CURRICULUM BASED MEASURES AND STUDENTS OF DIVERSE ETHNICITIES

Thesis
Submitted to
The School of Education and Allied Professions of the
UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for
The Degree
Educational Specialist in School Psychology

By
Autumn Nicole La Riche

UNIVERSITY OF DAYTON
Dayton, Ohio
August 2012
CURRICULUM BASED MEASURES AND STUDENTS OF DIVERSE ETHNICITIES

Name: La Riche, Autumn Nicole

APPROVED BY:

__________________________
Sawyer Hunley, Ph.D.
Faculty Advisor

__________________________
Susan Davies, Ed.D.
Committee Member

__________________________
Bobbie Fiori, Ed.S.
Committee Member
ABSTRACT

CURRICULUM BASED MEASURES AND STUDENTS OF DIVERSE ETHNICITIES

Name: La Riche, Autumn Nicole
University of Dayton

Advisor: Dr. Sawyer Hunley

This study evaluates the performance of students from diverse backgrounds on the curriculum-based measure “DIBELS Next.” While the literature supports the use of curriculum based measures to evaluate academic skill, research has been limited in the study of students with different backgrounds and their performance on curriculum based measures. Results of the present study indicated that DIBELS Next is a tool that does not discriminate among white non-Hispanic (WNH), black non-Hispanic (BNH), or Asian Pacific Islander (API) in the targeted schools. Thus, ethnicity does not have an effect on DIBELS Next oral reading fluency scores or weekly rate of change. Future research might evaluate students from other ethnicities to ensure DIBELS Next probes are an appropriate measure of oral reading fluency for all students.
ACKNOWLEDGEMENTS

I would like to sincerely recognize the help and support I have received from my colleagues in writing this thesis over the past three years, specifically my cohort. Extensive support has come from Emily Fox, Chris Powers, Allison Seaver, and Tera West. In addition, the support I have received from others has been tremendous. Without listing all of the names, thank you to those who have helped with edits, and more importantly, given me moral support.

I would also like to express gratitude to my professors in the School Psychology program at the University of Dayton because without their diverse backgrounds and interests I would not have such a well-rounded committee. Thank you Dr. Sawyer Hunley, Dr. Susan Davies, and Bobbie Fiori for the hours you have put into assisting me in more ways than I can recognize here.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>CHAPTER I INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER II LITERATURE REVIEW</td>
<td>3</td>
</tr>
<tr>
<td>Test Bias and Ethical Considerations</td>
<td>3</td>
</tr>
<tr>
<td>Curriculum Based Measures and Accountability</td>
<td>4</td>
</tr>
<tr>
<td>System-Wide CBM</td>
<td>6</td>
</tr>
<tr>
<td>Progress Monitoring</td>
<td>7</td>
</tr>
<tr>
<td>DIBELS and DIBELS Next</td>
<td>9</td>
</tr>
<tr>
<td>DIBELS Reliability and Validity</td>
<td>10</td>
</tr>
<tr>
<td>Technical Adequacy of DIBELS Next</td>
<td>13</td>
</tr>
<tr>
<td>DIBELS and Reading Comprehension</td>
<td>13</td>
</tr>
<tr>
<td>Diverse Students, English Language Learners, and CBMs</td>
<td>14</td>
</tr>
<tr>
<td>CHAPTER III METHODS</td>
<td>17</td>
</tr>
<tr>
<td>Participants</td>
<td>17</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

Schools are responsible for monitoring students’ learning across grades to ensure they are learning basic skills for future success. One way to measure student progress is with curriculum based measurement (CBM), a brief assessment given to students to assess their growth in a particular academic area (Dynamic Measurement Group, 2010). These measures give teachers and school staff an idea of how students are performing in comparison to themselves, their peers, and national norms.

DIBELS and DIBELS Next are CBMs used to monitor students’ progress in reading (Dynamic Measurement Group, 2010). DIBELS was designed to identify children who were at-risk for developing reading difficulties based on their acquisition of the necessary early literacy skills (Good III & Kaminski, 2009). DIBELS was created in the 1970s and research on its validity and reliability was conducted beginning in the 1980s at the University of Oregon. The Dynamic Measurement Group (2010) also conducted research to determine how well these assessments predict performance on statewide and national achievement tests.

Many studies have supported the use of DIBELS for monitoring students’
who come from ethnically diverse backgrounds has been limited. There is a
need to support the validity of CBM for students from different ethnicities,
specifically for DIBELS and DIBELS Next, in order to show the validity of the
assessment for diverse students. This study investigated the validity of DIBELS
Next oral reading fluency as it relates to student ethnicity.
CHAPTER II
LITERATURE REVIEW

Test Bias and Ethical Considerations

Recent research on how students of different backgrounds perform on curriculum-based measures is limited (Gofferda, DiPerna, and Pedersen 2009; Rouse and Fantuzzo, 2006 & Wayman et al., 2007). For example, there are very few investigations of predictive bias across various demographic groups using DIBELS, a commonly used assessment instrument for predicting students’ academic performance (Roehrig, Petscher, Nettles, Hudson, & Torgesen 2007).

Court cases have addressed the issue of biased assessments (Drasgow, 1987). Measurement bias is evident when subjects of different subpopulations score differently on an assessment, despite equal ability. For instance, if two students of equal ability but different subgroups, such as ethnicity, are given an assessment to measure oral reading fluency and the student from a subgroup differs significantly than the other student, the assessment may have a measurement bias. Tests and assessments must provide equitable scores for equitable abilities to be valid.
Although court cases (Whiting & Ford, n.d.) typically pertain to intelligence testing rather than to academic achievement testing, they show the importance of technically valid tests for specific subpopulations. Court cases including Larry P. v. Wilson Riles (1979) have ruled that intelligence tests are racially biased (Whiting & Ford, n.d.).

Therefore, there is a need to determine the validity of standardized assessments, comparing students of diverse backgrounds. As assessments and tests are routinely administered, educators must further evaluate what the test results mean and use caution when using the results to make high-stakes decisions, such as whether a student is eligible for special education services.

**Curriculum Based Measures and Accountability**

Schools are increasingly being held accountable for student learning. One way of measuring learning is through performance on curriculum-based measures (CBMs). These measures have been found to be correlated with high stakes achievement test scores (Keller-Margulis, Shapiro, & Hintze, 2008). In the Keller-Margulis study, correlations were found between AIMSweb, a reading fluency CBM, a Pennsylvania achievement test, and the Terra Nova (a national test that measures reading, language arts, writing, science, social studies, and math). Students who perform well on these brief CBMs are more likely to perform well on high stakes national and statewide tests.

Another study found that reading CBMs were correlated with other standardized tests of reading achievement (Reschly, Busch, Betts, Deno, & Long, 2008). This research took on a unique perspective, as it sought to summarize
evidence from a number of studies on the relationship of reading CBM scores and other standardized measures. The study showed a higher correlation among the CBM scores and national tests, such as the Minnesota Comprehensive Assessment, than between CBM scores and statewide achievement tests. The findings also supported the use of CBMs as predictors of future performance on national high stakes tests; CBM scores were correlated with national test scores over two or more years. Students with higher scores on CBMs were more likely to receive higher scores on future national high stakes assessments. This verifies that schools can use CBM data for identifying students who are at risk for learning problems and need assistance with reading.

Crawford, Tindal, and Stieber (2001) and Hunley, Davies, and Miller (in press) found that CBMs are also predictive of statewide reading achievement assessments. The study indicated that scores on these tests correlated at a higher level after one year rather than during the year the CBM was given. The authors also found that CBMs were moderately predictive of criterion-referenced tests, or tests that measure specific knowledge rather than general reading ability.

Most educators understand that knowing how to read is imperative for student success in other subjects. Reading is involved in so many other aspects of the curriculum that students need to know not only how to read, but to comprehend the material that they read. Wayman, Wallace, Wiley, Tichá, and Espin (2007) reviewed reading CBM literature and compared study results to
determine the best way to assess a student’s reading. They examined measures of oral reading fluency, maze selection, and word identification. In a maze, a student reads a passage that has omitted words and selects a word from a word bank to continue the passage (Dynamic Measurement Group, 2010). The Dynamic Measurement Group found that the best way to measure comprehension was oral reading fluency. One might assume that this is just a measure of “speed-reading,” but they found that oral reading fluency curriculum-based measures strongly correlated with overall reading proficiency.

An eight-year longitudinal study revealed consistent results about CBM predictive validity (McGlinchey & Hixson, 2004). For eight years, students were assessed using a reading CBM to determine if they were at risk for difficulties in reading achievement. Two weeks after the administration of the reading CBM the same students were given a statewide reading achievement test. A positive relationship was found between scores. This again supports the idea that CBMs accurately predict student performance on these high stakes tests and can identify students with significant needs in the area of reading.

**System-Wide CBM**

Response to Intervention (RtI) is a three-tiered approach to instruction and meeting students’ needs (Hunley & McNamara, 2010). Tier one represents instruction that is given to all students in the general education classroom. Tier two supports students who have some difficulty and need assistance in a small group setting. Tier three targets students who need even more assistance with a more one-to-one instructional approach.
Griffiths, VanDerHeyden, Skokut, and Lilles (2009) state that CBMs can drive instructional approaches in RtI, as they provide feedback which helps teachers determine who needs more assistance with reading and whether or not students are making progress. By assessing students’ progress using CBMs, teachers better understand which students are making adequate progress and which students are having difficulty. CBMs are also good tools to monitor progress during tier two and tier three interventions to see how much growth the students are making.

**Progress Monitoring**

Progress monitoring involves the administration of a CBM and comparing the scores to previous scores on similar CBMs. Thus, teachers can monitor student progress towards meeting goals.

Progress monitoring probes are typically administered once a week when students are involved in intense interventions. However, that can be time consuming for the teacher. Jenkins, Graff, and Miglioretti (2009) found that teachers could progress monitor every three weeks and gather enough data to estimate reading growth. This was determined by comparing growth slopes based on measurements every week, two weeks, three weeks, four weeks, and nine weeks. The researchers alternated the forms each student was administered so that no students were given the same two probes. The research suggests that rather than decreasing the number of CBMs given at one time, probes should be administered at less frequent intervals, thereby reducing the burden on teachers. The study found slightly inflated growth rate estimates were
present when students were given one probe per week rather than a few probes at other intervals. Results suggest that as long as teachers use multiple scores to obtain a median score, then progress monitoring can be used intermittently rather than weekly.

The order of probes and passages can change the results of a student’s growth trajectory, according to a study conducted by Francis, Santi, Barr, Fletcher, Varisco, and Foorman (2007). This effect was removed once the scores were equated according to the passage difficulty. The authors suggest that to progress monitor students and eliminate the possibility of interference due to order, interventions and progress monitoring should continue to be administered using three probes and taking the median score to equate the scores. Some order effects were evident, although they were not large effects. Scores on progress monitoring passages tended to correlate with the other scores on passages. In addition, Griffiths et al. (2009) found that CBM probes can be used for benchmarking and then again for progress monitoring without a reread effect, even when the same probe is used more than once. Students had similar results for novel passages and familiar passages.

A more recent study, conducted by Petscher and Kim (2011) examined the use of various Oral Reading Fluency score types to indicate which variation provided the most accurate predictions of future reading comprehension. The study used the median of all three passages, the mean of all three passages, the mean of passages 2 and 3, and the score from passage 3. Results showed that the most powerful indicator for low-performing students in grade one was the
mean of all three passages, but for older grade levels, grades 2 and 3, the median score was the best predictor for later success in reading comprehension.

**DIBELS and DIBELS Next**

Dynamic Indicators of Basic Early Literacy Skills (DIBELS) is a brief, standardized assessment which can be administered to students in Kindergarten through sixth grade to assess their early literacy skills. The assessment has different grade-based subtests. It assesses phonemic awareness, knowledge of alphabetic principle, phonics skills, word attack skills, accuracy and fluency of reading connected text, and reading comprehension. These skills reflect the National Reading Panel's Five Big Ideas in Reading, which are phonemic awareness, alphabetic principle, fluency with text, vocabulary, and comprehension (National Reading Panel, 2000). DIBELS Next is an updated version of the previous DIBELS assessment (Dynamic Measurement Group, 2011). The new assessment has new subtests, updated cut points, and improved passages and probes for both benchmarking and progress monitoring. Comprehension is now integrated into the oral reading fluency score by using retell. Another comprehension component has also been added, the maze.

There are four distinct purposes of curriculum-based measures and DIBELS, including screening, monitoring progress, diagnosing student's skills and needs, and measuring student outcomes. DIBELS can help with early detection of reading difficulties. It is quick, reliable and teacher-friendly, and it is easy to see when the student is not making adequate progress (Langdon, 2004).
DIBELS and DIBELS Next can be used to screen and identify at-risk students by using DIBELS nationally normed cut points. For example, pre-reading skills are assessed to determine where the student scores in relation to national norms. These skills include fluent decoding, word reading, and connected text reading, which are linked to comprehension. Subtests are also designed to assess basic phonological skills and include the ability to name letters, segment sounds, and decode nonsense words. These subtests are assessed in Kindergarten through second grade.

The assessments can also be used to monitor progress of at-risk students who are receiving interventions, specifically in schools using a Response to Intervention model. The literature supports DIBELS as a meaningful assessment in measuring early literacy skills (Coyne & Harn, 2006).

It is important to use curriculum-based measures to measure early reading skills. Findings by Martin and Shapiro (2011) suggest that reliance on teacher judgments of students’ early literacy skills tend to be insufficient in identifying students at risk for reading difficulties. Using a curriculum-based measure, such as DIBELS Next, helps derive concrete, objective data to drive decisions, such as students who are at-risk for reading difficulties.

**DIBELS Reliability and Validity**

The oral reading fluency component of DIBELS has been found to be valid in predicting outcomes of nationwide assessments (Goffeida, Diperna, & Pedersen, 2009). Other measures of DIBELS, such as phoneme segmentation and nonsense word fluency, were not as valid in predicting achievement on
national assessments. Oral reading fluency is a good screener for basic early literacy skills, but more research is needed on other components of DIBELS to determine their validity.

According to research done by Johnson, Jenkins, Petscher, and Catts (2009), three of the four DIBELS measures were not reliable for predicting progress at the end of the first grade. While ORF accurately predicted progress, Nonsense Word Fluency (NWF), Initial Sound Fluency (ISF), and Phoneme Segmentation Fluency (PSF) were not accurate measures. Compared with DIBELS, statewide assessments offer multiple choice answers instead of a measure of reading fluency. While it is important to compare how well students perform on DIBELS and the later achievement on statewide assessment, studies may not be comparing the same core area. Often performance is measured in different ways. For example, DIBELS measures reading with a one student to one administrator ratio, whereas statewide reading assessments are measured in a whole group. Using different administration practices can affect students’ scores.

This research also discusses specific cut-off scores and focuses on measuring outcomes with sensitivity. The researchers found that, for their specific district, most students did not fall in the DIBELS risk categories that were represented by their later achievement. They found a number of false positives and false negatives in their research. The authors suggest using different cut-off scores for students who fall in different ethnic subgroups, including districts that have a large number of diverse ethnicities in their schools. School norms may be
more predictive of later success. However, this may result in false positives. False positives would be, in this scenario, identifying students as “at-risk” when they really are not at-risk for reading difficulties. This may occur due to a number of variables including environmental concerns, stress, or administrative style.

DIBELS oral reading fluency has been found to have predictive validity for an individually administered achievement test, as well as Florida’s statewide assessment of reading comprehension (Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2007). This research found that performance on DIBELS was a more significant indicator of later performance on achievement tests than race/ethnicity, level of socioeconomic status, and language status, such as if the student was an English language learner.

A literature review was conducted that examined 27 articles that used DIBELS to determine reliability and validity. The results indicated that DIBELS showed strong reliability and validity for oral reading fluency (Goffreda & DiPerna, 2010). When oral reading fluency was used for screening and decision making, inter-rater reliability, test-retest, and alternate form were found to be reliable. Inter-rater reliability indicates that the scores across examiners are consistently the same. Test-retest reliability indicates that the student will score consistently when given the same measure again using a different probe at a different time but under the same conditions. Alternate form reliability indicates that the different probes will score similar results across different students. The difference between test-retest reliability and alternate form reliability is that in test-retest everyone is given the same two probes but in alternate form the
probes are different. Test-retest is under the same conditions where alternate form uses different probes.

**Technical Adequacy of DIBELS Next**

DIBELS Next released their technical adequacy manual in January of 2011. According to their data, most DIBELS subtests were comparable at the p<0.001 level when assessing reliability. This means that the subtests were reliable, measuring the skill consistently. They found that test-retest reliability for the second grade was not found to be reliable. Other measures of reliability were significant, including inter-rater reliability and test-retest reliability for other grades.

For third grade oral reading fluency, DIBELS criterion related validity indicated a moderate (an average of .56 for the year) level of validity in that it appropriately measured students’ oral reading fluency rate using a one minute probe. DIBELS assessments that utilized the Oral Reading Fluency measure appear to be the most adequate (Dynamic Measurement Group, 2011).

**DIBELS and Reading Comprehension**

Reading comprehension is imperative for students who are going to be successful in school, and it is one of the five big ideas of early literacy (Center on Teaching and Learning, n.d.). Riedel (2007) found that oral reading fluency predicted reading comprehension better than all the other DIBELS subtests, including the comprehension subtest. Oral reading fluency was also positively correlated with vocabulary scores. Even combining DIBELS subtests did not equate to more predictive power than oral reading fluency alone.
Diverse Students, English Language Learners, and CBMs

Some studies have examined how effectively CBMs measure diverse students' performance. Pearce and Gayle (2009) studied American Indian and Caucasian students to determine whether there was a difference in the predictive value across these ethnicities. Third graders were assessed using DIBELS and the statewide test in South Dakota. Results showed that DIBELS was indicative of performance on statewide assessments with both American Indian and Caucasian students. However, the researchers found that oral reading fluency on DIBELS was more predictive when identifying students with proficient skills rather than at-risk students. The authors stated that the research has found a number of false positives within reading intervention models, which would cause more students to receive interventions. This may be an issue for schools who already have limited resources. The study was limited in that it only examined one tribe of American Indian students; the researchers state that other tribes, as well as students of other ethnicities, should also be evaluated.

One study sought to determine if achievement on Kindergarten CBMs predicted achievement in second grade across ethnicities (Betts, Reschly, Pikart, Heistad, Sheran, & Marston, 2008). The researchers examined four ethnic subgroups: European Americans, African Americans, Asian Americans, and Hispanic Americans. They found that Kindergarten achievement was predictive of achievement in second grade across all ethnic subgroups. They also found a strong and significant association between the beginning of literacy skills and reading achievement across all ethnic subgroups. The researchers noted that
there was greater variability within subgroups than between them, thus ethnicity
was not a salient factor in predicting achievement with DIBELS.

Wiley and Deno (2005) sought to determine if adding a maze to oral
reading fluency measures increases the predictive ability of estimating a score on
statewide achievement tests for English language learners. They found strong
correlations between both oral reading fluency and maze assessments with
statewide achievement tests for non-English language learners. However,
adding the maze did not increase the ability to predict performance for English
Language Learners. For English Language Learners, CBMS that measure
reading in more than one way are more appropriate for monitoring progress and
benchmarking this particular group of students. The researchers suggest that
CBMs are imperative for vulnerable populations or populations who do not
historically perform well on statewide tests, such as English Language Learners.

Wayman et al. (2007) suggest that curriculum based measures may
overestimate the performance of diverse students. Rouse and Fantuzzo (2006)
indicate that there is a need for DIBELS correlational research to continue,
especially across various ethnic subgroups. Research conducted by Gofferda,
Diperna, and Pedersen (2009) reported that, relatively few studies have
examined the validity of DIBELS with standardized assessments in different
subgroups of race and ethnicity.

This research addressed these gaps by comparing DIBELS Next oral
reading fluency scores among minority students. The research will indicate
whether DIBELS Next is an adequate measure of oral reading fluency in groups
of students by comparing the raw scores and the weekly growth rate across ethnic groups.
CHAPTER III

METHODS

The purpose of this study was to determine whether there is a difference in DIBELS Next oral reading fluency median scores from culturally diverse students. There was an expectation that students from various ethnic groups will have scores that fall in different ranges than students who self-identify as White. Jenkins, Petscher, and Catts (2009) suggested that different cut scores are needed for racially diverse students due to differences in achievement. In addition, the weekly rate of reading fluency improvement was hypothesized to be significantly different according to ethnic background (Jenkins, Petscher, and Catts, 2009).

Participants

The sample was composed of third grade students in local public schools. Data were collected from 58 students across three local public schools. Participants who did not speak English at home were eliminated from the study to reduce confounding variables.

School A was a suburban school located south of Dayton, Ohio. The district has an Excellent rating on the most recent report card from the Ohio Department of Education. It met 9 state indicators of achievement,
did not meet adequate yearly progress, and did not have value added. The average daily attendance is 2,419 with the following percentages of ethnicities represented; Black, non-Hispanic (12.9%), American Indian (.3%), Asian (1.6%), Hispanic (3.7%), Multi-racial (4.8%), and White, non-Hispanic (76.7%). 64.7% of the district is economically disadvantaged, 0% are Limited English Proficient, and 18% are students with disabilities.

School B was part of an urban/suburban school district located east of Dayton, Ohio. This district has an Excellent with Distinction rating on the most recent report card. It met all 26 state indicators of achievement, made adequate yearly progress, and had value added. Test results were comparable to other similar school districts. The average daily student enrollment totaled 7571 with 3% being Black, non-Hispanic, 0.2% being American Indian or Alaska Native, 6.5% being Asian or Pacific Islander, 2.6% being Hispanic, 3.7% being multi-racial, and 84.0% being white non-Hispanic. 13.2% of the district is economically disadvantaged, 2.4% are Limited English Proficient, and 13.3% are students with disabilities.

School C was a suburban school located south of Dayton, Ohio. On the most recent report card from the Ohio Department of Education, the district was rated Excellent. It met all 26 state indicators of achievement, made adequate yearly progress, and had value added, which means that the school gained more than a year's worth of knowledge based on standardized testing. In addition, on a performance index, the school ranked at a 99.9. The average daily student enrollment totaled 5338 with 6.9% being Black, non-Hispanic, 2.4% being Asian
or Pacific Islander, 2.0% being Hispanic, 3.8% being multi-racial, and 84.8% being white non-Hispanic. 36% of the district is economically disadvantaged, 1% are Limited English Proficient, and 14.3% are students with disabilities. To see more information about school district expectations and the indicators of success, please see the Ohio Department of Education’s interactive local report card home at http://ilrc.ode.state.oh.us/

Instrument

Demographic information was gathered, such as ethnicity, using a data collection sheet (see Appendix). Information was recorded on the sheet and was kept confidential by using a number to identify each student.

Each student’s progress was assessed using DIBELS Next Oral Reading Fluency progress monitoring probes. These are available online at the DIBELS website (https://dibels.org/next/index.php).

Research Design

A quasi-experimental design was conducted with participants. The study was quasi-experimental, as students were not assigned to demographic groups. The first dependent variable, growth in oral reading fluency, was measured by using DIBELS Next Oral Reading Fluency probes protocols. The second dependent variable was weekly rate of change in oral reading fluency. The score from week 1 was subtracted from the score of week 10 and that amount was divided by ten, to calculate the weekly rate of improvement across the number of weeks the study took place. Students’ scores on DIBELS Next oral reading fluency were used to determine if there was a difference between raw oral
reading fluency score differences based on ethnicity. The difference was calculated by subtracting the median oral reading fluency score of the first week from the median oral reading fluency score of the last week. Ethnicity was categorized using the same demographics identified by the U.S. Census Bureau: White, Black, American Indian, Eskimo or Aleut, Asian or Pacific Islander, and Other races (2010). Ethnicity representation in this sample included white non-Hispanic, black non-Hispanic, and Asian or Pacific Islander. The second dependent variable is the weekly rate of change from the first week median score on DIBELS Next oral reading fluency assessments subtracted from the tenth week median score divided by 10 weeks.

Procedure

Students were assessed by trained school psychology graduate students using DIBELS progress monitoring beginning the fourth week of the school year, finishing ten weeks later. Students were given three passages to read in one minute and the median score was recorded each week over the course of ten weeks to determine if there was a difference. Progress monitoring probes from DIBELS Next were administered in a randomly selected order with some probes repeating, as necessary. There is no evidence of inflated scores due to the re-reading of passages (Griffiths, et. al, 2009). The probes are scored based on how many words participants’ read per minute and how many errors they make during that minute. See DIBELS administration instructions on the DIBELS webpage for more information (http://www.dibels.org/next.html).
Data Collection and Analysis

Demographic information was obtained using a data collection tool indicating the student’s ethnicity. The tool was a checklist indicating which category each individual falls under (see Appendix A). Data were collected for ten weeks from September 2011 through December 2011.

The raw scores of each student in all ethnic groups were compared using an ANOVA. The growth rate was found for each ethnic category by finding the difference between scores at week one and at week ten and then that number was divided by ten, the number of weeks data was collected, to find the weekly rate of change. The rate of change scores were then analyzed for differences between ethnic groups.

Weekly rate of change scores were calculated and compared among groups. The p-value was set at .05 to determine significance.
CHAPTER IV
RESULTS

The first hypothesis was tested by conducting a one-way analysis of variance (ANOVA) to determine if ethnic groups had significantly different changes in their oral reading fluency raw scores from the first week of data collection to the last week. The difference was computed by subtracting the median first week oral reading fluency raw score from the median last week oral reading fluency raw score. The ANOVA which compared ethnic group raw score difference revealed no statistically significant differences between ethnicities, \( F(3,54) = .920, p = .437 \). Thus, ethnicity did not have an effect on DIBELS Next oral reading fluency scores.

The second hypothesis was tested by conducting a one-way analysis of variance (ANOVA) to determine if groups had significantly different weekly rate of change scores when compared with other ethnic groups. The independent variable, ethnicity, had three levels; white non-Hispanic (WNH), black non-Hispanic (BNH), or Asian Pacific Islander (API). The dependent variable was the weekly rate of change scores, which was the median scores from week one subtracted from week ten median scores divided by 10 weeks. There were no
statistically significant differences between ethnic groups when comparing their rates of change, $F(3,54) = .844, p = .476$. Ethnicity did not have an effect on weekly rate of change scores for oral reading fluency.
CHAPTER V
DISCUSSION

The results of this study indicated that oral reading fluency is not significantly influenced by the ethnicity of third grade students. There were no significant differences between ethnic groups when measuring raw scores, or weekly rate of change scores. The present study shows that DIBELS Next is a valid tool that does not discriminate among various ethnicities represented in this study.

The study was conducted using a quasi-experimental design, as students could not randomly be assigned to groups. The student’s were given three probes each for ten weeks. Two one-way analyses of variance were used to determine if there were differences between scores based on ethnicity. The ANOVA was used to test two hypotheses. The first hypothesis asked if there was a difference in raw scores based on ethnicity. This difference was measured by subtracting the first week raw score from the last week raw score. The second hypothesis asked if there was a difference in the weekly rate of change from the first week to the last week. This was measured by taking the number from the first hypothesis and dividing it by the number of weeks data was collected, which
was ten weeks. The p-value was set at .05 to determine significance for both hypotheses. The first hypothesis was rejected and the null hypothesis was accepted, indicating that ethnicity did not have an effect on DIBELS Next oral reading fluency raw scores. The second hypothesis was also rejected, while the null hypothesis was accepted, indicating that ethnicity did not have an effect on DIBELS Next oral reading fluency weekly rate of change scores.

The present study provides evidence of technical adequacy for using DIBELS Next oral reading fluency probes as a tool to measure progress of students who are ethnically diverse. The study supports the use of DIBELS Next probes to measure students’ oral reading fluency progress.

There is limited research regarding student’s performance on curriculum based measures including DIBELS Next, for different ethnicities (Gofferda, Diperna, and Pedersen 2009; Rouse and Fantuzzo, 2006 & Wayman et al., 2007). Performance on these measures is important because research supports the idea that curriculum based measures accurately predict student performance on high stakes tests and can identify students with significant needs in the area of reading (McGlinchey & Hixson, 2004). A need for DIBELS correlational research to continue, especially across various ethnic subgroups has been indicated by research (Rouse & Fantuzzo, 2006).

Limitations and Future Research

This study, considered quasi-experimental, is limited in that random sampling was not possible in this research, groups already were categorized by
school, grade level and ethnicity. Data was also collected in a limited geographical range. Due to this, generalizability to other regions may be limited. Students who identified with groups that were not represented with a significant number were eliminated from the study. Groups that had a significant representation in this study included white non-Hispanic students, black non-Hispanic students, and Asian or Pacific Islander students. In addition, students who did not speak English at home and were identified as English Language Learners were also eliminated from the study because of the possible extraneous variables of lack of familiarity with the English language.

Future research that expands diversity across additional ethnicities and geography would be beneficial to ensure DIBELS Next probes are an appropriate measure of oral reading fluency for a wider range students. For instance, Pearce and Gayle (2009) studied American Indian and Caucasian students to determine whether there was a difference in the predictive value across these ethnicities and found that oral reading fluency on DIBELS was more predictive when identifying students with proficient skills rather than at-risk students. Another opportunity for future research would be to determine predictive value across other ethnicities than American Indian and Caucasian students.

In addition, students who are identified with a disability are often assessed using curriculum based measures such as DIBELS Next, and this assessment method should be further validated for these students.

Finally, it would be beneficial to assess other DIBELS Next measures, in addition to those that measure oral reading fluency, to determine if those are also
ethnically sensitive. Schools across the country use DIBELS Next to measure students’ progress and response to intervention. By validating the accuracy of all DIBELS Next measures, schools can be assured that they are using appropriate tools for all students.
REFERENCES


Center on Teaching and Learning. Big Ideas in Beginning Reading. Retrieved from http://reading.uoregon.edu/


&sid=2b58cb17-34ee-4913-a135-a83fcee2ef57%40sessionmgr12&bdata=JnNpdGU9ZWhvc3QtbGZlZVZQ%3
d%3d#db=a9h&AN=22734303

measures accurate predictors of reading achievement? *The Elementary
&sid=2b58cb17-34ee-4913-a135-a83fcee2ef57%40sessionmgr12&bdata=JnNpdGU9ZWhvc3QtbGZlZVZQ%3
d%3d#db=a9h&AN=24665317

progress in inclusive elementary settings. *Reading and Writing Quarterly,
22*, 91-97. doi:10.1080/10573560500203574


Literature synthesis on curriculum-based measurement in reading. *The
Journal of Special Education, 41*, 85-120. doi:
10.1177/00224669070410020401
http://www.education.com/reference/article/cultural-bias-in-testing/

of success for English learners on a state standards assessment.
Remedial and Special Education, 26, 207-214. Retrieved from
&sid=2b58cb17-34ee-4913-a135-
a83fcee2ef57%40sessionmgr12&bdata=JnNpdGU9ZWhvc3QtbnVt
&d%3d#db=a9h&AN=17805872
APPENDIX A
DATA COLLECTION TOOL

For use with every student assessed using DIBELS Next.

Collection date:

ORF scores:

Ethnicity:
  White
  Black
  American Indian, Eskimo or Aleut
  Asian or Pacific Islander
  Other

Free/reduced lunch?    Yes    No

Gender:                Male      Female

Age: