small, positive relationship ($\beta = .126$), but for EL students, there was a small, negative relationship ($\beta = -.111$) (Table 4). NWF and PSF display varied results between the two groups, with beta weights for the EL students suggesting that NWF still adds to the prediction of reading for EL students, but not for non-EL students. PSF appears to contribute the least predictive ability for both EL and non-EL groups.

Table 4

*Regression Analysis for Spring DIBELS Assessments and WKCE for non-EL and EL Groups*

<table>
<thead>
<tr>
<th></th>
<th>R</th>
<th>R²</th>
<th>F</th>
<th>Df</th>
<th>Sig. F</th>
<th>partial $\eta^2$</th>
<th>$\beta$ weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-EL</td>
<td>.65</td>
<td>.43</td>
<td>110.54</td>
<td>3</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.126</td>
<td></td>
</tr>
<tr>
<td>NWF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.065</td>
<td></td>
</tr>
<tr>
<td>ORF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.667</td>
<td></td>
</tr>
<tr>
<td>EL</td>
<td>.57</td>
<td>.32</td>
<td>8.56</td>
<td>3</td>
<td>.000</td>
<td>.48</td>
<td></td>
</tr>
<tr>
<td>PSF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.111</td>
<td></td>
</tr>
<tr>
<td>NWF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.227</td>
<td></td>
</tr>
<tr>
<td>ORF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.478</td>
<td></td>
</tr>
</tbody>
</table>

N=58 (EL Students), N=450 (non-EL Students)
CHAPTER V

This chapter will cover a discussion of the results of the study. An overall review of the findings, how the results are similar and different from previous research, implications for school psychologists in the school setting, limitations to the current study, and topics for future research will be covered.

Discussion

The population of students in the public school system in the United States is becoming more diverse, particularly as related to the level of English language proficiency each student demonstrates. The emphasis on achieving positive educational outcomes for all students, including minorities and students who are not proficient in English, is apparent now more than ever. It is becoming ever more imperative to assure that there are valid and reliable ways to identify students who are in need of additional academic support. Ensuring that there are ways to validly screen and monitor academic progress through an RtI system allows for teachers and educators to identify potential problematic areas and intervene at an early age to certify that all students are successfully meeting academic standards. With early and intense intervention, students with reading difficulties can achieve expected academic outcomes (Feazell, 2004).

As outlined by Fuchs in 2004, there are several paths that can be taken regarding CBM research and investigating the uses of CBM for RtI purposes. With this thesis reflecting on research questions asked within the first stage, predictive validity became the primary focus. Overall, conclusions from this study indicated that DIBELS is an overall valid predictor of future performance on the WKCE, both for EL and non-EL students. DIBELS assessments in first grade were effective in predicting third grade
reading performance for both EL and non-EL students. More specifically, Spring of first grade scores on DIBELS ORF were strongly associated with third grade performance on the reading portion of the WKCE. Additionally, Spring of first grade assessments with DIBELS NWF were moderately associated with 3rd grade performance on the WKCE. Findings are consistent with results reported by Riedel (2007) and further support the use of DIBELS assessments in identifying EL students at risk for underachievement in reading. The same measures used with native English speakers appear to be effective in predicting future reading performance for EL students (i.e., Baker & Good, 1995; Gersten et al., 2007). Ultimately, these findings are not surprising considering that a strong relationship between ORF and reading achievement measures has been established in previous research (Reschly et al., 2009; Shapiro et al., 2008).

**Predictive Validity of DIBELS Assessments**

As mentioned earlier, findings from this study builds on prior research completed that demonstrated the predictive ability of DIBELS. Buck and Torgesen (2003) found that although ORF was the strongest predictor of future performance on the Florida reading assessment in third grade for minority students, its relationship with future reading assessments was not as strong relative to the same predictions for non-minority students. Conversely, the authors found that minority students with scores below benchmark were more likely to score below expectations on the FCAT than non-minority students who scored below benchmark on ORF. The authors suggested that minority students typically have less developed reading skills, including vocabulary, which predicts reading comprehension and likely has a negative impact on oral reading fluency for EL students.
This idea is emulated with DIBELS NWF as well. With DIBELS NWF, it remains a valid predictor for EL students at the end of 1st grade while for non-EL students, NWF seems to lose its meaningfulness. This could be because NWF “ceilings out” for non-EL students and therefore is less sensitive to changes. However, EL students are still demonstrating a range in reading performance at that point in time, suggesting that NWF is still sensitive for EL readers. This could imply that the EL students’ reading skills are less developed than their non-EL counterparts, so they are in earlier stages of reading development, which results in early literacy measures like NWF still having meaning for them. Therefore, when reflecting on predictive validity of NWF specifically regarding the WKCE, it suggests that the NWF measure demonstrates a more significant and predictive relationship with EL students than non-EL students. This is a complementary finding with what was found regarding DIBELS ORF by Buck and Torgeson (2003) in that the reading assessments appear to provide different types of information when reflecting on performance of EL versus non-EL students.

In addition to the findings of this thesis that mirror previous studies, there are some unique aspects of this study that enhances the pool of data and information available. First of all, the criterion measure in this study differed from the criterion measure used in other studies, as it is specific to the Wisconsin state standards. Because of the positive correlation that was found, this suggests that DIBELS is robustly related to the Wisconsin statewide test. This extends the research to include Wisconsin state standardized tests to the list of statewide tests that DIBELS is correlated with. Also, something noteworthy and different about this thesis was the population of students studied. The primary home language for the majority of the EL students in this study was
Hmong, which is unique compared to much of the research that has been completed with EL students, who typically have spoken Spanish.

Muyskens, et al. (2009) sheds some light on the possible implications of having sample groups who speak distinct and different languages. They stated that the issue is much more complex than simply categorizing students into two subgroups: English-speakers and English-learners (or non-English speakers). This matters because each non-English language can be very different in itself. For example, both Spanish and English are considered to be “phonetic-based” languages, meaning that there is usually a direct relationship between the spelling of the word and the sound of the word. Therefore, Spanish and English may share several attributes, such as similar cognates. This would aid students who are not fluent in English to be able to build off of their own knowledge of their first language to develop a stronger vocabulary base in their second. But there are several languages that are spoken in the United States that are not comparable to English. For example, the Hmong language is distinctly a tonal, monosyllabic language. This means that pitch variations are used to distinguish between different words, even though the syllable itself is the same. Emphasizing different syllables will ultimately create different words, even though on paper, it may all look the same. The same is true for several other languages, such as Mandarin Chinese, Tagalog, and Vietnamese. Even though they may not be tonal languages to the extent that the Hmong language is, they still share many attributes of tonal language that make it incompatible with the English language (Muyskens et al, 2009). This thesis ultimately adds some research to the growing body of information that specifically reflects on the performance of Hmong-speaking students, which expands the utility of this study to a broader population.
Implications for School Psychology Practice

Teachers throughout the United States continue to face the challenge of providing effective reading instruction for an increasing population of EL students. Assessment is an integral component of effective reading programs, both to guide instruction and to screen for students at risk for developing reading problems. The current research demonstrates the feasibility of using first grade measures of basic early literacy skills, administered in English, to serve both those purposes with EL students. First grade DIBELS measures of Nonsense Word Fluency and Oral Reading Fluency can be effective in predicting third grade reading achievement on the WKCE for EL students. Teachers can have confidence that early literacy measures administered in English to EL students with possibly only minimal English proficiency have predictive validity.

There are several aspects of this study that could warrant deeper consideration when applying the findings to a school population, however. First, there are few studies that reflect specifically on the earlier DIBELS assessments with EL students relative to ORF, so further research is still needed to establish the validity of these measures with EL students. Presently, research suggests that there is a positive relationship between predictive ability of DIBELS and EL students, but there haven’t been enough studies conducted in this area to have this assertion be unequivocally accepted. Studies similar to this thesis would provide a stronger base of data to solidify the assertion that DIBELS is a valid predictor of future reading performance for all students.

Also, another implication for educational practice would be to consider the variety of tools that could be used along with the academic screeners, such as DIBELS. While screeners appear to be validly associated with performance on the state tests, teachers
should always consider other data, such as classroom performance, when making decisions about which students need additional intervention. After careful consideration using multiple avenues of information, educators should be able to confidently identify the students who are in need of academic intervention and then be able to support and build on the needed academic skill areas. The use of DIBELS to help with the identification of students who may be at-risk for academic struggles is only one step in creating a comprehensive plan to improve academic performance.

Limitations

There are several limitations that are present in this study. First, as touched on earlier, there is a question of whether or not there are differences between the various languages spoken and cultural backgrounds in terms of measuring EL students’ reading skills on an English-based measurement. Baker & Good (1995) stated that there are not only marked differences within EL groups, but there are significant differences between the numerous culturally and linguistically diverse EL groups as well. In this particular study, the EL population that was sampled was primarily Hmong-speaking, so applying these results to other language populations may not be accurate or appropriate. The findings vary as a result of design features and therefore, educators need to pay particular attention to how well research helps them consider the use of these tools with the populations of students they serve. Furthermore, the level of English proficiency was not established in this study. Students were combined into one complete group without controlling for the degree of their English-language development. This could have impacted the results.
Another limitation is that no formal reliability checks were conducted in this study. Although the examiners participated in DIBELS assessment trainings, the quality of the training and subsequent fidelity to standardized assessment procedures was not evaluated. Future studies should incorporate inter-rater data to ensure standardization of administration and accuracy of scoring. Another possible limitation to this study is that many of the students in the sample may have been provided with supplemental reading interventions which may have influenced performance on subsequent screenings. The type and intensity of reading instruction and reading intervention provided to each student may not have been similar. Therefore, it could be deduced that differences in scores achieved on the WKCE could be related to differences in educational experiences in the classroom and consequently DIBELS predictions of WKCE results could have been impacted.

Also, the 6th edition of DIBELS assessments was used to collect data and to compare psychometric properties. DIBELS has since come out with an updated assessment tool, named DIBELS Next (Dynamic Measurement Group, 2011). While there were conclusions that were drawn based on the 6th edition of DIBELS with classroom performance and state standardized testing performance, it cannot be assumed that those results can be generalized to the DIBELS Next assessments. This limits the utility of these findings as many schools will be transitioning to the DIBELS Next assessment system.

Furthermore, another limitation for this study would be the utility of these findings in light of the fact that the WKCE is no longer the state of Wisconsin’s standardized assessment used with students. Since the 2014-2015 school year, the
WKCE has been replaced by another state-wide standardized examination. This fact may in turn raise the question of the usefulness of these findings and how they can be applied to school psychology practice. Essentially, these findings add to the pool of information and knowledge regarding the use of DIBELS and their predictive validity for future reading performance. While the findings from this study cannot be replicated exactly due to the shift in Wisconsin’s state-wide assessment, it does add to the established research body that continues to demonstrate that DIBELS, and specifically Oral Reading Fluency, can be valid predictors of reading performance on state-wide assessments. This knowledge can be used by educators to confidently screen students using assessment tools like DIBELS to make educational decisions. Additionally, it would be useful to complete similar research with future mandated tests in the state of Wisconsin.

Finally, there were several aspects of this study that could not be controlled due to the archival nature of the data. The data that was analyzed was the only data that was available; further assessments and clarifying variables could not be introduced because the student population had graduated beyond the third grade and presenting new confounds would not have been possible. Several other factors, such as gender, cultural influences, exact age, prior school experiences, level of family involvement, family socioeconomic status, and invariably many other factors may have impacted these results. Furthermore, there was no data collected on the curriculum and instruction used in each elementary school. Because no data was collected on the quality and type of programming used with the students, it is impossible to know the level of impact that it may or may not have had on the performance of the different student groups.
Future Research

Further research on this topic ultimately begins with trying to control for some of the extraneous variables that were not controlled for in this study. For example, reflecting on results from a study by Hosp, et al. (2011), ORF demonstrated higher predictability for non-EL students, indicating that the measure was better at identifying which individuals in the monolingual, English-speaking group would meet or exceed the level needed for reading proficiency compared to the EL group. Considering the implications though, Hosp, et al. suggested that when using ORF to screen third grade students, more false positive errors would occur for the EL group and more false negatives would occur for non-EL group. More research is needed to identify if other factors outside of DIBELS performance can meaningfully add to the prediction of subsequent reading performance. Research also needs to be extended to examine predictive accuracy of assessments as well as the predictive validity.

Furthermore, comprehension of language and vocabulary development is imperative in English-learners’ overall success in the classroom. Cameron (2002) suggested that those students who are in the process of learning English are at a disadvantage because their vocabulary is not as extensive and strong as those who have been speaking that language their entire life. Cameron (2002) adds that research indicates that it takes an average of 2 to 3 years to achieve a basic conversational level of proficiency in a single language, and a minimum of 4 to 8 years to reach a higher, more academic level of proficiency. It is important to remember that not all EL students have minimal English proficiency, so another goal of future research would be to develop valid and reliable literacy measures so that EL students who speak no English or have very
limited English proficiency can have equal access to the early literacy screening so as to provide a gateway to early language and literacy programs. Additionally, future research could also investigate levels of language proficiency and how differing skill levels could impact the predictive utility of DIBELS assessments.

Finally, some skills that contribute to reading success are known to transfer across languages (August, Calderón, & Carlo, 2002; Durgunoglu, 2002; Proctor, August, Carlo, & Snow, 2006). What this means is that phonological awareness assessments in the student’s first language can be as good a predictor for reading in English as administering and interpreting English phonological assessments. Assessment in other languages would provide vital information for students being taught to read in English, as well as those being taught in their native language through bilingual programs or ESL services. Many of the assessments currently available in other languages are a translation of English assessments that may therefore lack psychometric integrity (Durgunoglu, 2002). Specifically in regards to this thesis, research is needed to examine how literacy develops in Hmong in both monolingual and bilingual Hmong-speaking populations in order to develop valid and reliable Hmong instruments to complement those already available in English. Only when we can provide research-based early literacy assessment to EL students in both their languages will we truly have a complete picture of their overall literacy development (Durgunoglu, 2002). This is especially important since we know that results of these early literacy assessments may vary by language.

Conclusion

Solidifying effective and appropriate screening measures within an RtI context for students of all language backgrounds is an important step in ensuring that students are
being properly monitored and targeted for academic instruction and intervention. This in turn allows for all students to improve academic skills so that they can demonstrate competency on important academic assessments and outcome measures. Specifically reflecting on reading and RtI, DIBELS Oral Reading Fluency assessments appear to be a consistent and strong predictor of future performance on standardized testing, both for students of EL and non-EL backgrounds. The methods of screening and monitoring students, especially EL students, need to continue to be investigated in order to ensure that programming can be designed to appropriately meet all needs and students can obtain the desired results. By continuous reflection and evaluation, schools can guarantee that they are providing the appropriate programs so that all students can succeed.
References


Journal, 107(2), 136-152.


doi:10.1097/00011363-200211000-00008


