EFFECTS OF PROGRESS MONITORING FEEDBACK ON EARLY LITERACY STUDENT ACHIEVEMENT

by

Jeremy Jon Lopuch

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Special Education

Charlotte

2016

Approved by:

____
Dr. Lindsay J. Flynn

____
Dr. Kristen D. Beach

____
Dr. Christopher P. O'Brien

____
Dr. Paola Pilonieta

ABSTRACT

JEREMY LOPUCH. Effects of progress monitoring feedback on early literacy student achievement.
(Under direction of DR. LINDSAY J. FLYNN)

The purpose of this study was to examine the effects of diagnostic formative assessment feedback on early literacy skills. The participants were 12 first-grade general education teachers and 51 of their students who were assigned to the following treatments, diagnostic feedback and skills feedback (control) which lasted for 10 weeks. During the study, participating teachers in both feedback conditions reviewed student response patterns on a one-minute assessment of word decoding to identify their students' phase of word reading development. Teachers of students in the diagnostic feedback condition were also provided instructional recommendations aligned to student word reading needs. Three measures were collected in the course of the study. First, all students were measured on word decoding for a total of four data waves. Second, teachers in both conditions recorded the number of instructional changes they planned to implement in response to changing student data. Third, a measure of oral reading fluency was collected at posttest. A multilevel growth-curve model indicated diagnostic feedback was a non-significant predictor of decoding skills. Analysis of covariance indicated diagnostic feedback group affiliation was not associated with oral reading fluency scores after controlling for initial decoding skills. With regards to instructional planning, skills feedback teachers reported a higher number of instructional changes. In addition, skills feedback teachers viewed progress monitoring feedback as significantly more helpful for planning compared to diagnostic teachers. Amount of instructional changes was unrelated to posttest decoding and oral reading fluency scores. Implications for practice and suggestions for future research will be discussed.

ACKNOWLEDGEMENTS

I would like to thank the members of my dissertation committee for their support on this project and in my growth as a special educator. Dr. Lindsay Flynn for her passion related to assessment and instruction. Her expertise in these areas helped me focus my interests to a career path of researching assessment and instruction for struggling readers. Dr. Christopher O'Brien for his boundless knowledge related to the field of special education. His insightful views have forever altered my understanding and awareness of individuals with disabilities. Dr. Kristen Beach for her valuable suggestions during the development and analysis of this project. Her willingness to donate her time to my many questions about HLM (and teaching!) are much appreciated. I would also like to thank Dr. Paola Pilonieta for her time and assistance on my study.

Finally, I wish to express my sincere gratitude to my loving wife and best friend, Brooke, for her continued support throughout my doctoral program and dissertation. Her motivation and caring encouragement these last five years provided me the drive to persevere during the most trying of times. She is truly deserving of an "honorary" doctorate.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Statement of the Problem

According to the National Center of Educational Statistics, 65% of students fail to read grade level text above a *Basic* level of comprehension (NCES, 2013). This means more than half of the students in the United States struggle to understand and make inferences from written text (NCES, 2013). Development of reading skills, especially during first-grade, often sets the stage for future reading success. In a longitudinal study, Cunningham and Stanovich (1997) found reading ability in first-grade associated with reading ability in middle elementary to 11th grade. Unfortunately, current data suggest many students fail to acquire the foundational reading skills necessary to become proficient readers. Historically, students who fall behind in first-grade rarely catch up (e.g., Juel, 1988; Phillips, Norris, Osmond, & Maynard, 2002). In a longitudinal study examining the reading achievement of 187 children from first through sixth grade, Norris and colleagues (2002) found that almost half of the children categorized as below average readers in first grade continued to demonstrate below average achievement in sixth grade. These data present evidence that reading skills between good and poor readers typically diverge in first-grade (Deno, Marston, & Shin, 2001; Speece & Ritchey, 2005) with this trend continuing until the end of elementary school and into adolescence (Francis et al., 1996; Roberts, Vaughn, Fletcher, Stuebing, & Barth, 2013; Vaughn et al., 2010). Given that higher levels of literacy skills are required for existing and future professions (e.g., Snow, Burns, & Griffin, 1998), persistent deficiencies in reading achievement reach beyond classroom performance and will likely impact future quality of life.

A primary reason for children falling behind in reading in first grade is poor attainment of early literacy skills, especially the alphabetic principle or phonics (Gough & Tunmer, 1986; Hoover & Gough, 1990). Adequate phonics skills consist of automatically identifying individual letter-sound correspondences, word parts and blending sounds together to form whole word units (Fuchs, Fuchs, Hosp, & Jenkins, 2001). Word level analysis skills are a good predictor of general reading ability in first grade (O'Conner & Jenkins, 1999; Speece et al., 2011). Acquisition of these literacy skills are rapid and time sensitive to reading development (Paris, 2005). Without these skills students struggle fluently reading words in written text, which impairs comprehension because the focus of the child's reading is recognizing the word and not understanding the text (Daane, Campbell, Grigg, Goodman, & Oranje, 2005; see LaBerge & Samuels, 1974). National data indicate more than a quarter of later elementary school students have difficulty reading grade level material accurately (Daane et al., 2005). This is dubious because students who struggle with reading tend to process decreased levels of text information compared to typical peers (e.g., Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003). The result of which is less motivation to read leading to poorly developed vocabulary skills impacting the ability to read and understand text (McNamara, Scissons, & Gutknecth, 2011; Stanovich, 1986)

Responsiveness to Instruction (RtI)

Historically, students exhibiting deficit decoding skills with problems reading and understanding written text often did not receive supplemental or special education support until second- or third-grade. This has been referred to in the literature as a wait-to-fail model because educators delayed the provision of intervention until struggling students were well behind their typically developing peers. This is problematic because research has consistently demonstrated that unaddressed early literacy skill deficits continue to persist into later grades (Ferrer et al., 2015).

In contrast to a wait-to-fail model, where students progress through school without receiving support, RtI models identify student problems early before academic difficulties become too deeply entrenched (Vaughn & Fuchs, 2003). Students who are not responsive to generally effective classroom instruction receive progressively intensive tiers of supplemental instruction designed to accelerate student learning. To measure the effectiveness of this supplemental instruction, learning progress is monitored more frequently (e.g., monthly or weekly) for readers receiving supplemental support than their average achieving reading peers. This system of measurement characterizes the rate of student learning across time and provides teachers with a database upon which to center instructional decision making according to individual student response to instruction (Fuchs & Fuchs, 2006).

Progress Monitoring to Formatively Assess Student Growth

With progress monitoring serving as an essential element in an RtI model for determining whether teacher pedagogy requires modification to improve student learning. The database should also provide information to the teacher about content and strategies to use with the student (Fuchs & Deno, 1994). Unsuccessful strategies should be discontinued and new activities should be implemented based on monitoring data. Progress monitoring, as a form of formative evaluation, has strong support in the literature (see Black & Wiliam, 1998; Fuchs & Fuchs, 1986) for improving student outcomes. In a synthesis of 138 teacher variables on student achievement, Hattie (2009) found progress monitoring and formative evaluation ranks as the third largest effect size (.90) on student achievement next to student expectations and teacher sincerity in the judgements of their students.

Progress Monitoring Early Literacy Skills

The *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS; Good & Kaminski, 2010) are a set of formative assessments developed specifically to measure early literacy skills. Test items consist of identified important behaviors in early literacy skill development, such as phonemic awareness, alphabetic principle (e.g. decoding), reading fluency, vocabulary, and reading comprehension (i.e., National Reading Panel, 2000). Educators are expected to use DIBELS to formatively assess beginning reading skills to provide teachers with information to plan instruction (Coyne & Harn, 2006). Research on the DIBELS subtests indicates students who achieve research-established benchmarks (i.e., low-risk status) have a high probability of demonstrating future reading proficiency (Good, Baker, & Peyton, 2009). Therefore, DIBELS provides teachers with progress monitoring feedback about the effectiveness of their early reading instruction, a key assumption related to the RtI model.

Linking Progress Monitoring to Improved Student Achievement: Teacher Feedback

Through the collection of progress monitoring data generally produces positive effects on student achievement (e.g., Hattie, 2009), simply monitoring student learning is not enough. In a review of the effects of progress monitoring on student achievement, Stecker, Fuchs, and Fuchs (2005) emphasize that is crucial for teachers to attend to the feedback provided by progress monitoring to realize the benefits of measuring student learning. Specifically, teachers have to use individual student's data to plan and change instruction as indicated through the student database. In addition, teachers need consultation to decide on content (i.e., skills) and pedagogy (i.e., instructional activities) to gain benefits of progress monitoring on student achievement. The fact that teachers need expert assistance to interpret measurement feedback is a common finding and reoccurring theme throughout the progress monitoring literature (e.g., Al Otaiba & Lake, 2007; Hoffman, Jenkins, & Dunlap, 2008; Roehrig, Duggar, Moats, Glover, & Mincey, 2008). Generally, research findings on progress monitoring indicate teachers do not find results useful for day-to-day instructional planning (Arthaud, Vasa, & Steckelberg, 2000; Wesson, King, & Deno, 1984; Yell, Deno, & Marston, 1992). In a recent study examining teacher use of progress monitoring data to plan instruction, Roehrig and colleagues (2008) found kindergarten and first grade classroom teachers struggled interpreting DIBELS subtest feedback to select appropriate pedagogy. Teachers in the study were only able to make a connection between student performance and instructional planning when provided expert instructional feedback from a consultant (i.e., reading coach). This finding supports the perception that teachers do not have the capacity to use DIBELS feedback to enhance student achievement. Furthermore, in a survey, Hoffman et al. (2008) revealed that 57% of teachers use DIBELS to monitor their students' learning; however, only 9% reported using progress monitoring feedback to inform their instruction. This suggests teachers are using DIBELS to formatively assess student learning, but are largely unaware of how to use the feedback to improve student learning. Since progress monitoring plays a key role in RtI and improving student achievement it would appear important for teachers to have the capacity to independently use student data to make educational decisions.

Research on Progress Monitoring and Student Achievement

The assumption that routinely measuring student progress has a positive impact on student achievement is varied in the literature (for reviews see Fuchs & Fuchs, 1986; Stecker, Fuchs, & Fuchs, 2005). In an early study, Stecker and Fuchs (2000) were the first to examine the effects of individualized progress monitoring on student achievement in math. Twenty-two special education teachers and forty-two of their students with disabilities in grades two through eight participated in the study. The participants were randomly assigned to either individualized or paired instruction. Students in the individualized condition received instruction based on individual student progress monitoring data, whereas students in the paired condition received instruction based on another student's data. Progress monitoring data were collected over 22 weeks and teachers received prompts from decision making software when an instructional change was needed. Results indicated that students whose teachers used individual data to make decisions demonstrated higher rates of growth and higher scores on a comprehensive math assessment compared to students who teachers used another student's data to plan instruction.

In a more recent study, Graney and Shinn (2005) examined the effects of progress monitoring feedback on individual and groups of students. Participants included 44 second grade teachers and 184 of their lowest preforming students. Teachers and students were randomly assigned to one of three conditions: group feedback, individual feedback, or no feedback (control). Teachers in the group feedback condition received graphs depicting the average oral reading growth for the group and individual students. Teachers in the individual feedback condition received a similar graph displaying the mean oral reading growth of an individual student. Results indicated higher mean rates of oral reading growth for groups of students compared to individual students. The authors noted overall mean improvement in oral reading suggesting potential benefits of progress monitoring on student achievement. However, students in the individual feedback group displayed mean reading slopes far less steep when compared to students in the group feedback condition. The authors concluded that higher rates of individual student growth probably necessitate more prescriptive feedback (i.e., content and pedagogy) provided to students' teachers.

In a similar study, Souvignier and Forester (2011) also investigated the effects of providing progress monitoring feedback on achievement for poor readers. This study differed because Souvignier and Forester (2011) examined the effects of both static (i.e., single time point) feedback and on-going (i.e., progress monitoring) feedback. Seventeen general education teachers and 144 of their poorest readers were randomly assigned to either static and progress monitoring feedback. Feedback was comprised of information about individual student reading (i.e., reading speed and accuracy) performance. Teachers in the progress monitoring group received feedback on their students' reading achievement across 24 weeks using computer-based reading assessments, while teachers in the static group were only provided information about their students' initial performance. Results indicated that teachers of students in the progress monitoring group significantly outperformed their static group counterparts. In addition, participating teachers rated the progress monitoring feedback as feasible and useful for general education instructional planning.

Most recently, Forester and Souvignier (2015) conducted a follow-up study examining the effects of varying the levels of progress monitoring class-wide feedback on third-grade student performance in reading fluency and comprehension. Forty-three general education classroom teachers and 958 of their students were randomly assigned to one of three conditions: (a) teachers in the progress monitoring condition were provided bi-weekly feedback on student achievement, (b) teachers in the progress monitoring plus teacher training received feedback at the same frequency and were also offered additional training how to interpret feedback for instructional planning, and (c) teachers in the static group only received initial achievement information about their students. Students completed bi-weekly computer-based reading tests across 6 months and feedback was provided through the computer program. Both progress monitoring groups displayed steeper slopes of improvement compared to students in the static group. The effects of additional teacher training were of a small magnitude. Similar to the findings of Souvignier and Forester (2011), general education teachers reported classwide feedback useful for planning instruction.

Research on Progress Monitoring Feedback on Early Literacy Achievement

In the first study of progress monitoring feedback on early literacy skills, Iannuccilli (2003) investigated the effects of DIBELS formative assessment feedback on student achievement. Six first-grade general education classroom teachers and 54 of their students were randomly assigned to one of two groups including feedback (*n* = 26) and no feedback (*n* = 28). Teachers in the feedback condition were provided charted student data and a skills profile of their phonological awareness skills. Beginning in the winter of first-grade, students were assessed bi-weekly for nine weeks on the Phoneme Segmentation Fluency (PSF) and Nonsense Word Fluency (NWF) subtests of the DIBELS suite. Teachers of students in the feedback group received the feedback after each assessment was conducted. Results of a multivariate analysis of covariance (MANCOVA) at posttest indicated no significant differences between groups on both measures. In addition, students in the feedback group displayed similar rates of learning compared to students whose teachers received no feedback.

In a similar study, Ball and Gettinger (2009) examined the effects of early literacy progress monitoring feedback on student achievement. Eight Kindergarten teachers and 103 of their students were randomly assigned to either feedback (*n* = 55) or no feedback (*n* = 48) groups. Feedback consisted of the explanation of DIBELS scores and risk levels. Students were assessed using four DIBELS subtests in the fall, winter, and spring of their Kindergarten year. Feedback was provided only after the fall and winter data collection waves. Students of teachers who received feedback demonstrated significantly higher gains on three of the four DIBELS subtests from fall to spring compared to students whose teachers did not receive feedback. Yet, only half of the students in the feedback group achieved the spring DIBELS benchmark. The authors concluded that feedback had a minimal effect on student achievement due to teachers reporting the provided feedback had minimal to modest support informing instruction.

Limitations in Evidence to Support Early Literacy Progress Monitoring

Currently, there are few published studies on the use of early literacy progress monitoring feedback on the academic achievement of first-grade students. Almost all studies investigating the effects of progress monitoring feedback in reading have been conducted in grades 2 and beyond using oral reading fluency. Most of the recent studies examining the utility of DIBELS feedback on improving student outcomes have focused on the diagnostic accuracy of the measures, meaning how well the measure can predict performance on distal measures of reading achievement. The only published study (Ball & Gettinger, 2009) directly examining the effects of progress monitoring on student achievement was conducted in Kindergarten and consisted of a small sample of teachers (n = 8). Additional limitations stand out impacting the utility of the findings. First, the study procedures provided limited frequency (i.e., twice) of feedback throughout the school year. Next, project staff provided token feedback content to the classroom teacher. Finally, sample limitations, such as lack of diversity in participants further limited the generalization of study outcomes.

Teachers who do receive explicit progress monitoring feedback will not reap benefits of improved student achievement on their own. Review of literature suggests teachers cannot interpret this feedback without technical assistance (Stecker et al., 2005). Perhaps that is why most teachers report progress monitoring data not useful for instructional planning (Ball & Gettinger, 2009). Presently, more than 15,000 schools use

DIBELS (University of Oregon Center on Teaching and Learning [UO-CTL], 2015). If teachers are expected to administer and interpret the DIBELS, it would be helpful to have the ability to apply feedback generated by the measure to enhance student achievement. Ignoring this point will continue to result in critical instruction time wasted administering assessments that are of no use for teacher and student.

Purpose and Research Questions

This study will extend the literature by examining the effects of early literacy progress monitoring feedback on student achievement. As part of the investigation, general education classroom teachers in the feedback group will be trained to interpret feedback from the DIBELS NWF subtest. Then, feedback group teachers will apply the results from the feedback to develop instruction sensitive to individual student needs concerning reading skill level and appropriate research-based instructional strategy. The purpose of this study, therefore, is to examine the effects of diagnostic progress monitoring feedback on early literacy achievement for students in first grade. This study will address three research questions:

1. What is the difference in growth of decoding, as measured using DIBELS NWF, between students whose teachers receive diagnostic feedback and those who receive skills feedback from fall to winter in first grade?
2. What is the difference between oral reading fluency (ORF) scores on the winter benchmark, as measured by DIBELS ORF (DORF), for students whose teachers receive diagnostic feedback and students whose teachers receive skills feedback?

3. What is the difference in instructional changes (i.e., number of reported instructional changes) between students whose teachers receive diagnostic feedback and those who receive skills feedback over fall to winter in first grade?
4. What is the relation between the number of instructional changes and student achievement on DIBELS NWF and DORF for students whose teachers use diagnostic or skills feedback to plan instruction?
5. Which type of feedback (i.e., diagnostic or skills) do teachers report as being more helpful for planning instruction?

Significance of Study

This study will contribute to the literature in several ways. First, study results will contribute to the limited knowledge base of the effects of early literacy formative assessment on student achievement. As mentioned earlier, the literature is virtually bereft of studies examining the effects progress monitoring feedback on student achievement in reading. The results of this study will provide support for the use of progress monitoring, especially for the use of DIBELS NWF, to enhance early literacy instruction for classroom teachers. Second, this study will add to the established knowledge base related to the broader topic of progress monitoring and student reading achievement within the context of RtI. The RtI model assumes that progress monitoring will provide teachers with the necessary feedback to make decisions regarding the modification of instruction, as needed, to enhance student outcomes (Fuchs & Deshler, 2007). Furthermore, this study will test this assumption and provide the field guidance related to the utility of formative evaluation of early literacy to improve student reading achievement within an RtI model.

Delimitations

This study will address the effects of diagnostic feedback related to progress monitoring performance on instructional planning and student achievement. There are delimitations of the study that need to be addressed in order to assist with interpreting the results. First, just one school in the southeastern United States was selected to participate in the study and generalization to other parts of the country with different demographics is discouraged. Second, the school already uses DIBELS subtests to measure early literacy student achievement in grade K-3 and receive general feedback. Third, the participants are limited to first grade general education classroom teachers and their students, which speaks to a limited portion of the population using DIBELS measures. Finally, teachers in the experimental group will be provided feedback regarding student performance on NWF; however, the teachers will administer other measures that assess early literacy skills throughout first grade and results of the other measures may influence instructional planning, in turn impacting achievement.

Definitions

Alphabetic principle- The alphabetic principle is the understanding that letters and letter patterns form the sounds of written language (NRP, 2000). The understanding of this principle allows children to anticipate pattern letters and words make to read unfamiliar words (Adams, 1990)

Decoding – Decoding is the skill in using letter-sound sound patterns to fluently read new words (Florida Center for Reading Research [FCRR], 2007). Decoding words is a goal of effective alphabetic principle/phonics instruction (Armbuster, Lehr, & Osborn, 2001).

Diagnostic Feedback – Diagnostic feedback is feedback derived from assessments that is explicitly prescriptive in terms of instructional content (i.e., skills) and teacher pedagogy (i.e., strategy). In the progress monitoring literature, diagnostic feedback typically involves a profile of individual skill needs and recommended instructional activities for remediation. (e.g., Capizzi & Fuchs, 2005).

Dynamic Indicators of Basic Early Literacy Skills (DIBELS) - DIBELS is a suite of brief measures that have been shown through research to be key indicators of pre-requisite early literacy skills to reading (Good et al., 2001).

Feedback – Feedback is the data produced from an individual completing a task/assessment. The resulting data informs the teacher where the student is in relation to a goal (Wiggins, 2012). Hattie (2012) defines effective feedback as information that teachers can use to close the gap between current student knowledge/skills and the instructional goal.

Formative assessment – Formative assessment refers to any assessment that provides information to help teachers change instruction based on their students' evolving academic needs (Kaminski & Cummings, 2008).

Pedagogy – Pedagogy refers to a strategy or activity used to teach content or a skill. For example, a pedagogy for teaching how to decode words in written text (i.e., skill) could be to underline the words (i.e., activity) that require decoding (FCRR, 2007).

Responsiveness to Instruction (RtI)- RtI is a multi-tiered system of supports developed to prevent academic difficulties through the use of screening to identify struggling students, routinely progress monitoring student learning, and use student data to make educational decisions (National Center on Response to Intervention [NCRTI], 2010).

Progress Monitoring- Progress monitoring is the process of collecting student data across time to assess the effect of instruction on the rate of student learning (NCRTI, 2010).

Skills Feedback – Skills feedback is feedback from an assessment that creates a skills profile unique to an individual student. Skills feedback in the progress monitoring literature generally consists of student specific academic needs so teachers can design and implement targeted instruction (Stecker et al., 2005).

CHAPTER 2: REVIEW OF LITERATURE

The purpose of this chapter is to review and synthesize the literature on the effects of formative assessment teacher feedback on student achievement. Domains specific to this topic include (a) response to intervention, (b) curriculum-based measurement, (c) formative assessment of early literacy skills, and (d) progress monitoring derived teacher feedback and student achievement.

The following chapter will be divided into four sections. First, the theory and essential components of response to intervention will be reviewed. Second, curriculum-based measurement of reading will be reviewed. Third, the purpose of DIBELS and related research about the DIBELS NWF subtest will be reviewed. Finally, the literature will be reviewed regarding formative assessment feedback of reading (i.e., CBM) and early literacy skills (i.e., DIBELS) on student achievement.

Responsiveness to Intervention

Responsiveness to instruction (RtI) is a multi-level prevention system of assessment and instruction with the goal of enhancing achievement for all students (National Center on Response to Intervention [NCRTI], 2010). Through this system, struggling students are identified and provided progressively intensive instruction (Fuchs & Fuchs, 2009). Student learning is monitored recurrently and student outcome data are used to support educational decision making (NCRTI, 2010).

As stated above, the fundamental principle of RtI assessment and instruction is *prevention* (Fletcher & Vaughn, 2009). This principle is closely aligned with the history (National Reading Panel [NRP], 2000; Snow, Burns, & Griffin, 1998) and current knowledge (e.g., Lane, 2014) surrounding effective early literacy instruction. Beginning reading instruction should consist of explicit instruction on hearing and manipulating the individual sounds of oral language, mapping sounds to letter symbols to blend words, reading words, and using comprehension strategies to understand written text (NRP, 2000; Lane, 2014). The majority of children provided these components of scientific research-based early literacy instruction will acquire these skills with little difficulty (National Early Literacy Panel, 2008).

Unfortunately, a subset of students will be unresponsive to generally effective beginning reading instruction for a variety of reasons. Research consistently shows that students enter elementary school with varying levels of the above early literacy skills (O'Connor & Jenkins, 1999; Al Otaiba et al., 2011). In addition, some students have substantial trouble manipulating and matching sounds to letters (Schatschneider & Torgesen, 2004). These students are significantly at-risk for developing future reading problems if their academic needs are not immediately addressed (Torgesen, 2004).

Initial inquiries on early identification and intervention systems provided broad evidence to support a prevention framework of assessment and instruction (e.g., O'Connor, 2000; Torgesen, et al., 1999; Vellutino et al., 1996). Struggling students in these studies displayed stronger reading skills after receiving tiered instruction of graduated intensity. The emphasis on *prevention*, as suggested by these early studies, provided educators with a basis for not having to wait until students demonstrate repeated academic failure before providing additional academic help (Vaughn & Fuchs, 2003). In addition, this provided educators the opportunity to change struggling students' learning trajectories in foundational literacy skills before they become too intractable (Simmons et al., 2008).

Components of RtI

As opposed to waiting for student failure, RtI places the emphasis of educators' energies and actions on preventing reading problems (Fletcher & Vaughn, 2009). This is contributed through a service delivery system of progressively intensive instructional support, assessment, and decision making (Glover & DiPerna, 2007). As part of this system, NCRTI (2010) has identified four key components of RtI including (a) universal screening, (b) progress monitoring, (c) multi-tiered systems of prevention and (d) data-based educational decision making.

Universal screening is recognized as the first phase in the RtI system (NCRTI, 2010). All students in a school are tri-annually assessed using brief screening measures that possess technical adequacy (i.e., reliability, validity, diagnostic accuracy) to predict future learning outcomes. Students who perform poorly on the screen have a high probability of reading difficulties (Fuchs & Fuchs, 2009). Students who perform poorly on the screen are the primary candidates for progress monitoring.

The purpose of progress monitoring is to formatively assess student response to varying forms of instruction across time by measuring the rate student learning (NCRTI, 2010). Measures used for progress monitoring are typically brief, sensitive to small amounts of student growth, and can be administered repeatedly over time. Progress monitoring is used at different frequencies depending on a student's level of support. For example, students provided supplementary academic support receive progress monitoring more frequently than students who do not receive extra assistance. Therefore, progress monitoring is the “R” of the RtI model (Shinn, 2012).

Multi-tiered systems of support consist of the strategies and settings in which instruction occurs. Most RtI systems in the literature are comprised of three tiers of instructional support (Berkeley, Bender, Peaster, & Saunders, 2009). In these systems, the primary focus is on general education in Tiers 1 and 2, with special education in Tier 3 (Fuchs, Fuchs, & Stecker, 2010). Tier 1 instruction is provided in the general education classroom using research-based curricula that address the needs of all students (NCRTI, 2010). Teacher classroom strategies include differentiated instruction, accommodations devised to enable access to the curriculum for students with disabilities, and annual screening and episodic progress monitoring (Fuchs & Fuchs, 2009). In Tier 1, the emphasis is on strong general education instruction to reduce the number of children at risk for reading difficulties (Al Otaiba et al., 2011; Wanzek, Roberts, Al Otaiba, & Kent, 2014).

Unfortunately, even the best designed Tier 1 programs will not prevent students from struggling (Al Otaiba & Fuchs, 2006). Students unresponsive to general education strategies alone are provided brief, supplemental intervention in Tier 2. The instruction in Tier 2 differs from Tier 1 instruction in several ways (Fuchs & Fuchs, 2009). First, Tier 2 instruction typically relies on evidence-based intervention programs which are validated through rigorous research (Fuchs & Fuchs, 2009). These programs are often scripted to ensure systematic replication by practitioners and a higher probability of treatment fidelity. The instructional procedures of these programs provide more explicit instruction through the use of more modeling and practice afforded in many general education classrooms. Second, Tier 2 interventions are primarily adult directed, in small group (3-6 students) formats, several days a week for 20 to 40 minutes (Gersten et al., 2009). A key feature of Tier 2 services is frequency of monitoring compared to Tier 1 (Fuchs & Fuchs, 2009). Students who require Tier 2 support have their progress monitored more frequently (i.e., at least once a month) to alert educators on the effectiveness of their efforts. Students who display poor growth in response to empirically validated Tier 2 intervention then begin to receive Tier 3 intervention.

Tier 3 interventions are the highest level of support students can receive in an RtI system and often fall under the umbrella of special education services (Fuchs et al., 2010). Instructional support in this tier is typically described as highly individualized (Fuchs, McMaster, Fuchs, & Al Otaiba, 2013). This is accomplished through the use of instructional materials that may not align with the student’s chronological grade curriculum, but matches the student’s current level of performance (Fuchs et al., 2010). In addition, materials and content are continually adapted directly in response to more frequent assessments of student learning. Tier 3 interventions differ from subsequent tiers because the interventions may also be provided to the student across longer durations of time including an entire school year (Denton, Fletcher, Anthony, & Francis, 2006; Wanzek & Vaughn, 2010).

Embedded within the RtI tiered system are a series of assessment and instructional decisions. These data-based decisions rely on student outcomes to make judgements regarding instructional effectiveness. Thus, decision making is directly informed by student progress. Decisions can consist of changing instructional materials, content, group size, intervention time, tier placement, and frequency of progress monitoring (NCRTI, 2010). For example, prior to beginning supplemental intervention for a small group of struggling readers, the interventionist decides that students who do not improve at a certain rate (e.g., 1 word read correctly per week) may require a more intense approach for teaching reading. This example illustrates how through the use of a set of data-based decision rules, educators can fluidly move students between tiers of increasingly intensive instruction if their rate of learning is poor and continue to display a relative low level of academic performance (Fuchs & Fuchs, 2009).

In summary, through the use of screening and progress monitoring, educators identify and monitor student learning in response to more intensive pedagogical strategies. In addition, educators use assessment results to engage in data-based decision making for the purposes of instructional (e.g., altering group size or intervention time) or program planning (i.e., tier placement or special education eligibility). Overall, the emerging body of evidence on multi-tiered service delivery systems suggests improvement on student academic outcomes. The following describes a system of measurement (i.e., CBM) that has been widely employed in RtI systems for screening and progress monitoring.

Curriculum-Based Measurement

Measurement, especially progress monitoring, represents an essential assumption within the RtI system. As previously noted, progress monitoring data provide educators with information about the effectiveness of instruction in promoting student learning. Coordinated prevention teams operating without adequate progress monitoring measures will not make accurate decisions regarding instruction, which will likely result in poor student outcomes (Margolis, 2012). Thus, RtI systems require valid and reliable instruments to formatively assess student learning in response to effective teacher pedagogy (Lembke, McMaster, & Stecker, 2010). Using curriculum-based measurement (CBM) tools within an RtI system represents a reliable and valid measurement and decision making system.

Originally conceived as a framework for special education teachers to measure the efficacy of their instructional practices over time (Deno & Mirkin, 1977), modern CBM has developed into a set of uniform procedures designed to formatively evaluate student learning in reading, spelling, and math (Tindal, 2013). These procedures are repeated across a period of time to (a) measure the rates of individual student learning, (b) alert teachers when student learning is inadequate and requires an instructional change, and (c) individualize instruction tailored to particular student needs (National Center on Intensive Intervention [NCII], 2014). For the purposes of this review, the remaining portion of this section will discuss the characteristics and uses of reading CBM (i.e., progress monitoring) within an RtI system.

Characteristics of CBM

CBM is a type of curriculum-based assessment (CBA). CBA is a formative assessment system consisting of (a) test materials that sample the curriculum, (b) repeated measurement occasions, and (c) use of student outcome data to inform instruction (NCII, 2014). Although both CBA and CBM superficially appear to share measurement qualities, Deno (2003) describes CBM as differing from CBA (e.g., informal skill inventories, chapter tests) in several key ways. First, CBM in reading demonstrates technically adequate psychometric properties for decision making. Reading

CBM displays high levels of reliability (Good & Jefferson, 1998; Marston, 1989) to provide confidence in trustworthiness of results. In regards to validity, over 30 years of research confirms that performance on CBM in reading is strongly related to performance on broader measures of reading (Reschly, Busch, Betts, Deno, & Long, 2009; Wayman, Wallace, Ticha, & Espin, 2007). In a synthesis of the literature examining the technical adequacy of CBM in reading, Wayman and colleagues (2007) found a strong relation between reading CBM procedures and comprehensive measures of reading achievement. The authors concluded CBM in reading can serve as a useful index for overall reading proficiency for students in elementary school.

In contrast, the literature suggests CBA may not possess the same level of evidence (as CBM) in terms of decision making (Fawson, Reutzel, Smith, Ludlow, & Sudweeks, 2006; Fuchs, Fuchs, & Deno, 1982; Spector, 2005). For example, Spector (2005) reviewed the reported reliability data from nine commonly used informal reading inventory manuals. A critical analysis of the manuals indicated poor confidence in assessment results. In particular, only 44% ($n = 4$) of the instruments reported reliability data. Manuals that did report reliability, employed less rigor in methodologies (i.e., poorly described sample characteristics, small sample sizes, single form of reliability reported) than commonly found in the CBM literature (e.g., Reschly et al., 2009).

Next, CBM uses standardized administration, scoring procedures. The purpose for these features is two-fold (Deno, 2003). Primarily, uniform procedures provide educators with confidence related to CBM outcomes. Secondly, reliable outcomes ensure that student data are useful for educational decision making. For example, a third-grade teacher uses CBM to measure a student’s progress in reading. Since the teacher follows the standardized procedures, the measured performance can be ascribed to the student and not a confounding variable. To further illustrate, if the teacher provided the student with extra time to take the CBM or additional guiding prompts, it would be impossible to determine whether student progress could be ascribed to increases in student learning or the unstandardized procedures. As a result of the standardized procedures, the teacher can be reasonably assured the data generated from the measurement process is trustworthy and beneficial in quantifying student growth in response to instruction.

Whereas CBM instills educators with confidence in describing student performance, comparative methods of formative assessment may not provide such assurance (e.g., Denton, Ciancio, & Fletcher, 2006). This case is illustrated through a study by Denton and colleagues on the technical properties of the *Observation Survey in Early Literacy Achievement* (Clay, 2013), a commonly used informal formative assessment in early elementary grades. Study authors indicated that test administration procedures violated standardization through the use of non-uniform prompting. For example, test administration directions allowed the examiner to unsystematically encourage and probe students to elaborate on their responses. Since examinees are provided prompting that varies depending on the characteristic of the examiner and examinee there really is no way of making comparisons among individual examinees. Moreover, the tests results will also likely generate unreliable outcomes as that it is impossible to assure constancy of scores across time or examiners if administration procedures vary with testing situations. As a consequence of this violation, Denton et al. (2006) advised educators against using informal assessment results for educational decision making, especially as evidence of whether a student adequately responded to instruction.

Furthermore, CBM serves as an educational thermometer for instructional effectiveness (Kaminski & Cummings, 2007). In other words, CBM tools are sensitive to small, significant quantities of student change in learning (Hosp, Hosp, & Howell, 2007). The result of this feature is the ability to administer CBM repeatedly across time and infer the increase in scores as the student making progress. For example, when a second-grade teacher uses CBM to monitor a student's progress in reading, only second-grade passages would be administered. This distinguishing feature is not present with other commonly used formative assessment systems in reading. For example, running records, uses materials from different text levels to measure student growth in reading. Due to this feature of the running record assessment, it is impossible to determine if and how much growth a student is making (Paris, 2002).

Finally, CBM tools measure learning *across* the curriculum; not just progress *within* the curriculum. For example, the results of a chapter test will only demonstrate whether students have learned the content specific to that chapter of the curriculum (Kaminski & Cummings, 2007). This represents an example sampling narrowly *within* curriculum measure. Conversely, since CBM samples broadly from the curriculum, the results of any one CBM indicate whether the student is making progress *across* the curriculum. Hence, CBM answers the question of whether a student is making progress toward the long-term goal of being successful in the curriculum (Kaminski & Cummings, 2007).

That being said, CBM as a global indicator, has several limitations in its usefulness. Due to the fact CBM does not sample all of the skills of an academic domain, CBM does not measure all of the subordinate skills of an academic domain (Hosp, et al., 2007). This limitation can bound the usefulness for instructional planning for all of the skills of an academic behavior

CBM Progress Monitoring and RtI Logic

The rationale of RtI is to prevent academic problems. In order to accomplish this educators should formatively evaluate student learning to prevent reading failure (Hunley & McNamara, 2010). As research demonstrates (e.g., Compton et al., 2006, 2010; Deno et al., 2009), CBM fulfills the need for progress monitoring in RtI. CBM procedures are composed of alternate tests that are relatively equivalent. This allows educators to model student learning across time (Deno et al., 2001) and assist educators in decision making (Busch & Reschly, 2007). CBM can readily detect small amounts of change over brief periods of time in student learning (Stecker, Lembke, & Foegen, 2008) to reliably use progress monitoring data to evaluate the relative effectiveness of tiered instruction on student learning (Stecker, Fuchs, & Fuchs, 2008). In addition, the use of these data provide educators with parameters of acceptable growth (Fuchs, Fuchs, Hamlett, Walze, & Germann, 1993). Most importantly, CBM progress monitoring in an RtI system provide treatment validity (Fuchs, Fuchs, & Speece, 2002). To demonstrate, progress monitoring data should provide information to the teacher about treatment effectiveness by measuring student response. Thus, CBM provides the measurement and evaluation system in RtI (Tindal, 2013).

CBM in an RtI System of Educational Decision Making

Initially imagined as a decision making framework for evaluating the effectiveness of special education instruction (Deno & Mirkin, 1977), the present day CBM Problem Solving Model (Shinn, 2008) is a four phase model for addressing student academic concerns through an iterative process of problem solving. In the first phase, problem identification, actual student performance is compared to expected student performance. The quantitative difference between actual and expected student achievement is defined as the problem (Shinn, 2008). For example, if Student A reads 20 words correct and the grade level expectation is 50 words read aloud, the problem for Student A is a difference of 30 words read aloud. This phase of the model is similar to the universal screening stage in RtI systems where struggling students are identified.

Next, in the problem certification phase, student problems are compared to established benchmarks of achievement. Students with performance substantially different from their peers are marked for more intensive instruction. For example, Student A's problem of 30 words represents a large divergence from other students in the grade-level. Thus, Student A is at-risk for continued below level performance unless an instructional change is made.

Then, the exploring solutions phase of the model is initiated. During this phase, instructional program options are investigated for the student. Monitoring instruments (i.e., CBM) are implemented to measure the student's rate of learning in response across time. Finally, at the end of a decided upon period of time, the effectiveness of the instruction for the student is determined to be adequate or needs to be changed. This is the evaluating solutions phase of the model. Educators can also use evidence of student progress to determine whether more intensive instruction, such as delivered by special education, is appropriate. To illustrate, in response to Student A’s data, the teacher decides to implement a small group tutoring program to improve reading proficiency. The teacher decides to implement this instruction for 10 weeks with monitoring using reading CBM. After 10 weeks, the teacher reviews the student data and decides to discontinue the intervention because Student A has made unsatisfactory progress toward the goal. Instead, the teacher has decided that a more intense adaptation (e.g., more instructional time) is required to accelerate Student A’s learning. This stage of the model is closely aligned with the progress monitoring, and data-based decision making in RtI.

There are definitive conceptual differences between CBM and CBA measurement systems. Clearly, CBM fits seamlessly and parsimoniously in an RtI system as is evidenced by the applied example demonstrating how educators use CBM to problem solve individual student academic difficulties. The following describes the extension of CBM research into the measurement of early literacy skills.

Dynamic Indicators of Basic Early Literacy Skills (DIBELS)

The ensuing section will describe (a) the rational and practice of formative assessment in early literacy skills, (b) the relationship between CBM and DIBELS as systems of measurement and evaluation, and (c) how an early literacy outcomes-driven model of prevention converges with RtI.

Rationale

As mentioned earlier, the teaching of reading skills to an adequate level of literacy represents a social imperative (Deno, 1989). Students who do not read with comprehension often display poor performance in high school content area classes

(O’Connor, Beach, Sanchez, Bocian, & Flynn, 2015) and may struggle securing meaningful employment due to present-day occupations continuing need for higher levels of literacy (e.g., Snow et al., 1998). To monitor reading proficiency, educators need a measurement and decision making system, such as CBM, for identifying and resolving student reading problems. Although CBM represents an effective approach, the methodology is primarily limited to reading aged children in Grades 2 and beyond. This presents a problem because early literacy skills taught in beginning elementary (i.e., Kindergartern-1st grade) establish the foundation for later reading skills. For example, identifying and manipulating sounds in oral language (i.e., phonological awareness) and letter-sound knowledge (i.e., alphabetic knowledge) are highly predictive of future reading proficiency (National Early Literacy Panel, 2008; O’Connor & Jenkins, 1999). Measurement of skills such as these is important because students who struggle to learn how to read typically display deficits in phonological awareness and alphabetic knowledge (Al Otaiba & Fuchs, 2002; 2006). Therefore, a system of measuring student learning in early literacy skills is vital to identify and provide intervention addressing emerging reading problems (Coyne & Harn, 2006; McAlenney & Coyne, 2011).

Dynamic Basic Early Literacy Skills (DIBELS; Kaminski & Good III, 2002) is a collection of fluency-based subtests used to formatively measure student attainment of early literacy skills in kindergarten through sixth-grade. The subtests measure a set of key early literacy skills, such as phonological awareness, alphabetic principle, and reading fluency which have been identified as necessary for later reading success (Lane, 2014; NRP, 2000; Snow et al., 1998). The early research on DIBELS established the measures and task materials (Good & Kaminiski, 1996; Kaminiski & Good, 1996) to be consistent with research on early literacy skills. The resulting subtests include Letter Naming Fluency (LNF), First Sound Fluency (FSF), Phoneme Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), and Oral Reading Fluency (ORF).

The NWF task is a subtest comprised of vowel-consonant (VC) and consonant-vowel-consonant (CVC) pseudo-words (e.g., ib, nat) used as a measure of a student's knowledge of the alphabetic principle and simple phonics skills. The alphabetic principle is the awareness that words are composed of letters and letter sounds (Joseph, 2015). According to Ehri (2005), children acquire the alphabetic principle through five overlapping phases. In the initial phase, the Pre-alphabetic phase, children are essentially non-readers because of lack of awareness of alphabetic principle (Ehri, 2005). Children read words utilizing visual features of letters and picture cues. For example, a child says *McDonald's* after viewing a McDonald's store sign.

Next, in the partial-alphabetic phase, children begin to learn the names and sounds of some letters which they use to read words. Typically, initial and final sounds are learned first due to ease of acquisition (Ehri, 2005). For instance, a child reads *cut* when presented the word cat. As the example illustrates, children in this phase make errors with words that have similar features (i.e., beginning and final sounds) and have difficulty decoding unfamiliar words.

Then, in the full-alphabetic phase, children become familiar and use most letter-sound correspondences to read words. Decoding unfamiliar words and remembering sight words are characteristic of this phase of development. Word reading is often slow through the use of one-by-one (e.g., /*b*/ /*i*/ /*g*/) letter-sound correspondences (Ehri, 2005). After children have mastered all of the letter-sound associations, they enter the consolidated-alphabetic phase of development. In this phase, children recognize words that have similar spelling and sound patterns (Ehri, 2005). For instance, children can read the word *mat* because they are familiar with the words *cat* and *bat*. The consolidated-alphabetic phase is characterized by children reading a larger number of new words and doing so more fluently (Ehri, 2005). Lastly, children enter the automatic-alphabetic phase of word reading. In this phase, children automatically identify common word patterns and larger units of words. Ehri (2005) describes members of this phase as having developed unitization because now they can now read words as whole units (e.g., *cat*); whereas children in lesser phases typically read words using word parts (e.g., /*c*/ /*at*/) and individual letter-sounds (e.g., /*c*/ /*a*/ /*t*/).

The ability to automatically recognize letter-sound correspondences is a prerequisite skill to efficient word reading (Snow et al., 1998). The alphabetic principle has been identified as a key early literacy skill (Lane, 2014; NRP, 2000). Students who struggle automatically identifying letter-sound associations very likely will struggle with word recognition (Lane, Pullen, Hudson, & Konold, 2009; Vellutino, Fletcher, Snowling, & Scanlon, 2004). NWF, as a formative assessment of basic phonics skills, is sensitive to instruction on the alphabetic principle (Good et al., 2009). Inadequate progress as measured by NWF alerts teachers when instruction is not effective and indicates a change needs to be made before wasting too much time (Good III, Kaminski, Fien, Powell-Smith, & Cummings, 2012). It is important to progress monitor the effects of early phonics instruction because students at-risk for reading difficulties in the fall of first-grade who make strong gains in phonics, make additional gains in word reading, reading fluency of connected text, and reading comprehension in response to early, supplemental instruction (O'Connor, Harty, & Fulmer, 2005).

CBM and DIBELS

CBM by design shares many features with DIBELS. Since CBM measure student performance in Grades 2 and beyond, DIBELS represents an expansion of CBM procedures into early elementary grades (Kindergarten and first grade). Each approach are strong indicators of global student performance (Cummings & Kaminski, 2007). Both exemplify acceptable levels of reliability and validity on high-stakes tests of accountability (Good et al., 2001). Most importantly, both measure basic skills and are sensitive to small increments of important student growth (Shinn & Bamonoto, 1998). This provides educators with confidence that changes in raw score should be associated with general progress (Kaminski & Cummings, 2008). Consequently, DIBELS can be categorized as an early literacy CBM.

Although both measurement tools are academic performance indicators, Kaminski and Good III (1998) articulated several important differences between CBM and DIBELS. One, the measurement and decision making models of CBM and DIBELS diverge in terms of system purpose. The rationale for CBM is based on a framework of remediating academic problems that deviate significantly from typically developing peers (Kaminski & Good, 1998). Conversely, DIBELS was conceptualized with the explicit objective of preventing reading difficulties. The underlying principle behind this logic is that it is far more expedient, with respect to time and personnel, for educators to avert reading problems rather than correcting later (Lembke et al., 2010). This feature of the DIBELS logic is strongly supported by the literature on early identification and intervention (Foorman et al., 1998; O'Connor et al.,, 2005; O'Connor & Klingner, 2010; Scanlon, Vellutino, Scanlon, Small, & Fanuele, 2006; Vellutino, Small, Fanuele, & Sweeney, 2005).

Second, the degree of risk related to educational decision making is much higher for CBM compared to DIBELS (Kaminski & Good III, 1998). CBM is typically used to progress monitor students who have established significant reading problems. This could be considered a high-stakes decision because these students require a measurement system possessing a high level of reliability and validity to determine whether the instruction presented leads to improvement. The rationale is time cannot be wasted through the provision of ineffective instruction for students with significant and persistent learning needs. These students are already well behind their typically developing peers in reading and the more time they waste receiving instruction not responsive to their needs will result in falling hopelessly further behind. On the other hand, the rationale for DIBELS is to prevent reading problems by formatively assessing the effectiveness of early literacy instruction. Since the technical properties of DIBELS are relatively weaker compared to CBM, some students might incorrectly be targeted for additional instruction. This is a low-stakes decision because if some students are erroneously provided supplementary early literacy instruction that was not necessary, the relative effects on the students and their families will not be deleterious (Kaminski & Good III, 1998).

Third, the duration of skill measurement for CBM contrasts from DIBELS. Both CBM and DIBELS are indicators of proficiency on important literacy skills. This measurement feature allows educators to reasonable predict sometime into the future whether students will be struggling readers. CBM can be used as a long-term outcome measure because it measures skills across years. DIBELS measures important skills that are more transitory (Kaminski & Good III, 1998). For example, while oral reading fluency is a good predictor of general reading proficiency (Wayman et al., 2007), the relation between NWF performance and across grade reading ability is limited (Reidel, 2007).

Lastly, the nature of skills measured by CBM and DIBELS differs. CBM directly measures important reading skills. DIBELS measures skills that are prerequisites of early reading skills; however, the skills in and of themselves are not enough for reading competency (Kaminski & Good III, 1998). Good et al. (2012) provided an appropriate example using NWF. The authors stated that while reading pseudo-words is an important indicator of basic phonics skills, the goal of reading is not limited to reading simple words. The overall objective of reading is comprehension. Therefore, low performance on NWF should not be the focus of teacher instruction. Instead, a component of effective instruction should be based on general skills in phonics to accelerate the learning of the student. The reason is high rates of word reading fluency and accuracy are necessary for adequate levels of reading proficiency (Pikulski & Chard, 2005).

Outcomes-Driven Model and RtI

As has been shown, the basis and purpose of DIBELS as a downward extension of CBM to avoid reading problems is necessary. However, an adequate measurement system also requires a coherent framework to improve student outcomes. The Outcomes-Driven Model is a prevention-oriented framework aligned with the RtI logic (Kaminski, Cummings, Powell-Smith, & Good III, 2008). The model extends on the work of using CBM in a problem-solving model (Deno, 1989; Shinn 2008) into the earlier grades. The model also builds on an original DIBELS decision making framework which relied upon the problem-solving model (Kaminski & Good, 1998). The Outcomes Driven Model diverges from CBM because problem-solving with this approach is associated with significant learning problems and evaluating special education instruction (Shinn, 2008). Instead, the Outcomes-Driven Model rationale is to avoid reading problems through a series of iterative phases (Cummings, Atkins, Allison, & Cole, 2008). Both the Outcomes Driven Model and CBM problem solving model ask the same questions related to student performance: (a) what is the problem? (b) why is the problem occurring? (c) what are potential solutions? and (d) did the solution reduce the problem (Shinn, 2008).

There are four phases in the Outcomes Driven Model: (a) identify need for support, (b) validate need, (c) plan and implement support, and (d) evaluate and modify support. The initial phase of the Outcomes-Driven Model is to identify need for support. The purpose of this phase is to determine which students will need of some form of intervention (Kaminski et al., 2008). To accomplish this objective, all students are screened at least three times per school year. The function of the multiple assessment waves is two-fold. First, it prevents student problems from going undetected for prolonged periods of time. Second, it provides an opportunity to measure all students' growth. Students scoring below the DIBELS benchmarks when assessed in this initial phase could be targeted for further problem-solving.

The DIBELS benchmarking system was developed to provide goals against which educators may evaluate the effectiveness of instruction (Kaminski & Cummings, 2008). Student scores are grouped by likelihood of future reading difficulty. Students scoring in the low-risk category have a high probability of reaching future benchmarks and reading success without additional support. Students scoring in the at-risk category have an extremely high chance of not attaining early literacy skills to adequate level to avoid future reading problems. Students in the at-risk category will require intensive intervention strategies to rapidly accelerate their learning. Finally, students scoring in the some-risk category are at an elevated level of danger for poor future reading outcomes. These students will require additional support to improve; however it will not be at the same level (time, interventionist expertise, duration) of students who performed in the at-risk group.

In the validate need phase, educators determine why the student did poorly ona benchmark screening measure. Common reasons, other than early literacy skill deficits, for poor performance include inadequate testing environment (e.g., noisy room) and student behavior (e.g., sick, misunderstood task directions) (Kaminski et al., 2008). Since young children's performance can be highly variable on these tasks, Kaminski and colleagues (2008) recommend short-term formative assessment to validate need for support. In addition, the students' teacher can review student response patterns on several DIBELS tasks. This way individual student behavior may be indicative of where a student is in the skill and provide ideas for instructional strategies (Kaminski & Good III, 2011).

In the plan and implement support phase, educators begin to develop an instructional plan to support student learning. First, a clear learning goal and timeline are established. In the case of a fictional student, John, his teacher decided he should produce 30 CLS by the middle of October. Next, educators must select the early literacy instructional strategies to implement. For example, John's teacher decided to provide

John additional fluency practice of letter sounds. This information should be provided from the validate need phase by means of reviewing a DIBELS probe and targeting specific skills the student is lacking. To illustrate, John's performance on a NWF probe demonstrated that he knew all of the letter-sounds, but he did not recognize them automatically. Therefore, his teacher choose to increase his fluency of letter-sounds to facilitate his word reading skills. Then, a discussion regarding the intensity of support will be necessary for the student. For example, John, who performed in the at-risk category of the benchmarking system will likely require more intensive support (time, instructor expertise) than Bobby who is slightly below benchmark. In the example here, the decision is made to provide John additional support in a small group (five students) of 30 minutes daily from the school reading specialist. Finally, a decision is made in terms of how to measure and monitor student progress. This choice is assisted by student needs and instructional strategy. For example, since John demonstrates an adequate level of phonemic awareness and he is receiving instruction of the alphabetic principle, his teacher decided to monitor his progress using NWF.

Although there are a number of research supported instructional procedures to improve the rate of learning for most children, individual student response to research-based instruction varies dramatically (Deno, 1990). Torgesen (2000) estimates that 2 – 6% of the general population will be generally unresponsive to research-based literacy curricula. This problem is especially magnified for the above mentioned subgroup of struggling students who fail to attain early literacy skills and continue to struggle with acquiring adequate later reading skills. As the literature indicates, the need for a progress monitoring system is vital to ensure students are learning and mastering skills taught through early literacy instructional efforts at every level (Good, Simmons, & Smith, 1998; Kaminski & Cummings, 2008). Thus, in the evaluation and modify support phase, educators review the outcomes of the instructional plan. This includes analyzing student learning across time and asking (a) does the student data trend up? (b) were the efforts effective? and (c) if not, what instructional variables (i.e., time, group size, materials) need to be adapted. In the case of John, he did not demonstrate an increase to his stated goal of 30 CLS by October. In response to John's outcome data, his teacher concluded that John needs a more intensive intervention than previous thought. John's teacher decided to decrease the intervention group size from five students to three. The rationale for this decision is to increase the amount of opportunities to respond and be provided immediate feedback.

Several important themes emerge from the literature. First, the acquisition of decoding skills represents a series of overlapping phases. These phases are evanescent and represent important milestones for typically developing readers. Second, the measurement of these early literacy skills need to be ongoing and precise to confirm learning trajectories trend upward. In addition, the accurate measurement in early literacy skills, such as DIBELS, verifies the effectiveness of beginning reading instruction to prevent most long-term reading problems. Finally, a practical example demonstrated how DIBELS, similar to CBM, can be used within an outcomes-drive model to problem solve early literacy student troubles. The phases of the model recurrently student information to plan and implement effective instruction. The closing section describes how specific components of student data (i.e., feedback) provides teachers with information that is central to the improvement of student early literacy achievement.

Formative Evaluation Feedback and Student Achievement

The concluding section will include (a) a definition of feedback, (b) a review of the literature on progress monitoring feedback and student achievement in reading, and (c) a review of the literature on formative assessment feedback and student achievement in early literacy.

Feedback

Hattie and Timberley (2007) define *feedback* as the performance data (i.e., student responses) on a task which generate information about one's learning. Feedback includes data regarding a student's current level of learning. These data come in the form of asking questions or giving an assessment (Hattie & Timberley, 2007). In the present study, *feedback* was therefore defined in terms of the range or pattern of answers students provide to formative assessments (e.g., test items).

Feedback provides insight for the teacher to evaluate any gaps between present student learning and expectation for student learning. The feedback should provide information on how to enable a student to reach future learning objectives. Therefore, feedback is important for enhancing student learning because it provides teachers with a roadmap for tailoring instruction to meet student needs.

Research on feedback has continually demonstrated its effectiveness on student performance. In a synthesis of 900 meta-analyses, Hattie (2009) found feedback to be a powerful teacher influence on student achievement with a large (.73) effect size. For example, the use of formative assessments to generate feedback about instruction can provide teachers with important information, such as who has and has not learned the content well, to improve student learning (Hattie, 2012).

Components of Effective Feedback

Chan et al. (2014) identified three key characteristics of effective feedback. First, feedback should be delivered *immediately*. The temporal feature of effective feedback is that it has a short life-span (Shute, 2008). Since individuals are dynamic and constantly changing, feedback that is out-of-date will not be pertinent to present levels of learning. For instance, if a student takes a test in the fall and the teacher does not receive the results until several months later, the information will not support the student's learning because the student's learning has potentially changed dramatically since that time period. Thus, the delayed feedback missed a critical window of opportunity to forward the student's learning.

Second, feedback should be *specific* (Hattie & Timperley, 2007). For example, a first-grade teacher wants to know if the reading instruction he provides is improving student decoding skills. The teacher decides to administer an assessment measure that directly assesses student word decoding skills. The information from this assessment measure should inform the teacher if student learning is adequate because feedback is specific to word decoding skills.

Finally, feedback must have practical *instructional value*. The practicality refers to how the teacher can use the feedback to move the student along in their learning (Wiliam, 2006). Hattie (2012) describes this process as a continuum of three phases geared toward directing a learner toward an objective. First, feedback should provide a clear learning target for the student. Next, feedback should evidence student progress toward the goal. Finally, effective feedback informs teacher pedagogy to reach the goal. Using the example of the first grade teacher above, the results of the word decoding assessment indicated that most students could blend word parts (e.g., /c/ /at/) to read words and only a few still struggled with recognizing letter-sound associations (e.g., /c/ /t/ /a/) . Since the teacher had recently provided instruction on blending word parts (instructional goal), it was clear the instruction was fairly effective because the majority of the students mastered the content; although the assessment showed (evidence toward goal) a handful of students struggled to meet the teacher's expectations. Consequently, the teacher decided to provide the struggling students with additional support to improve their word decoding using instructional strategies (teacher pedagogy) designed to build on their current level of performance.

Feedback as a Part of Progress Monitoring

Feedback is an important factor that enhances student achievement. Progress monitoring, using CBM or DIBELS, generates all of the features of effective feedback. Measures such as CBM or DIBELS are brief and easy to administer and interpret. Therefore the feedback to educators is available relatively quickly. Together each directly measure the effects of curriculum or instruction implementation. This provides specific feedback to the teacher regarding instructional effectiveness with respect to student learning. Finally, the feedback produced by progress monitoring supplies practical information to the teacher. Since CBM and DIBELS can be administered repeatedly across a period a time, teachers can formatively assess their students' progress and make instructional changes when needed. In addition, student response patterns on formative assessments can direct the teacher's attention to student instructional needs (Capizzi & Fuchs, 2005). For example, a student who performs poorly on NWF because of individually sounding out each letter-sound (e.g., /c/ /a/ /t/) may require additional instruction on blending sounds together to fluently read words (e.g., cat).

Feedback through on-going and targeted formative assessment fills in the information in decision making frameworks such as the Outcomes-Driven Model and RtI model. An essential component of each model is progress monitoring to answer questions about student achievement. These questions include (a) what goals do students need to reach? (Good et al., 2001), (b) what is the rate of student learning? (Good et al., 2009) and (c) what pedagogy will lead to enhanced student learning? (Kaminski & Cummings, 2008). Formative assessment feedback assists educators in making instructional decisions that define the RtI model.

The three key characteristics of effective teacher feedback included the immediacy, content, and instructional significance of feedback. Through these three features educators quickly gauge expected and actual student learning of skills/content. Then, teachers plan and implement targeted instruction geared to student academic needs. This is an ongoing and recursive process which continues until an effective teaching strategy is discovered and student learning progresses toward performance expectations. CBM or DIBELS progress monitoring provides this level of teacher feedback to educators. Thus, teachers use CBM or DIBELS formatively assess students reading skills, compare current and expected performance, plan, and then implement instruction. Discussion of results and findings of research on formative assessment on student achievement in reading follows.

Research on Progress Monitoring Feedback on Student Achievement in Reading

The primary function of formative assessments is to facilitate data-based educational decision making. This includes direct, on-going measurement of broad academic outcomes and feedback alerting teachers regarding the effectiveness of their pedagogy (Deno, 1985, 2003). Implicit in these procedures is the feedback generated from CBM should enhance instructional planning as well as improve student achievement (Fuchs & Deno, 1994). The nature of the feedback is both quantitative and qualitative. The quantitative element provides an observable and measureable indicator of student performance, whereas qualitative information refers to a descriptive analysis of student responses. Therefore, research on CBM procedures should consist of two components, including (a) measurement and (b) instructional utility.

To address issues related to research on CBM, L. S. Fuchs (2004) proposed a three phase research agenda for validating CBM procedures. Phase One research consisted of demonstrating a reliable and valid CBM performance level score to predict important future academic outcomes. In Phase Two, research would focus on establishing whether increases in a student’s score over time (i.e., slope or growth) are related to broad change for important academic outcomes. These phases explicitly center on of the measurement system of CBM. Finally, Phase Three studies should address whether the use of CBM results enhanced teacher pedagogy and corresponding growth in student achievement. The function of this stage of research is to validate the instructional utility of CBM procedures for teachers and their students.

Stecker, Fuchs, and Fuchs (2005) synthesized the literature to examine the effects of Phase Three CBM research as it relates to student achievement. The authors reviewed studies from 1981 to 2000 pertaining to the use of CBM procedures in reading, math and spelling for students with and without disabilities in grades two through nine. Results indicated that teachers' use of CBM lead to improved academic achievement for middle to late elementary-aged students. It is, however, important to note that the authors' concluded the use of the CBM measurement system *alone* did not dramatically lead to increases in student academic performance. If teachers are to produce increases in student achievement, the use of program evaluation components (i.e., visual analysis of graphed student data and decision rules), skills analysis (i.e., content) and instructional recommendations (i.e., pedagogy) must be implemented. These findings, when placed within the context of feedback for informing instruction, suggest that multiple forms of CBM feedback are essential for developing meaningful instruction in order to bring about positive academic performance in students. This is especially salient when considering that feedback from the assessment-teaching loop is an important component in planning effective instruction (Chan, et al., 2014)

The question then remains: what barriers prevent the resultant data from assisting planning, if teachers know how to implement the measurement and evaluative system? Research on the use of CBM procedures in teacher decision making and planning indicates it is function of perceived lack of utility (Arthaud, Vasa, & Steckelberg, 2000; Wesson, King, & Deno, 1984; Yell, Deno, & Marston, 1992). Yell and colleagues (1992) reported that most teachers do not use CBM data to assist in decision making because they do not find the data useful. In a more recent investigation, Arthaud et al. (2000) documented similar findings, wherein teachers reported not using CBM to assess reading and further reported the results as not useful for planning. These findings suggest an expanded role of CBM in teacher planning should occur only if practitioners can see the relevance and practicality of these procedures.

To address the issues of perceived efficacy, Phase Three research on CBM has progressed through a series of overlapping stages. These stages included (a) using quantitative visual analysis of student data to make educational decisions, (b) using skills analysis of student response patterns to plan instruction, and (c) explicitly suggesting instructional modifications to enhance instructional planning (Fuchs, 1998).

Quantitative Visual Analysis of Student Data

The following section includes studies attending to methodology that consisted of graphing student data and using visual analysis/data-based decision rules to alter instruction. All of the studies explicitly or implicitly measured student achievement as the dependent variable. Independent variables consisted of CBM procedures for making educational decisions (i.e., graph displaying a student scores over time, data decision rules for identifying when a change to a student's instructional program is needed) or no CBM.

Wesson, Skiba, Sevcik, King, and Deno (1984) investigated the effects of low versus high levels of CBM implementation on student reading achievement. Thirty-one special education resource teachers and 117 of their students between first- and seventh-grade participated in the study. Teachers received training in CBM procedures which consisted of using the CBM measurement system, writing student goals and objectives, graphing and using a data decision rule to cue when to make a program change. Students were assessed on the number of words read correctly in one minute on three 3rd-grade reading passages weekly. At the end of the study, the authors' rank-ordered students by observed level of CBM implementation fidelity of the teachers. To examine the effects of varying levels of implementation fidelity, the total group was split into the top 27% and bottom 27% of correct implementation and used in the final analysis. Results indicated teachers who maintained a high level of CBM implementation fidelity had students who scored higher in reading fluency. Wesson and colleagues concluded that ongoing CBM and using data for decision making may produce higher student academic achievement. The authors cautioned that teachers will need on-going technical assistance and preparation to make sure CBM data inform instruction.

Fuchs and Fuchs (1986) conducted a meta-analysis to investigate the impact of CBM procedures on student achievement. The authors reviewed 21 studies consisting primarily of students with learning disabilities, representing 98% of all participants. Study features included (a) on-going CBM data collection in reading, writing, spelling, and/or math; (b) data collection occurred at least twice weekly; and (c) teacher's use of a data-decision making rules to conclude whether a pedagogical change was needed. Results of the meta-analysis indicated the use of systematic and repeated measures of student performance resulted in higher student achievement with an unadjusted large effect size of .70. The authors concluded the use of formative evaluation to monitor the progress of students with mild disabilities may lead to increases in student achievement. In addition, the use of decision rules and graphing of data may result in further increases in academic achievement.

In a similar study to Wesson et al. (1984), Jones and Kruse (1988) examined the effects of training student teachers to use CBM procedures on student achievement in reading. Twenty-six special education student teachers and 21 of their students in grades

3 through 6 took part in the study. Student teachers were randomly assigned to either a data-based experimental or control group. In the data-based experimental group, teachers collected and graphed student reading data (i.e., words read correctly) for at least 8 weeks and made decisions regarding whether to change instruction based on data. Teachers in the control condition did not use systematic data collection nor data-based decision rules to inform instructional changes. At the end of the study, teachers from both groups completed questionnaires describing what information they used for monitoring student progress. Experimental group teachers completed an additional questionnaire regarding their perception of the data-based decision making procedures. Participating students were assessed using the Passage Reading Test (Fuchs et al., 1982a) and a researcher adapted test of literal reading comprehension. Results of a repeated one-way analysis of variance (ANOVA) revealed the students of experimental group teachers displayed significantly higher posttest scores in reading fluency and reading comprehension than the students in the control group. Experimental group teachers reported data-based evaluation procedures informed instruction, facilitated progress monitoring, and easily communicated results

In their seminal study, L. S. Fuchs, Deno, and Mirkin (1984) examined the effects of CBM procedures versus traditional methods of evaluation (e.g., chapter tests) on instructional planning and student achievement in reading. The study consisted of 39 special education teachers and 121 of their students (teacher selected). Participating teachers were randomly assigned to either CBM or no CBM. In the CBM group, teachers developed student goals in reading, assessed students using CBM at least twice weekly, graphed student data (i.e., words read correct in one minute), and used a data-decision rule that required them to make an instructional change as necessary. Teachers in the no CBM or control group developed student IEP goals and monitored student progress using preferred measures (e.g., teacher constructed and chapter tests). Students were assessed at least weekly using reading fluency CBM over 18 weeks. Teachers in both conditions completed a questionnaire, which reported teacher perception of student attainment of goals, changes in reading goals, and a present level of functioning in reading. Results of the study showed that students whose teachers used CBM procedures displayed higher rates of growth in reading fluency, decoding, and comprehension compared to students in the control group. With respect to instructional practices, CBM teachers changed goals more frequently and were more knowledgeable regarding student goals and whether the goals were reasonable for students to achieve.

Using procedures outlined by Fuchs et al. (1984), L. S. Fuchs, Fuchs, and Hamlet (1989a) examined the effects of different goal configurations using CBM in math on student achievement. Where prior studies relied on teacher analysis of visual data, this study was the first to use computer-based decision making software to cue teachers when an instructional modification was warranted. Thirty special education teachers and pairs of their students (*N* = 60) between grades 2 and 9 participated in the study. Teachers were randomly assigned to one of three conditions for 15 weeks. In the varying goal group, teachers changed student CBM goals and teacher pedagogy in response to student data. Teachers in the fixed goal condition used data-based decision rules regarding instructional changes based on the initial goal. In the no CBM or control group, teachers did not use CBM to create goals. Students were assessed using the Math Computation Test, a math CBM and Concepts of Number subtest from the Stanford Achievement Test

(Gardner, Rudman, Karlsen, & Merwin, 1982). Both groups of CBM teachers received computer-based reports that notified teachers when it was time to make a change based on student data. An analysis of covariance (ANCOVA) on the Math Computation Test demonstrated significantly higher achievement (effect size = .52) in math for students in both groups of CBM than those in the control group. With respect to occurrence of instructional planning, varying goal teachers increased goals more frequently. The authors concluded that changing instructional programming in response to student CBM data, when the amount of curriculum content is held constant, would likely result in increased learning.

In a follow-up study, Fuchs, Fuchs, and Hamlet, (1989b) compared the effects of providing teachers with a prompt to change instruction versus only providing them a graph of student progress. Study participants included 29 special education teachers and pairs of their students (N = 53) with disabilities in second- through ninth -grades. Teachers were randomly assigned to CBM and no CBM conditions. After 15 weeks of CBM data collection, the CBM group was further split into two groups: CBM plus evaluation (CBM + E) and CBM only (CBM). Teachers in the CBM group received a graph visually depicting student data, whereas the CBM + E group received a graph and notifications when an instructional program change should be initiated. Students were assessed using the reading comprehension subtest of the Stanford Achievement Test (Gardner et al., 1982) and two CBMs (i.e., number of words correctly recalled from a passage [recall] and number of words correctly supplied for a sentence [cloze]). ANOVA indicated students in the CBM + E group demonstrated a significantly higher rate of reading achievement compared to the CBM group. The authors summarized the study findings by stating that CBM data collection and evaluation procedures should produce improved teacher pedagogy and higher rates of learning for students with disabilities. Fuchs and colleagues followed this conclusion with a warning that general education teachers would need on-going technical support from special educators to sustain CBM in the general education setting.

In summary, teachers who implemented direct and repeated measures of student performance, who displayed high levels of implementation fidelity, and who applied data-based decision rules for instructional programming produced higher levels of student achievement. A related finding was teachers will need on-going technical assistance and training to effectively implement all components of CBM. However, it is important to note the outcomes from this line of research are restricted to students in grades two through eight.

Skills Profile of Student Response Patterns

As a consequence of early exploration of CBM, researchers were able to construct a set of procedures for measuring and evaluating program effectiveness. Yet, as mentioned earlier, teachers reported not being able to use the resulting data to inform instruction (Wesson et al, 1984). To resolve this problem, a program of research was initiated to determine whether CBM could provide qualitative details to the teacher about individual student skill profiles. This phase of research focused on finding *what* to teach (Fuchs, 1998). *What* refers to the content or skills (i.e., spelling word patterns, math facts, parts of a story in reading) of an instructional domain for struggling students. Therefore, reviews of the following studies will address findings of research related to the independent variable of graphed student data *and* skills analysis with student achievement and/or instructional planning serving as dependent variables

Fuchs, Fuchs, and Hamlett (1989c) contrasted two forms of feedback on teacher pedagogy and student achievement. Twenty-two special education teachers and 41 of their students in grades 3 through 9 were randomly assigned to one of two conditions: CBM graphed data only (CBM only) or graphed data plus skills analysis (CBM + SA). In the CBM only condition, teachers received graphs visually illustrating student progress across time on total number of content words retold. Teachers in the CBM + SA condition were provided graphed student data along with a specific analysis of student responses. Students were assessed using the Standford Achievement Test reading comprehension subtest (Gardner et al., 1982), the Passage Reading Test (Fuchs et al., 1982a) measuring reading fluency, and a written retell of a read aloud passage. Results indicated teachers in the CBM + SA condition identified higher numbers of story elements (i.e., characters, setting, problem, events, ending) for instruction, changed instruction more frequently, and produced higher levels of student achievement in reading. The authors concluded CBM procedures offers teachers diagnostic feedback on student errors to assist in planning for instruction. The authors cautioned, however, that teachers need to modify their pedagogy in response to CBM feedback to generate increased student achievement.

In a follow-up study, Fuchs, Fuchs, Hamlett, and Stecker (1990) compared several types of teacher feedback on instructional planning and student achievement in math. The study consisted of 30 special education teachers and 91 of their students in grades three through nine. Participants were randomly assigned to one of three conditions: CBM only, CBM + Skills Analysis (CBM + SA), and no CBM. Teachers in the CBM only group received a graph showing student progress over time and prompting when an instructional change was needed. In the CBM + SA group, teachers received the student graphs and instructional change prompts, but also received feedback describing specific levels of mastery in math objectives. Students in the CBM groups were assessed weekly over 15 weeks using a math CBM test. The number of goals, instructional changes, and specific math skills cited was recorded across the length of the treatment. Results indicated that CBM + SA teachers enacted more specific instructional modifications and increased student achievement at higher levels than CBM only or no CBM.

In an attempt to extend the findings of Fuchs et al. (1989c; 1990) to spelling, Fuchs, Hamlett, and Allinder (1991b) compared the effects of several types of teacher feedback on pedagogy and student achievement. The study included 30 special education teachers and 90 of their students in grades three through nine. Teachers were randomly assigned to one of three conditions: CBM Skills Analysis (CBM + SA), CBM and no CBM. Teachers in the CBM + SA group received weekly graphical displays of correct letter sequences across time and specific spelling student errors. In the CBM group, teachers were provided graphs visually depicting correct letter sequences over time. Teachers were provided bi-weekly instructional consultation regarding program changes from research assistants. Students were assessed twice weekly using the CBM spelling tests over 15 weeks. Teachers documented the number of goal changes and instructional modifications implemented and specific spelling skills mentioned. Results of the study indicated both CBM groups made a similar number of instructional and goal changes.

Teachers in the CBM + SA group referenced a higher number of specific spelling skills for intervention and their students displayed higher levels of achievement in spelling compared to the CBM and control groups. The authors concluded the act of examining student responses leads to enhanced instructional planning and improved student achievement.

More recently, Stecker and Fuchs (2000) examined the effects of individualized progress monitoring on student achievement in math. Whereas prior studies examined groups of students' responses to monitoring data across time, Stecker and Fuchs matched pairs of students in one of two conditions: instruction based on a target student's (treatment) data and instruction based on matched peer's data. Twenty-two special education teachers and pairs of their students (N = 42) in grades two through eight participated in the study. All Students were administered a math CBM over 22 weeks to monitor their progress, and a pretest and a posttest using a broad achievement test of math to evaluate increases in student performance. For CBM treatment students, instruction was informed by computer-generated graphs displaying student progress, a monthly prompt to teachers when a pedagogical change was needed based on rate of student learning, and a bi-weekly skills analysis profile similar to that in Fuchs et al. (1990) study. Results indicated that CBM treatment students and matched peers both displayed increases in math CBM scores; however, CBM treatment students improved significantly on a test of overall math achievement compared to matched peers. The authors concluded that when teachers used CBM data that were specific to student needs, students reached higher levels of achievement than students whose instruction was planned based on data from another student.

According to the above review of literature addressing skills analysis of student response patterns, teachers who used skills analysis significantly demonstrated improved awareness in identifying individual student skill deficits. Compared to participants in control groups, CBM and skills analysis teachers also changed instruction more frequently and the resulting instructional changes focused on specific skill deficits of the students. Additionally, students of CBM skills analysis teachers displayed increased levels of student achievement.

Still, these results are limited to CBM in reading, spelling, and math for middle to later elementary-students. In regards to early literacy CBM, the literature is limited in the number of studies directly examining the effects of skills analysis on teacher instructional planning and student achievement. The existing studies that attempt to test these assumptions do so implicitly. For example, several recent studies (i.e., Cummings, Dewey, Latimer, & Good, 2011; Harn, Stoolmiller, Chard, 2008) examined the predictive utility of student growth in decoding using the DIBELS Nonsense Word Fluency subtest in Fall of first grade on the growth of oral reading fluency in Spring of first grade. However, this line of inquiry is consistent with Phase Two research (i.e., measurement) and failed to directly evaluate the instructional utility (i.e., enhanced teacher pedagogy and improved student achievement) of the measure. Subsequently, the next time period of CBM research would focus on more efficiently connecting student skill profiles to instructional suggestions.

Diagnostic Feedback for Instructional Recommendations

Prior studies (e.g., Fuchs et al., 1989c; Fuchs et al, 1990) emphasized using skills analysis and data decision rules to assist teachers in *what* to teach. This included using graphical analysis of student data and examining individual student responses to inform instruction. Teachers, in response to specific skills data, demonstrated more regular changes to instruction, but still struggled to directly translate CBM feedback to effective instruction (Yell et al., 1992). To address this issue, the most recent advance in CBM studies focused on supporting teachers in selecting instructional practices or in *how* to teach (Fuchs 1998; Fuchs, Fuchs, & Hamlett, 1994). This phase of CBM feedback research is referred to in the literature as expert systems or diagnostic feedback. This type of feedback includes the use of quantitative analysis of student data along with classwide and individual student skill profiles and suggestions of intervention strategies for teachers to implement (Fuchs, Fuchs, & Hamlett, 2007; Fuchs, Fuchs, Hosp, & Hamlett, 2003). Studies reviewed in the next section used expert or diagnostic feedback consisting of intervention recommendations generated from student responses. Dependent variables targeted in the studies included student achievement and quantitative instruction features (e.g., number of changes, number of minutes of instruction, number of goal changes).

Fuchs, Fuchs, and Stecker (1989) were the first to compare the effects of diagnostic and quantitative feedback on instructional planning. Thirty special education teachers and their students in grades two to nine were monitored for 12 to 15 weeks using reading CBM. Teachers were randomly assigned to one of three groups, including: CBM plus computer assistance (CBM + CA), CBM non-computer (CBM - NC), and no CBM (control). Teachers in the CBM - NC condition received graphs visually displaying student words read correct across time. In the CBM + CA condition, teachers were provided a report containing a student’s graphed data, a notification when an instructional change is needed, and intervention recommendations. Both CBM groups collected data on their students. All teachers completed an instructional planning questionnaire during the study, which included the total number of instructional changes and student goals. Study results indicated both groups of CBM teachers developed specific goal statements and changed instruction more frequently than teachers in the control group. The authors prefaced results within the context of data-based decision making by stating objective data from CBM can result in improved instructional planning for teachers.

Fuchs, Fuchs, Hamlett, and Stecker (1991) examined the effects of differing levels of teacher feedback on instructional planning. Thirty-three special education teachers and pairs of their students across grades 2 through 8 were randomly assigned to one of their groups: CBM plus instructional feedback (CBM + IF), CBM only (CBM), and no CBM (control). In the CBM + IF condition, teachers received a graph of CBM student data, a skills analysis of student response patterns, and instructional recommendations for making a change in pedagogy. Student outcome measures included pretest and posttest standardized math tests and math CBM growth or slope. Throughout the study, teachers documented the number of instructional changes made in response to feedback systems. Results of the study showed both CBM groups altered instruction more often than the non-CBM group. In regards to student achievement, only CBM + EF displayed significantly higher academic achievement. Fuchs and colleagues summarized some important findings associated with the study outcomes. First, the cue provided by CBM to alert teachers when a change is needed is not enough to increase student achievement. Second, the act of making an instructional change alone will not lead to improved student performance. Therefore, teachers will need specific feedback related to student needs. This feedback must provide teachers the content/skills (*what*) to teach and what pedagogy (*how*) to teach it.

Fuchs, Fuchs, Hamlett, and Allinder (1991a) were the first to conduct an investigation on the effects of diagnostic feedback on planning and student achievement in spelling. Participants included 30 special education teachers and 60 of their students in grades two through eight. The teachers were randomly assigned to one of three conditions, including: CBM plus expert systems (CBM + ES); CBM only (CBM), and non CBM. Both CBM groups received graphs visually depicting student progress in spelling across time and alerts when a pedagogical modification was necessary based on student data. In addition to graphs and prompts, teachers in the CBM + ES condition received instructional suggestions to implement while the CBM group relied upon professional judgment. Students were assessed twice weekly for 18 weeks in spelling (correct letter sequences) and the Word Spelling Test as a pretest and posttest. Teacher planning behaviors were recorded during the study, including the total number of goal increases and instructional modifications, as well as the nature of pedagogical changes. Study results revealed that both CBM groups demonstrated significantly higher levels of student achievement compared to students in the control group. No differences in achievement were documented between CBM groups. In regards to instructional planning, teachers in the CBM groups initiated a significantly higher number of program changes compared to teachers in the control group. Teachers in the CBM + ES group, when given the choice, more frequently delivered pedagogy that was easier to implement (i.e., drill and practice), whereas CBM-NES teachers delivered more individualized instruction (e.g., direct instruction on spelling skills). Fuchs and colleagues suggested that

CBM procedures represent a strong option to improve instructional planning for students with disabilities. However, the authors ended with a warning related to the need for ongoing technical support related to the nature of instructional changes.

Whereas prior studies analyzed the effects of different forms of feedback (e.g., enhanced CBM feedback or CBM only), Wesson (1991) examined the effects of different levels of consultation feedback on student achievement in reading. Fifty-five special education teachers and their students were randomly assigned to either CBM or traditional, teacher-developed measurement systems. The groups were further split into either receiving individual expert feedback regarding student data or feedback shared among teachers in groups. Both conditions collected student data across 20 weeks. At the conclusion of the study, students of CBM teachers demonstrated statistically significantly higher rates of reading fluency compared to students of teachers that used teacher-developed assessment and evaluation systems. In addition, students of CBM teachers in the group consultation condition displayed higher levels of reading fluency compared to students of CBM teachers who received individual consultation. Results suggested that group teacher instructional planning using CBM reading data could be more effective than individual planning sessions, which may be more time consuming.

In an attempt to extend the findings of Fuchs et al. (1991a) to reading, Fuchs, Fuchs, Hamlett, and Fergusson (1992) examined the effects of CBM feedback on teacher pedagogy and student achievement. Thirty-three special education teachers and 63 of their students in grades two to eight were randomly assigned to one of three groups: CBM plus instructional feedback (CBM + IF), CBM with no feedback (CBM), and no CBM (control). Teachers in the CBM conditions implemented CBM procedures across

17 weeks. Results indicated both groups of CBM teachers changed instruction more frequently than teachers in the control group. Students of CBM teachers demonstrated higher levels of achievement in reading fluency and comprehension. CBM + IF group differentiated instruction more frequently by focusing on more skills in reading (e.g., decoding, fluency, comprehension) and students outperformed students in the CBM only group on a measure of recall. Additionally, teachers in the CBM only group made changes to instruction that were similar to the assessment task.

Fuchs, Fuchs, and Bishop (1992) conducted the first study to compare the effects of placement (i.e., special versus general education) and measurement system (i.e., CBM versus unsystematic observations) on instructional planning. Thirty-seven special education and 25 general education teachers who taught at least one student with a learning disability in their class participated in the study. Twenty-five of the special education teachers were assigned to use CBM procedures (CBM) and the remaining special and general education teachers (no-CBM) monitored student progress using subjective, unsystematic data (e.g., observation, chapter tests). Results indicated teachers in the CBM group more frequently changed student goals and used individual CBM data to develop goals than teachers in the no-CBM group who more frequently used peer comparisons to develop goals. Regarding the nature of instruction, CBM teachers more frequently altered the features of instruction (i.e., group size, instructional practice) than non-CBM teachers. The study also revealed the differences of planning between placement settings. Specifically, special education teachers reported specific instructional strategies for student intervention programs, whereas general educators reported a higher use of accommodations (e.g., manipulatives and audio/visual).

Unlike previous studies which examined instructional planning between general and special education teachers in response to CBM, Fuchs, Fuchs, Hamlett, Phillips, and Bentz (1994) investigated the effects of CBM teacher feedback at the class level. Forty general education classroom teachers, who had at least one student in their class with math goals on their individualized education programs (IEPs), were randomly assigned to one of three conditions: classwide report + instructional feedback (CBM + IF), classwide report (CBM-C), and no CBM (control). Student achievement was measured using a standardized broad measure of math achievement (name the test) and CBM math slope across 25 weeks. All of the teachers documented key instructional planning variables across 3 weeks, including total number of skills taught, isolated skills taught, minutes for instruction, and minutes for different instructional grouping (i.e., whole- or small-group, or one-to-one tutoring). Teachers were also asked to report their perception of CBM procedures (i.e., approval, impact on student performance, usefulness of feedback reports). Results indicated CBM + IF students outperformed control students on a global measure of math achievement. Students of CBM + IF teachers also demonstrated higher rates of learning in math compared to students of CBM – C and control teachers. Compared to CBM – C and control teachers, CBM + IF teachers reported teaching more skills and in general delivering more individualized instruction more frequently. Teachers from both CBM conditions also reported a generally positive view of CBM procedures. In reviewing key findings, Fuchs and colleagues concluded feedback from CBM can be adjusted to broadly apply to classwide instruction and may enhance student performance across levels of achievement. The authors tempered their findings by suggesting that general education teachers will need specific instructional suggestions to effectively plan and produce higher levels of student achievement.

Capizzi and Fuchs (2005) investigated the effects of CBM feedback on teacher instructional planning for classes and individual students. This study differed from previous research because it also examined the effects of CBM feedback on instructional planning for individual students of differing achievement levels (i.e., low, average, and high). Study participants included 19 second grade general education teachers and their 309 students. Sixteen special education resource teachers and 127 students across first- and fifth-grade also took part in the study. The authors selected one low- (LA), average- (AA), and high-achieving (HA) student from each teacher. Teachers were randomly assigned to CBM only (class report which rank-ordered students), CBM plus diagnostic feedback (CBM + DF), or no CBM (control). In addition to the class report provided to CBM only teachers, CBM + DF teachers received reports indicating student skills analysis patterns and explicit pedagogical suggestions for reading instruction (i.e., word analysis, fluency, comprehension). Classroom- and individual -level planning sheets were completed by participating teachers for LA, AA, and HA students. At the conclusion of the study, CBM + DF teachers documented fewer weekly objectives than teachers in the control condition. On average, CBM + DF general education teachers created more suitable objectives for AA achieving students than general education teachers in the CBM only group. Special education teachers in the CBM + DF group instituted more objectives for LA and AA achieving students than special education teachers in the control or CBM only groups. In regards to CBM reports, Capizzi and Fuchs concluded explicit instructional feedback may support teachers in attending to a smaller number of important skills in classwide planning. In addition, CBM reports and instructional feedback to special education teachers may inform instructional planning for LA and AA achieving students.

Whereas Capizzi and Fuchs (2005) examined the effects of feedback from static CBM on general and special education teacher planning at the individual and class level, Graney and Shinn (2005) investigated the effect of group and individual levels of feedback on student progress monitoring data for general education teacher planning and student achievement. The study consisted of 44 second grade teachers and 184 of their lowest performing students in reading. Participating students were assessed at least weekly across 11 weeks in reading (i.e., words read correctly in 1 minute). After 5 weeks, teachers were randomly assigned to group feedback, individual feedback, or no feedback. Teachers in the feedback condition received graphs visually depicting the average reading growth for the group and individual students. Individual feedback teachers were provided charted CBM data for a randomly selected individual students in their group. During week 6, research staff briefly consulted individually with teachers to review student growth and perceived teacher need for instructional changes. Results indicated that neither group- or individual-level feedback enhanced student achievement, although achievement improved across all groups suggesting potential benefits of monitoring student progress in reading. The authors concluded that general education teachers will require more diagnostic feedback than graphs with slope data and broad pedagogical recommendations if student achievement is to be increased.

In another study examining low performing students, Souvignier and Forster (2011) investigated the effects of computer-based progress monitoring on fourth-grade students. Twenty-four general education teachers and 144 of their lowest performing students in reading were randomly assigned to either the progress monitoring or control condition. Students of teachers in the progress monitoring condition were administered eight reading cloze tests across six months and the teachers were provided information about their learning trajectories. Teachers in the control condition only received initial performance information about their students. Results indicated that students in the progress monitoring feedback condition exemplified significantly higher growth (d = .50) compared to control group students. Participating teachers perceived the feedback as helpful and used the information to inform instruction.

While Souvignier and Forster (2011) and Graney and Shinn (2005) evaluated the impact of progress monitoring feedback on lower performing students, Forster and Souvignier (2015) examined the effects of progress monitoring teacher feedback on student achievement for classrooms of third-grade students. Forty-three general education 3rd grade teachers and 958 of their students were randomly assigned to one of three conditions: static reading level of performance (control), progress monitoring growth, and progress monitoring growth with additional training in selecting appropriate reading interventions. Students completed a short reading test every two weeks and the results were available to teachers through a computer-based software program. Teachers of students in the control group were provided with their students' beginning levels of reading fluency and comprehension skills. Teachers of students in the progress monitoring groups received feedback every two weeks regarding student reading growth, while the additional training group was provided extra training on how to select reading interventions. At the conclusion of the study, the teachers completed a questionnaire assessing their perception of usefulness of the feedback. Using a multilevel means-as-outcomes analysis, results indicated no significant differences on reading achievement between the progress monitoring groups. Students of teachers who received feedback and additional training significantly outperformed students whose teachers received only initial status feedback in reading fluency (z = 1.90, p = .029) and reading comprehension (z = 1.81, p =.036). In addition, teachers reported the usefulness of the feedback for instructional planning as high.

In summary, the results of studies on teachers who received expert systems/diagnostic progress monitoring feedback reported some consistent results with previous phases of CBM feedback research. CBM + D teachers more frequently instituted pedagogical changes and whose students generally displayed increased achievement. This finding was stable across classes and individual students. However, an important limitation to the impact of feedback was also consistent across studies and stages of research. This finding was teachers need on-going technical assistance to make these positive effects a reality.

Yet again, the above mentioned results are limited to both instructional planning and student achievement on middle to later elementary-aged students using CBM in reading, spelling, and math. A review of the literature reveals an absence of studies which examine the effects of early literacy formative assessment diagnostic feedback on teacher pedagogy and student achievement. As discussed earlier, prior studies have extensively examined the predictive utility of early CBM initial status and growth on future reading outcomes (e.g., Fein, Park, Baker, Mercier-Smith, Stoolmiller, & Kame'enui, 2010; Good et al., 2009; Harn et al., 2008). However, none of the studies have explicitly investigated whether early literacy CBM feedback enhances student achievement. This stage of research is clearly necessary to validate the use of early literacy CBM for informing reading instruction for beginning elementary-aged students.

Research on DIBELS Progress Monitoring Feedback on Early Literacy Student Achievement

A search of the literature for studies using DIBELS formative assessment feedback on student achievement resulted in one doctoral dissertation (Iannuccilli, 2003) and one published study (Ball & Gettinger, 2009). Iannuccilli (2003) was the first to examine the effects of progress monitoring teacher feedback on students' early literacy skills. Study participants consisted of a convenience sample of 54 first-grade students from six classrooms in a single school in the North East of the United States. The percentage of students in the district who received free-and-reduced lunch was 9.8%. The study also reported the percentage of students who received English as second language (0.3%), and special education (14.5%) services. Students were randomly assigned to one of two conditions: feedback and no feedback (control). Teachers whose students were in the feedback condition (n = 26) were provided with a graph charting student data over time and a general profile of student early literacy skills. As part of the study, target students were screened in December of first-grade using the DIBELS PSF and NWF. Beginning in January, target students were assessed using PSF and NWF bi-monthly for 9 weeks. Results on PSF and NWF posttest group scores using a multivariate analysis of variance (MANCOVA) indicated no significant differences ($F(2,49) = .521, p>.05$) between the groups on either measure.

While Iannuccilli (2003) provided feedback to teachers on first-grade students' early literacy skills, Ball and Gettinger (2009) examined the effects of formative assessment feedback on beginning reading skills of children in kindergarten. Eight kindergarten teachers and 103 of their students were randomly assigned to feedback and non-feedback groups. Teachers of students in the feedback group were provided feedback in the form of a description of the DIBELS measurement system and the risk status (i.e., low risk, some risk, at-risk) indicating the probability a student will meet future early literacy skill goals. Students were assessed in fall, winter, and spring of kindergarten using the DIBELS ISF, LNF, PSF, and NWF subtests. Teachers were provided feedback after the fall and winter assessment waves. The authors also collected social validity data regarding teacher perception of the usefulness of the feedback. Results of a repeated measure ANOVA indicated statistically significant differences in group scores from September to May favoring feedback on the ISF ($F(1, 101) = 5.60$, $p<.02$, $n^2 = .02$), PSF ($F(1, 101) = 8.02$, $p<.006$, $n^2 = .07$) , and LNF ($F(1, 101) = 12.96$, $p<.001$, $n^2 = .11$) DIBELS subtests. Although the students of teachers who received feedback made more progress across the school year, slightly more than half failed to attain recommend DIBELS benchmarks in spring of kindergarten. In addition, classroom teachers reported that the provided feedback was not generally helpful for improving student performance.

In summary, the two studies which examined the effects of formative assessment teacher feedback on early literacy student achievement offered mixed results. Iannuccilli (2003) did not observe a difference between feedback groups' scores on PSF and NWF in first-grade. In contrast, Ball and Gettinger (2009) found a difference in mean level of learning on PSF, ISF, and LNF for kindergarten students whose teachers were provided with feedback.

Summary

In summarizing the literature on progress monitoring feedback on student achievement, several important findings emerge. First, the use of feedback for students with and without disabilities from middle elementary school generally leads to improved student achievement. In all of the studies, CBM teachers more frequently implemented pedagogical changes and documented increased student academic performance.

Second, to obtain optimal benefits of formative assessment feedback, teachers must adhere to all of the feedback. For example, teachers must repeatedly measure student performance, receive and react to prompts recommending a change to instruction, use CBM derived skill profiles to assist in planning *what* to teach, and receive and select pedagogical recommendations on *how* to teach. Consequently, general and special education teachers will need continuous technical support in the form of instructional recommendations. The outcomes of not using all the feedback include developing instructional plans that reflect the measurement task, and not individualized to student needs. Finally, research examining the utility of progress monitoring feedback for enhancing student achievement has been relatively scant in the past decade (i.e., Ball & Gettinger, 2009; Forster & Souvignier, 2015; Souvignier & Forster, 2011). Instead, the majority of research has attended to the adequacy of the psychometric properties (i.e., reliability and validity) related to initial status scores (i.e., performance level) and slope (i.e., growth) in predicting important reading outcomes (e.g., Fuchs & Vaughn, 2012). Early literacy formative assessments, such as the DIBELS (Good & Kaminski, 2002) have demonstrated technically adequate reliability and validity for identifying students at risk for reading difficulties and subsequently monitoring the effectiveness of interventions (e.g., Cummings et al., 2011; Good et al., 2009). While this line of research is important to establish confidence in measurement results, a clear and present need is justified for Phase Three research, especially for improving teacher pedagogy and increasing achievement for early elementary-aged students (i.e., first-grade). Since the current emphasis in beginning literacy instruction strongly recommends prevention as a means to mitigate future reading problems (e.g., Lane, 2014; Snow, Burns, & Griffin, 1998), this line of inquiry is clearly essential.

CHAPTER 3: METHOD

The purpose of the current study is to examine the effects of assessment generated teacher feedback on early literacy student achievement. The study employed a quasi-experimental research design (Shadish, Cook, & Campbell, 2002) to compare the benefits of providing teachers with diagnostic feedback (i.e., skills profile and instructional suggestions) versus skills feedback (i.e., skills profile) derived from NWF progress monitoring scores (*N* = 4) across 10 weeks of first-grade. The study examined the effects of feedback on NWF growth scores and oral reading fluency performance level scores in winter of first-grade. The study also examined whether diagnostic progress monitoring feedback leads teachers to implement instructional changes more frequently than skills feedback. The following research questions were addressed.

Research Questions

1. What is the difference in growth of decoding, as measured using DIBELS NWF, between students whose teachers receive diagnostic feedback and those who receive skills feedback from fall to winter in first grade?
2. What is the difference between oral reading fluency (ORF) scores on the winter benchmark, as measured by DIBELS ORF (DORF), for students whose teachers receive diagnostic feedback and students whose teachers receive skills feedback?

3. What is the difference in instructional changes (i.e., number of reported instructional changes) between students whose teachers receive diagnostic feedback and those who receive skills feedback over fall to winter in first grade?
4. What is the relation between the number of instructional changes and student achievement on DIBELS NWF and DORF for students whose teachers use diagnostic or skills feedback to plan instruction?
5. Which type of feedback (i.e., diagnostic or skills) do teachers report as being more helpful for planning instruction?

Design

This investigation employed a two-group quasi-experimental time series group design (Shadish, et al., 2002). Participating teachers and students were from a convenience sample of intact classrooms. Teachers were assigned by the author to condition by school to prevent contamination across treatments. The dependent variables for the study were the DIBELS NWF and DORF scores. The independent variables were the conditions (i.e., skills only and diagnostic) of progress monitoring teacher feedback.

Participants and Setting

The study was conducted in three elementary schools in a small rural/suburban school district in North Carolina. Approximately 70 - 90% of all students enrolled in the schools were eligible to receive free and reduced lunch.

Teachers. Participants included 12 first-grade female general education classroom teachers from three elementary schools. All of the participating teachers were Caucasian. Four teachers possessed a Master’s degree. Two teachers obtained National Board Certification. Mean years of teaching experience across the sample was 11.17 years with a range of 1 to 27 years. Skills teachers had higher a mean years of experience (*M* = 12.86; range = 1 – 27) compared to diagnostic teachers (*M* = 8.80; range = 1 – 13); however, a chi-square test revealed no significant differences (p>.05) between groups.

Students. Out of the 250 parental consent forms sent home, 74 were signed and returned. Twenty student were excluded from the study because they did not meet inclusion criteria. Three students moved at the beginning of the leaving a total of 51 students who participated in the study. For inclusion in the study, students must have received at least 30 minutes of their reading instruction from the general education first-grade classroom teacher. In addition, participating students must have met at least one of the following measurement criteria. These included (a) below fall benchmark goal NWF score of 27 correct letter sounds, (b) below Rigby (2007) book level of D, or (c) their classroom teacher had concerns about their overall reading. Skills students had slightly higher mean NWF scores (*M* = 28.13; *SD* = 13.14) than diagnostic students (*M* = 25.33; *SD* = 7.84). In addition, skills students had marginally higher median book level (*Mdn* = C) compared to diagnostic students (*Mdn* = B). A t test for independent means found no statistically significant differences between groups on NWF scores $t(49) = -.87$, $p = .38$. A Chi-square test did not show any significant differences (p>.05) between groups on Rigby book levels.

The sample was comprised of a larger percentage (*N* = 59%) of male students. Students from minority backgrounds (i.e., Latino and African American) made up slightly more than half the sample. Students who spoke another language other than English comprised a quarter of the total sample. Students identified by the school district as having a disability made up nearly 12% of the total sample. Two of these students were identified with a specific learning disability with the remaining four as speech/language impaired. Chi-square tests resulted no significant differences (p>.05) between groups on gender, race/ethnicity, English language learner, and special education status.

Classroom Reading Instruction. In general education, schools in the participating district implemented a workshop model for delivering daily literacy instruction (120 minutes) consisting of a mini-lesson, teacher read-aloud to class, small group instruction (strategy group/guided reading group), conferencing, and independent reading. Phonics instruction was provided supplementary to the workshop model using the commercially published first-grade *Letterland* (Letterland International, 2014) curriculum. The *Letterland* curriculum consists of specific daily lessons designed to improve learning of phonics for typical and at-risk readers. The curriculum uses multi-sensory learning strategies to teach word families, blends, vowel teams, diphthongs, diagraphs, prefixes, and suffixes. Lesson materials include songs, flashcards and decodable stories. Approximately 15 minutes are expended on phonics, 15 minutes spent on reading and sorting words, and 5 minutes used on sentence reading instruction.

Measures

Nonsense word fluency (NWF). NWF is a subtest of the *DIBELS Next* (Kaminski & Good III, 2011) assessment suite and measures an examinee's knowledge of the alphabetic principle and basic phonics skills. The raw score is number of correct letter sounds (CLS) identified in one-minute. The examinee is presented with a model and a practice item. Then, the examinee is given a list of vowel-consonant and consonant-vowel-consonant pseudo-words (e.g.., ib, nat) and asked to read them. The directions ask the examinee to read the make-believe words the best they can by reading the whole word or saying letter sounds the examinee knows. For instance, if the word is nat, the student can respond /n/ /a/ /t/ or “nat”. Correct individual and blended letter sounds are underlined by the examiner. If an examinee reads the whole word without saying individual letter sounds the whole word is underlined. Self-corrected responses before 3-seconds are marked as correct. If an examinee does not respond in 3-seconds, the examiner provides the correct letter sound or whole word. The benchmark score for beginning (fall) of first-grade is 27 CLS. The alternative-form single form reliability for NWF was .85 in winter of first-grade (Good III et al., 2011). NWF displays moderate concurrent validity with the Group Reading Assessment and Diagnostic Evaluation (GRADE; Williams, 2002) Total Test in fall (.43) and winter (.51) of first grade (Good III et al., 2011). NWF shows moderate predictive validity (.43) with the GRADE Total Test in fall and winter (.51) of first-grade (Good III et al., 2011).

Oral Reading Fluency (DORF). DORF is a subtest of the *DIBELS Next* (Kaminski & Good III, 2011) assessment suite that measures advanced phonics, accurate and fluent reading connected text, and reading comprehension skills. Examinees are presented with a grade level passage and directed to read aloud for one-minute. If a student hesitates for longer than three seconds in pronouncing a word the examiner tells the student the word and marks the provided word as an error. The raw score is median number of words read correct (WRC) aloud in one minute from three passages. The benchmark score for middle (winter) of first-grade is 23 WRC. Winter of first-grade DORF displays moderate predictive validity (.64) on the GRADE Total Test. Alternate form reliability for first-grade DORF is high with a coefficient of .98 (Good III et al., 2011).

Teacher Questionnaire. At the conclusion of the study, teachers of students in both feedback groups completed a five item questionnaire (see Appendix A) measuring their perceived usefulness of the progress monitoring teacher feedback. An example is "*the formative assessment feedback was helpful for planning instruction.*" The measure used a five-point Likert scale (0 = *strongly disagree*; 1 = *disagree*; 3 = *neutral*; 4 = *agree*; 5 = *strongly agree*). This information was used as a social validity measure to document the teachers' view of the utility for feedback to improve student achievement. The target is for teachers to report an overall rating of 18 to 25. This target was determined based on prior research documenting relatively high levels of utility when teachers use progress monitoring feedback for planning instruction (e.g., Forester & Souvignier, 2015).

Instructional Modification Log. Participating teachers indicated the frequency of instructional changes for individual students through the use of a teacher log. Each teacher was provided a log (see Appendix B) containing their class list and a blank column to indicate an instructional change. After each progress monitoring feedback (n = 5) period, the teacher logged (by marking "Yes") if the teacher decided to change instruction for individual students. For the purposes of the study, an instructional change consisted of changes in (a) targeted skills (e.g., letter-sound knowledge fluency to blending word parts), (b) amount of time in minutes reserved for instruction, (c) group size (number of students), and (d) instructional materials (e.g., letter flashcards to word part tiles and place mat). Similar procedures have been used in prior studies to document teachers' change of instructional components (Capizzi & Fuchs, 2005; Wesson & Deno, 1989). To document whether expected implement changes are actually implemented, teachers also indicated the actual date an instructional change occurred on the instructional modification log.

NWF Diagnostic Feedback (Experimental)

Across 10 weeks, each teacher in the experimental group used DIBELS NWF to generate diagnostic teacher feedback to inform instruction. The diagnostic feedback consisted of two important features: skills profile and instructional recommendations.

Skills Profile. Data collection using the DIBELS NWF benchmark and progress monitoring probes occurred at least bi-monthly from November through January. After receiving the scored probes from project staff, the teacher reviewed each student's NWF response pattern and completed the *NWF General Performance Pattern and Instructional Recommendation* chart (see Appendix C) (University of Oregon Center for Teaching and Learning [UOCTL], 2010). The chart consisted of four columns representing the student's current level progress on letter-sound identification and word blending. The four columns included students who (a) pronounce individual sounds (/f/ /e/ /k/), (b) pronounce individual sounds and then recode into a single word ([/f/ /e/ /k/] fek), (c) partially blend word parts (/f/ /ek/), and (d) read the whole word automatically (e.g., fek). A student was placed in a column based on >50% of student responses corresponding to that column. In the case that student responses are spread uniformly across the four columns (e.g., 20% in A, 30% in B and C, and 20% in D), teachers were instructed to select the column corresponding to the highest point in the skills hierarchy (e.g., D for the example). The reason for this is the example student demonstrates adequate prerequisite skills to complete higher level word reading activities.

Instructional Recommendations. At the bottom of each column were 2-3 instructional activities teachers could select to implement. Activities were selected for inclusion in the chart based on several factors including (a) appropriateness to the student's decoding stage of development, and (b) feasibility and ease for implementation. The activities were included as part of the original *NWF General Performance Pattern and Instructional Recommendation* (UOCTL, 2010) document.

NWF Skills Feedback (Control)

Teachers in the skills feedback only group created a skills profile for their students using the same general procedures as the diagnostic feedback group; however, they did not receive the instructional recommendations as provided to the diagnostic feedback group. The version of the *NWF General Performance Pattern and Instructional Recommendation* (UOCTL, 2010) chart furnished to the control group had the instructional recommendations row intentionally deleted prior to provision.

Teacher Feedback Study Preparation and Training

All participating teachers were prepared to generate feedback from student progress monitoring through a series of three steps. Table 1 charted project activities from start to end of the study. Upon receiving institutional review board (IRB) approval, the study author met with district teachers and explained the study and potential benefits. Next, the study author collected signed research participation consent forms from classroom teachers and their students' parents (see Appendix D). Teachers in the both feedback conditions received a 60 minute in-service designed to instruct teachers how to generate teacher feedback from NWF progress monitoring probes. Both in-services provided an overview and rationale for the NWF and formative assessment. Teachers in both conditions received training on how to independently use the *NWF General Performance Pattern and Instructional Recommendation* chart. Training consisted of a model, guided and independent practice, and feedback on how to complete the skills portion of the chart. The criterion for proficiency for using the *NWF General Performance Pattern and Instructional Recommendation* chart was set at a teacher matching >90% of their students to the correct performance pattern (i.e., skill) column. Prior to leaving the in-service, all teachers from both conditions met the designated criterion after an additional round of guided practice. Teachers in the diagnostic feedback condition were provided instructional recommendations based on specific student skill profiles. In addition, diagnostic feedback teachers were provided with short (i.e., 4-5 minutes) videos modeling the recommended instructional activities on the chart. Teachers in the skills feedback condition did not have access to the instructional recommendation portion of the chart. Videos were available for diagnostic feedback teachers to view at their convenience as a link to an online folder that was disseminated at the conclusion of the training. In November, all teachers were provided study materials (i.e., *NWF General Performance Pattern and Instructional Recommendation* charts, teacher logs) in a 3-ring binder. The binder containing the study materials was collected at in the last week of January.

Students in participating classrooms were assessed by the study author using a first grade NWF progress monitor probed during the second week of November 2015. After data collection in November and December, the study author reviewed *NWF General Performance Pattern and Instructional Recommendation* charts for both groups for minimum levels of proficiency. If proficiency was below 90% on the chart, the study author held a short booster session to improve proficiency to an acceptable level. The Instructional Modification Log was reviewed in November and December to (a) remind teachers the log's purpose for documenting instructional changes and (b) and to confirm actual instructional changes are indicated on the log. In the case of a teacher not documenting instructional changes, a brief review was conducted with the teacher on the benefits and importance of noting instructional modifications. During the course of the study, all participating teachers from both groups correctly entered student names on the chart. One diagnostic and skills teacher required a short booster session to remind them the importance for completing feedback charts during the week of data collection.

Table 1: Timeline for study from preparation to conclusion

July, 2015	Obtain letter of support from superintendent of schools and school administrators to conduct study
October, 2015	Obtain university IRB approval for study
	Discuss study with teachers and obtain signatures on consent forms
	Send out student consent forms
November, 2015	Conduct 60 minute workshops with both feedback groups to show how to generate teacher feedback from NWF probes
November, 2015 through December, 2015	Booster sessions for teachers displaying <90% proficiency in identifying student skills profiles (study author)
November, 2015 through January, 2015	Administered DIBELS NWF progress monitoring probes monthly to students for whom consent is received (study author)
	Teachers in both condition completed feedback document

January, 2016	Administered DIBELS NWF and DORF benchmark probes to all students for whom consent is received (classroom teacher)

Data Collection

Classroom teachers collected the NWF benchmark and DORF benchmark data in January. DIBELS benchmark data collection for NWF and DORF is mandated by the school district as a responsibility of the general education classroom teacher. Teacher data was collected using iPads and stored on mClass (Amplify Education Inc, 2015), a commercially-available on-line database. The study author collected progress monitoring data using NWF progress monitoring probes on a district issued iPad and stored the data on mClass. These formative assessment data were collected at bi-weekly from November to December for a total of 3 progress monitoring NWF data points. The collection of benchmark and progress monitoring data resulted in a grand total of 4 data points on NWF. Teacher questionnaires were distributed in the second week of January and collected 2 weeks later. Teacher logs were collected when questionnaires were also collected.

Data Analysis

The study used a quasi-experimental design with data collected and entered into Microsoft Excel. Due to the hierarchical nature of the data (i.e., responses nested in students, students nested in study conditions and classrooms), the study employed a multi-level analysis framework using Hierarchical Linear Modeling version 7 (HLMv7;

Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011) to create a three-level growth model with repeated measures (i.e., time) at level 1, student covariates (i.e., gender, special education eligibility status and feedback affiliation) at level 2, and classroom level effects at level 3. HLM was selected for the proposed study because multilevel analysis procedures have been used in prior research to model student change in decoding rates (e.g., Compton, 2000).

Research Question One: 1.What is the difference in growth of decoding, as measured using DIBELS NWF, between students whose teachers receive diagnostic feedback and those who receive skills feedback from fall to winter in first grade?

In order to answer research question one, NWF progress monitoring data was collected from individual target students at least bi-weekly from November to January, for a total of 4 data points. The study examined the effect of feedback on the growth of student decoding (i.e., NWF) by fitting a growth model to measure how NWF scores vary *within* and *between* individuals across time (Raudenbush & Bryk, 2002).

A random-intercept and -slope model was employed to examine how level 2 variables vary between individuals and classrooms (Raudenbush & Bryk, 2002). The unconditional model (without predictors except for time at level 1) was examined to investigate mean and variance of between student and classroom decoding initial status and growth and to provide baseline statistics for comparing to the conditional models (Raudenbush & Bryk, 2002). The conditional model added three predictors of growth at level 2: gender, special education eligibility and feedback group membership. All three variables were dummy coded. Gender and special education eligibility status were selected for inclusion in the model based on prior research (e.g., Deno et al., 2001) indicating effects on literacy skill growth rates. Level 3 (classroom) variance was included in the model to provide more precise estimates for level 2 covariates. No variables were introduced in level 3.

Unconditional model. The below model presents an overview of the empty or unconditional model. This model contains no predictors and was used to estimate statistics for investigating individual student growth rates, which included the dividing of variability in the slope parameters into level-2 and level-3 elements (Raudenbush & Bryk, 2002). The model is as follows:

(Level 1 repeated measures model) $Y_{tij} = \pi_{oij} + \pi_{1ij}\alpha_{ti} + e_{tij}$;

where Y_{tij} is NWF scores (dependent variable) at time *t* for person *i* (i.e., 1…53) in classroom *j* (i.e., 1…12), π_{oij} is the November intercept for person *i* at $\alpha_{ti} = 0$ in classroom *j*; π_{1ij} is the NWF growth rate for person *i* in classroom *j* across November to January; α_{tij} is the week (i.e., 0…4) at time *t* for person *i* in classroom *j*, and e_{ti} represents the difference between person *i*'s NWF score at time *t* from model predicted score at that time point.

(Level 2 unconditional model) $\pi_{1ij} = \beta_{10ij} + r_{1ij}$

where β_{10ij} is the overall slope in NWF of students within classroom j and r_{1ij} is a random effect that represent the difference of person *i*'s NWF growth from the overall classroom slope (β_{10j}).

(Level 3 unconditional model) $\beta_{10j} = \gamma_{100} + \mu_{10j}$

where γ_{100} is the overall slope in NWF classroom *j* and μ_{10j} is a level-3 random effect that represent the difference of classroom *j*'s NWF growth from its predicted value based on the classroom-level model.

Student Covariate Model. The model below illustrates the full or conditional model. This model includes predictors that allows estimation of the effects of student gender, special education status, and feedback level on individual decoding learning rate. The level-1 model remains the same as the unconditional model with time as a covariate. The level-2 models now include predictors.

(Level 2 conditional model) $\pi_{1ij} = \beta_{10j} + \beta_{11j}$ (GENDER) + β_{12j} (SPED) + r_{1ij}

where β_{10j} is mean NWF growth rate for males who do not receive special education services, β_{11j} is the difference in the NWF growth rate for gender groups in classroom *j*, β_{12j} is the difference in NWF growth rate for students with and without disabilities in classroom *j*, 06and r_{1ij} is a random effect.

The level-3 model will continue to be unconditional (no predictors) for the purpose of modeling classroom level variance to obtain more precise level 2 covariate β slope coefficients. The below model hypothesized that

(Level 3 unconditional model)

$\beta_{10j} = \gamma_{100} + \mu_{10j}$

where γ_{100} is the mean NWF growth rate across all classrooms. The error (μ_{0j}) represents the difference between classroom *j*'s NWF score from the model expected score and will be allowed to vary.

Feedback Model. The model below illustrates the final model. This model includes predictors that allows estimation of the effects of feedback group affiliation on individual decoding learning rate. The level 1 and level 3 model remains the same as the student covariate model. The level-2 model now includes the feedback predictor.

(Level 2 model) $\pi_{1ij} = \beta_{10j} + \beta_{11j}$ (FEEDBACK) + r_{1ij}

where β_{10j} is mean NWF growth rate for students whose teachers didn't have access to diagnostic feedback for planning, β_{11j} is the difference in the NWF growth rate for feedback groups in classroom *j*, and r_{1ij} is a random effect. Gender and special education status will be removed from the model if analysis reveals the student covariates aren't significant predictors of NWF score growth rates.

Research Question Two: What is the difference in winter oral reading fluency (ORF) scores, as measured by DIBELS Oral Reading Fluency between students whose teachers receive diagnostic feedback and students whose teachers receive general feedback?

To address the second question, DORF benchmark data will be collected from individual participating students in January, for a total of 1 data point. Analysis of covariance (ANCOVA) to compare posttest DORF scores to determine whether a significant difference in group mean level of performance. Pretest NWF scores will serve as the covariate to control for initial level of performance in decoding.

The equation for the model is:

$DORF_{ij} = \mu + \alpha_i + \beta mNWF_{ij} + e_{ij}$

where $DORF_{ij}$ is the DORF score for student *i* in group *j* (j = 0, 1) with initial mNWF scores NWF_{ij}, μ is the grand mean, α_i is the group effect, β is the regression effect for the covariate (NWF), and e_{ij} is the residual

Hypotheses:

Null hypothesis *Ho*: There is no relationship between levels of feedback on DORF scores controlling for fall of first grade NWF scores. Mean DORF scores for diagnostic feedback = mean DORF scores for skills feedback.

Alternative hypothesis *Ha*: There is a relationship between levels of feedback on DORF scores, controlling for fall of first grade NWF scores. Mean DORF scores for diagnostic feedback ≠ mean DORF scores for skills feedback.

Statistical significance *p* value and partial eta n^2 effect size will be reported the CHAPTER 4.

Research Question Three: What is the difference in frequency (i.e., discrete) of instructional changes between students whose teachers receive diagnostic feedback and those who receive skills feedback over the fall to winter benchmark in first grade?

To address the third question a Mann-Whitney's U test was used to compare teacher indicated instructional changes between the feedback groups to determine whether a significant difference in group mean number of changes exists. For the purposes of the analysis, a response of "Yes" was coded as 1 and "No" as a 0. Statistical significance *p* value will be reported in CHAPTER 4.

Research Question Four: What is the relation between the number of instructional changes and student achievement on NWF and DORF?

To address the fourth question a Pearson's *r* correlation was used to measure the magnitude of the relation between the number of instructional changes and student achievement (i.e., NWF HLM growth estimate and posttest DORF scores). Statistical significance *p* value and *r* coefficient will be reported in CHAPTER 4.

Research Question Five: Which type of feedback (i.e., diagnostic or skills) do teachers report as being more helpful for planning instruction?

To address the fifth question a Mann-Whitney's U test was used to compare teacher questionnaire responses between the feedback groups to determine whether a significant difference in group mean existed. Statistical significance *p* value will be reported in CHAPTER 4.

Summary of Methods and Procedures

First-grade students (*N* = 53) from 12 classrooms were assigned to either the skills feedback (control) or diagnostic feedback (experimental). After 10 weeks of progress monitoring using the DIBELS NWF, five NWF scores were used to model growth in decoding to compare the effects of varying feedback levels on growth of decoding. Posttest DORF scores of the groups were compared using ANCOVA with pretest NWF scores as a covariate to determine the effects of varying feedback on reading fluency when initial level of decoding is controlled. Instructional modification log data were used to determine whether differences in number of instructional changes exist between diagnostic and skill feedback groups using a Mann-Whitney' U test. The relationship between the number of instructional changes and student achievement on NWF and DORF was investigated using a Pearson's r correlation.

CHAPTER 4: RESULTS

The purpose of this study was to examine the effects of diagnostic formative assessment feedback on decoding and oral reading fluency for first graders. The current study also addressed several additional research questions. First, did teachers who received diagnostic feedback implement more frequent instructional changes? Next, did documented instructional changes resulted in higher student achievement in early reading skills? Finally, was there a significant difference between groups in their reporting of feedback as helpful for planning instruction? This chapter describes the results of the data analyses used to evaluate each of the five research questions. The statistical procedures used in the study are outlined in this chapter. This section is followed by a description of statistical analysis used to examine each research question.

Research Questions

Research Question One: What is the difference in growth of decoding, as measured using DIBELS NWF, between students whose teachers receive diagnostic feedback and those who receive skills feedback from fall to winter in first grade?

A three-level model was conducted to determine if decoding growth rates (dependent variable) could be predicted from feedback group affiliation (independent variable). HLM was used to address research question one to account for the hierarchical nature (i.e., responses nested in students and students nested in classrooms) of the data (Raudenbush & Bryk, 2002). HLM was also selected because the procedure can handle missing occasion level (i.e., level-1) data without having to delete student level (i.e., level-2) cases (Raudenbush & Bryk, 2002). Prior to the analysis, the data were screened for several assumptions including adequate sample size, missingness, independence, normality, homogeneity of variance, and multicollinearity.

To determine the necessary sample size, a priori power analysis was conducted using *Optimal Design Plus Version 3.0* software (Spybrook et al., 2011). The program is designed to compute statistical power and minimum effect size for randomized experiments (Spybrook et al., 2011a). To detect statistical significance (p>.05) and an effect size of at least .50 using 4 measurement points, approximately 51 participants would be necessary. An effect size of at least .50 is expected based on prior research (e.g., Ball & Gettinger, 2009) that found η^2 ranged from .45 to .77. This assumption was met as data for a total of 51 participants was collected during the study.

To determine missingness, Statistical Package for the Social Sciences (SPSS) EXPLORE calculated the total of observed and missing cases. Only 1% of the total sample had missing data. There were two (1.7 %) skills feedback and one (1.2%) diagnostic feedback groups missing cases at Level-1 which were subsequently deleted when the analysis was run. This resulted in N = 201 cases at Level-1. There was no missing data at Level-2.This is in line with the literature on missing data which suggests missingness should be below 30% for person-level data (Newman, 2014).

To examine the assumption of independent and normal distribution of residuals, SPSS residual files were created in HLM7 for Levels-1 and -2. The residual files were plotted on quantile-quantile plots (Q-Q plots) in SPSS. Both plots indicated an approximately normal distribution for residuals. As shown in Figure 1, the Q-Q plot

Level-1 residuals identified six data points that departed from their predicted values. The analysis was run with and without these 6 data points. These cases were ultimately retained in the final analysis as non-significant differences were observed when they were excluded. Review of Q-Q plots of Level-2 residuals identified one data point that deviated from the expected value. Thus, the assumption of normality was met.

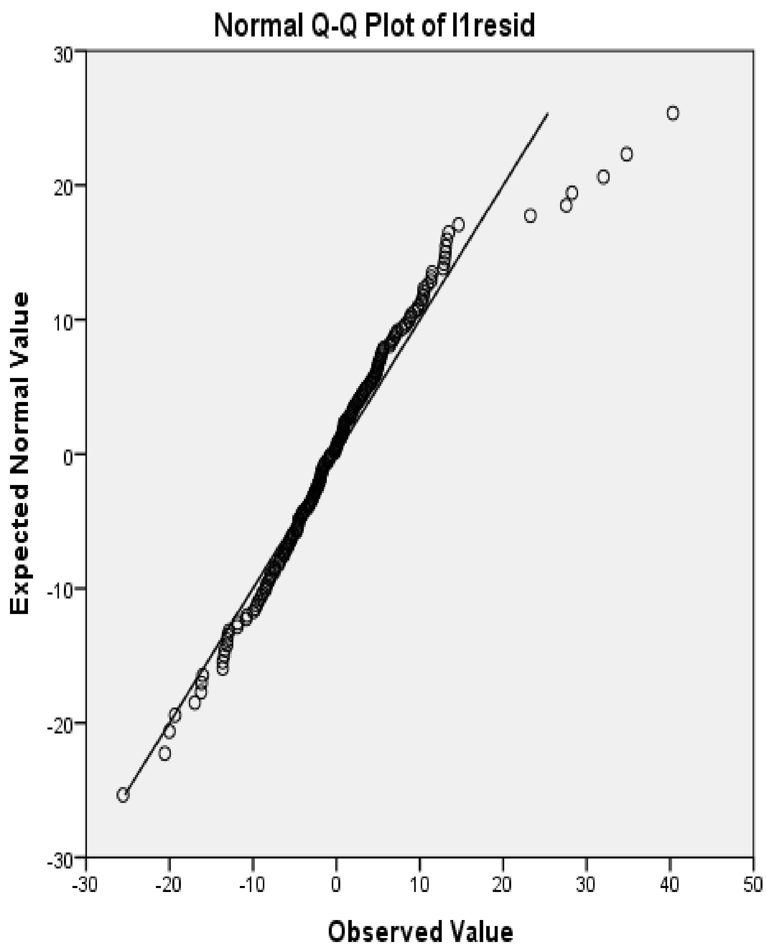

Figure 1: Level 1 Residual Q-Q Plot

A chi-square (χ^2) test was conducted to examine the assumption of level 1 homogeneity of error variance. Results of the chi-square test indicated heterogeneous level 1 variance, χ^2 (50, N = 51) = 92.27, p = .00. When the data suggests that level 1 residual variance is variable within time points, Raudenbush and Bryk (2002) recommend specifying a log-linear model. In other words, the log-linear model (σ^2_t) assumes level 1 variance (σ^2) might be specific for each time point (i.e., 0…3) versus across time. The log-linear model was then compared with the original model to determine model fit. A model comparison suggested that the original model with level 1 residual variance across time points appeared appropriate (χ^2 = 5.07, df = 1, p =.02). In addition, fixed effects coefficients and standard errors weren't substantially different than zero across models, suggesting the assumption was met.

Lastly, SPSS CORRELATE was used to examine for multicollinearity. A bivariate Pearson correlation for the independent variables was less than .90, suggesting that multicollinearity was not an issue.

Descriptive Statistics

Table 2: Means and standard deviations of NWF scores across time

	Diagnostic			Skills		
Variable	*M*	*SD*	*n*	*M*	*SD*	*n*
Time 1	35.50	10.54	21	39.04	17.78	30
Time 2	45.70	18.31	21	44.04	18.85	29
Time 3	52.15	25.30	20	48.75	25.34	29
Time 4	50.40	22.33	21	51.07	28.23	30

Table 2 shows the descriptive statistics for NWF scores at each time point by feedback group. As can be seen from the table, on the average, both groups improved in decoding over time.

Model 1 (Unconditional model)

The level 3 unconditional model consisted of a random-intercept and–slope model. As was presented in the METHOD, the following are the equations at Level-1, -2 and -3:

Level 1 $Y_{ti} = \pi_{oi} + \pi_{1i}\alpha_{ti} + e_{ti;}$

Level 2 $\pi_{1i} = \beta_{10i} + r_{1i}$

Level 3 $\beta_{10j} = \gamma_{100} + \mu_{10j}$

The above equations represent a linear growth model that examined within student growth in NWF scores across time. This model contained no student-level (e.g., gender, feedback group affiliation) predictors that could account for between student variability. Table 3 shows the fixed and random effects from the model. On the average, the predicted initial NWF score was statistically significant (p<.00) for all students at 39.40 CLS. Participating students are predicted to increase by a statistically significant (p<.00) 4.37 CLS every two weeks. Variance component coefficients of level 2 random effects on the intercept (p<.00) and growth rate (p<.02) were statistically significant. This indicates significant variability in initial status and growth rates within students across time. Level 3 variance component coefficients indicated significant variation (p>.05) in growth rates, but not for initial status (p<.50) between classrooms.

As the first step in the model building process, the intra-class correlation (ICC) was calculated to examine the ratio of between student slope variance to total slope variance. The following several sentences will outline the equation for calculating the ICC. The symbol σ^2 signifies the total slope variance in NWF scores across students (level 2) that can be accounted by the unconditional model, whereas τ_{00} is the total slope variance (i.e., between classrooms). Thus, the ICC equation for the unconditional model is $\tau_{00} / \tau_{00} + \sigma^2$. The ICC ranges from 0 to 1 and provides the proportion of the variance in outcome within and between students which serves as the effect size for HLM (Hox, 2000). For the unconditional model the formula is 12.31/ (12.31 + 83.68) = .13. Thus, for the level-3 unconditional model, 13% of the variability in decoding growth rates was due to between student differences, and the remaining 87% was due to variability within students across time. Given significant variance in slopes across students, the next step was to predict the variance in slopes at level 2 by introducing student level predictors.

Table 3: Model 1 (Unconditional model)

Fixed Effect	*Coefficient*	*SE*	*t ratio*	*p-value*
Intercept	39.40	2.10	18.74	.00
NWF Change Score	4.37	0.97	4.47	.00
Random Effect	*Variance Component*	*df*	χ^2	*p-value*
Level 1 variance				
Time e	83.68			
Level 2 (student within classroom)				
Intercept	160.44	39	143.60	.00
Slope	12.31	39	59.15	.02
Level 3 (between classrooms)				
Intercept	1.21	11	10.86	.50
Slope	4.15	11	19.33	.05

Model 2 (Covariate Model)

Since the primary research question addressed growth, a slopes-as-outcome model was specified using student level covariates (i.e., gender and special education status) and the Level 2 predictor of interest: receipt of intervention. The model building rationale was based on a two-step process where covariates were added to the unconditional model and removed from subsequent models if not statistically significant (Compton, 2000). The final model (reported later) includes significant covariates and the predictor of interest. The results of the covariate model is presented in Table 4. Gender (p = .51) and special education status (p = .10) did not have a statistically significant effect on NWF growth. On average, male students read .97 CLS more than females every 2 weeks. Students who received special education services read 3.98 fewer CLS every two weeks compared to students who did not receive special education services. The introduction of the level 2 predictors explained (model 1 variance component [12.31] - model 2 variance component [9.68]/model 1 variance component [12.31] = .21) an additional 21% of the variation in NWF growth scores between students after accounting for the covariates.

Table 4: Model 2 (Student covariate model)

Fixed Effect	*Coefficient*	*SE*	*t ratio*	*p-value*
NWF Change Score				
Male NWF Change Score	.97	1.48	.65	.51
SPED NWF Change Score	-3.98	2.35	-1.69	.10
Random Effect	*Variance Component*	*df*	χ^2	*p-value*
Level 1 variance				
Time e	83.64			
Level 2 (student within classroom)				
Intercept	156.51	39	143.67	.00
Slope	9.68	37	53.76	.03
Level 3 (between classrooms)				
Intercept	5.17	11	11.06	.43
Slope	5.21	11	21.78	.02

To examine the goodness of fit of Model 2 to the unconditional model (i.e., Model 1), a chi-square (χ^2) test was used to measure the amount of reduction in deviance from the first model. This step in the model building process is important when determining whether the original model or a more complex model is a better fit. A decrease in deviance that is not statistically significant would suggest both models demonstrate similar fit with the original model being ideal (O'Connell, Logan, Pentimonti, McCoach, 2013). The inclusion of the student-level covariates didn't result in model improvement over the unconditional model. The difference among the models was not significant (p>.05), thus the unconditional model remained the ideal model. Since neither gender nor special education status were significant predictors of NWF growth, both were excluded from final model. In addition, variance components for this model increased above the original model. Raudenbush and Bryk (2002) have indicated non-significant predictors may cause anomalies in statistics. Thus, further evidence to support the exclusion on the student covariates from the final model.

Model 3 (Feedback Model)

The results of the final model are presented in Table 5. Feedback group affiliation was added to the level 2 of the final model. On the average, students whose teachers used diagnostic feedback increased 1.53 CLS higher than students whose teachers used skills feedback; however, the difference was not statistically significant (p>.05). Adding the level-2 predictor *feedback* accounted (model 1 variance component [12.31] - model 3 variance component [12.21]/model 1 variance component [12.31] = ._) for only an additional 1% of the variance of growth rates. In addition, the feedback model didn't significantly reduce deviance (p>.05) when compared to the unconditional model suggesting the unconditional model was the best fit for the data.

Table 5: Model 3 (Feedback model)

Fixed Effect	*Coefficient*	*SE*	*t ratio*	*p-value*
NWF Change Score				
Feedback NWF Change Score	1.53	1.89	.83	.42
Random Effect	*Variance Component*	*df*	χ^2	*p-value*
Level 1 variance				
Time	83.48			
Level 2 (student within classroom)				
Intercept	160.41	39	143.96	.00
Slope	12.21	38	59.29	.01
Level 3 (between classrooms)				
Intercept	1.62	11	10.87	.50
Slope	4.06	11	19.06	.06

Figure 2 illustrates the development of decoding skills across time for both diagnostic and skills feedback groups. As can be seen from the figure, the growth curves for both groups overlap at several times growth across time suggesting similar rates of growth across both groups. The amount of reduction in deviance between Model 1 and Model 3 wasn't statistically significant from zero suggesting the original model was the best fit.

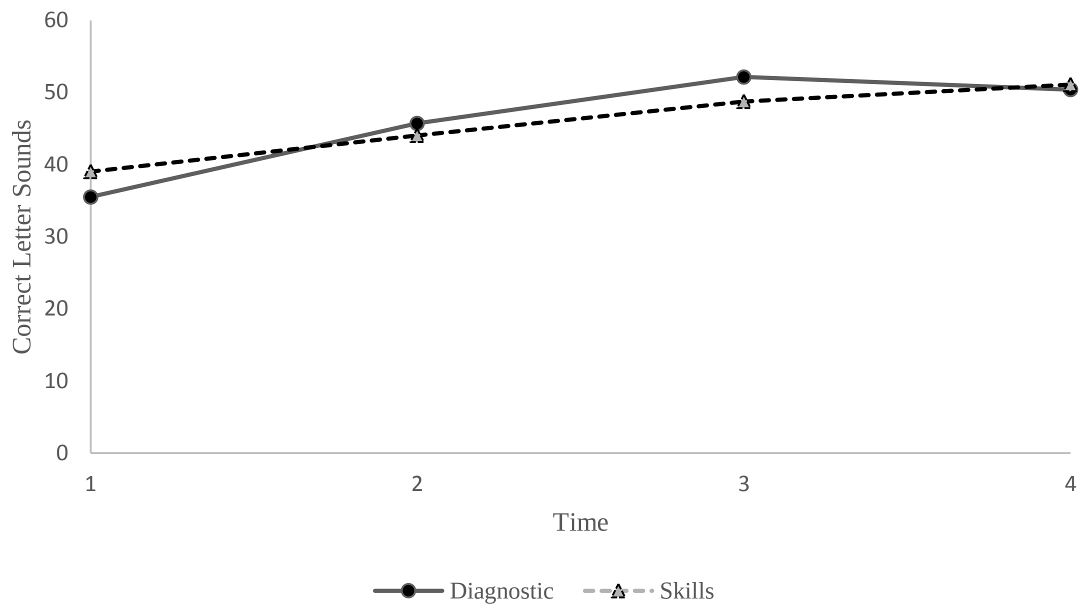

Figure 2: Feedback Growth Curves

Research Question Two: What is the difference in winter oral reading fluency (ORF) scores, as measured by DIBELS Oral Reading Fluency between students whose teachers receive diagnostic feedback and students whose teachers receive general feedback?

In response to research question two, a one-way ANCOVA was conducted. The independent variable, feedback affiliation, included two levels: diagnostic and skills feedback. The dependent variable was winter DORF scores and the covariate was initial NWF scores. Before conducting the analysis, DORF and initial NWF scores were examined for missing values and fit between their distributions and assumptions of multivariate analysis. The variables were examined separately for the 30 skills feedback students and 21 diagnostic feedback students.

SPSS EXPLORE indicated that all 51 cases contained no missing data. Next, SPSS DESCRIPTIVES calculated standardized *z*-scores for each variable of interest. One case in the skills feedback group was a univariate outlier. By using SPSS REGRESSION

Mahalanobis distance with $p < .001$, two cases from the diagnostic feedback group were identified as multivariate outliers. These cases were retained in the final analysis because removing them did not affect a significant change in results. The final analysis consisted of 30 cases in the skills feedback group and 21 in the diagnostic group ($N = 51$ cases).

The SPSS EXPLORE function showed information about the normality assumption. Skewness and kurtosis were both within the range of acceptable values from -2.0 to +2.0 for NWF and DORF scores for the skills feedback group. In regards to the diagnostic feedback group, skewness was within the above mentioned range of acceptable values, but kurtosis (3.83; $SE = .97$) exceeded those values. A histogram of diagnostic feedback DORF scores indicated a platykurtic distribution (i.e., low degree of peakedness). This was not considered problematic because prior research has documented non-normal distributions for first grade winter DORF scores (Catts, Petscher, Schatschneider, & Bridges, 2009).

To address homogeneity assumptions, a homogeneity-of-regression (slope) assumption was met indicating no significant interaction between the independent variable (i.e., feedback membership) and the covariate (i.e., initial NWF scores), $F(1, 46) = .74$, $p = .39$. Similarly, Levene's Test of Equality of Variances was not significant ($p > .05$). To examine the multicollinearity of the variables, SPSS REGRESSION LINEAR revealed correlation between variables below .90 and fairly small (2.5) variance inflation factor (VIF) results.

Posttest DORF scores were similar for students in the skills and diagnostic feedback group. There was not a significant effect of feedback group membership on the level of performance for first grade winter DORF after controlling for the effect of initial

NWF scores, $F(1,46) = .31$, $p = .57$ (See Table 4). The amount of variance winter DORF scores explained by feedback group affiliation controlling for initial NWF scores was small ($\eta^2 = .007$) by conventional standards for effect size magnitude (Cohen, 1992).

Table 6: Means, standard deviations, skewness, and kurtosis

Variable	*M*	*SD*	Skewness	Kurtosis
Diagnostic				
NWF	35.67	10.30	1.18	1.92
DORF	19.33	10.72	1.60	3.83
Skills				
NWF	39.17	17.41	.48	.49
DORF	19.62	14.81	1.07	.22

Table 7: Analysis of co-variance for DORF scores by feedback group

Source	*SS*	*df*	*MS*	*F*	*p*
Feedback	24.31	1	24.31	.31	.57
NWF Score	3648.53	1	3648.53	47.70	.00
Error	3518.28	46	76.48		
Total	8450.50	50			

Research Question Three: What is the difference in frequency (i.e., discrete) of instructional changes between students whose teachers receive diagnostic feedback and those who receive skills feedback over the fall to winter benchmark in first grade?

In response to research question three, a Mann-Whitney *U* test was conducted to evaluate the hypothesis that diagnostic feedback teachers would more frequently, on the average, alter their instruction compared to skills feedback teachers. Table 8 presents the summary statistics for the analysis. Overall, skills teachers implemented more instructional changes than diagnostic teachers. All diagnostic teachers implemented at least one instructional change; in contrast, one skills teachers didn’t implement a change.

However, on the average, the difference was not significant between the feedback groups (U = 13, p = .46).

Table 8: Teacher documented changes to instruction

	M (SD)	*SD*	*Range*
Diagnostic Group $n = 5$	6	3.08	2-9
Skills Group $n = 7$	8.57	5.88	0 - 17

Research Question Four: What is the relation between the number of instructional changes and student achievement on NWF and DORF?

In order to assess the relation between instructional changes and early reading skills a Pearson product-moment correlation was calculated. The correlation between posttest NWF scores and instructional changes was not statistically significant, r = -.47, p = .13. The correlation between DORF scores and the amount of instructional changes was not statistically significant, r = -.26, p = .44. These results suggest a relative weak relation between beginning reading skills and the amount of instructional changes a teacher implements.

Research Question Five: What is the difference of perceived usefulness of progress monitoring feedback between diagnostic feedback teachers and skills feedback teachers?

In response to research question five, a Mann-Whitney U test was conducted to evaluate the hypothesis that diagnostic feedback teachers would rate the usefulness of progress monitoring feedback higher, on the average, than skills feedback teachers. Table 9 presents the summary statistics for the analysis. Overall, teachers rated the usefulness of progress monitoring feedback as moderately high. Teachers in the skills feedback group rated the usefulness of the feedback, on the average, significantly higher than teachers in the diagnostic feedback group (*U* = 5, *p* = .041).

Table 9: Teacher ratings of feedback usefulness for instructional planning

	Usefulness of Feedback		
	M	*SD*	*Range*
Diagnostic Group (*n* = 5)	19.80	0.83	19 - 21
Skills Group (*n* = 7)	22.57*	3.15	16 - 25

p<.05*

Summary of Results

The results in this chapter indicate that diagnostic progress monitor feedback is not related to early literacy student achievement after controlling for gender and special education eligibility. The results of a multi-level analysis revealed no significant differences between growth rates in decoding as predicted by feedback group affiliation. In addition, feedback group membership didn't predict oral reading fluency skills after controlling for initial level of word decoding.

In regards to planned instructional changes, skills feedback teachers reported more overall changes to their instruction than diagnostic feedback teachers. A Mann-Whitney *U* test did not show any significant differences between groups. The association between the amount of instructional changes and early literacy skills showed weak negative correlations. Finally, responses from a social validity measure were analyzed to evaluate differences of perceived usefulness of progress monitoring feedback between groups. Both groups reported, on average, that feedback was useful for planning instruction. A Mann-Whitney U test showed there was a significant difference ($p < .05$) between the two groups favoring skills feedback teachers.

The next chapter moves on to discuss findings in terms of important findings, how the current study contributes to the existing literature on progress monitoring feedback, implications for practice, and suggestions for future research.

CHAPTER 5: DISCUSSION

The purpose of this study was to examine the effects of formative assessment teacher feedback on early literacy student achievement. The results indicate students of teachers in the diagnostic (i.e., skills profile and instructional recommendations) feedback group did not differ significantly on NWF scores to students whose teachers only had skills feedback. In addition, groups did not differ significantly in DORF scores in winter of first grade after controlling for initial decoding skills. In terms of decoding fluency, a salient initial finding is students with disabilities lag behind typically developing peers quite substantially over time. With respect to instructional planning, skills feedback teachers reported a higher mean number of instructional changes; however, the groups did not vary significantly. A correlation between instructional changes and early literacy skills indicated a negative relation between the variables. Furthermore, on average, the reported the usefulness of progress monitoring feedback for skills feedback teachers was significantly higher than for diagnostic feedback teachers.

The subsequent final chapter is divided into three parts. First, a discussion on each of the key findings of the current study is presented. Next, findings are placed in the context of previous research on CBM, progress monitoring and RtI. Finally, implications for future research, practice and limitations are also discussed.

CBM and Diagnostic Feedback on Decoding and Reading Fluency

The first research question, after controlling for student-level characteristics (i.e., gender and special education status), sought to examine whether diagnostic feedback group membership was related to NWF growth rates for first grade students. The second research question in this study sought to investigate the association between diagnostic feedback group affiliation and DORF scores for first grade students when holding initial decoding fluency constant.

With respect to the first and second research questions, results indicated no clear evidence that diagnostic progress monitoring feedback is associated with increases in decoding and oral reading fluency for first grade students. Decoding fluency for both groups, on the average, did improve across the 10 weeks of the study; though as Figure 2 illustrated, the rate between groups was not markedly different. Mean posttest NWF scores across groups were above the recommended cut point for low risk (> 43 CLS) and both groups appeared to exit the study with, on average, minimal difficulty mapping letters-to-sounds. This finding was not surprising because mean pretest NWF scores across conditions were only slightly below the DIBELS benchmark goal for fall of first grade.

Similarly, at the conclusion of the study, diagnostic feedback was not related to improved oral reading fluency performance. Posttest DORF scores for both groups, on average, were below established benchmarks for low risk (< 23 words read correct per minute). Therefore, most students were not on track in passage reading and would require intensive instructional support. These outcomes are alarming because research has consistently documented that oral reading fluency is a strong indicator of reading proficiency (Wayman et al., 2007; Rechsley et al., 2009). Specifically, students who struggle to learn to read at this critical point in their beginning literacy development continue to do so beyond first grade (e.g., Juel 1988, Ferrer et al., 2015).

The question now begs why was diagnostic feedback as compared to skills feedback not associated with increased beginning reading skills for first graders? There are several possible explanations for these results. First, the schedule of the study. Across the 10 weeks of the study, two lengthy holiday breaks and several snow days interrupted the typical instructional schedule. This contracted timetable did not allow teachers to provide typical amounts of literacy instruction. It is possible, therefore, that if staff provided instruction on a normal continuous schedule different results may have been observed. This explanation might on the surface appear speculative, but is consistent with limited research on the deleterious academic effects of interrupted reading instruction for elementary-aged students (e.g., Marcotte & Hemelt, 2007).

Second, it may be that both groups of students benefited from receiving some form of the treatment. In other words, both groups were impacted by the feedback (i.e., skills and diagnostic) and all students made growth because of skills feedback rather than needing the more detailed diagnostic feedback. Prior studies that reported a difference between groups (Förster & Souvignier, 2015; Stecker et al., 2005) employed a business-as-usual control where at least one group didn't receive any form of feedback. The current study employed a "treated" control with the skills feedback teachers receiving information about their students. Therefore, the possibility exists that providing teachers with *only* a skills profile of their students' word reading development may have led to the increases for both groups observed in the present study.

Another potential explanation is diagnostic teachers did not have the knowledge, skills, and resources to adapt the recommendations into their classroom routines. Diagnostic teacher training and ongoing technical assistance in the current study was minimal. Teacher training was limited to a 50 to 60 minute after school in-service, which included instructional strategy/intervention recommendations, and several brief (i.e., 5-10 minutes) follow up sessions. The single purpose of the follow-up sessions was to verify teachers completed study documents accurately. No supplemental assistance was provided on the selection of instructional strategies/interventions. Perhaps if teachers had received more extensive and continuing support in selecting instructional strategies/interventions different results may have been observed. For example, diagnostic teachers may have been unfamiliar with the research base supporting the recommended instructional strategies/interventions aligned to student needs. Diagnostic teachers were provided a short video model for each instructional strategy/intervention, but received no additional training on the activities. Similarly, if teachers were provided supplemental activity suggestions more closely aligned with the core reading program (i.e., *LetterLand*), perhaps observed effects may have been stronger.

This line of reasoning is consistent with the research on CBM and progress monitoring (Hoffman et al., 2009; Roehrig et al., 2008; Stecker et al., 2005). In a qualitative investigation on the aids and challenges to using progress monitoring to plan instruction, Roehrig and colleagues (2008) discovered that early elementary teachers were inadequately trained to interpret DIBELS progress monitoring data. Consequently, teachers in the study were unable to translate the progress monitoring feedback to instruction without assistance. Likewise, teachers surveyed by Hoffman and colleagues (2009) indicated limited training on translating DIBELS feedback to meaningful pedagogy. In addition, survey responses suggested confusion about the connection of progress monitoring and designing instruction. Overall, these findings suggest the need for extensive training and ongoing support for general education teachers to effectively use progress monitoring feedback.

Divergent findings

Results of the present study are somewhat mixed when compared with the literature on progress monitoring and CBM at large. Stecker et al. (2005) found providing progress monitoring feedback that included student graphs, skills profiles, and instructional recommendations was associated with improved academic skills for students. Souvignier and Förster (2011) and Förster and Souvignier (2015) included charted student data and a summary of student reading skills (i.e., reading rate, reading accuracy, and text comprehension). Results across studies indicated strong effects on measures of reading fluency and comprehension for students whose teachers used the detailed feedback to plan instruction. Several explanations exist for these conflicting results. First, grade level might act as a moderator on the impact of progress monitoring feedback. Stecker and colleagues reviewed studies of students in Grades 1 thru 8. In addition, Förster and Souvignier (2015) and Souvignier and Förster (2011) examined the effects of progress monitoring feedback for middle grade (i.e., third- and fourth-grade) students in Germany. The students in these studies were in a different phase of reading development than students in early elementary grades. It may be that since early literacy skills are recognized as transitory (Kaminski & Good, 1998), first grade students may not show as much growth in decoding fluency when the overall focus of instruction shifts (Good et al., 2012). For example, the focus of beginning first grade instruction tends to be identifying and blending individual letter sounds. In contrast, instruction from the middle to spring tends to focus on reading sentences and short paragraphs. In other words, students may have reached a ceiling in their simple word decoding fluency, therefore resulting in limited observable growth.

Second, length of the study may explain divergent findings. The duration of studies showing an effect of progress monitoring feedback on student achievement varies dramatically. Förster and Souvignier (2015) formatively assessed students and provided feedback bi-weekly for 6 months. In contrast, Stecker et al. (2005) reviewed studies that collected progress monitoring data at weekly intervals from 7 to 28 weeks. It may be, had the present study lasted longer than 10 weeks and included more than four data collection points, better outcomes may have been achieved by students in their overall reading.

Convergent findings

Despite divergent findings with literature focused on grade levels beyond the scope of this study, the present results are consistent with data obtained in studies conducted in early elementary settings. Iannucio (2003) found that providing classroom teachers with skills feedback of individual student phonemic awareness and decoding fluency did not result in higher growth over 9 weeks compared to non-skills feedback students. Similarly, Ball and Gettinger (2009) reported kindergarten teachers who received feedback (i.e., is the student at-risk for not meeting the next benchmark?) did not show significant differences in NWF scores from students whose teachers received no feedback. In addition, more than half of the feedback group students failed to reach the spring NWF cut point (Ball & Gettinger, 2009).

Although there was not a particular research question connected to the decoding fluency growth of students with disabilities, an important preliminary finding related to question one is striking. Students with disabilities, in general, grew at a much lower rate than students without disabilities. NWF growth rates indicate students with disabilities lagged behind peers by almost 4 CLS every two weeks. This finding is startling when one extrapolates student decoding trajectories across time. To illustrate, after 10 weeks of instruction, students with disabilities were predicted to be behind typical peers by approximately 20 CLS. So a typical developing peer would read 50 CLS in a minute, while a student with a disability only 30 CLS. This suggests that even over a relatively short period of time (i.e., 10 weeks), the learning gap between students with disabilities and their typically developing counterparts can enlarge dramatically. A note of caution is due here given the sample size (n = 6) for students with disabilities was small.

In summary, this study found that providing general education teachers with information about their students' needs and aligned instructional strategies/interventions may not lead to improved student outcomes. This could be due to general education teachers' lack of knowledge and skills to translate progress monitoring feedback into effective instruction. Taken together, these results suggest general education teachers may need protracted technical assistance to make an impact.

With respect to research question three, 11 of 12 teachers in both groups reported changing instruction at least once during the study. Surprisingly, on average, skills feedback teachers documented more instructional changes than diagnostic teachers. This finding was unexpected because skills teachers, unlike diagnostic teachers, did not have access to strategies directly aligned with student needs. This result may be explained by the fact that skills teachers reported a high number of changes involving the group size over the course of the study. Modification of intervention dosage, time and materials were observed to occur less frequently for both groups.

These results are in agreement with those obtained by Ball and Gettinger (2009). The authors used an observation rating scale to measure the quality of classroom literacy environments across the beginning to end of kindergarten. Scale items included questions about classroom organization, oral language activities, and literacy materials. Higher scores indicated better instructional practices for increasing student achievement. Both feedback groups displayed increased scores across the year with no group demonstrating significantly higher scores.

With respect to research question four, the major finding was that a higher number of instructional changes was not related to higher achievement. A bidirectional relation was observed between the number of instructional changes and early literacy skills. In other words, more instructional changes resulted in lower student scores for NWF and DORF. A possible explanation for these results may be the lack of adequate time for instruction to be effective before making a change (Jenkins, Hudson, & Lee, 2007). Jenkins and colleagues (2007) have posited that it may only take a few intervention sessions to judge a student's acquisition of single skills (i.e., letter-sound relationships), but it may take much longer (i.e. > 9 weeks) to evaluate general reading skills such as decoding and reading fluency. The study duration was a comparatively short 10 weeks which may have made it difficult to observe a measureable change with the high number of reported changes.

The higher number of instructional changes leading to lower scores for NWF and DORF from the current study does not align with previous research findings. Ball and Gettinger (2009) used the Early Language and Literacy Classroom Observation (ELLCO) rating scale to monitor observable improvements in classroom literacy activities. Higher scores on the ELLCO indicate the quality of classroom reading activities are increasing over time. Correlations between spring of kindergarten ELLCO and DIBELS scores resulted in low to moderate range relations (r = .17 - .31). This result may be explained by the fact that study duration was substantially longer (i.e., 9 months) than the present study. This lengthier interval of time may have allowed for more variation in student achievement thus leading to higher scores at the final study measurement point. In addition, since the ELLCO measured the features of changes in classroom instruction those changes were observed and included in the analyses. The present study did not include a tool for measuring the quality of instructional change. Therefore it is difficult to quantify the impact instructional modifications had on the classroom environment outside of student scores.

The fifth research question in this study investigated teacher perceptions of progress monitoring feedback for planning instruction. Teachers across both groups reported feedback as generally helpful. Group mean scores on the questionnaire were above the threshold of 18 established at the outset of the study. Unexpectedly, skills teachers, on average, described feedback as pointedly more helpful than diagnostic teachers. This finding was unanticipated because it was assumed that diagnostic teachers would rate feedback higher due to the inclusion of intervention recommendations.

There are several potential reasons for teachers finding the feedback provided as helpful. First, teachers in both conditions may have viewed their student progress as related to using feedback to plan instruction because all students across groups generally improved in decoding fluency over time. Although, it is important to note the decoding gains were modest. Another explanation is that teachers, overall, reflected on the value of the feedback in terms of improved scores regardless of the magnitude of the gains when completing the questionnaire. This seems possible due to the conflicting accounts of high usefulness ratings and modest improvements in decoding fluency rates across time. Finally, and most likely, teachers viewed the feedback as helpful, but did not know how to translate the information to effective instruction. As discussed earlier in this chapter, diagnostic teachers were provided instructional strategies to target student skill needs. Yet, at the end of the study, diagnostic students scored approximately the same as their skills counterparts in decoding or reading fluency. It is possible the teachers needed additional training and technical support to make meaningful changes to their instruction. Perhaps if additional training and assistance were provided, teachers might have implemented pedagogical strategies with more impact for their students.

These results are consist with data obtained by Förster and Souvignier (2015). The authors found teachers viewed feedback as useful for instructional planning. Similar to the present study, feedback was frequent (i.e., bi-weekly) and fairly descriptive (i.e., reading fluency and comprehension skills). In addition, the delivery method was through computer-based reports which simply required the teacher to log in the system to access.

Although, the teacher perception results are consistent with the work of Förster and Souvignier (2015), they are different from several earlier studies. Ball and Gettinger (2009) reported classroom teachers' finding progress monitoring feedback as limited for planning instruction. The following explanations may provide insight regarding the conflicting results. Classroom teachers in the present study were provided feedback bi-weekly compared to three times annually. This shorter latency between feedback points might have been viewed as informative by teachers because they could more closely monitor student progress in decoding. Another explanation relates to the content of feedback. Both groups of teachers in the present study were provided information on the phases of reading development (Ehri, 2005) for individual struggling students. The feedback was based on Ehri's theory of word reading development in which children advance through a series of four overlapping phases (e.g., sound-by-sound, sound-by-sound and recode, partial, and full word reading). Ball and Gettinger (2009) teachers were provided, in comparison, the relatively limited information of whether their students were at or below DIBELS benchmarks.

Graney and Shinn (2005) noted similar responses to Ball and Gettinger (2009) among second grade teachers. Participating teachers were asked if their judgement changed regarding whether an instructional modification was needed after viewing a graph depicting student progress across time. Teacher perception did change for several teachers whose students' pre-feedback did not indicate a need to adjust instruction. In contrast, teachers whose students were judged to need a change pre-feedback didn't change their views when feedback confirmed they did not need to make a change.

The findings from the present investigation suggest first grade teachers view both forms (i.e., skills profile and/or diagnostic information) of feedback as helpful under certain conditions. First, feedback content needs to be descriptive. The more particular the information about student skills the better for planning. Second, teachers reported feedback as favorable when it can be accessed in a short time frame from the assessment. Lastly, teachers viewed having data on individual struggling students as helpful. The caveat in this finding is that, yet again, diagnostic (and skills) teachers did not take feedback provided and develop instruction to accelerate student learning.

Contributions of Findings to the Literature

The findings from the current study extend the knowledge and literature base in several areas. For example, prior research demonstrates providing general education teachers with limited feedback on student performance on progress monitoring measures over time is typically associated with weak effects on early literacy skills (Ball & Gettinger, 2009; Iannucio, 2003). The present study extends these findings by offering additional evidence regarding the provision of diagnostic feedback for general education teachers. Observed effects on decoding and oral reading fluency were smaller in magnitude when compared to conventional standards for effect sizes (Cohen, 1992). Likewise, effects were weaker than previous research on formative assessment (Fuchs & Fuchs, 1986; Hattie, 2009; Stecker et al., 2005). The current state of the evidence suggests general education teachers need more than knowledge of student risk status (i.e., is student's early reading skills off-track?), summary skills, and instructional recommendations. Instead, it seems teachers need further training and ongoing support surrounding feedback to enhance instructional planning and accelerate student learning.

Technical assistance should be in the form of explicit data-based (Stecker et al., 2005) decision making (i.e., make a change after X points below target line) and coaching from specialists (e.g., reading coaches, school psychologists) in their school buildings (Roehrig et al., 2008).

In addition, results provide preliminary evidence on the ineffectiveness of progress monitoring feedback for students with disabilities in early elementary school. Prior studies have not included (Ball & Gettinger, 2009; Iannucio, 2003) students with disabilities in their samples. Initial results of the current study extend the literature by suggesting early literacy progress monitoring feedback, in general, may not be enough to support general education classroom teachers. Prior intervention research consistently demonstrates that students with significant learning needs require more intensive teaching strategies than may be reasonably provided by general education teachers alone (e.g., Al Otaiba & Fuchs, 2006; Fuchs et al., 2010; O'Connor & Klinger, 2010). For that reason, general and special education teachers need to cooperatively review progress monitoring data. This collaboration should lead to a plan where special education teachers can support students through intense, personalized interventions. Some examples of individualized modifications include smaller group size to allow for more opportunities to respond and corrective feedback, longer treatment and session durations, and materials closely aligned with students' present levels of performance, which will most likely be below grade-level (NCII, 2014). Through this supportive collaboration, students with disabilities should benefit more from progress monitoring feedback.

The results of the present study challenge the assumptions of RtI logic about progress monitoring for early literacy skills. For example, progress monitoring plays an essential role in measuring the effectiveness of beginning reading instruction (NCRtI, 2010). The purpose is to develop a data-base across time to support decision making for changing tiers and presumably, instruction. The evidence on progress monitoring for struggling students for these purposes on middle to later elementary grade children is emerging (Gersten et al., 2009). This study's finding that diagnostic progress monitoring feedback was not related to improvement puts into question the value of this practice in beginning reading environments. This should not be viewed as proclamation for the elimination of progress monitoring from RtI systems. Instead, the field of education will benefit with more clarity surrounding empirically supported *purposes* of progress monitoring. To be specific, progress monitoring (i.e., slope coefficient) has been shown to identify students who are not benefiting from general and supplemental instruction in early elementary settings (e.g., Compton et al., 2006; Compton et al., 2010; Gilbert et al., 2013). In contrast, the literature is nearly bereft of studies supporting the concept of using progress monitoring feedback (i.e., student responses) to tailor instruction to student specific needs. Thus, researchers need to continue to devise research-validated practices to formatively assess student progress and simultaneously modify instruction for struggling students in early elementary-age children.

The results of the current study highlight a gap in the literature associated with an optimal frequency of progress monitoring for struggling early elementary-aged students. Existing studies employ vastly different frequencies for collecting progress monitoring data. For example, Ball and Gettinger (2009) used a loose schedule of at least three to four months (i.e., annually) apart, whereas Iannucio (2003) used a more intensive or dense schedule (i.e., bi-weekly across 9 weeks). The current study extends the literature on progress monitoring frequency by suggesting a short, dense schedule may not provide adequate information to facilitate data-based decision making. Instead, a monitoring schedule of bi-weekly data collection that is more long-term (12-16 weeks) might be appropriate. This will allow teachers to collect a sufficient amount of formative assessment points to develop a data base for decision making. This is line with Kaminski and Good (1998) who suggest collecting multiple data points across time since young children are erratic in their responding.

Limitations

Several limitations constrain the current study's results. First, the dependent variable and outcome measure were DIBELS subtests. The use of different assessment battery subtests measuring varying dimensions of reading (i.e., timed or untimed) may result in divergent effects. In addition, the research design did not use random assignment to conditions which limits the external validity of study results. In addition, the sample size (N = 51) and number of data collection waves (N = 4) were comparably small to prior studies. This makes the results less generalizable to settings outside of the context of the current study. Next, participating teachers were relatively homogenous in ethnicity and background experience and the overall student sample consisted of a small percentage (n = 9%) of students identified through the school district as being eligible for a special education services. Therefore, these findings cannot be assumed to generalize to other settings with a more diverse population of teachers and students with disabilities. The primary data collector (e.g., study author) was not blind to student condition thereby potentially limiting the internal validity of the study. Furthermore, the study design did not control for other data teachers collected or accessed to plan instruction. These data included current and past DIBELS scores and running records. Thus, it remains unknown how teachers in the present study triangulated additional formative assessment data to assist in facilitating instructional decisions. Finally, students in the feedback group did not explicitly receive feedback on their learning. Perhaps if students were episodically provided with visual depictions of their progress and goals, student improvement may have increased to levels observed in studies that utilized this feedback component of progress monitoring.

Future Research

Based on the present study's findings and limitations, several recommendations for future research are warranted. First, to develop a fuller picture of diagnostic progress monitoring feedback additional replication studies will be needed in early elementary settings (i.e., kindergarten and first grade). To expand the research, these replications may use a different design methodology (e.g., business-as-usual control) and intervention agents (i.e., special education teachers). Second, future studies should use different outcome measures. These may include a range of standardized measures of reading components (i.e., word reading, decoding, oral reading fluency, and reading comprehension) and widely used early reading CBMs (i.e., AIMSweb, easyCBM). In addition, considerably more work needs to be done to determine adequate frequencies of progress monitoring to improve student learning. Therefore, research should focus on determining whether more (e.g., weekly; twice weekly) or less (e.g., monthly; bi-monthly) regular formative assessment results in enhanced early literacy student achievement. Third, the current study's sample was comprised of small number of students with disabilities. Thus, future research should therefore concentrate on the investigation of the effects of beginning reading progress monitoring feedback on students with disabilities. Fourth, another suggestion for future research is path modeling of the moderating effect of teacher response to feedback (i.e., number of instructional changes) on early literacy student achievement may provide some insight into the impact of assessment feedback on teacher planning. Finally, future studies should provide students with feedback on their word reading development. This feedback should be comprised of goal planning, visual analysis of charted progress, and a skills profile of word reading skills.

Implications for Practice

The findings of the current study have a number of important implications for future practice. For example, feedback generated from formative assessment needs to be specific enough to provide teachers with potential instructional strategies/interventions. This should provide teachers the ability, on some level, to diagnosis and prescribe instruction for their students. Teachers should also view the feedback as helpful in engineering differentiated classroom environments.

In addition, classroom teachers should continue to use CBM in beginning reading to formatively assess their students. This should occur for several reasons including prevention and identification. Repeatedly examining student progress will inform teachers on the effectiveness of their classwide pedagogy. Students not responsive to general effective classroom instruction should then be provided more targeted instruction. This instruction should focus on a more narrow set of skills (i.e., phonological awareness, decoding, reading fluency, vocabulary, or reading comprehension) in addition to skills learned through the general curriculum. The purpose being to identify a subset of students who need supplemental support beyond the general curriculum and prevent students from falling behind their typically developing peers. Students receiving targeted instruction should have their achievement monitored to alert teachers regarding the effectiveness of supplemental supports.

Finally, echoing the sentiments of preceding authors on the issue (i.e., Graney & Shinn 2005; Stecker et al., 2005), general education teachers need systematic and continuous technical assistance to interpret and develop targeted instruction based on CBM data. This implication emerges from data that teachers reported feedback as useful, but strong effects in student achievement were not documented. Presently, some teachers may spend upwards of 20% of their time administering assessments (VanDerHeyden et al., 2016). The expectation is that teachers can translate these data into meaningful instruction. Unfortunately, this does not appear to be actual practice as most schools appear to be “data rich” but “interpretation poor” (Slotnik & Orland, 2010). To meet this expectation of effective data-based decision making, support will be needed. Support can be provided by converting student data into a consumable form that teachers understand. Most data collection systems currently used by school districts (e.g., DIBELS, mClass, and AIMSweb) compile and organize student data into charts and graphs by individual student, small group, and class across a building. However, these visual aids alone may not be interpreted accurately by busy classroom teachers. Administrators, literacy coaches, and school psychologists can help facilitate the process of reviewing and transforming multiple pieces of data into a logical and sensible document. Using a communication tool such as a graphic organizer (e.g., Capizzi & Barton-Arwood, 2009) may assist in translating student data into meaningful instructional practices. The purpose of the organizer is to ensure educators follow several steps to group and plan instruction. First, early literacy indicator scores (e.g., letter-sound fluency, phoneme segmentation fluency, nonsense word fluency) for students are compared with the assessment suite's benchmarks. Then, student names are placed in boxes matched to their anticipated academic needs. Finally, the teacher can then provide targeted instruction to increase student skills.

Conclusions

The ability to read has been documented as the most essential skill students will need for success throughout school (O'Connor et al., 2015). Unfortunately, a significant number of school-age children struggle to learn the foundational skills of reading. According to RtI logic, teachers of struggling students must identify their academic needs, implement targeted instruction, and monitor progress. However, the results of the current study suggest further work is needed to determine the correct combination of feedback and teacher support, which has a strong effect on student achievement. Continued work in this line of research may yield robust outcomes in the search for optimal feedback and technical assistance.

REFERENCES

Adams, M. J. (1990). *Beginning to read: Learning and thinking about print*. Cambridge, MA: MIT Press.

Al Otaiba, S., & Fuchs, D. (2006).Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *Journal of Learning Disabilities, 39,* 414–431.

Al Otaiba, S., & Lake, V. E. (2007). Preparing special educators to teach reading and use curriculum-based assessments. *Reading and Writing*: *An Interdisciplinary Journal*, *20*, 591-617.

Al Otaiba, S., Folsom, J. S., Schatschneider, C., Wanzek, J., Greulich, L., Meadows, J., Li, Z., & Connor, C. M. (2011). Predicting first-grade reading performance from kindergarten response to tier 1 instruction. *Exceptional Children*, *77*, 453-470.

Amplify Education Inc, (2015). mCLASS. New York: Brooklyn

Armbruster, B., Lehr, F. & Osborn, J. (2001). *Put reading first: The research building blocks for teaching children to read*. The Partnership for Reading: National Institute for Literacy; National Institute of Child Health and Human Development; and U.S. Department of Education.

Arthaud, T. J., Vasa, S. F., & Steckelberg, A. L. (2000). Reading assessment and instructional practices in special education. *Assessment for Effective Intervention*, *25*, 205-227. doi: 10.1177/073724770002500302

Ball, C., & Gettinger, M. (2009). Monitoring children's growth in early literacy skills: Effects of feedback on performance and classroom environments. *Education and Treatment of Children*, *32*, 189-212.

Berkeley, S., Bender, W. N., Peaster, L. G., & Saunders, L. (2009).Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities*, *42*, 85–95.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*: *Principles, Policy & Practice*, *5*, 7-72.

Busch, T.W., & Reschly, A.L. (2007). Progress monitoring in reading: Using curriculum-based measurement in a Response to Intervention Model. *Assessment for Effective Intervention*, *32*, 223-230.

Capizzi, A. M., & Barton-Arwood, S. M. (2009). Using a curriculum-based measurement graphic organizer to facilitate collaboration in reading. *Intervention in School and Clinic*, *45*, 14-23.

Capizzi, A. M., & Fuchs, L. S. (2005). Effects of curriculum-based measurement with and without diagnostic feedback on teacher planning. *Remedial and Special Education*, *26*, 159-174.

Catts, H. W., Petscher, Y., Schatschneider, C., Bridges, M. S., Mendoza, K. (2009) Floor effects associated with universal screening and their impact on the early identification of reading disabilities. *Journal of Learning Disabilities*, *42*, 163–176.

Clay, M. M. (2013). *An observation survey of early literacy achievement* (3rd ed.). Auckland: Pearson.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.

Compton, D. L. (2000). Modeling the growth of decoding skills in first-grade children. *Scientific Studies of Reading*, *4*, 219-259.

Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology*, *98*, 394–409. doi:10.1037/0022-0663.98.2.394

Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Gilbert, J. K., & Barquero, L. A. (2010). Selecting at-risk first-grade readers for early intervention: Eliminating false positives and exploring the promise of a two-stage gated screening process. *Journal of Educational Psychology*, *102*, 327–340. doi:10.1037/a0018448

Coyne, M. D., & Harn, B. A. (2006). Promoting beginning reading success through meaningful assessment of early literacy skills. *Psychology in the Schools*, *43*, 33-43. doi:10.1002/pits.20127

Cummings, K. D., Atkins, T., Allison, R., & Cole, C. (2008). Response to intervention: Investigating the new role of special educators. *TEACHING Exceptional Children*, *40*, 24-31.

Cummings, K. D., Dewey, E. N., Latimer, R. J., & Good, R. H. (2011). Pathways to word reading and decoding: The roles of automaticity and accuracy. *School Psychology Review*, *40*, 284-295.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*, 934-945.

Daane, M. C., Campbell, J. R., Grigg, W. S., Goodman, M. J., & Oranje, A. (2005). *Fourth-Grade Students Reading Aloud: NAEP 2002 Special Study of Oral Reading* (NCES 2006-469). U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. Washington, DC: Government Printing Office.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*, 219-231.

Deno, S. L. (1989). Curriculum-based measurement and alternative special education services: A fundamental and direct relationship. In M. R. Shinn (Ed.), *Curriculum-Based Measurement: Assessing special children* (pp. 1–17). New York: Guilford Press.

Deno, S. L. (1990). Individual differences and individual difference: The essential difference of special education. *The Journal of Special Education*, *24*, 160-173.

Deno, S. L. (2003). Curriculum-based measures: Development and perspectives. *Assessment for Effective Intervention*, *28*, 3-12. doi:10.1177/073724770302800302

Deno. S. L., & Mirkin. P. K. (1977). *Data-based program modification: A manual.* Reston, VA: Council tor Exceptional Children.

Denton, C. A., Ciancio, D. J., & Fletcher, J. M. (2006). Validity, reliability, and utility of the Observation Survey of Early Literacy Achievement. *Reading Research Quarterly*, *41*, 8-34.

Denton, C. A., Fletcher, J. M., Anthony, J. L., & Francis, D. J. (2006). An evaluation of intensive intervention for students with persistent reading difficulties. *Journal of Learning Disabilities*, *39*, 447-466.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, *9*, 167-188.

Fawson, P. C., Ludlow, B. C., Reutzel, D. R., Sudweeks, R., & Smith, J. A. (2006). Examining the reliability of running records: Attaining generalizable results. *The Journal of Educational Research*, *100*, 113-126.

Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K. E., Michaels, R., & Shaywitz, S. E. (2015). Achievement gap in reading is present as early as first grade and persists through adolescence. *The Journal of Pediatrics*, *167,* 1121-1125. Retrieved from http://dx.doi.org/10.1016/j.jpeds.2015.07.045

Fien, H., Baker, S. K., Smolkowski, K., Smith, J. L., Kame'enui, E. J., & Beck, C. T. (2008). Using nonsense word fluency to predict reading proficiency in kindergarten through second grade for English learners and native English speakers. *School Psychology Review*, *37,* 391-408.

Fien, H., Park, Y., Baker, S. K., Mercier Smith, J. L., Stoolmiller, M., & Kame'enui, E. J. (2010). An examination of the relation of nonsense word fluency initial status and gains of reading outcomes for beginning readers. *School Psychology Review*, *39*, 631-653.

Fletcher, J., & Vaughn, S. (2009). Response to intervention: Preventing and remediating academic difficulties. *Child Development Perspectives*, *3,* 30–37.

Florida Center for Reading Research [FCRR] (2007). *Comprehensive glossary of terms*. Retrieved from http://www.fcrr.org/Curriculum/glossary/glossaryOfReading.pdf

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, *90*, 37.

Förster, N., Souvignier, E. (2015). Effects of providing teachers with information about their students' reading progress. *School Psychology Review*, *44*, 60-75.

Francis, D. J., Steubing, K. K., Shaywitz, S. E., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, *88*, 3-17.

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review*, *33,* 188-192.

Fuchs, L. S., & Deno, S. L. (1994). Must instructionally useful performance assessment be based in the curriculum? *Exceptional Children*, *61*, 15-24.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. (1984). The effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement and student awareness of learning. *American Education Research Journal*, *21,* 449-460.

Fuchs, D., & Deshler, D. D. (2007). What we need to know about responsiveness to intervention (and shouldn't be afraid to ask). *Learning Disabilities Research & Practice*, *22*, 129-136.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, *53*, 199-208.

Fuchs, D. & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, *41*, 93-99.

Fuchs, D., & Fuchs, L. S. (2009). Responsiveness to intervention: Multilevel assessment and instruction as early intervention and disability identification. *The Reading Teacher*, *63*, 250-252.

Fuchs, L. S., Fuchs, D., & Bishop, N. (1992). Teacher planning for students with learning disabilities: Differences between general and special educators. *Learning Disabilities Research & Practice*, *7*, 120-128.

Fuchs, L. S., Fuchs, D., & Deno, S. L. (1982). Reliability and validity of curriculum-based informal reading inventories. *Reading Research Quarterly*, 6-26.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1994). Strengthening the connection between assessment and instructional planning with expert systems. *Exceptional Children*, *61*, 138-146.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review*, *22*, 27-48.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*, 239-256.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989a). Effects of alternative goal structures within curriculum-based measurement. *Exceptional Children*, *55*, 429-438.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989b). Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education*, *10*, 43-52.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989c). Monitoring reading growth using student recalls: Effects of two teacher feedback systems. *Journal of Educational Research*, *83*, 103-111.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1994). Strengthening the connection between assessment and instructional planning with expert systems. *Exceptional Children*, *61*, 138-146.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (2007). Using curriculum-based measurement to inform reading instruction. *Reading and Writing: An Interdisciplinary Journa*l, 20, 553-567. doi 10.1007/s11145-007-9051-4

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Allinder, R. M. (1991a). Effects of expert system advice within curriculum-based measurement on teacher planning and student achievement in spelling. *School Psychology Review, 20*, 49-66

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Allinder, R. M. (1991b). The contribution of skills analysis to curriculum-based measurement in spelling. *Exceptional Children*, *57*, 443-452.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement using a reading maze task. *Exceptional Children*, *58,* 436-450.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., & Bentz, J. (1994). Classwide curriculum-based measurement: Helping general educators meet the challenge of student diversity. *Exceptional Children*, *60*, 518-537.

Fuchs, L.S., Fuchs, D., Hamlett, C. L., & Stecker. P. M. (1990). The role of skills analysis to curriculum-based measurement in math. *School Psychology Review*, *19*, 6-22.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal*, *28*, 617-641.

Fuchs, L. S., Fuchs, D. & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly*, *25*, 33-45.

Fuchs, D., Fuchs, L. S., & Stecker, P. M. (2010). The "blurring" of special education in a new continuum of general education placements and services. *Exceptional Children*, *76*, 301-323.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Hamlett, C. L. (2003). The potential for diagnostic analysis within curriculum-based measurement. *Assessment for Effective Intervention*, *28*, 13-22. doi:10.1177/073724770302800303

Fuchs, L. S., Fuchs, D., & Stecker, P. M. (1989). Effects of curriculum-based measurement on teachers' instructional planning. *Journal of Learning Disabilities*, *22*, 51-59.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Allinder, R. M. (1991). The contribution of skills analysis to curriculum-based measurement in spelling. *Exceptional Children*, *57*, 443-452.

Fuchs, D., McMaster, K. L., Fuchs, L. S., & Al Otaiba, S. (2013). Data-based individualization as a means of providing intensive instruction to students with serious learning disorders. . In Swanson, H. L. Harris, K. R., and S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed.). (pp. 526-544). New York, NY: Guilford Press.

Fuchs, L. S., & Vaughn, S. (2012). Responsiveness-to-intervention: A decade later. *Journal of Learning Disabilities*, *45*, 195-203. doi: 10.1177/0022219412442150

Gardner, E. F., Rudman, H. C., Karlsen, B., & Merwin, J. C. (1982). *Stanford achievement test*. Iowa City, IA: Harcourt, Brace, Jovanavich.

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2009). *Assisting students struggling with reading: Response to intervention and multitier intervention for reading in the primary grades. A practice guide* (NCEE 2009-4045). Washington, DC: U.S. Department of Education, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/

Gilbert, J. K., Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Barquero, L. A., & Cho, E. (2013) Efficacy of a first-grade responsiveness-to-intervention prevention model for struggling readers. *Reading Research Quarterly*, *48*, 135-154.

Glover, T. A., & DiPerna, J. C. (2007). Service delivery for response to intervention: Core components and directions for future research. *School Psychology Review*, *36*, 526-540.

Good, R. H., Baker, S. K., & Peyton, J. A. (2009). Making sense of nonsense word fluency: Determining adequate progress in early first-grade reading. *Reading & Writing Quarterly*, *25*, 33-56.

Good III, R. H., & Jefferson, G. (1998). Contemporary perspectives on Curriculum-Based Measurement validity. In M. R. Shinn (Ed.) *Advanced applications of Curriculum-Based Measurement* (pp. 61-88). New York, NY: Guilford Press

Good, R. H., & Kaminski, R. A. (2010). *DIBELS Next Assessment Manual*. Eugene, OR: Dynamic Measurement Group.

Good III, R., Kaminski, R. A., Dewey, E. N., Wallin, J., Powell-Smith, K. A., & Latimer, R. J. (2011). *DIBELS Next Technical Manual*. Eugene, OR: Dynamic Measurement Group.

Good, R. H., III, Kaminski, R. A., Fien, H., Powell-Smith, K. A., & Cummings, K. D. (2012). How progress monitoring research contributed to early intervention for and prevention of reading difficulty. In C. A. Espin, K. L. McMaster, S. Rose, and M. M. Wayman (Eds.), *A measure of success: The influence curriculum-based measurement on education*. Minneapolis, MN: University of Minnesota Press.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading*, *5*, 257–288.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial & Special Education*, *7*, 6-10.

Graney, S. B., & Shinn, M. R. (2005). Effects of reading curriculum-based measurement (R-CBM) teacher feedback in general education classrooms. *School Psychology Review*, *34*, 184-201.

Harn, B. A., Stoolmiller, M., & Chard, D. J. (2008). Measuring the dimensions of alphabetic principle on the reading development of first graders: The role of automaticity and unitization. *Journal of Learning Disabilities, 41*, 143-157.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.

Hattie, J. (2012). Know thy IMPACT. *Educational Leadership*, *70*, 18–23.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*, 81–112.

Hoffman, A. R., Jenkins, J. E., & Dunlap, S. K. (2009). Using DIBELS: A survey of purposes and practices. *Reading Psychology*, *30*, 1-16. doi:10.1080/02702710802274820

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and writing*, *2*, 127-160.

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2007). *The ABCs of CBM. A practical guide to curriculum based measurement*. New York, NY: Guilford.

Hunley, S., & McNamara, K. (2010). *Tier 3 of the RTI model: Problem solving through a* *case study approach*. Bethesda, MD: National Association of School Psychologists & Corwin.

Iannuccilli, J. A. (2003). *Monitoring the progress of first-grade students with Dynamic Indicators of Basic Early Literacy Skills*. Unpublished doctoral dissertation, Indiana University of Pennsylvania.

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, *95*, 719-729.

Jenkins, J. R., Hudson, R. F., & Lee, S. H. (2007). Using CBM-Reading assessments to monitor reading progress. *Perspectives on Language and Literacy, 33*, 11-18.

Jones, E. D., & Kruse, J. P. (1988). The effectiveness of data-based instruction by student teachers in classrooms for pupils with mild handicaps. *Teacher Education and Special Education*, *11*, 9-19.

Joseph, L. (2015). *Understanding, assessing, and intervening on reading* (2nd ed.). Bethesda, MD: National Association of School Psychologists.

Juel, C. (1988). Learning to read and write: A longitudinal study of fifty-four children from first through fourth grades. *Journal of Educational Psychology, 80*, 437-447.

Kaminski, R.A. & Cummings, K. D. (2007). Assessment for learning: Using general outcomes measures. *Threshold, Winter 2007,* 26-28. Retrieved from http://ciconline.org/threshold

Kaminski, R.A., Cummings, K.D., Powell-Smith, K.A., Good, R.H. (2008). Best practices in using Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in an Outcomes-Driven model. In A. Thomas and J. Grimes (Eds.), *Best Practices in School Psychology V*. Bethesda, MD: National Association of School Psychologists.

Kaminski, R. A., & Good III, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review*, *25*, 215-227.

Kaminski, R. A., & Good, R. H., III, (1998). Assessing early literacy skills in a problem-solving model: Dynamic Indicators of Basic Early Literacy Skills. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 113-142). New York: Guilford Press.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychologist*, *6*, 293-323.

Lane, H. (2014). *Evidence-based reading instruction for grades K-5* (Document No. IC-12). Retrieved from University of Florida, Collaboration for Effective Educator, Development, Accountability, and Reform Center website: http://ceedar.education.ufl.edu/tools/innovation-configurations/

Lane, H. B., Pullen, P. C., Hudson, R. F., & Konold, T. R. (2009). Identifying essential instructional components of literacy tutoring for struggling beginning readers. *Literacy Research and Instruction*, *48*, 277-297.

Lembke, E. S., McMaster, K. L., & Stecker, P. M. (2010). The prevention science of reading research within a Response-to-Intervention model. *Psychology in the Schools*, *47*, 22–35.

Letterland International (2014). *Letterland*. Surrey, UK: Letterland International.

Margolis, H. (2012). Response to intervention: RTI's linchpins. *Reading Psychology, 33*, 8-10. doi: 10.1080/02702711.2011.630600

Marston, D. B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why we do it. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18-78). New York: Guilford.

Marcotte, D. E., & Hemelt, S. W. (2007). Unscheduled school closings and student performance. *Institute for the Study of Labor Discussion Papers*, *29*, 1-28. Retrieved from http://hdl.handle.net/10419/34654

McAlenney, A. L., & Coyne, M. D. (2011). Identifying at-risk students for early reading intervention: Challenges and possible solutions. *Reading & Writing Quarterly*, *27*, 306-323.

McNamara, J. K., Scissons, M., & Gutknecth, N. (2011). A longitudinal study of kindergarten children at risk for reading disabilities: The poor really are getting poorer. *Journal of Learning Disabilities*, *44*, 421-430.

National Center for Education Statistics. (2013). *The Nation's Report Card: A First Look: 2013 Mathematics and Reading* (NCES 2014-451). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

National Center on Intensive Intervention. (2014). *NCII glossary of terms*. Retrieved from http://www.intensiveintervention.org/sites/default/files/NCII_Glossary.pdf

National Early Literacy Panel. (2008). *Developing early literacy*: *Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy. Retrieved from http://www.nifl.gov/earlychildhood/NELP/NELPreport.htm

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Bethesda, MD: National Institute of Child Health and Human Development.

Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational Research Methods*, *17*, 372-411.

O'Connell, A. A., Logan, J. A. R., Pentimonti, J. M., & McCoach, D. B. (2013). Linear and quadratic growth models for continuous and dichotomous outcomes. In Y. M. Petscher, & C. Schatschneider (Eds.), *Applied quantitative analysis in education and the social sciences*, (pp. 125-168). New York: Routledge.

O'Connor, R. E. (2000). Increasing the intensity of intervention in kindergarten and first grade. *Learning Disabilities Research & Practice*, *15*, 43–54.

O'Connor, R. E., Beach, K. D., Sanchez, V. M., Bocian, K. M., & Flynn, L. J. (2015). Building BRIDGES: A design experiment to improve reading and United States history knowledge of poor readers in eighth grade. *Exceptional Children*, *81*, 399-425.

O'Connor, R. E., Harty, K. R., & Fulmer, D. (2005). Tiers of intervention in kindergarten through third grade. *Journal of Learning Disabilities*, *38*, 532-538.

O'Connor, R. E., & Jenkins, J. R. (1999). Prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading*, *3*, 159-197.

O'Connor, R. E., & Klingner, J. (2010). Poor responders in RTI. *Theory Into Practice*, *49*, 297-304.

Paris, S. G. (2002). Measuring children's reading development using leveled texts. *The Reading Teacher*, *56*, 168-170.

Phillips, L. M., Norris, S. P., Osmond, W. C., & Maynard, A. M. (2002). Relative reading achievement: A longitudinal study of 187 children from first through sixth grade. *Journal of Educational Psychology*, *94*, 3-13. doi:10.1037//0022-0663.94.1.3

Pikulski, J. J., & Chard, D. J. (2005). Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher*, *58*, 510-519.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R., & du Toit, M. (2011). *HLM7*: *Hierarchical Linear and Nonlinear Models*. Lincolnwood, IL: SSI.

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, *47*, 427-469.

Riedel, B. W. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade students. *Reading Research Quarterly*, *42*, 546-567.

Rigby PM Ultra Benchmark assessment guide. (2007). Austin, TX: Harcourt.

Roberts, G., Vaughn, S., Fletcher, J., Stuebing, K., & Barth, A. (2013). Effects of a response-based, tiered framework for intervening with struggling readers in middle school. *Reading Research Quarterly*, *48*, 237-254.

Roehrig, A. D., Duggar, S. W., Moats, L., Glover, M., & Mincey, B. (2008). When teachers work to use progress monitoring data to inform literacy instruction: identifying potential supports and challenges. *Remedial and Special Education*, *29*, 364-382. doi: 10.1177/0741932507314021

Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities evidence for the role of kindergarten and first-grade interventions. *Journal of learning disabilities*, *39*, 157-169.

Schatschneider, C., & Torgesen, J. K. (2004). Using our current understanding of dyslexia to support early identification and intervention. *Journal of Child Neurology*, *19*, 759-765.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Shinn, M. (2008). Best practices in curriculum-based measurement and its use in a problem-solving model. In J. Grimes & A. Thomas (Eds.), *Best practices in school psychology V* (pp. 547–568). Silver Spring, MD: National Association of School Psychologists

Shinn, M. R. (February 2, 2012). *Progress monitoring in RTI/MTSS: Critical tools-emerging school practice*. Retrieved from http://www.doe. virginia.gov/instruction/Virginia_tiered_system_supports/training/cohort/2012/jan/progress_monitoring_mtss_1_per_pg.pdf

Shinn, M. R., & Bamonto, S. (1998). Advances in Curriculum-Based Measurement and its use in a Problem-Solving model: Big ideas and avoiding confusion. In M.R.

Shinn, (Ed.) *Advanced applications of Curriculum-Based Measurement* (pp. 1-31). New York, NY: Guilford.

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, *78*, 153-189.

Simmons, D. C., Coyne, M. D., Kwok, O., McDonagh, S., Harn, B. A., & Kame'enui, E. J. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities*, *41*, 158-173.

Slotnik, W. J., & Orland, M. (2010). Data rich but information poor. *Education Week*, *30*, 13-14.

Souvignier, E., & Förster, N. (2011). Effects of curriculum-based measurement on reading achievement in fourth graders. *Empirical Special Needs Education*, *3*, 243-255.

Spector, J. E. (2005). How reliable are informal reading inventories? *Psychology in the Schools*, *42*, 593-603.

Speece, D. L., & Ritchey, K. D. (2005). A longitudinal study of the development of oral reading fluency in young children at risk for reading failure. *Journal of Learning Disabilities*, *38*, 387-399.

Speece, D. L., Schatschneider, C., Silverman, R., Case, L. P., Cooper, D. H., & Jacobs, D. M. (2011). Identification of reading problems in first grade within a response-to-intervention framework. *The Elementary School Journal*, *111*, 585-607.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407.

Stecker, P. M., & Fuchs, L. S. (2000). Effecting superior achievement using curriculum-based measurement: The importance of individual progress monitoring. *Learning Disabilities Research and Practice*, *15*, 128-134.

Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools*, 795-819. doi:10.1002/pits.20113

Stecker, P. M., Fuchs, D., & Fuchs, L. S. (2008). Progress monitoring as essential practice within response to intervention. *Rural Special Education Quarterly*, *27*, 10.

Stecker, P. M., Lembke, E. S., & Foegen, A. (2008). Using progress-monitoring data to improve instructional decision making. *Preventing School Failure: Alternative Education for Children and Youth*, *52*, 48-58.

Tindal, G. (2013). Curriculum-based measurement: A brief history of nearly everything from the 1970s to the present. *ISRN Education*, *1*, 1-29. doi: 10.1155/ 2013/958530

Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice*, *15,* 55-64.

Torgesen, J. K. (2004). Avoiding the devastating downward spiral: The evidence that early intervention prevents reading failure. *American Educator*, *28,* 6-19.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual response to instruction. *Journal of Educational Psychology*, *91*, 579-593.

University of Oregon Center for Teaching and Learning. (2010). *NWF assessment general performance pattern and instructional recommendations*. Retrieved from http://oregonreadingfirst.uoregon.edu

University of Oregon Center on Teaching and Learning. (2015). Retrieved from https://dibels.uoregon.edu/market/assessment/dibels

VanDerHeyden, A., Burns, M. K., Brown, R., Shinn, M. R., Kukic, S., Gibbons, K., Batsche, G., & Tilly, W. D. (2016). Four steps to implement RtI correctly. *Education Week*. Retrieved from http://www.edweek. org/ew/articles/2016/01/06/ four-steps-to-implement-rti-correctly.html

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education*, *41*, 85-120.

Wanzek, J., Roberts, G., Al Otaiba, S., Kent, S. C. (2014). The relationship of print reading in tier 1 instruction and reading achievement for kindergarten students at risk of reading difficulties. *Learning Disability Quarterly*, *37*, 147–159.

Wanzek, J., & Vaughn, S. (2010). Tier 3 interventions for students with significant reading problems. *Theory Into Practice*, *49*, 305-314.

Wesson, C. L. (1991). Curriculum-based measurement and two models of follow-up consultation. *Exceptional Children*, *57*, 246-256.

Wesson, C. L., & Deno, S. L. (1989). An analysis of long-term instructional plans in reading for elementary resource room students. *Remedial and Special Education*, *10*, 21-28.

Wesson, C. L., King, R. P., & Deno, S. L. (1984). Direct and frequent measurement of student performance: If it's good for us, why don't we do it? *Learning Disability Quarterly*, 7, 45-48.

Wesson, C., Skiba, R., Sevcik, B., King, R. P., & Deno, S. (1984). The effects of technically adequate instructional data on achievement. *Remedial and Special Education*, *5*, 17-22.

Wiggins, G. (2012). Seven keys to effective feedback. *Educational Leadership*, *70*, 18-23.

Wiliam, D. (2006). Formative assessment: Getting the focus right. *Educational Assessment*, *11*, 283-289.

Williams, K. T. (2001). *Group Reading Assessment and Diagnostic Evaluation* (GRADE). New York: Pearson.

Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice*, *18*, 137–146. doi:10.1111/1540-5826.00070

Vaughn, S., Cirino, P., Wanzek, J., Wexler, J., Fletcher, J., Denton, C., . . . Francis, D. (2010). Response to intervention for middle school students with reading difficulties: Effects of a primary and secondary intervention. *School Psychology Review*, *39*, 3–21.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*, 2-40.

Vellutino, F., Scalon, D., Silpay, E., Small, S., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, *88*, 601-638.

Scanlon, D. M., Vellutino, F. R., Small, S. G., Fanuele, D. P., & Sweeney, J. M. (2005). Severe reading difficulties—Can they be prevented? A comparison of prevention and intervention approaches. *Exceptionality*, *13*, 209-227.

Yell, M., Deno, S., & Marston, D. (1992). Barriers to implementing curriculum-based measurement. *Assessment for Effective Intervention, 18*, 99. doi:10.1177/153450849201800109

APPENDIX A: TEACHER QUESTIONNAIRE

Teacher:____________________

Date:_____________

Please circle your response for items 1-5 by indicating the extent to which you agree or disagree with each statement. Please respond to items 6-7.

Questions	Responses				
1. Using formative feedback would be feasible to implement in my classroom.	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
2. Formative assessment feedback is important and appropriate for improving student achievement.	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
3. Formative assessment feedback is important and appropriate for students exhibiting problems acquiring early literacy skills.	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4. Formative assessment feedback is helpful deciding when to make an instructional change.	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
5. I noticed meaningful increases in the students' early literacy skills after the implementation of the progress monitoring feedback.	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree

6. What changes/additions would you suggest for formative assessment feedback?

7. Additional comments:

APPENDIX B: INSTRUCTIONAL MODIFICATION LOG

Teacher:			School:			Month:	
Student ID	Instructional Change Needed? (*mark Y/N*)	Actual Change Implementation Date	Focus /Skill	Time	Teaching Strategy	Material	Group Size

APPENDIX C: NONSENSE WORD FLUENCY ASSESSMENT GENERAL PERFORMANCE PATTERN AND INSTRUCTIONAL RECOMMENDATIONS

Strategy	Sound Only (/f/ /e/ /k/)		Sound by Sound and then Recode (/f/ /e/ /k/ /fek/)		Partial Blend (/f/ /ek/)		Whole Word or Unit Reading (/fek/)	
	Not Accurate (< 90% accuracy)	*Accurate (>90% accuracy)*	*Not Accurate (< 90% accuracy)*	*Accurate (>90% accuracy)*	*Not Accurate (< 90% accuracy)*	*Accurate (>90% accuracy)*	*Not Accurate (< 90% accuracy)*	*Accurate (>90 % accuracy)*
List Student Names								
Instructional Implications	- Focus on accuracy instruction at the letter-sound level -	- Focus on blending fluency practice at the word level	- Focus on accuracy instruction at the letter-sound level and then accuracy instruction at the blending level	- Focus on blending fluency practice at the word level -	- Focus on accuracy instruction at the letter-sound level and then accuracy instruction at the blending level	- Focus on blending fluency practice at the word level -	- Focus on accuracy instruction at the letter-sound level and then accuracy instruction at the blending level	- Focus instruction on accuracy and fluency in connected text
Example Activities	- *Continued Phonics Instruction* - Fluency with known sounds * 1 Minute Sound Dash * Rapid Read Sounds	- Instruction in continuous blending of words with known sounds (i.e., Card 9) followed by re-reading the blended words as whole words (i.e., Card 3)	- *Continued Phonics Instruction* - First, Fluency with known sounds * 1 Minute Sound Dash * Rapid Read Sounds - Instruction in continuous blending	- Blending practice in reading words accurately as whole units * No Peeps (e.g., Sound it out in your head…Say the whole word) - Fluency with known words * 5 x 5 matrix	- *Continued Phonics Instruction* - First, Fluency with known sounds * 1 Minute Dash * Rapid Read Sounds - Instruction in continuous blending of words with known sounds	- Blending practice in reading words accurately as whole units * No Peeps (e.g., Sound it out in your head…Say the whole word) - Fluency with known words * 5 x 5 matrix	- *Continued Phonics Instruction* - First, Fluency with known sounds * 1 Minute Dash * Rapid Read Sounds - Then, fluency practice in reading words as whole units	- Fluency building activities in connected text * Repeated Reading Strategies *

APPENDIX D: CONSENT FORMS

UNC CHARLOTTE
College of Education

Department of Special Education and Child Development

9201 University City Blvd, Charlotte, NC 28223-0001
t/ 704.687.8828 f/ 704.687.2916 www.uncc.edu

October, 2015

Dear Parent/Guardian,

We would like permission for your student, ______________________________, to participate in the *Progress Monitoring Feedback* research project with Kannapolis City Schools and the University of North Carolina at Charlotte. As part of this study, your student's teacher will receive individualized feedback on the specific skills he/she needs to be successful. The purpose of this study is to determine the effectiveness of a progress monitoring feedback, which includes your student being assessed multiple times across time. This study will help identify how to best improve the early literacy skills of first grade students. This is important, because early literacy skills are directly tied to understanding what is read. With your permission and your student's agreement, your student will:

- Complete brief (1-2 minutes) tests of reading ability and achievement, during the school day, at school. The tests will be conducted at least once per month and will begin in September 2015 and will continue for 16 weeks.

The tests of ability and achievement your student completes are private and will not be made part of your student's record without your permission. Your student's teacher supports this project and will be looking at your student's work and reading progress reports with the UNC Charlotte staff. We will be happy to share these updates with you at any time. Participation in this study is voluntary. The decision to participate in this study is completely up to you. If you decide to allow your student to be in the study, you may stop at any time. You and your student will not be treated any differently if you decide not to participate the study, or if you stop once the project has started. Your decision for your student to participate (or not) in the project will not affect any future services the district may provide for your student.

Although some risks are unforeseeable, the risks of participation in this study are minimal to none. The only area of risk could potentially be breach of confidentiality. Identifying student information will be removed and replaced with a number. When presenting findings, names will be replaced by numbers in publications.

The potential benefits of participation in the study include determining components of feedbackto increase academic achievement for students in first grade. In addition to the potential global benefits of this study, your student may make achievement gains bringing him/her to benchmark levels, reducing the potential risk for reading difficulty, and eliminating the need for supplemental intervention.

Please feel free to contact your school principal, your student's teacher, ____________________, or the Project Coordinator, Mr. Jeremy Lopuch, at (704) 560-1202 if you have questions. If you have any comments or questions regarding the conduct of this research, please contact Dr. Lindsay J. Flynn at (704) 687-8829. UNC Charlotte wants to make sure you and your student are treated in a fair and respectful manner. Contact the university's Research Compliance Office (704-687-3309) if you have questions about how your student is treated as a study participant. Please complete the attached form if you would like to participate in the study and return it to your student's teacher immediately. Thank you.

Sincerely,

Jeremy Lopuch

Doctoral Student, UNC Charlotte

I have read the information on this consent form. I have had the chance to ask questions about this study, and those questions have been answered to my satisfaction. I am at least 18 years of age and allow my child to participate in this research project. I understand that I will receive a copy of this form after it has been signed by me and the principle investigators of this research study.

____

Child’s name (PLEASE PRINT)

____

Parent’s name (PLEASE PRINT)

____

Parent’s signature & Date

____

____

Investigators’ signatures & Date

UNC CHARLOTTE

College of Education
Department of Special Education and Child Development

9201 University City Blvd, Charlotte, NC 28223-0001
t/ 704.687.8772 f/ 704.687.2916

October, 2015

Dear Teacher,

I am currently pursuing a Ph.D. in Special Education at the University of North Carolina at Charlotte. As part of the requirements for my dissertation in the doctoral program I will be conducting *Effects Progress Monitoring Feedback on Early Literacy Students Achievement*. As part of my research, I will provide first grade teachers with training on how to use DIBELS NWF progress monitoring to improve student achievement.

At the conclusion of the study I would like to have first grade teachers complete a questionnaire about using progress monitoring feedback in their classrooms. All information collected will be kept anonymous and your answers will be discussed in the research article once the study is complete.

You are a volunteer. The decision to participate in this study is completely up to you. UNC Charlotte wants to make sure that you are treated in a fair and respectful manner. Contact the University's Research Compliance Office (704-687-1871) if you have any questions about how you are treated as a study participant. If you have any questions about the project, please contact Jeremy Lopuch at 704-560-1202 or Dr. Lindsay Flynn at 704-687-8829.

I sincerely appreciate you time. If you have any questions regarding this form or the study, please do not hesitate to call me.

Sincerely,

Jeremy Lopuch

If you want to participate in this study, please sign your name below.

____ ____________

Signature of Teacher Date