MEASURING RATES OF READING GROWTH ASSOCIATED WITH GENDER USING DYNAMIC INDICATORS OF BASIC EARLY LITERACY SKILLS

Thesis

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

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

By
Nicholas J. DeGrazia
Dayton, Ohio
August, 2013
MEASURING RATES OF READING GROWTH ASSOCIATED WITH GENDER
USING DYNAMIC INDICATORS OF BASIC EARLY LITERACY SKILLS

Name: DeGrazia, Nicholas J

APPROVED BY:

_____________________________
Sawyer Hunley, Ph.D.
Advisory Committee Chairman
Associate Professor
Department of Counselor Education & Human Services

_____________________________
Susan Davies, Ed.D.
Committee Member
Assistant Professor
Department of Counselor Education & Human Services

_____________________________
Elana Bernstein, Ph.D.
Committee Member
Clinical Faculty
Department of Counselor Education & Human Services
ABSTRACT

MEASURING RATES OF READING GROWTH ASSOCIATED WITH GENDER USING DYNAMIC INDICATORS OF BASIC EARLY LITERACY SKILLS

Name: DeGrazia, Nicholas J.
University of Dayton

Advisor: Dr. Sawyer Hunley

With the increased emphasis on data collection in schools, educators need efficient, reliable systems for progress monitoring. Dynamic Indicators of Basic Early Literacy Skills (DIBELS) is one such system for basic reading skills. Previous studies demonstrated differences in growth rates between males and females with respect to oral reading fluency. The present study examined if, despite previous researches noted differences in growth rates, fewer data points would give appropriate growth estimates in males and females. Results indicated that growth estimates in weeks 1 and 5 gave a comparable growth estimate to those taken in weeks 1 and 10. This comparison held true for both male and female students. These important results can help school psychologists guide school staffs on how to efficiently gather oral reading fluency progress monitoring data.
To the best third-grade class from a suburban Midwestern town
ACKNOWLEDGMENTS

I would like to sincerely recognize the help and support I have received from many people while writing this thesis. Thank you to Steve who allowed me to come into his school and Molly who allowed me to come into her classroom and work with her students. Also, thank you to Amy and Jen for their patience in my pulling students during instruction. Support in writing has come from Autumn La Riche (professional) and Kolleen Hryb (emotional).

I would also like to express my gratitude to all members of my thesis committee at the University of Dayton. I would not have produced the work I did without the help and support of Dr. Sawyer Hunley, Dr. Susan Davies, and Dr. Elana Bernstein. All three put in more time than I would have expected.
# TABLE OF CONTENTS

ABSTRACT ...................................................................................................................... iv
DEDICATION ................................................................................................................... v
ACKNOWLEDGEMENTS .............................................................................................. vi
LIST OF TABLES ........................................................................................................... x
CHAPTER I: INTRODUCTION ....................................................................................... 1
CHAPTER II: LITERATURE REVIEW ........................................................................... 3
  Curriculum Based Measurement ............................................................................. 4
  Background ............................................................................................................... 4
  Implementation ......................................................................................................... 5
  Dependability ........................................................................................................... 6
  Oral Reading Fluency and CBM ............................................................................. 8
    Words Read Correctly ......................................................................................... 8
  ORF as an Indicator of Reading Achievement .................................................... 9
  Statewide Testing and ORF .................................................................................... 9
  Oral Reading Fluency and CBM ............................................................................. 10
  ORF Probes ............................................................................................................ 11
  Maximizing Validity and R-CBM ........................................................................ 11
  Progress Monitoring Intervals and Growth Rates .............................................. 13
Appendix A: Sample DIBELS Next Probe ................................................................. 44
Appendix B: DIBELS Next Administration Instructions ........................................ 46
Appendix C: Data Collection Tool ........................................................................ 47
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Frequencies and Percentages of Ethnicity</td>
<td>25</td>
</tr>
<tr>
<td>Table 2</td>
<td>Correlations of Growth Estimates for the Entire Sample</td>
<td>28</td>
</tr>
<tr>
<td>Table 3</td>
<td>Correlations of Growth Estimates for Males</td>
<td>28</td>
</tr>
<tr>
<td>Table 4</td>
<td>Correlations of Growth Estimates for Females</td>
<td>29</td>
</tr>
<tr>
<td>Table 5</td>
<td>Correlations of Growth Estimates Using More than Two Data Points</td>
<td>30</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

When the No Child Left Behind Act of 2001 (NCLB, 2001) was enacted, assessment and accountability in the classroom were critical issues. A key component of NCLB was classroom assessments that allow teachers to measure student progress in certain subject areas. If teachers know which students are not making adequate progress in their class they can adjust their teaching and curriculum to best suit the child (U.S. Department of Education, 2001). As such, instruments that provide accurate results in an appropriate amount of time are required (Griffiths, VanDerHeyden, Skokut, & Lilles, 2009). One type of such assessment is curriculum-based measurement (CBM).

One common use of CBM is monitoring of reading growth. Reading growth requires two data points that are typically taken 10 weeks apart for the general population at points called benchmarks. Research has historically demonstrated a gap in reading achievement between males and females, with girls consistently scoring higher. This can be seen early on with females scoring higher in reading achievement beginning in kindergarten. More recent studies have shown that this gap may not be as significant and that males and females develop at the same rates with respect to reading (Below, Skinner, Fearrington, & Sorrell, 2010).
This study will examine gender differences and the possibility of decreasing the number of data collections while maintaining similar progress monitoring results. The reading fluency curriculum based measure used was Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Next, which were originally titled Dynamic Indicators of Basic Early Literacy Skills (DIBELS) and created by Roland H. Good, III and Ruth Kaminski (DMG’s Philosophy and Mission, 2009).
CHAPTER II
LITERATURE REVIEW

This literature review will focus on curriculum-based measurement, specifically measurements for oral reading fluency. It will also address research focused on gender differences in reading.

The No Child Left Behind (NCLB) Act of 2001 focused on the growing need for assessments that could be conducted by teachers and staff in the classroom. The intention was for teachers to use ongoing assessment data to guide their instructional practices to ensure all children maintained certain standards for achievement (U.S. Department of Education, 2001). NCLB 2001 was followed by the reauthorization of the Individuals with Disabilities Education Improvement Act (IDEA) in 2004. IDEA 2004 allowed for the identification of a learning disability if a student did not improve performance after researched based interventions were implemented for a period of time (IDEA, 2004). This change in frame of reference for identification of students with learning disabilities introduced the concept of Response to Intervention (RtI) into schools (Wright, 2007). RtI employs tiers (or levels) of instruction and intervention to help identify students’ needs and assist in their instruction. As a child’s needs increase, so does the level of intervention (Hunley & McNamara, 2010).
One essential component to the RtI framework is curriculum-based measurement (CBM; Deno et al., 2009) because of its usefulness in combination with other data (Fewster & Macmillan, 2002). CBM helps identify students at risk for academic problems who would benefit from additional support (Deno et al., 2009). It also predicts potential academic problems that students might encounter in the future (Fewster & Macmillan, 2002) because CBM data can provide growth rate estimates (Ardoin & Christ, 2008). CBM is also an important factor in eligibility determination for special education, because of the use of RtI in the special education identification process (Compton, 2000).

Curriculum-Based Measurement

**Background.** Curriculum-based measurement (CBM) was developed and researched in the late 1970’s and early 1980’s at the University of Minnesota by Stanley Deno. The assessment was developed as a measurement and evaluation procedure that could be routinely used by teachers helping to inform if and when they should change a student’s instruction (Deno, 1985). Initially, studies focused on elementary school teachers and providing them with support to increase the performance of students struggling in basic reading, writing, and arithmetic skills (Deno, 2003). During the initial development of CBM, research focused on creating measures that were reliable and valid, simple and efficient, easily understood, and inexpensive (Deno, 1985). Since the development and initial research on CBM, the data derived were used for or contributed to screening individual students, making placement decisions for special education, setting Individualized Education Plan (IEP) goals, conducting pre-referral evaluations, and making reintegration and inclusion decisions (Deno, 1985; Deno, 2003).
CBM is unique because even though it is based in traditional methods of gathering data, it also employs concepts not normally associated with collecting information. This includes visual/graphic representations of data, multiple samplings of performance, designated amounts of time for testing, and qualitative descriptions of how the student performed during the assessment. Estimates of growth use slope and a snapshot of current level of performance (Deno, Fuchs, Marston, & Shin, 2001).

**Implementation.** CBM is used for monitoring students’ progress in reading, spelling, written expression, and math. When fluency performance is assessed in these areas, the person administering the probe examines the number of responses correct within a given period of time (i.e., words read correctly in 1 minute of oral reading assessment). All types of CBM have a set of standardized procedures and an assumption that the measure is designed to continually monitor student progress throughout the year to determine proper instruction (Deno, 1985; Deno 2003). Overall, the procedure is designed to maximize student learning by providing data that can be used to guide instruction (Fewster & Macmillan, 2002).

CBM has high acceptability among educators because it is shown to be a reliable and valid progress monitoring tool (Deno, 2003; Deno et al., 2009), it targets individual students while attending to groups, and it produces results that can be referenced in three different areas: criterion-referenced, individually referenced, and peer-referenced. Criterion-referenced results show how the student is progressing in relation to the general classroom curriculum, while individually referenced and peer-referenced results show how the child compares to his own previous score and his classmates scores, respectively (Deno, 1985). While targeting individual students, a teacher can build a database which
shows whether the classroom instruction is resulting in adequate growth of students’ abilities (Hosp & Hosp, 2003). This estimation of growth using slope was shown as a reliable with both general and special education students (Deno et al., 2001). CBM was also supported for general screening of students and their current level of performance (Deno et al., 2009; Keller-Margulis, Shapiro, & Hintze, 2008).

Deno (1985) describes five advantages of CBM: improved communication, increased sensitivity, improved data base, peer referencing, and cost-effectiveness. Improved communication can be seen in the visual representations that the graphs of student performance provide. Because of the simplicity of the graphs, student progress is easily understood by regular and special education teachers who must make instructional decisions for the student. CBM is sensitive to growth over short periods of time. For example, data from probes administered in time periods as short as weeks or even days help teachers monitor progress over time as compared to data from other types of assessments that take place less frequently. Data can be used to predict a student’s growth along with current level of performance. In addition, educators can evaluate a student’s progress as it directly relates to the rest of the class, based on teacher generated local norms. Finally, CBM allows for data collection as often as necessary without the high cost of many standardized tests (Deno, 1985).

**Dependability.** Research has determined what to measure with CBM, how to measure with CBM, and how to use a database derived from CBM. In each of these three areas, studies have shown that CBM is technically adequate, valid, and feasible (Deno et al., 2001).
CBM is appropriate for making both low- and high-stakes decisions. Low-stakes decisions are decisions that do not have a great impact on the student or school, such as the amount of extra time given for reading instruction. High-stakes decisions are used to make choices such as funding for schools or eligibility for special education. For example, Christ, Johnson-Gros, and Hintze (2005) found that in curriculum-based measurement in mathematics (M-CBM), one- and four-minute administrations of computation assessments can be used to compare a student to the rest of the class for both low- and high-stakes decisions (respectively). They also found that four- and 13-minute administrations were appropriate for low- and high-stakes decisions (respectively). The extra time was only necessary when more information could be obtained from the additional time. When used in conjunction with other data (ie. record reviews, norm-referenced assessments, and observations) bi-weekly CBM data is useful for making high stakes decisions (Ardoin & Christ, 2008).

Studies have also shown that CBM is a valid predictor of performance on statewide tests (Atkins and Cummings, 2011; McGlinchey & Hixson, 2004; Nese, Park, Alonzo, and Tindal, 2011; Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2007; Silberglitt, Burns, Madyun, and Lail, 2006). With the current emphasis on standardized tests (U.S. Department of Education, 2001), it is helpful for educators to have CBM data to help in curriculum planning.

Single-point CBM data is an acceptable predictor of performance on certain statewide tests along with assessments that have a national norm. This holds true for data that was taken up to two years before the test was given, both in reading and math. Growth estimates tend to show the same results as single-point data, but are more reliable
in the younger grades. This is significant because many standardized tests begin in 3rd grade, which means CBM data can be relevant as a predictor of growth as early as 1st grade (Keller-Margulis et al., 2008).

**Oral Reading Fluency and CBM**

Among the different subject areas (reading, spelling, writing, and math) in which CBM is used, reading (R-CBM) has received the majority of attention from researchers (Wayman, Wallace, Wiley, Ticha, & Espin, 2007). More specifically, oral reading fluency (ORF) in grades 2-6 was the focus of most research, thereby giving it the most credibility as a progress monitoring tool (Fuchs, Fuchs, & Hamlett, 2007). Adams (1990) determined that the most marked measure of a proficient reader is the speed and ability to turn printed text into spoken language. What Adams was describing is also known as ORF. Further studies have shown ORF is a strong indicator of overall reading competence. It is important to note that oral reading fluency is a stronger indicator of reading performance than silent reading fluency (Fuchs, Fuchs, Hosp, & Jenkins, 2001).

**Words Read Correctly.** ORF is generally measured as words read correctly (WRC) in a given time period, typically one minute. WRC is a good measure of reading growth in CBM progress monitoring (Fuchs et al., 2001). The number of WRC by students in sixth- and seventh-grade has a strong correlation with their grades in eighth-, ninth-, and tenth-grade (Fewster & Macmillan, 2002). The standard error of measurement (SEM) of WRC can be reduced to provide a more accurate representation of the child’s actual score on the given assessment. This can happen in two ways. The
first method is controlling passage variability and the second is administering multiple alternate passages (Poncy, Skinner, & Axtell, 2005).

**ORF as an Indicator of Reading Achievement.** One of the many benefits of ORF is that it is a very reliable measure of overall reading competence (Fuchs et al., 2001; Hintze, Callahan, Matthews, Williams, & Tobin, 2002). Even in relation to other reading measures (i.e. MAZE), ORF provides more accurate data (Madelaine & Wheldall, 2004). Studies have also shown ORF to be an accurate predictor of reading comprehension across ethnicities (Pearce & Gayle, 2009) and Hintz et al. (2002) showed that ORF was non-biased with regard to ethnicity or socio-economic status (SES). To determine the best prediction of year end achievement, educators should use ORF results from administration in the fall, with the students’ words per minute (WPM) indicating future performance (Schilling, Carlisle, Scott, & Zeng, 2007).

**Statewide Testing and ORF.** As more and more emphasis is placed on standardized testing and statewide assessments, it is beneficial to know that ORF can predict student outcomes on these tests. In a study conducted by McGlinchey and Hixson (2004), a positive correlation was demonstrated between fourth-grade students’ performance on an ORF assessment two weeks before the Michigan Educational Assessment Program’s (MEAP) fourth-grade reading assessment. This is even more impressive when it is revealed that the ORF assessment required only one minute for administration (McGlinchey & Hixson, 2004). Silberglitt, Burns, Madyun, and Lail (2006) showed the same outcome with respect to state testing in Minnesota. They expanded their testing to grades 3, 5, 7, and 8 and discovered the highest correlation with third-graders and a lower correlation as the students reached upper grades. Research has
also shown that in a third-grade population, there is a strong positive correlation between an ORF measure and performance on the Florida Comprehensive Assessment Test (FCAT-SSS) along with the Stanford Achievement Test (SAT-10) reading comprehension measures. These results held true between all subgroups (Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2007). Atkins and Cummings (2011) used data from three points in the school year (fall, winter, and spring) to predict achievement on the Montana Comprehensive Assessment System. Despite lower correlation with upper grades in some studies, Hunley, Davies, and Miller (2013) showed a correlation between passage of the 7th grade Ohio Reading Achievement Test and students who read at least 100 WPM. This is promising because it shows that ORF can help guide instruction in preparation for statewide high-stakes testing.

**Oral Reading Fluency and CBM.** Oral Reading Fluency curriculum-based measures (R-CBM) have the appeal of being efficient, effective, and providing clear data (Deno et al., 2009). Fuchs et al. (2007) describe the following general standards for using R-CBM. A teacher begins by selecting 30 probes (passages) of the same degree of difficulty. For example, for third grade students, the teacher would generally choose a passage appropriate for the third grade. Each week the target students read one of the probes aloud for 1 minute. While each student is reading, the teacher follows along and at the end of the 1 minute the number of WRC is scored. These scores are then graphed and rate of growth over time is established from the number of WRC in the given time period.

As with any implementation of progress monitoring in the classroom, results need to guide instruction. After five weeks of R-CBM and teacher feedback on results, there
was an increase in classroom performance of low achieving students (Graney & Shin, 2005). When R-CBM data are used to inform instruction, educators get better results. For example, when special education teachers use the results efficiently, greater success occurs in resource rooms than in general education classes (Fuchs et al., 2007).

**ORF Probes.** With regard to R-CBM, the same passage can be used for progress monitoring or RtI decision making (Griffiths et al., 2009). The probes do not necessarily need to come from the student’s instructional material. The probes can have a lesser or greater degree of difficulty than the current level of instruction; however, growth rate estimates can be affected if the material is too difficult (Waymann et al., 2007). The currently accepted practice is to use the median value of WRC of three probes each time the student is assessed (Ardoin & Christ, 2008). It is also advisable to find a balance in the probes so there is not a large variance in level of difficulty (Hintze & Christ, 2004) and establish equivalence between passages when the results will be used to make more important/high-stakes decisions (Wayman et al., 2007).

When Jenkins, Graff, and Miglioretti (2009) looked at the number of CBM scores for each monitoring session, they found that when only one score was collected, the reliability of data went down. This was the case for establishing a baseline and estimating growth. When these ideas are combined, the authors claim that educators should not decrease the number of scores collected at each monitoring session.

**Maximizing Validity of R-CBM.** Even with all of the research in support of CBM, and specifically R-CBM, errors still occur in gathering and interpreting the data. Error is the difference between the performance that is observed and the performance that is expected. This error can be a problem in progress monitoring because proper decisions
might not be made if the student’s performance is over- or under-estimated. When error is reduced in CBM progress monitoring, prediction of student performance is more reliable (Hintze & Christ, 2004). To reduce this standard error of measurement (SEM), Hintze and Christ (2004) recommend constructing probes with a conscious effort to control for difficulty of the text. This reduction of error by controlling passage difficulty can be seen at both single point plots of reading performance and growth estimates using the slope of multiple points. Christ, Zopluoglu, Long, and Monaghen (2012) recommend the use of a good or very good passage set.

Error is a legitimate concern with any classroom assessment. Because R-CBM is used to make decisions such as special education eligibility or intervention selection, it is necessary to know what types of errors occur and how this error can affect the confidence interval (CI) in a study (Christ, 2006). CI gives an estimate of reliability in a study. When scores fall near cutoff points, there is a higher likelihood for misclassification (Coulter, Shavin, & Gichuru, 2009). R-CBMs produced a median estimate of standard error of measure (SEM) of 10 WRC per minute. SEM is the standard deviation of errors of measurement that are associated with test scores. Grade level, measurement conditions, and sample differences could all affect the amount of error. This suggests that the SEM for R-CBM will probably not be zero. Therefore, school psychologists should help teachers understand how to incorporate SEM and CIs while using CBM data to guide instruction. Only after SEM is considered can CIs be constructed and future direction of instruction within the classroom determined (Christ & Silberglitt, 2007). Another common form of error is the standard error of the slope (SEb). SEb is how much the found slope differs from the actual slope. This SEb is used in place of SEM when
slope is used to assess R-CBM data with regards to student growth (Hintze & Christ, 2004). To best reduce SEb, longer progress monitoring durations are more appropriate (Christ, 2006). The growth should be determined by using ordinal least squares (OLS) regression, which is accepted as the strongest predictor of actual growth (Christ, Zopluoglu, Long, & Monaghen, 2012; Good & Shin, 1990).

The standard error of the estimate (SEE) is defined as the accuracy of predictions made. It can be reduced when measurement conditions are controlled (Christ, 2006; Christ & Silberglitt, 2007). Controlling measurement conditions is best done by ensuring a quiet testing area, the same administrator giving the assessment in the same fashion, the same location, probes of equal difficulties, and standardized directions (Christ, 2006). Consistency of administration is a key component of attaining accurate results. These conditions can be controlled by the administrator so the accuracy of the results does depend on their skill. Timing errors are frequently made by inexperienced administrators. Best results with administrators are seen in lower grades, but this may be due to the fact that reading is performed at a slower rate by these students. It may be beneficial for teachers and other administrators to review the administration manual before administering assessments to students. Finally, it is important to make sure administrators are aware that even a one-minute assessment needs to be given with fidelity because of the importance of the results (Coulter et al., 2009). Christ, Zopluoglu, Long, and Monaghen (2012) supported the necessity of standardized administration procedures.

**Progress Monitoring Intervals and Growth Rates.** Jenkins, Graff, and Miglioretti (2009) conducted a study with elementary students to determine how to best
maximize results from R-CBM. They examined frequency of collecting data and number of scores recorded at each monitoring point. When it comes to frequency of progress monitoring, the researchers found that growth estimates from points taken every three weeks gave the most accurate prediction of actual growth rate. When estimates were taken every two weeks, there was a very low correlation between predicted growth rate and actual growth rate. This has a strong implication for educators because with further study, educators may be able to decrease how often they assess students using CBM, while still getting valid results.

A minimum of two data points is necessary to establish growth estimates. Typically in the classroom these estimates are taken as benchmarks three to four times a year. Ardoin and Christ (2008) claim that data points taken from the same semester provide the best estimate of growth. If benchmarking occurs on a quarterly basis throughout the school year, this would provide for nine to ten weeks between data point collections. For optimal growth estimates Christ, Zopluoglu, Long, and Monaghan (2012) suggest one data point a week for 14 weeks. Because of time constraints, this may not be possible in the classroom.

Christ, Silbergitt, Yeo, and Cormier (2010) showed that growth rates were higher for the general education population than the special education population. They also reported that growth was larger in early elementary grades and there is a seasonal effect with rates of growth. Specifically, there is more growth in the fall than in the spring, and the seasonal effect is not as prominent with the special education population.
Dynamic Indicators of Basic Early Literacy Skills

The Dynamic Indicators of Basic Early Literacy Skills (DIBELS) is a set of procedures and measures for assessing the acquisition of early literacy skills from kindergarten through sixth grade produced by the Dynamic Measurement Group. They are designed to be short (one minute) fluency measures used to regularly monitor the development of early literacy and early reading skills (Good & Kaminski, 2009). The three main reasons teachers cite for using DIBELS are to identify at-risk learners, develop appropriate interventions, and monitor progress (Hoffman, Jenkins, & Dunlap, 2009).

DIBELS were developed and researched in the late 1980s at the University of Oregon, were based in the CBM ideas developed by Deno (1985), and were designed to be efficient and economical. DIBELS measure Phonological Awareness, Alphabetic Principal, Accuracy and Fluency with Connected Text, Vocabulary, and Comprehension. These five areas are considered the main components of reading instruction (National Reading Panel, 2000). They are not specific to any one curriculum, district, or state; they can be used across the country. DIBELS are not designed to give precise measurements for all components of the five main areas of reading; however they are good indicators of general levels of success in any of the 5 measured areas (Good & Kaminski, 2009). DIBELS measures are beneficial because of the short time needed for assessment, the fact that it can easily be repeated, and that it can be easily integrated in to the current curriculum. They are easy to score and fairly easy to administer (Elliott, Lee, & Tollefson, 2001).
DIBELS provides benchmark goals to evaluate student progress. These goals are the lowest level at which a student can perform and still be progressing towards being a proficient reader. Eighty to 85 percent of students who achieve benchmark goals successfully reach the next goal. DIBELS also provides cutoff scores which reveal the expectation for the student to meet early literacy goals. Only 20% of students who score at the cutoff level go on to reach the next goal without additional help. When scores fall between the benchmark goal and cutoff score, 50% of students reach the projected goals. These students can also benefit from additional assistance in the classroom to help them achieve to their potential. The benchmark goals and cutoff scores are both research-based and criterion referenced (Dynamic Measurement Group, 2011).

A large number of teachers use DIBELS in their classroom; volunteers and paraprofessionals often help with administration in order to expedite the assessment process. It is important to make sure those in charge of administration receive the proper training for DIBELS. DIBELS is generally used with whole classes and not just for at-risk students. In fact, DIBELS can be a major part of any three-tiered RtI system. Even though DIBELS is traditionally being used at Tier 1, it is also being used at Tiers 2 and 3 (Hoffman, Jenkins, & Dunlap, 2009). Tier 1 of the RtI system involves the entire student population and focuses on core instruction. Tier 2 impacts a smaller portion of the students in a school who are generally considered to be at-risk of academic failure. Students in Tier 3 require highly specialized instruction, using specific school resources.

Teachers tend to appreciate that each administration lasts only a short amount of time, but may be frustrated by the fact that administrations only occur three times a year and would like results on progress sooner (Hoffman, Jenkins, & Dunlap, 2009). DIBELS
can be an integral part of the classroom because teachers may overestimate students’ early literacy skills and may not accurately identify students at risk (Martin and Shapiro, 2011). Thus, a tool like DIBELS can help teachers in knowing which students may require additional help with reading.

Multiple studies have examined DIBELS’ validity. Nelson (2008) found DIBELS more accurately identified students with adequate reading skills as opposed to those students at-risk for reading problems. Another study found DIBELS ORF measures were the best predictors of year-end achievement, particularly the fall administration scores (Schilling et al., 2007). A modified version of DIBELS (DIBELS-M) for kindergarten children assesses Letter Naming Fluency, Sound Naming Fluency, Initial Phoneme Ability, and Phonemic Segmentation Ability, all of which are areas of early literacy development. DIBELS-M reliably targets kindergarten students who would benefit from extra instruction. It can also be used to monitor progress of reading development of all students, and appraise how well instruction is working in a classroom (Elliot et al., 2001).

DIBELS can effectively predict students’ success on standardized tests. For example, first graders’ DIBELS ORF scores significantly predicted success on other measures of achievement, such as the TerraNova California Achievement Test (CAT) Assessment in second-grade and the Pennsylvania System of School Assessment (PSSA) in third-grade (Goffreda, DiPerna, & Pedersen, 2009). When students’ performance on DIBELS ORF was compared to their scores on FCAT-SSS and SAT-10, there was a strong correlation observed regardless of race, ethnicity, language, and socio economic status (SES) (Roehrig et al., 2008).
To decrease the standard error of the slope of a progress line (SEb) when using DIBELS, Ardoin and Christ (2008) provide several recommendations. First, even though only one probe is necessary for universal screening, better results and estimates of growth are attained when using the median performance score of three probes. They also state that the semester (i.e. fall, winter, or spring) of administration has an effect on growth estimates. The growth estimates with the highest reliability are from scores taken from within a semester as opposed to between semesters, and when probe sets are consistent across administrations. More growth is typically seen between fall and winter than between winter and spring. There is a stronger chance of a higher SEb when difficulty level is not controlled than when students are given the same probe multiple times. Current studies have shown that using the mean of the three probes may be a more accurate indication of a students’ words per minute (WPM) because of practice effects (Pescher and Kim, 2011).

**DIBELS Next.** In 2010 the Dynamic Measurement Group released DIBELS Next. Following the same overall ideas and concepts on the first six editions of DIBELS, DIBELS Next had changes that included new forms and passages, new directions for assessor and student which were more clearly understood, and retell that went with the ORF measure. Also, there were two new procedures not included in past versions of DIBELS: First Sound Fluency (which replaced Initial Sound Fluency) and the Daze measure, which is based on the maze procedures (DIBELS Next, 2010). According to the DIBELS Next Technical Manual (2011, p 13), “The goal of DIBELS Next was to improve DIBELS without fundamentally changing how it functions.” The Dynamic
Measurement Group completed studies to show the validity and reliability for all measures of DIBELS Next (DIBELS Next Technical, 2011).

The specific changes made to DIBELS ORF measures include the following: a new font, new layout and directions, new reading passages, a checklist of typical response patterns, and new scores for both error and retell (DIBELS Next, 2010). Powell-Smith, Good, and Atkins (2010) completed a study supporting reliability and validity of DIBELS Next ORF. They also studied the passage difficulty of the measure as well and found difficulty and readability of DIBELS Next ORF were appropriate for each grade level.

**Gender Differences and Reading Ability**

Over the years, research has shown a significant difference in reading ability between males and females in different populations, age levels, grade levels, and with different measurement tools. Moreover, reading skills developed by boys are lower than that of girls when they are entering kindergarten (Below, Skinner, Fearrington, & Sorrell, 2010). Based on analysis of the National Assessment of Educational Progress (NAEP), females consistently score higher than males. This holds true for scores from fourth-, eighth-, and twelfth-grade (Klecker, 2006). One possible reason for this is that males exhibit a higher rate of reading disabilities (RD). RD rates for males were likely inflated, but there is still a possibility that males demonstrate reading difficulties more often than females (Liederman, Kantrowitz, & Flannery, 2005). Reading abilities in males are measured by achievement outcomes in reading, which can be influenced by expectations of the male students. These expectations and perceptions could have an impact on the level of performance of boys (Coddington & Guthrie, 2009).
Another explanation for lower reading performance in boys than girls is that boys may have less motivation to read. This motivation or desire to read plays a role in the literacy development of any child. Boys and girls who are average readers are equally self-confident when it comes to their reading abilities. However, average achieving boys are less motivated to read when compared to average achieving girls. Boys tend to value reading less than girls, even when there are minimal differences in the level of motivation with respect to self-concept (Marinak & Gambrell, 2010). Boys are at risk for decreased motivation to read if they begin with low phonemic skills (Coddington & Guthrie, 2009). A popular belief is that boys will read more if they have an interest in the material that is being read. The majority of boys will pick a book that is non-narrative or non-fiction reading when given the choice. However, the majority of required reading in school is narrative. Because schools struggle to interest boys in reading, it is recommended that teachers offer material that was found to be interesting to boys (Merisuo-Storm, 2006).

Skinner, Fearrington, and Sorrell (2010) recently conducted a study using DIBELS as a measurement tool to assess for gender differences in reading. This was the first published study addressing gender differences using DIBELS. The study was conducted with kindergarten through fifth-grade students; it addressed four early reading skills: phonemic awareness, phonics, orthography, and fluency. As with previous studies, girls achieved higher scores on all kindergarten pre-literacy skills. This study was unique because it did not show that this gap grew as students went through school. Any differences that were significant in kindergarten for the Letter Naming Fluency (LNF), Phoneme Segmentation Fluency (PSF), and Nonsense Word Fluency (NWF) tests became insignificant when the students were in first-grade. When testing ORF there
were no significant gender differences in first- through third-grade. There was a slight difference in fourth-grade, but this was no longer apparent by the time the students reached fifth-grade. Boys in fifth-grade actually showed larger gains in ORF than girls from winter to spring.

The authors suggest that these findings support the idea that males tend to mature later than females. They also claim that school environment, curricula, and teachers do not have an effect on promoting one gender over another. Current practices in schools do not favor females over males and observed differences in reading abilities may be related to typical development of both sexes. It should be noted that kindergarten teachers should be aware of differences, but realize that the differences are small and can be overcome. The authors also suggest a change in reading material may help male readers develop an interest in reading early on (Below et al., 2010).

Summary

Studies have shown that CBM is a valuable tool in general and special education classrooms. Specifically, R-CBM can be used to assess reading progress and give an indication of reading ability. While older studies have shown a gender gap in reading achievement and reading growth, more recent studies have shown this to not be true. Few recent studies have examined the issue of progress monitoring with relation to gender. With greater emphasis put on collection of data on all students, it is important that data are gathered in a timely and efficient manner. This can be difficult because of the variety of skill levels of the students. The purpose of this study is to determine if it is possible to reliably reduce the number of weeks required for data collection and analysis of R-CBM, specifically DIBELS, while still obtaining the same results for both genders.
Because of noted gender differences in past research, it is important to address progress monitoring with males and females as individual groups to increase the validity of the assessment.
CHAPTER III
METHOD

Research Questions and Hypotheses

The current study answered the following research questions: 1) Can progress monitoring data using two data points collected over a shorter period of time (i.e. weeks 1 and 5) yield the same growth rate as progress monitoring data using two data points collected over ten weeks (weeks 1 and 10)?; 2) Can progress monitoring data from males, using two data points collected over a shorter period of time (i.e. weeks 1 and 5), yield the same growth rate as progress monitoring data using two data points collected over ten weeks (weeks 1 and 10)?; 3) Can progress monitoring data from females, using two data points collected over a shorter period of time (i.e. weeks 1 and 5), yield the same growth rate as progress monitoring data using two data points collected over ten weeks (weeks 1 and 10)?; 4) Can progress monitoring data from three data points over ten weeks yield the same growth rate as progress monitoring data collected weekly for ten weeks?

For research questions 1-3, it was hypothesized that the rate of change data collected in a time period shorter than ten weeks will yield the same results with respect to growth in oral reading fluency (ORF) as data collected every week over a ten week period of time, and this will be true for both male and female populations. The null hypothesis was that rate of change data collected in a time period shorter than ten weeks
would not yield the same results with respect to growth in oral reading fluency (ORF) as data collected every week over a ten week period of time.

For research question 4, it was also hypothesized that rate of change data from three data points over ten weeks will yield the same results with respect to growth in oral reading fluency (ORF) as data collected every week over a ten week period of time. This prediction was based on a study by Jenkins, Graff, and Miglioretti (2009), who found that data collected every three weeks gave a better prediction than data collected every two weeks. The null hypothesis was that rate of change data from three data points over ten weeks would not yield the same results with respect to growth in oral reading fluency (ORF) as data collected every week over a ten week period of time.

**Participants and Setting**

The sample consisted of (n = 98) third grade students from local public schools. Data were collected from five different schools in the area surrounding a midsize city in the Midwest of the United States. Approximately half of the participants (53%) were female (n = 52) and the rest were males (n = 46). Table 1 shows the frequencies and percentages for ethnicity. Eleven other students participated in the study; however, the data for these students was not used because they had three or more missing data points or had a missing data point in week 1 or 10.
Table 1  
*Frequencies and Percentages of Ethnicity*

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>White, Non-Hispanic</td>
<td>74</td>
<td>75.5</td>
</tr>
<tr>
<td>Asian, Pacific Islander</td>
<td>11</td>
<td>11.2</td>
</tr>
<tr>
<td>African American</td>
<td>8</td>
<td>8.2</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>2.0</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>Not Identified</td>
<td>2</td>
<td>2.0</td>
</tr>
</tbody>
</table>

**Research Design**

This study involved quasi-experimental research which examined relationships between variables. For research questions 1-3 the dependent variable was the rate of change of ORF, which is also the slope of the line of best fit for data points collected, and the independent variable was the weeks between collections of the data points. For research question 4 the dependent variable was the rate of change of ORF, which is also the slope of the line of best fit for data points collected, and the independent variable was the number of data points collected.

The study examined if the slope was different when data points attained closer to one another were used in comparison to data points collected nine weeks apart. For example, the slope, as determined by data points in the first and third week, was compared with the slope from data points in the first and tenth week of data collection. Additionally, the sample was separated by gender and the comparison was made again. Finally, a growth estimate across three weeks was compared to a growth estimate across nine weeks and ten data points.
Materials

DIBELS Next Oral Reading Fluency probes (see Appendix A for a sample probe) were used to obtain oral reading fluency data. Grade-specific probes in DIBELS Next are supported as reliable and valid indicators of ORF for students in the appropriate grade (DIBELS Next Technical, 2011; Powell-Smith, Good, and Atkins, 2010). The probes can be found online at https://dibels.org/next/index.php.

Procedures

Initial approval from the Institutional Review Board (IRB) at the University of Dayton was gained in August 2011 and re-approval was gained in the summer of 2012. Five trained school psychology student interns were placed at five different area schools. Consent was obtained from the principal of each building as well as the parents of all the children involved. Assent was obtained from all participating students.

Data collection started the third week of school in each building. The interns each administered DIBELS Next to one third-grade class once a week for 10 weeks. Three probes were administered to each student using the standardized administration guidelines (see Appendix B), and the median score of the three probes was used for the weekly data point. Along with DORF data, demographic information was also collected for all participants (see Appendix C), including gender, ethnicity, qualification for free and reduced lunch, and age. The data were collected in the Fall of 2011 and two in the Fall of 2012.
CHAPTER IV
RESULTS

The results for research questions 1-3 produced interval data. Multiple growth estimates were conducted using two data points from the ten that were collected. Correlational statistics were used to analyze the data.

Research Question 1

Correlational coefficients were computed among the different growth rates. Using the Bonferroni approach to control for Type I error across the 8 correlations, a $p$ value of less than .006 ($0.05/8 = 0.006$) was required for significance. The results of the correlational analyses presented in Table 2 show that 7 out of 8 correlations were significantly correlated and were greater than or equal to .35, with the exception being growth calculated from weeks one and two. These all had moderate correlation with the highest correlation to growth from weeks one to ten associated growth from weeks one to seven (.652) and weeks one and six (.588).
Table 2  
*Correlations of Growth Estimates for the Entire Sample (N = 98)*

<table>
<thead>
<tr>
<th></th>
<th>Weeks One and Ten</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weeks One and Two</td>
<td>.32*</td>
</tr>
<tr>
<td>Weeks One and Three</td>
<td>.46*</td>
</tr>
<tr>
<td>Weeks One and Four</td>
<td>.51*</td>
</tr>
<tr>
<td>Weeks One and Five</td>
<td>.51*</td>
</tr>
<tr>
<td>Weeks One and Six</td>
<td>.59*</td>
</tr>
<tr>
<td>Weeks One and Seven</td>
<td>.65*</td>
</tr>
<tr>
<td>Weeks One and Eight</td>
<td>.56*</td>
</tr>
<tr>
<td>Weeks One and Nine</td>
<td>.45*</td>
</tr>
</tbody>
</table>

* p < .006 (2-tailed)

**Research Question 2**

Correlational coefficients were computed among the different growth rates for the male population. Using the Bonferroni approach to control for Type I error across the 8 correlations, a *p* value of less than .006 (.05/8 = .006) was required for significance. The results of the correlational analyses presented in Table 3 show that 6 out of 8 correlations were significantly correlated and were greater than or equal to .35. These all had moderate correlation with the highest correlation to growth using weeks one and ten coming from growth using weeks one and eight (.669) and weeks one and seven (.612).

Table 3  
*Correlations of Growth Estimates for Males (N = 46)*

<table>
<thead>
<tr>
<th></th>
<th>Weeks One and Ten</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weeks One and Two</td>
<td>.15</td>
</tr>
<tr>
<td>Weeks One and Three</td>
<td>.36</td>
</tr>
<tr>
<td>Weeks One and Four</td>
<td>.43*</td>
</tr>
<tr>
<td>Weeks One and Five</td>
<td>.48*</td>
</tr>
<tr>
<td>Weeks One and Six</td>
<td>.52*</td>
</tr>
<tr>
<td>Weeks One and Seven</td>
<td>.61*</td>
</tr>
<tr>
<td>Weeks One and Eight</td>
<td>.67*</td>
</tr>
<tr>
<td>Weeks One and Nine</td>
<td>.43*</td>
</tr>
</tbody>
</table>

* p < .006 (2-tailed)
Research Question 3

Correlational coefficients were computed among the different growth rates for the female population. Using the Bonferroni approach to control for Type I error across the 8 correlations, a $p$ value of less than .006 ($0.05/8 = .006$) was required for significance. The results of the correlational analyses presented in Table 4 show that 8 out of 8 correlations were significantly correlated and were greater than or equal to .35. These all had moderate correlation with the highest correlation to growth using weeks one and ten coming from growth using weeks one and seven (.697) and weeks one and six (.630).

Table 4

| Correlations of Growth Estimates for Females (N = 52) |
|-----------------|-----------------|
| Weeks One and Two | .44* |
| Weeks One and Three | .53* |
| Weeks One and Four | .57* |
| Weeks One and Five | .52* |
| Weeks One and Six | .63* |
| Weeks One and Seven | .70* |
| Weeks One and Eight | .51* |
| Weeks One and Nine | .49* |

* $p < .006$ (2-tailed)

Research Question 4

Multiple growth estimates were conducted using three data points from the ten that were collected and one growth estimate was conducted using all ten data points. Ordinal least squares (OLS) regression was used to estimate growth. Correlational statistics were then used to analyze the data. The results from the testing produced interval data.

Correlational coefficients were computed among the different growth rates. Using the Bonferroni approach to control for Type I error across the 6 correlations, a $p$
value of less than .008 (.05/6 = .008) was required for significance. The results of the correlational analyses presented in Table 5 show that 5 out of 6 correlations were statistically significant and were greater than or equal to .35. These all had moderate to high correlation with the highest correlation to growth using all ten data points coming from growth using weeks one, six, and ten (.741), one, five, and ten (.733), and one, five, and nine (.722).

Table 5

<table>
<thead>
<tr>
<th>Correlations of Growth Estimates Using More than Two Data Points (N = 98)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Ten Data Points</td>
</tr>
<tr>
<td>Weeks One, Two, and Three</td>
</tr>
<tr>
<td>Weeks One, Three, and Five</td>
</tr>
<tr>
<td>Weeks One, Four, and Seven</td>
</tr>
<tr>
<td>Weeks One, Five, and Nine</td>
</tr>
<tr>
<td>Weeks One, Five, and Ten</td>
</tr>
<tr>
<td>Weeks One, Six, and Ten</td>
</tr>
</tbody>
</table>

* p < .008 (2-tailed)
CHAPTER V
DISCUSSION

Review of Purpose and Major Findings

As the demand for the use of data in education increases, so does the demand for ways to accumulate the data. CBM is one form of progress monitoring that can be used easily in schools. The information from data collected with CBM can be used to give feedback on the instruction in the classroom. Sometimes time constraints do not allow progress monitoring on a regular basis or over the course of ten weeks. If the number of data points needed for a reliable growth estimate can be collected at shorter intervals or reduced, schools can make a better use of time and devote more resources to instruction. It also benefits schools to know that a reduction in data points collected would be possible to do with both males and females even though these two groups have historically shown a difference in reading growth.

Results from the current study show that it may be possible to produce a reliable growth estimate in reading when data points are collected less than ten weeks apart. This holds true for both male and female students with respect to CBM-R. Specifically, data points from weeks one and seven of a ten week period will yield approximately the same rate of growth as data points from weeks one and ten.
The current study also indicates that using linear regression with three data points will produce a similar growth estimate as using ten data points. The highest correlation was observed with data points taken from weeks one, six, and ten.

**Interpretation of Findings Relative to Predictions**

Based on results from correlational analyses of multiple growth rates derived from two data points with differing time between collection of the data points of the entire sample; it was determined that growth estimates using data points from weeks one and six, weeks one and seven, and weeks one and eight are more closely related to the growth estimate produced by data from weeks one and ten than other combinations of weeks.

Based on results from correlational analyses of multiple growth rates derived from two data points with differing time between collection of the data points of the male sample; it was determined that growth estimates using data points from weeks one and seven and weeks one and eight are more closely related to the growth estimate produced by data from weeks one and ten than other combinations of weeks.

Based on results from correlational analyses of multiple growth rates derived from two data points with differing time between collection of the data points of the female sample; it was determined that growth estimates using data points from weeks one and six and weeks one and seven are more closely related to the growth estimate produced by data from weeks one and ten than other combinations of weeks.

Based on results from correlational analyses of multiple growth rates derived from three data points taken from different combinations of weeks over the ten weeks; it was determined that growth estimates using data points from weeks one, five, and nine,
weeks one, five, and ten, and weeks one, six, and ten are more closely related to the
growth estimate produced by data collected weekly than other combinations of weeks.

In all four cases, statistical analysis produced a $p$ value < .01; thus, the four null
hypotheses were rejected, which indicated the initial predictions were true. It is possible
to produce reliable growth estimates for both males and females in a time period shorter
than ten weeks when only using two data points to calculate the rate of growth. Also, it is
possible to produce growth estimates similar to those produced by weekly data collection
when only three weeks are used for the estimate.

The present study provides evidence for validity of DIBELS Next oral reading
fluency probes as a tool for measuring reading growth. Previous research suggests that
feedback after five weeks of instruction can positively impact classroom outcomes
(Graney & Shin, 2005). The present study shows that valid growth estimates can be
obtained after five weeks. Previous studies also suggested that it may be possible to
collect fewer data points while still obtaining valid growth estimates (Jenkins, Graff, &
Miglioretti, 2009), which was confirmed by the present study.

Past research showed gender differences in reading abilities and growth of male
and female students (Below, Skinner, Fearrington, & Sorrell, 2010; Klecker, 2006;
Liederman, Kantrowitz, & Flannery, 2005), but more recent studies showed that this may
not be the case (Below et. al., 2010; Skinner, Fearrington, & Sorrell, 2010). While the
present study did not look at differences in reading abilities between males and females,
it did provide evidence that valid growth estimates can be obtained for each of these
populations when fewer data points are collected and when two data points are collected
closer together.
Limitations

There are several limitations to this study. First, the data were collected using only one type of curriculum-based measurement, (CBM)-DIBELS Next. Second, the study was conducted in the fall. Time of year can impact testing results; therefore, it would be ideal to have data from throughout the school year. Also, the data were taken over a ten week period of time. Collecting data for up to fourteen weeks may have produced a more exact estimate of the students’ actual growth. Finally, the testing was only conducted at five elementary schools, producing a small sample size with narrow demographics.

Implications for Future Research

This study can be expanded in different ways. For example, other types of curriculum based measurements could be used. These might include reading comprehension probes (i.e. DIBELS Next DAZE) or mathematics progress monitoring probes. Likewise, different CBM brands, such as AIMSweb could be studied in the same manner.

The same study could be conducted at different point in the school year (i.e. winter or spring) to see if results would hold true. Finally, researchers might expand the length of the study to determine if fewer progress monitoring data points would still yield the same rate of growth.

Conclusion

With economic and time constraints faced by school personnel, staff and administration must make changes in delivery of services without compromising the validity and integrity of instruction. The use of CBMs to monitor student progress and
growth is important in schools because of how results correlate with high stakes test scores and the greater use of RtI.

This study showed that the number of data points collected can be reduced and still yield the same results (rate of growth) as data collected on a more frequent basis. Also, data collected over a shorter time period will produce a comparable growth estimate to that of a typical benchmarking cycle (ten weeks). With conflicting information on differences in reading ability of males and females, it is important to note that collecting growth estimates can be done more efficiently for both groups.

While it is tempting to reduce the frequency of all data collection and shorten progress monitoring periods, it is important to remember that a higher number of data points are needed for high-stakes decisions such as special education identification. This study provides a starting point for schools to adapt their practices to changing times and climates in the modern educational environment.
REFERENCES


and dependability of curriculum-based outcomes within the context of educational decisions. *Psychology in the Schools, 42*(6), 615-622.


doi:10.1080/02702710802274820


A New Ball Game
On the first day of school, Roy’s teacher asked him to write a letter about himself. Roy was glad to have the chance to talk about his life in Africa. Roy had been born in the United States, but his family had lived in a small town in Africa for three years. Now his family had moved back to the United States. Roy’s stepmom was a doctor. She worked in a clinic, where she treated sick people and gave immunizations. His dad taught music at the school Roy and his brother attended. Roy and his friends played sports together and practiced playing the instruments his dad taught them. Football was his favorite sport and there always seemed to be a game going after school.
In his letter, Roy wrote about his life and that he missed playing football the most. He had seen American football and it was a very different game. It was hard to figure out why they were chasing each other and when to cheer.
He handed the letter to his teacher the next morning. That afternoon as he was leaving, his teacher called him over. Another boy was standing next to her. “Roy, this is Spencer,” his teacher said. “He’s going to introduce you to the soccer team. I think you’ll enjoy it.”
Spencer smiled at Roy and led him outside to the field, where a group of kids were playing. “The teacher said you call this football in Africa, but here it’s called soccer,” said Spencer. Roy looked around. The kids were playing the same game he had played back in Africa! He couldn’t wait to join them.

APPENDIX B

DIBELS NEXT ADMINISTRATION INSTRUCTIONS

Make sure you have reviewed the directions in the DIBELS Assessment Manual and have them available. Say these specific directions to the student:

*I would like you to read a story to me. Please do your best reading. If you do not know a word, I will read the word for you. Keep reading until I say “stop.” Be ready to tell me all about the story when you finish.* (Place the passage in front of the student.)

Begin testing. *Put your finger under the first word* (point to the first word of the passage).

*Ready, begin.*

**Timing** 1 minute. Start your stopwatch after the student says the first word of the passage. Place a bracket ( ) and say *Stop* after 1 minute.

**Wait** If no response in 3 seconds, say the word and mark it as incorrect.

**Discontinue** If no words are read correctly in the first line, say *Stop*, record a score of 0, and do not administer Retell.

If fewer than 40 words are read correctly on any passage, use professional judgment whether to administer Retell for that passage.

**Reminders** If the student stops (not a hesitation on a specific item), say *Keep going.* (Repeat as often as needed.)

If the student loses his/her place, point. (Repeat as often as needed.)

## APPENDIX C
### DATA COLLECTION TOOL

*For use with every student assessed using DIBELS Next.*

Collection date:  
ORF scores:  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

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

Free/reduced lunch?  
- Yes  
- No  

Gender:  
- Male  
- Female  

Age: