BUILDING TUTOR CAPACITY FOR IMPLEMENTING CURRICULUM-BASED MEASURES OF ORAL READING FLUENCY

A Capstone Report

Presented to

The Faculty of the School of Education and Human Development

University of Virginia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

---

by

Lauren E. Hauser, M.T., Ed. S.

May 2023

Abstract

Recent trends in reading assessment data indicate that many elementary-school students are not getting the instruction they need to ensure adequate progress. (U.S. Department of Education, 2019). Emerging evidence supports data-based individualization (DBI) as an effective way to prevent or reduce reading difficulties. (Austin & Filderman, 2020; Filderman et al., 2018; Filderman & Gesel, 2022; Filderman & Toste, 2017; Jung et al., 2018; Lemons et al., 2019). Curriculum-Based Measures, or CBMs, are a commonly-used data source for DBI and must be administered with fidelity. At a mid-Atlantic university-based reading center, previous anecdotal observations suggested that tutors experience difficulty administering CBMs with fidelity, making it difficult to gauge student progress. This qualitative case study sought to begin to address this problem through a systematic examination of the efficacy of tutor training practices. Data were collected in the form of training module artifacts, observation checklists with field notes, and semi-structured interviews to learn about common challenges to fidelity, how tutors' varied backgrounds impact fidelity, and how DRC might improve tutor training practices. Major findings include:

1. Challenges in achieving assessment fidelity often stem from tutors' care for their tutees' comfort and confidence.
2. While tutors were already invested in their tutees' success, the training process heightened their investment in measuring success through CBMs.
3. Many tutors made relevant connections across the data they collected, the instruction they provided, and next steps.

These findings inform recommendations about future practices in tutor training as well as implications for DBI in the context of the university-based reading center.

Keywords: tutoring, literacy, data-based individualization, curriculum-based measures, progress monitoring, oral reading fluency, assessment fidelity, clinical practice

Table of Contents

Dedication

For my Dad—always finish what you start.

Acknowledgments

I am humbled by the many people who have given their time and energy to help me grow as an educator and as a scholar.

To my committee chair, Dr. Tisha Hayes—words cannot express my gratitude for your support, your insightful feedback, and your humor throughout the capstone process. I have been learning from you for more than a decade, and I am so fortunate to continue to learn from you every day. Thank you for being a role model, guide, and friend.

To my committee members, Dr. Lysandra Cook and Dr. Gail Lovette—thank you for your careful attention to my work and for helping me navigate the challenges I encountered along the way. You both demonstrate how advocacy for children can infuse teaching and scholarship, and I am grateful for your example.

To my peers, especially Emma Pearson and Austen Hecker—you all have been the most wonderful thought partners throughout the course of this program. I am thankful for all your moral support and wisdom.

To the leadership and staff in EHD and the Sheila Johnson Center—it takes a village to sustain our tutoring work, and I am grateful for your support with every detail, from reserving space to organizing registration forms to hiring tutors to fundraising. Thank you for your work that impacts the community and enables my scholarship.

To the tutors in our program—thank you for being open to learning new practices to better serve children. Through your hard work and dedication, you truly make a difference for children and their families, and you contribute to my learning as well.

To the children, and to the families who share them with us—I value our partnership, and I am grateful for the energy, determination, and joy that you bring to our work together.

To my family, especially Zach and Jude—you are simply the best. Thank you for the boundless support, sacrifice, and snacks.

List of Figures

List of Tables

Chapter 1: Introduction

The ability to read proficiently enables access to knowledge and ideas. When children learn to read, they acquire tools for learning about the world around them and the perspectives of others; they also experience the power of imagination and narrative in new ways. As they grow, literacy forms their foundation for advanced educational opportunities and full social participation. Because learning to read is not a natural process, children rely on instruction to obtain these crucial literacy skills. They also rely on educators to determine whether their instruction meets their students' needs and to adjust accordingly to ensure their continued growth.

For some students, evidence-based classroom instruction is sufficient to enable reading proficiency. However, others require more intensive support to make adequate progress. For those students, the process of data-based individualization (DBI) holds promise (See Appendix A for key terms as defined in this project). DBI is a research-based process that utilizes systematic data collection to personalize instruction and intensify it according to student needs (National Center on Intensive Intervention [NCII], 2013). To yield data useful for DBI, tools and methods must be appropriate to the context and implemented with fidelity, or in accordance with standardized procedures (Bundock et al., 2018; Christ et al., 2013; Hosp et al., 2016; Reed & Sturges, 2013). However, achieving assessment fidelity is a challenge, both globally and in the local context of this capstone, a university-based reading center (e.g., Reed & Sturges, 2013). This capstone explores the process of training uncertified tutors at Douglass Reading Center (DRC)[1] to collect high-quality data using curriculum-based measures, or CBMs (Hosp et al., 2016). By learning more about the factors impacting tutors' administration and scoring of CBMs,

[1] Names of all places and people are pseudonyms

DRC can improve training practices, laying the foundation for long-term organizational goals: (a) using data to communicate tutees' progress with key stakeholders and (b) adjusting instruction to improve tutee learning outcomes.

Trends in Literacy Achievement

Across national, state, and local levels, reading achievement data suggests that current instructional practices are not sufficient for all students to become proficient readers. In Virginia, for example, little change is evident in reading scores on the National Assessment of Educational Progress (NAEP) for fourth graders over the last decade, and average scores for students who identify as Black or Hispanic are consistently lower than those of their white and Asian peers (U.S. Department of Education, 2022). For students with disabilities, prevailing instructional practices have yielded an average score that falls short of even "basic" proficiency in reading (U.S. Department of Education, 2022). Further, scores on the Phonological Awareness Literacy Screening (PALS) K-2 show that, statewide, an elevated percentage of children began the Fall 2021 school year at high risk for reading difficulties as compared to 2019 and 2020 (University of Virginia School of Education and Human Development, 2021). These findings demonstrate that the COVID-19 pandemic and associated disruptions to learning had a significant impact on students' reading development.

Addressing the Challenge through Tutoring

Given the challenges that students are facing, alternative settings for providing instruction beyond the regular school day could provide valuable support. Tutoring programs have been shown to positively impact tutee achievement (Allor & McCathren, 2004; Baker et al., 2000; Denton et al., 2010; Jacob et al., 2016; Neitzel et al., 2022; Samson et al., 2015; Vadasy et al., 2002). This is especially true when tutoring occurs with a small tutee-to-tutor ratio, especially

one–to-one (Bloom, 1984; Kirschner & Hendrick, 2020). In a one-to-one tutoring environment, tutors are able to appropriately scaffold their instruction by recognizing which step the tutee is ready to take next and providing the necessary guidance for them to take that step, ultimately helping the tutee take individual steps toward a desired outcome. (Kirschner & Hendrick, 2020; Wood et al., 1976). Because this process is highly individualized, the tutorial setting supports this type of instruction. Along with scaffolding, tutors are also able to individualize pacing, allowing their tutees to learn content to mastery (Bloom, 1984; Kirschner & Hendrick, 2020).

These positive effects of tutoring can be obtained even when the tutors are not certified teachers. In a 2022 meta-analysis, Neitzel et al. examined studies of tutoring programs across a series of parameters, including tutor qualifications, in an attempt to identify actionable recommendations for practitioners. For one-to-one tutoring, they found that impact was comparable across tutor groups with different levels of expertise, whether they were certified teachers, teaching assistants, or paid volunteers (e.g., AmeriCorps members and tutors whose time was donated by employers). A one-to-one tutoring program requires a dedicated tutor for every tutee served. Therefore, the finding that uncertified tutors, even paid volunteer tutors like those at DRC, can impact their tutees' achievement offers promise for programs seeking to provide effective services within personnel and budget constraints.

Factors that Make Tutoring Successful

For successful programs employing uncertified tutors, such as paid volunteers, Neitzel et al. (2022) credit extensive tutor training and a well-structured tutoring model. Other researchers also emphasize the importance of tutor training and a supportive lesson structure. For example, in Allor and McCathren's (2004) effective reading intervention, college student tutors received extensive training, and the training model was refined by the researchers across two iterations.

Additionally, because the lessons themselves were highly structured and contained key components such as scaffolding and cumulative review, tutors without a literacy background were able to implement instructional best practices. In a synthesis of studies that employed paraprofessionals as tutors, Samson et al. (2015) also highlight the need for extensive training and supervision, with training taking the form of modeling, co-teaching, and a review of the components of reading. Like Allor and McCathren (2004), Samson et al. (2015) concluded that a highly structured program, which removes the need for guesswork among uncertified tutors, is a key characteristic of successful one-to-one tutoring efforts.

Another key characteristic of successful tutoring is responsiveness of instruction to tutees' needs. Across studies, researchers attribute tutee gains to the ability of the tutor to adapt instruction based on the individual learners before them (Allor & McCathren, 2004; Morris, 2022; Wasik & Slavin, 1990). Through their QAIT model of best practices in tutoring, Wasik and Slavin (1990) explain the importance of four key elements for success: Quality of instruction, Appropriate level of instruction, Incentive, and Time. To paraphrase, tutors must present appropriate content in a format that is clear and accessible for the tutee. They must also allocate time appropriately to balance new material, practice, and review within each lesson as well as pace content across the days and weeks. When tutors have the capacity to meet these criteria, tutoring programs, even when they employ uncertified tutors, can help their tutees build proficiency in reading.

Addressing the Challenge through Data-Based Individualization

Fulfilling the QAIT criteria for effective tutoring is possible only when tutors have a sense of what their tutees know, what they are ready to learn next, and how they are responding to current instructional practices. Without the ability to gauge a tutee's current learning and adapt

instruction accordingly, tutoring practices may not have the desired impact on progress. One process for approaching this is DBI (Austin & Filderman, 2020; Filderman et al., 2018; Filderman & Gesel, 2022; Filderman & Toste, 2018; Jung et al., 2018; Lemons et al., 2014). DBI is derived from the process of data-based decision making (DBDM). Many sources use the terms interchangeably (Filderman et al., 2018); however, DBI is distinct because of its narrower focus on instructional decision making for a particular student. Typically, when applied to literacy, the process of DBI begins with identifying students who require personalization or intensification of their instructional program (NCII, 2013). Their instructional program then serves as the "platform," or starting point, that the tutor will modify through the individualization process. The tutor uses diagnostic data alongside ongoing progress monitoring data to determine the student's response to the instruction, deciding whether to make modifications and what those modifications should be. Aligning instruction to the student's individual needs and responding as those needs change enables improved outcomes (Bambara & Kern, 2005; Fuchs et al., 2008).

Evidence Base for DBI

In general, modifying instruction in response to data is a promising practice, but few studies exist to empirically validate the process as a whole. In the abstract, DBI may sound like a straightforward, logical process, but it cannot be studied in a vacuum, and its key components cannot be taken for granted in the context of real educational settings (Barnes et al., 2022). True DBI implementation assumes that educators have access to high-quality data that is suited to their instructional purpose and that they have the knowledge and skill to interpret that data and adapt practices accordingly. Further, DBI is contingent upon context-based systems and structures that support data collection and use. Creating a foundation for DBI requires a high level of commitment from an educational context, a unified sense of purpose among the staff,

and a robust professional development agenda to support implementation. Within complex systems, this foundation may not be in place, which compromises efforts to research the practice; therefore, drawing conclusions about the effectiveness of DBI as a process can be challenging.

Despite these challenges, evidence is emerging that DBI can improve student reading outcomes (Filderman et al., 2018; Filderman & Toste, 2021). This interpretation is offered with caution, because effect sizes in empirical studies of DBI tend to be modest, and the overall evidence base may be described as "tenuous" when interpreted relative to general guidance in the social sciences (Cohen, 1988; Filderman et al., 2018). However, frameworks for interpreting effect size do not offer absolutes, and a simple categorization of an effect size as "small," and subsequent dismissal of an intervention, could obscure practical benefits for students (Cook et al., 2009). The students participating in studies of DBI have experienced substantial challenges on their path to reading proficiency and have not responded to typical instructional practices. For these students, the relative practical impact of any progress they might make is significant.

Such progress was evident across studies in a 2018 synthesis and meta-analysis by Filderman et al. The researchers identified 15 studies of DBDM practices (DBDM) in reading for students in grades K-12. Of these studies, only six included control conditions to isolate the impact of DBDM. Across these studies, the weighted mean effect size (.27) was modest but could translate into real gains for students, particularly for secondary students whose progress is especially hard-won.

Further evidence in support of DBDM emerged from a randomized controlled study by Filderman and Toste (2021) which found a significant positive impact when employing the practice with upper elementary students. One group of students received an intervention focused on multi-syllabic word reading, and, for them, instruction was adjusted partway through the

intervention period based on their progress. These students were compared with students who received the same intervention but without any mid-stream adjustment, as well as students who did not receive the intervention at all (business as usual or BAU condition). The researchers found that, on a standardized measure of nonsense word decoding, which mirrored the content of the intervention, students whose instruction was modified in response to data significantly outperformed their peers in the other intervention group. In fact, the students whose intervention was not adjusted performed comparably to the students in the BAU condition on this measure. In other words, it was not the content of the instruction that made the difference; rather, the key was the individualization of that instruction to align with students' specific needs. While further research is needed to determine the impact of DBI across skills, age groups, and contexts, studies such as this one demonstrate that it is a promising avenue for boosting reading achievement.

CBMs and DBI

Within a DBI framework, curriculum-based measures, or CBMs, are one common source of data for monitoring student progress toward instructional goals (Hosp et al., 2016). In reading, CBMs serve as a general outcome measure (GOM) by assessing areas that correlate with overall reading proficiency, such as nonsense word decoding or oral reading fluency. CBMs have highly standardized administration protocols and are designed to be administered quickly and frequently. Because each measure has many equivalent forms, tutors can use them to obtain multiple comparable data points and gauge student performance over time.

As is the case for DBI, the evidence base for use of CBMs within a DBI framework is slim but promising. CBMs in and of themselves have demonstrated technical adequacy, but most studies utilizing them focus on their use in identifying students at risk for long-term reading difficulties, rather than monitoring their progress. Encapsulating this point, in their 2018

synthesis and meta-analysis, Filderman and Toste found that, of the 15 high-quality DBI studies they identified, only three actually incorporated CBMs for their data collection. That said, a 2018 meta-analysis (Jung et al.) found that incorporating CBMs into DBI procedures had a moderately positive effect size (0.37) on student reading outcomes. According to these authors, the challenge with determining the impact of CBMs is that myriad variables affect their efficacy, including the appropriateness of the selected CBM to the task and student population; the frequency and duration of data collection; and the degree of support provided to the tutors administering and interpreting them. Ultimately, the use of CBM data within a DBI framework has the potential to improve student reading achievement; however, success is contingent on a series of factors that are highly specific to individual contexts, and practitioners must employ them with a reflective lens.

Statement of the Problem

Given the promise of both the DBI process and one-on-one tutoring frameworks to improve reading outcomes for tutees, it would be of benefit to DRC to gather data on the process of training uncertified tutors to collect progress monitoring data. According to both DRC's Director and Lead Clinician, this practice would align with the organizational goal of improving reading outcomes for children using evidence-based practices in instruction and assessment (L. Hayes, personal communication, July 2022).

However, administering CBMs with fidelity presents a challenge, both globally and within the local context, particularly with the recent scaling up of tutoring services and the time constraints on hiring and training uncertified tutors. For the DBI process to function as intended, data collection procedures must yield information that accurately represents tutee progress; therefore, these procedures must be implemented with fidelity to standardized protocols set forth

by the test developers. Although previous attempts at progress monitoring in this way have been made at DRC, it was observed at the time that tutors experienced difficulty administering and scoring CBMs. Further, although they collected data consistently for an initial period, the regularity of the administration tended to lapse for some tutors as the semester progressed, resulting in an incomplete picture of tutee progress. A systematic investigation of this problem could assist reading center staff in improving the quality of their data collection and gaining a better understanding of tutee progress tutees so they can make adjustments to their instructional program in accordance with DBI procedures. Further, findings from the study would lay the groundwork for efforts to draw larger conclusions about program efficacy to be communicated with university and community stakeholders.

Local Setting

The context of this study is an after-school tutoring program offered by DRC, a university-based reading center in the mid-Atlantic region. DRC has a Director, who provides insights on organizational goals, as well as a Lead Clinician, the researcher, who coordinates tutoring services. Each semester, approximately 50 school-age tutees (representing grades One through Eight) participate in one-on-one lessons with a consistent tutor. The lessons occur twice per week and last approximately 50 minutes. There are sessions in the fall and in the spring that correspond with academic semesters, and each session lasts approximately nine weeks. During the fall and spring sessions, each tutee works with their tutor for a total of 18 lessons per session, or 36 lessons in total across the academic year.

Prior to the COVID-19 pandemic, all tutoring took place in person, and all tutees were local to the university; although the focus of this inquiry is the in-person tutoring, it should be noted that DRC now offers a virtual option and serves tutees from a wider geographic range. The

program serves tutees who have instructional needs at the word level (i.e., challenges in phonemic awareness and word knowledge that impair word recognition and overall reading fluency); although some tutees may also engage in targeted comprehension work based on their assessed needs, the core instructional focus of the lessons is code-based, rather than meaning-based.

With the advent of the pandemic, more families requested reading intervention services, and the number of tutees served has increased by 200%. As more parents and families have become concerned about their children's interrupted learning opportunities, they have sought additional support, and demand for tutoring through DRC has risen. DRC personnel have worked to meet the demand by expanding the size of the tutoring staff from approximately a dozen to over 30 individuals. Historically, the tutors at DRC have been full-time, on-campus students in the Reading Education program at the university, and they participated in coursework to support their intervention work. However, over the last several years, enrollment patterns in the Reading Education program have shifted, and many students have opted to complete the program part-time through a distance-education model.

Because of this change, current tutors are best classified as what Neitzel et al. (2022) refer to as "paid volunteers." While a few are hired directly through DRC, most applied to and were hired through America Reads, a Federal Work-Study program. These tutors were awarded work-study as part of their financial aid package to attend college or graduate school, and they elected to earn that work-study award through hourly work in educational settings. The America Reads program coordinator places the tutors at a variety of school and community-based sites according to their schedules and their interests; DRC is one such site.

The tutors hired by or placed at DRC represent a range of backgrounds and qualifications. While some are enrolled in graduate programs, others are undergraduates. Some are pre-service teachers or are studying in a closely related field, such as school counseling or speech/language pathology (SLP). At the same time, others are majoring in other disciplines such as psychology, public policy, or engineering. Some have extensive experience working with children in formal instructional settings, while others have more limited experience (e.g., as a camp counselor or after-school program assistant). Generally, the tutors hired by or placed at DRC are selected based on experience working with children, a professional demeanor, and an enthusiasm for the high level of commitment required to work one-on-one with a tutee experiencing challenges on their literacy journey.

Once selected, the tutors undergo training with a reading specialist about how to implement scripted structured literacy lessons from evidence-based programs that include tutor modeling, guided practice, and intentional scaffolding (Denton et al., 2016; Solari et al., 2018). These lessons build word-level skills and oral reading fluency in alignment with each tutee's areas of need. Periodically, tutors complete a brief assessment of mastery that provides a snapshot of whether the tutee is learning the content of the lessons. Based on tutee performance, the tutor decides to either move on to new material or to reteach material and provide additional practice. In this way, the tutor individualizes pacing and ensures that the tutee is learning content to mastery. However, while these mastery measures indicate that the tutee is learning the specific material taught, they do not show whether the tutee is generalizing the instruction to progress in the overall domain of reading, or where they stand relative to grade-level benchmarks. For this reason, supplementation with additional data obtained through CBMs would be of use in this context.

Stakeholder Impact

Obtaining high-quality progress monitoring data would be of benefit to stakeholders both within the university and the broader community, including tutees, families, tutors, and the program overall. The most important outcome of this inquiry would be a foundation for engaging in DBI practices, which impacts all stakeholders. If the data indicates that a tutee is making adequate progress, the tutor could adjust the pacing with which they move through the scope and sequence, offer more challenging texts to the tutee, or decrease the amount of instructional scaffolding they provide. The tutor could also engage in conversations with the tutee around goal setting and the progress they are making toward their goals. Tutees and tutors would be motivated by seeing progress (Al-Bataineh et al., 2019; Cabral-Márquez, 2015), and this increased motivation could result in greater investment in the tutoring sessions, both on the part of the tutors and the tutees. Families would benefit from more explicit and data-based feedback that their child's reading has improved to the extent that tutoring with a code-based focus is no longer the best fit for their instructional needs. Families commit time and financial resources to obtaining academic support services for the tutees, and demand for services is high; having a clear process by which a tutee can "graduate" from the tutoring program would therefore benefit multiple stakeholders.

In many ways, this project is even more important for the tutees who are not making adequate progress in the tutoring program. For these tutees, there is a mismatch between the instruction they need and the instruction they are receiving, and a process must be in place to determine this as quickly as possible through data collection and analysis. For these tutees, instructional delivery intensifications may be utilized (Filderman & Gesel, 2022; Lemons et al., 2014; Wanzek et al., 2019). For example, the Lead Clinician at DRC could engage in a more

rigorous cycle of lesson observations with a fidelity checklist and coaching sessions with tutors to ensure the implementation fidelity of the program. The Lead Clinician could also recommend additional adjustments, such as more extensive cumulative practice, additional time in decodable text, or use of manipulatives. Along with intensification of instructional delivery, organizational intensifications such as additional sessions per week could be suggested to families.

Importantly, if data indicates a need for intensification, available options for intensification within the tutoring program are limited by the expertise of the uncertified tutors, scheduling considerations, and supervisory capacity. Ultimately, the student may have instructional needs that cannot be met within the program structure. In that case, data would be useful for explaining to parents why they might consider engaging a private tutor who may have greater expertise, a more flexible schedule, and more opportunity to interface with the student's teachers to ensure instructional alignment. Further, some families may be working with their child's school to establish an Individualized Education Program (IEP), and this process includes an examination of the student's achievement trends. If assurance can be provided that progress monitoring data were collected according to standardized procedures, the center could provide the data to families to aid in the IEP process. This type of support is only possible in the local context if a system is developed to obtain high-quality data on student progress.

Finally, the center and the university reading program would benefit from improved data collection procedures to help demonstrate the overall efficacy of the program. Scaling up from the level of the individual student, collecting these data could help the center detect and adapt to any larger achievement patterns that may be present. Before enrolling their students in the program, families sometimes inquire about what growth they should anticipate, and having student progress data to report could help clarify expectations. Similarly, when schools or other

community organizations inquire about potential partnerships, sharing data could help increase the legitimacy of services. Given the potential impact of data collection on this range of stakeholders, the center should explore the feasibility of training tutors to administer progress monitoring measures in accordance with standardized procedures. This inquiry would align with the organizational goal of the tutoring center: to establish a system of data collection foundational for DBI and, therefore, to support students' reading achievement through the responsive use of evidence-based practices.

The Problem

While engaging in practices of data collection and analysis would be of benefit to these various stakeholders, ensuring that the assessments are administered according to standard procedures is a significant challenge, particularly in the local context of DRC. To ensure that the data collected through CBMs accurately represents the tutees' progress, the CBMs must be administered with fidelity, or in alignment with the practices designed by the developers of the assessment. However, achieving assessment fidelity requires extensive training and ongoing support. For DRC tutors, who do not generally have background knowledge about assessment, achieving and maintaining assessment fidelity poses a particular challenge, undermining the organizational goal of building a foundation for DBI practices at DRC.

Fidelity Defined

The challenges of implementation fidelity are evident not just in the local context, and not just with progress monitoring, but more broadly across settings and practices. Generally, the term "fidelity" refers to "the extent to which an enacted program is consistent with the intended program model" (Century et al., 2008, p. 2). Within the context of a study, if an intervention is not implemented according to the intended practice, the relationship between the program and

the outcome becomes difficult to establish (Harn et al., 2013). The concept of fidelity is relevant across research in fields beyond education; for example, in the medical field, researchers examining treatment outcomes must establish "treatment integrity," or the degree to which the treatment was implemented as intended; otherwise, data showing the efficacy or lack thereof of a given treatment is called into question (Perepletchikova & Kazdin, 2005).

Fidelity is not just important within the context of a study; it also becomes important when considering how to bridge the findings of an intervention study to practices in authentic contexts. In education, for many years, researchers examining the results of an intervention did not examine the implementation fidelity of their findings in schools; rather, they operated under the assumption that adopters of practices were passive recipients, rather than active modifiers of those practices (O'Donnell, 2008; Rogers, 2003). However, beginning in the 1970s, researchers began to recognize that fidelity of implementation could not be assumed and that educators took liberties in adapting an intervention to meet the needs in their contexts (O'Donnell, 2008; Rogers, 2003). Therefore, over time, bodies such as the U.S. Department of Education began to place more of a premium on not only identifying evidence-based practices to improve student achievement, but also on achieving those same outcomes outside of a tightly controlled research setting (Harn et al., 2013).

Assessment Fidelity

Amidst widespread attempts to define and measure implementation fidelity, Reed and Sturges (2013) have argued that one particular facet of implementation fidelity in research has been largely ignored: the degree to which the administration and interpretation of pre- and posttests aligns with standardized procedures. They refer to this facet as assessment fidelity. While the concept of assessment fidelity applies to general educational contexts, Reed and

Sturges' (2013) study is of particular relevance to the DRC context because they examined Oral Reading Fluency (ORF) CBMs, which are commonly used and extensively researched. Although CBMs have many advantages for progress monitoring, their utility diminishes when they are not administered according to recommended procedures, because the validity of the data obtained are jeopardized (Bundock et al., 2018; Christ et al., 2013; Hosp et al., 2016; Reed & Sturges, 2013). Therefore, users must attempt to implement them with as much fidelity to the recommendations as possible.

With ORF CBMs, researchers have noted that, even though all passages within a set (e.g., all individual passages within a set of Third Grade passages) are meant to be equivalent, some variability is evident in student performance (Cummings et al., 2013; Hintzie & Christ, 2004; Poncy et al., 2005). However, the degree of variability is diminished when the passage sets are used with their intended population, readers experiencing difficulty, as opposed to when they are used with a broader population (Bundock et al., 2018; O'Keeffe et al., 2017; Tindal et al., 2016). In the context of this study, the tutoring students represent an appropriate population for the use of ORF CBMs, so that aspect of implementation fidelity is not of concern.

Of greater relevance to the local study context are what Bundock et al. (2018) refer to as setting factors, which comprise where the measures are administered; who administers them; and to what extent the administrators utilize the standard protocol for administration and scoring. In one study, which compared ORF administration and scoring across several administrators and settings, student performance varied significantly (Derr-Minneci & Shapiro, 1992). Further, even among trained administrators, errors in administration procedure and scoring were common (Reed & Sturges, 2013). Importantly, even when administrators were trained to fidelity, adherence to standardized procedures was found to diminish over time if refresher training was

not provided (Reed & Sturges, 2013). As the quality of the data declines, so too does the utility of that data for presenting an accurate picture of student progress.

Assessment Fidelity at DRC

While the challenges of assessment fidelity have been documented across a variety of settings, several exacerbating elements are present in the local context such as the recent scaling-up of tutoring services and the time constraints on hiring and training uncertified tutors. As previously mentioned, the number of students served by the reading center increased by 200% in 2020, and, since that time, student enrollment has held steady. Prior to the pandemic, center staff were engaging in elements of DBI and working to develop a comprehensive process to do so in a sustainable way. However, scaling up these procedures to the current program size has proven difficult, especially without a commensurate increase in supervisory capacity to establish and maintain a system as well as to oversee the execution of the process.

Further, the DRC Lead Clinician is not involved in hiring tutors as the hiring process runs through the America Reads work study program, and the timing of the hiring process often puts a time stressor on sufficient, adequate training for new tutors. While evidence demonstrates that uncertified tutors are capable of getting results with the proper supports (Allor & McCathren, 2004; Baker et al., 2000; Jones et al., 2016, Neitzel et al., 2022), thorough training and structure is crucial to their success. Unfortunately, the varying and sometimes competing schedules of the tutors limit opportunities for synchronous training, and new tutors have a steep learning curve to build background knowledge about literacy development, lesson implementation, and basic DRC logistics.

Because the tutors are undergraduate or graduate students, many from fields outside of education, they do not come to DRC with prior knowledge of progress monitoring and why it is

important. They have chosen to work in the center because they want to engage with children, and elements of CBM administration may seem overly clinical to them, especially if the student they work with experiences undue difficulty with the passage reading. Without a full understanding of the purpose and procedures of administering the ORF CBM, tutor fidelity of administration can falter, as has been evident in prior attempts at employing CBMs in the context of the DRC (L. Hayes, personal communication, July 2023) Without addressing these challenges, the center will not be able to achieve the organizational goal of establishing a process for data collection that will enable instructional adjustments and clear communication with families around student progress.

Research Questions

To learn more about how to make progress toward this organizational goal, I will investigate the following research questions:

1. What challenges do DRC tutors experience when administering and scoring CBMs?
2. In what ways do a DRC tutor's knowledge, skills, and beliefs impact their administration and scoring of CBMs?
3. In what ways can DRC improve current CBM training practices to build DRC tutors' capacity for administering and scoring CBMs?

Conceptual Framework

Because my study focused on the tutor training process, I approached the inquiry through the lens of adult learning theory. In addition, I drew on literature related to progress monitoring, educator professional development, and implementation fidelity. This body of theory and research guided my methodological choices, data collection and analysis, and subsequent recommendations.

Theoretical Framework: Adult Learning Theory

Disciplines such as educational leadership and professional development widely use "adult learning theory" as a term, and it is often alluded to in recommendations for programming in these areas (Alford, 2013). However, the term does not necessarily represent a unified, cohesive theory; rather, it is "a mosaic of theories, models, and explanations that, combined, compose the knowledge base of adult learning" (Merriam, 2001, p. 3). A unified theory of adult learning, or andragogy, is challenging to construct because the nature of adult learning varies widely across contexts and individuals (Merriam, 2001). Further, some theorists argue that the distinction between how adults learn and how children learn, or pedagogy, is flawed, or even false (e.g., Jarvis, 1992; Knudson, R. S., 1980; Merriam, 2001; Rachal, J. R., 1983).

That said, as it is most commonly employed, the term adult learning theory represents the ideas proposed by Malcolm Knowles in the 1950s-1980s. Knowles himself was influenced by scholars and theorists who emphasized the importance of adult educators being "facilitators of learning" rather than "teachers" (Alford, 2013, p. 126). Ideas from Paolo Freire, including that education should be "consciousness raising," feature in Knowles' work (Freire, 1970, p. 81). Knowles was also influenced by John Dewey's ideas about experiential and interactive learning, though, unlike Dewey, Knowles maintained that these traits were distinctly important for andragogy rather than for pedagogy (Alford, 2013).

Knowles' notion of andragogy also was derived directly from his experiences working with adult learners (Alford, 2013). Noting that the instructional strategies most often used with children in the early to mid-twentieth century were ineffective with adult learners, resulting in a high dropout rate, he came to believe that andragogy stood in contrast to pedagogy. He noted that adults need to know the "why" of what they are learning and that they bring a wide range of

prior knowledge and experiences to the table that impact their learning. Ultimately, Knowles' (1984) observations coalesced into six major assumptions regarding adult learners:

1. The need to know: Adult learners need to recognize a gap in their knowledge or skill and see the value in closing that gap; this encourages them to be self-directed, or to take an active role in the learning process.
2. The learner's self-concept: When working with adults, instructors must acknowledge their need to be self-directed, rather than trying to impose their will on their students. Knowles acknowledges that, in some contexts, this may not be possible, but the instructor should nonetheless look for opportunities to allow for students' independence even within those contexts.
3. The role of the learner's experience: Because adults come to the learning event with a diverse range of backgrounds and experiences, it is incumbent upon an instructor to view diversity as an asset and incorporate it into the learning process.
4. Readiness to learn: Adults learn best when met where they are developmentally with appropriate instructional scaffolding.
5. Orientation to learning: Problem-based or task-based learning is helpful for adults so that they can see the relevance of their learning and apply it to "real-life" situations.
6. Motivation: Adults are intrinsically motivated and more apt to learn for the sake of achieving a goal or for their own sense of self-worth.

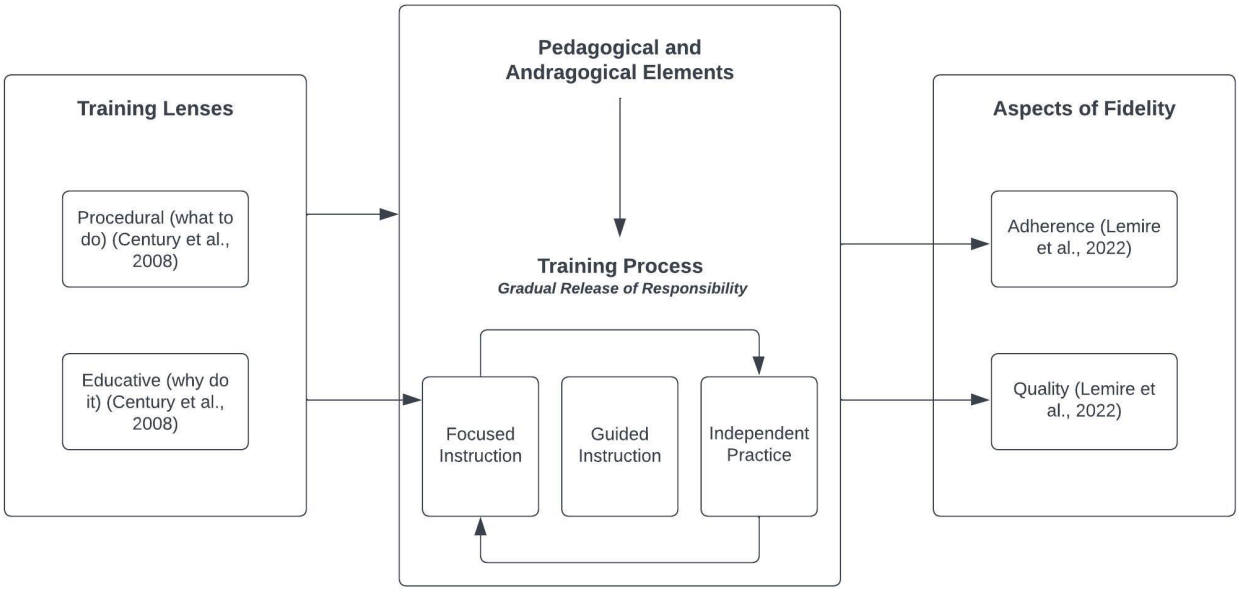
Based on these assumptions about adult learners, Knowles and educators who adopt this set of assumptions have espoused particular principles for designing adult-centered instruction. To capitalize on the strengths and dispositions of adult learners, adult-centered instruction should be clearly connected to relevant tasks; problem-based; encompassing of the prior experiences of the

learners; and self-paced (Halpern & Tucker, 2014). As he refined his thinking across his career, Knowles softened his stance that these andragogical elements must be used exclusively with adults while pedagogical elements should be reserved for educating children; rather, elements of andragogy and pedagogy should be used flexibly but thoughtfully, when appropriate, regardless of the age of the students (Alford, 2013). Subsequent critics of Knowles take this further, arguing that these andragogical assumptions simply align with a constructivist philosophy of education prevalent in K-12 settings and that little daylight is visible between andragogy and pedagogy (Halpern & Tucker, 2014).

Regardless of terminology, attention to these assumptions about adult learners can support the development of learning experiences that tap into the dispositions and assets of a diverse range of learners. Given that the tutors at DRC represent an array of backgrounds and experiences, and that they attend a university with rigorous academic standards, adult learning theory is the lens through which I frame this study. By considering adult learning theory as I examined the impact of tutor training on assessment fidelity, I collected and analyzed data that will be relevant for improving future iterations of the training process, leading to actionable recommendations for DRC and supporting organizational goals. The assumptions and instructional principles of adult learning theory are interwoven within my conceptual framework (Figure 1.1). By considering how adults learn, I developed a conceptual framework mapping the relationship between tutor training and assessment fidelity.

Figure 1.1

A Framework for Building Tutor Capacity for Implementing CBMs with Fidelity



To explain my conceptual framework, the following sections will describe each element from Figure 1.1 and discuss the relevance of adult learning theory for each component, as relevant.

Aspects of Assessment Fidelity

To begin, it is important to consider the goal of the tutor training, which is to build tutor capacity for high-quality data collection using CBMs. Successful training will culminate in tutors administering CBMs with fidelity, which, in the realm of assessment, is generally defined as performing and scoring assessments as intended by the developer (Reed & Sturges, 2013). However, some researchers who consider implementation fidelity in general, rather than assessment fidelity in particular, have developed frameworks that broaden the meaning of fidelity beyond this common conception by distinguishing "adherence," or following the standard procedures of an intervention, from "quality," or conveying investment in that intervention (Century et al., 2008; Dane & Schneider, 1998; Lemire et al., 2022).

Adherence

In these frameworks, following the procedures of an intervention as intended is just one aspect of fidelity, generally termed "adherence." In general, fidelity checks for administering assessments attend to only the adherence aspect of fidelity. This makes sense because elements of adherence are (a) directly observable and (b) generally objective. In the context of progress monitoring at DRC, adherence pertains to tutors accurately performing each step of administering and scoring a CBM. For this study, adherence will be examined through scoring checks and observations with checklists.

Quality

Along with adherence, considering the "quality" aspect of fidelity may enable personnel at DRC to enrich the effectiveness with which tutors collect data on tutees' progress. Quality has to do with the enthusiasm with which someone implements a program and their perception of its effectiveness (Dane & Schneider, 1998; Lemire et al., 2022). Although they do not use the term "quality" as an aspect of fidelity, Century et al. (2008) allude to perceptions and attitudes as factors that could potentially impact fidelity, so there is broad agreement that these factors may be worthy of examination. Although hints about quality might be observable in a CBM administration, more in-depth conversation with tutors is necessary to fully unpack it. For this reason, along with tutor observations, the current study will use tutor interviews to learn about factors impacting administration quality.

At first glance, for CBM administration, it may seem like adherence is the only aspect of fidelity that matters—after all, the procedures that tutors follow are very constrained, and they yield quantitative data about a tutee's performance, so there is little room for interpretation. Even so, it is possible that tutors' beliefs and attitudes about the task could impact their administration

in subtle ways, even if, on paper, they adhere to the standard procedures. For this reason, building tutor capacity for CBM administration demands a specific focus on quality as well as adherence.

Training Lenses

In their conceptual framework for evaluating implementation fidelity, Century et al. (2010) refer to the "procedural" and "educative" elements of an intervention as critical to implementation fidelity. These elements are inherent to the intervention, and they convey to the user the developer's intentions about what exactly they should do (procedural) and why they should do it (educative). In other words, an intervention must have organizing elements that convey to the user what to do (procedural critical components) and built-in training that conveys the developers' expectations to the user (educative critical components). If these organizing elements are not made clear to the user, implementation fidelity will suffer.

These terms, procedural and educative, serve as helpful lenses for developing tutor training at DRC. Importantly, Century et al. (2010) are addressing implementation fidelity in the context of an intervention, while I am concerned with assessment fidelity. Unlike with an intervention, the procedural and educative components needed for assessment fidelity are not inherent to the CBMs used by the DRC tutors. Therefore, those two components must come from an external source-- the training I have developed for the DRC tutors. In other words, in building the tutor training, I infused the process with both procedural and educative elements to support the tutors in working toward assessment fidelity.

Procedural

Procedural components can be thought of as conveying, "in the simplest sense, *what to do*" (Century et al., 2010, p. 205). An intervention can implicitly or explicitly suggest to a user

what they should do in terms of the basic steps of the procedures. In the context of training DRC tutors to administer CBMs, I view procedural learning as an aspect of the training that fosters tutors' knowledge of the basic steps to administer a CBM. Procedural components of the training cover topics such as access to materials, steps in the process to administer and score the CBMs, and where to record the data.

Educative

In contrast, educative components entail what the user needs to *know* to enact the intervention with fidelity. Century et al. (2010) explain that the educative components of an intervention "reflect the developers' understandings that users need a particular body of knowledge to enact the intervention as intended (p. 205). They explain that, for instance, implementation of an intervention may require a basic level of content and pedagogical knowledge; some teachers may be equipped with that knowledge already, but others may not. Therefore, to be implemented with fidelity, the intervention may be structured to include built-in professional development or training to ensure that teachers implementing it can do so in accordance with the developer's expectations. Further, some interventions may require external professional development or training components to support implementation fidelity. In working toward assessment fidelity at DRC, I used an educative lens in designing tutor training with the goal of building the tutors' knowledge base about the rationale for administering CBMs. By drawing on andragogical principles of the "need to know" and the "orientation toward learning" (Knowles, 1984), I included elements in the training designed to help tutors adopt a mindset around CBM administration grounded in their purpose and that would support administration fidelity.

By conceiving the training through both procedural and educative lenses, I made purposeful decisions about when to adopt a more pedagogical vs. an andragogical stance toward the tutors and their learning. For example, learning the procedures of CBM administration aligns with a more instructor-centered approach. However, educative aspects of the training allow more room for the tutors to consider the relevance of their learning to their tutoring work and to co-construct meaning alongside the instructor.

Training Process

The stages of the tutor training process align with the model of gradual release of responsibility (Pearson & Gallagher, 1983; Pearson et al., 2019). The model is a sound choice for all learners, but it resonates with principles of adult learning theory in particular and Knowles' assertion that adult learners benefit from instruction that meets them where they are and scaffolds appropriately (Knowles, 1984). In the gradual release model, responsibility for learning begins with the instructor and, over time, shifts to the learner (Pearson & Gallagher, 1983; Pearson et al., 2019). As the learner deepens their knowledge, they grow in their ability to make that knowledge their own, and they can become more self-directed.

Focused Instruction

The first phase in the training process aligns with the "focused instruction" phase of the gradual release model. In this phase, students are "alerted and primed" for learning (Fisher & Frey, 2013). It may include elements such as establishing a purpose for learning, direct instruction, modeling, and asking guiding questions to activate prior knowledge. For tutors at DRC, the focused instruction phase takes place in a two-part asynchronous module in the Canvas online learning-management system.

Educative Lens. Part 1 of the module, which has an educative focus, begins with a set of guiding questions designed regarding tutors' conceptions about tutees' progress, how it might be measured, some of the challenges of measuring it, and why it might be useful to measure (See Appendix D). From there, the module provides a brief overview of the DBI process and the advantages of CBMs. In alignment with andragogical principles, the goal of Part 1 is to help tutors identify a problem with current tutoring practices—specifically, a lack of systematic information about tutees' progress—so that they will be invested in participating in a possible solution to that problem.

Procedural Lens. Moving on from the educative lens, Part 2 of the module is more procedural, explaining the steps of administering a CBM and calculating a tutee's score. This part of the module includes a video example of a CBM administration meant as a model. Tutors watch the video example twice. The first time they watch, they score the CBM and compare their scoring to the instructor's scoring provided in the module. On their second viewing, tutors consult a fidelity checklist to observe how the model fulfills the administration criteria for the CBM. Part 2 of the module culminates with a brief formative assessment, a scoring check. The purpose of the scoring check is twofold: (a) it "chunks" the CBM administration process for the tutors, allowing them to focus on scoring alone before requiring them to score *and* administer in the same setting and (b) it enables me to begin collecting data on the efficacy of the modules to establish fidelity in scoring.

Guided Instruction

Having built a knowledge base of both educative and procedural learning of CBM administration, the tutors move into the "guided instruction" and "independent practice" phases of the training process. In these phases, tutors administer and score a CBM. First, they administer

it to me in a practice administration. If they feel comfortable, they can try administering and scoring it according to the procedure, from start to finish. However, if they do not feel prepared by their experience with the online module, they have the opportunity to pause throughout the process to ask questions or clarify as needed. Offering this option honors the role of the learner's experience an enables their self-direction in selecting the level of scaffolding they need. Following the practice administration, the tutors are provided with immediate corrective and suggestive feedback and the opportunity to discuss any lingering concerns or questions with the instructor. As this is their first practice opportunity, most tutors are still focused on adherence, and the feedback offered uses more of a procedural lens. However, an educative lens may be employed if a tutor seems confused about the corrective feedback, or if the tutor asks a question related to the rationale for a specific step or for the process as a whole.

Independent Practice

Finally, in the "independent practice" phase, the tutors administer the CBM to a tutee; again, this administration is followed by an immediate opportunity to debrief with the instructor. In this phase of the training, procedural and educative lenses are both present, but, as in the guided practice phase, the dominant lens adapts based on the needs of each tutor, because the instruction is co-created in the feedback sessions with the tutors. For example, one tutor may make errors in administration or scoring, or have more questions about the adherence aspects of fidelity. Another tutor may administer it with adherence to procedures, but their affect during administration or the questions they might ask could relate to their level of investment in the process, which could impact quality.

Summary

Guided by adult learning theory (Knowles, 1984), my conceptual framework shows the relationship between tutor training and assessment fidelity in the context of DRC. Tutors participate in a training process that gradually releases responsibility and that is designed to convey not only procedural knowledge about CBM administration, but also educative knowledge so that tutors can understand the purpose and rationale. The goal of this process is to build tutor capacity for administering CBMs with fidelity, not just in terms of adherence to procedures but also in terms of quality. In order to provide recommendations to DRC to improve the process, I undertook a study to address the following research questions:

1. What challenges do DRC tutors experience when administering and scoring CBMs?
2. In what ways do a DRC tutor's knowledge, skills, and beliefs impact their administration and scoring of CBMs?
3. In what ways can DRC improve current CBM training practices to build DRC tutors' capacity for administering and scoring CBMs?

Chapter 2: Literature Review

For students experiencing reading difficulties, one-to-one tutoring is an effective intervention (Austin et al., 2007; Elbaum et al., 2000; Slavin et al., 2011; Wanzek et al., 2019; Wanzek & Vaughn, 2007). Because the one-to-one format requires a tutor for each student served, using uncertified tutors can offset some of the costs and enable more efficient service delivery for students (Kirschner & Hendrick, 2020; Neitzel et al., 2022). Uncertified tutors can positively impact reading outcomes for children when they receive appropriate support to deliver structured, high-quality instruction that is responsive to students' needs (Allor & McCathren, 2004; Baker et al., 2000; Jones et al., 2016; Neitzel et al., 2022).

To enable this responsiveness, tutoring programs must engage in systematic monitoring of student progress through data collection and analysis. Because uncertified tutors do not have background knowledge about literacy assessment, they require intensive training and supervision to administer curriculum-based measures (CBMs) with fidelity to yield meaningful data (Bundock et al., 2018; Hosp et al., 2016; Reed & Sturges, 2013). Globally, even certified teachers and highly trained assessors experience difficulty administering CBMs commensurate with standard protocol, both in terms of adherence and quality (Bundock et al., 2018; Derr-Minneci & Shapiro, 1992; Reed & Sturges, 2013). In the local context, where tutoring services have expanded rapidly and training time and resources are stretched thin, achieving assessment fidelity with CBMs is an even greater challenge. To explore this problem, I review literature regarding one-to-one tutoring programs that utilize uncertified tutors and identify key features of such programs that lead to student growth, highlighting the need for responsiveness to students' instructional needs. Next, I discuss how progress monitoring within a data-based individualization (DBI) framework enables that responsiveness. Finally, I explore what is known

about how to train examiners to administer CBMs with fidelity, in terms of both adherence and quality, to obtain high-quality data for use in the DBI process.

One-to-One Tutoring Using Uncertified Tutors

One-to-one tutoring is an effective intervention for improving the literacy achievement of students in the elementary grades who experience reading difficulties (Austin et al., 2007; Elbaum et al., 2000; Slavin et al., 2011; Wanzek et al., 2019; Wanzek & Vaughn, 2007). However, the format places an intense demand on financial resources, particularly if employing highly qualified tutors who would require a higher rate of pay; therefore, one-on-one tutoring can carry significant limitations on the number of students who can be served as well as the duration of those services (Kirschner & Hendrick, 2020). For this reason, tutoring approaches increasingly make use of teaching assistants and volunteers, both paid and unpaid, as tutors (Neitzel et al., 2022).

In hopes of tapping these alternative tutor pools, researchers have investigated whether tutors who are not certified teachers can obtain similar results as teacher-delivered interventions and found that, in many cases, results are comparable. In a 2022 meta-analysis, Neitzel et al. applied rigorous inclusion standards to their study selection in order to draw conclusions about the efficacy of tutoring programs, including those that do not employ certified teachers, for accelerating achievement for readers in elementary schools. Prior to Neitzel et al. (2022), several studies (e.g., Austin et al., 2017; Gersten et al., 2020; Scamacca et al., 2015; Slavin et al., 2011; Vaughn et al., 2009; Wanzek et al., 2019; Wanzek & Vaughn, 2007) determined that supplemental literacy tutoring is effective. However, these prior studies identified few factors that mediated those effects. For example, although previous meta-analyses determined that impacts were stronger for code-based rather than comprehension skills, their focus on overall

effects rather than on more specific characteristics of the instruction prevented them from drawing conclusions about grade level, group size, or the expertise level of the person delivering the instruction. Further, these previous meta-analyses included studies with very small sample sizes (e.g., Vadasy et al., 2006, n=21) and studies where the research team provided tutoring (e.g., Case et al., 2010) and/or designed a measure to closely align with the instruction (e.g., Schwartz, 2005). A program that appears to be effective under these tightly controlled conditions may not have the same impact in more naturalistic school settings, so the utility of these findings is limited when gauging applicability to real-world contexts.

In contrast, Neitzel et al. (2022) designed their review to identify key features that impact reading outcomes for students that could more directly translate into action for practitioners designing tutoring programs. Of particular relevance, they found that effect sizes for one-to-one tutoring were comparable whether delivered by certified teachers (0.38); teaching assistants (0.44); or paid volunteers (e.g., AmeriCorps members and tutors whose time was donated by employers) (0.46). For the current inquiry, the most relevant among the studies analyzed by Neitzel et al. (2022) pertain to the use of paid volunteers, which is the category most closely mirroring the tutors at McGuffey (Allor & McCathren, 2004; Baker et al., 2000; Jones et al., 2016). Across these studies, Neitzel et al. (2022) credit "well-structured tutoring models, with extensive training" (p. 172) as the basis for success. They also note that the paid volunteers were more successful than unpaid volunteers because, likely, the paid tutors attend more regularly over a longer period of time; this continuity enables more extensive investment in building relationships with children. Overall, Neitzel et al.'s (2022) findings offer promise for organizations who aim to increase access to the benefits of one-to-one tutoring for students but

who operate with limited financial resources for compensating tutors directly and rely instead on paid volunteers.

Certainly, there are studies that call into question the efficacy of using uncertified tutors to achieve long-term impact on student reading achievement through one-to-one tutoring (e.g., Fives et al., 2013; Jacob et al., 2016; Lee et al., 2012; Miller & Connolly, 2013; Nichols et al., 2022; Ritter, 2000; Villiger et al., 2018). However, for the most part, these studies examined programs staffed by what Neitzel et al. (2022) classify as "unpaid volunteers." The exception among these studies is a randomized-controlled trial of a program called "Time to Read" (Miller & Connolly, 2013), which was not included in the Neitzel et al. (2022) synthesis. The tutors participating in the one-to-one tutoring program investigated by Miller and Connolly (2013) were recruited by their workplaces and would therefore fall into the category of "paid volunteer tutors" as defined by Neitzel et al. (2022). Miller and Connolly (2013) found little to no impact on the outcomes assessed, which tended to be more related to affect (e.g., self-esteem, enjoyment of reading) as well as a cloze assessment of reading comprehension. The researchers posit two possible explanations for the lack of impact. First, given the developmental stage of the students and the instruction provided in the program, measures of decoding and fluency would have been more likely to capture reading growth than a measure of comprehension. Second, the program's dosage (one hour per week) may have been insufficient to impact students' progress in the areas that *were* assessed. Overall, although the results from Miller and Connolly (2013) appear to undermine the Neitzel et al. (2022) conclusion that paid volunteer tutors can be effective in one-to-one tutoring contexts, issues specific to the "Time to Read" program appear to bear greater responsibility for the lack of impact.

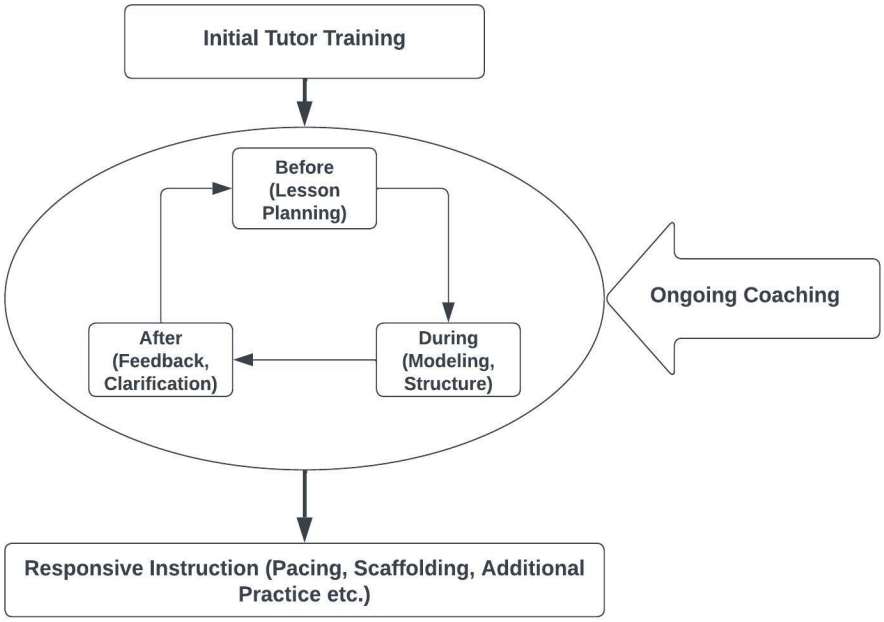
Factors Enabling Success of One-to-One Tutoring

While Neitzel et al. (2022) highlight the importance of "well-structured tutoring models, with extensive training" (p. 172) for uncertified tutors, they do not articulate the precise nature of those models and training. Interestingly, the programs in two of the studies they analyze took almost opposite approaches, despite achieving similar positive results. In the tutoring program investigated by Allor and McCathren (2004), tutors, who were college students, received extensive training that was refined by the researchers across two iterations. The lessons were highly scripted and structured; key content knowledge, instructional scaffolding, and systematic review were built into the lessons so that tutors could implement them without extensive background knowledge. In contrast, for tutors participating in the SMART program investigated by Baker et al. (2000), training was kept to a minimum, and the tutors were essentially left to their own devices to determine lesson content. Based on these studies alone, it is difficult to draw conclusions about what factors enabled the success of the one-on-one tutoring programs under study.

That said, using a wider lens, research on one-to-one tutoring programs emphasizes the importance of extensive training and supervision, as represented in Figure 2.1 (Allor & McCathren, 2004; Al Otaiba & Lake, 2007; Morris, 2006; Samson et al., 2015).

Figure 2.1

Key Components of Successful Tutoring



In his 2006 synthesis of five studies using uncertified tutors, Morris identified the role of a coach as critical to developing the model replicating their success; although the coach would provide only a small amount of up-front training for the tutors, they would provide support throughout the tutorials by assisting with lesson planning, modeling instructional techniques, providing feedback to tutors, and guiding decisions about pacing. Similarly, in their synthesis of studies that employed paraprofessionals as tutors, Samson et al. (2015) highlight the need for extensive training and supervision, with training taking the form of modeling, co-teaching, and a review of the components of reading. Like Allor and McCathren (2004), Samson et al. (2015) also concluded that a scripted program, which removes the need for guesswork among uncertified tutors, is a key characteristic of successful one-to-one tutoring efforts.

An additional important factor in tutoring success is responsiveness; regardless of who the tutors are, changing instruction in response to data is key. In their synthesis of literature on best practices in tutoring, Wasik and Slavin (1990) developed the QAIT model to delineate four elements of successful tutoring: Quality of instruction, Appropriate level of instruction, Incentive, and Time. They argue that the elements are multiplicatively related to reading gains, meaning that all are necessary and that improvement in any one factor has the potential to greatly enhance student achievement. To summarize, tutors must present the student with information in a format that is easily understood; the material must be neither too easy nor too difficult; the student must be motivated to engage in the instructional tasks; and the student must have sufficient time to learn the material being taught. Arguably, none of these elements are possible without an awareness of a student's current level of performance, a way of gauging student responsiveness to the instruction provided, and then the tutor being responsive in turn.

In published articles about tutoring studies, few authors actually discuss tutor responsiveness to students' changing instructional needs at great length; however, they often imply that adapting instruction in response student need is inherent to providing good instruction. For example, Wasik and Slavin (1990) refer to "the degree to which tutors actually adapt instruction to the needs of tutees" as a determining factor of student success (p. 9). Morris (2022) similarly articulates the need for a tutor who can "adapt instruction to the needs of an individual learner" (p. 10). In the context of building a relationship between tutor and supervisor, Allor and McCathren (2004) note that frequent communication yields tutors with "more confidence in their work and decision-making abilities" (p. 126), implying that the most effective tutors feel prepared to adapt their instruction in response to their students' needs. One study that specifically examined tutors' adapting literacy instruction in a one-to-one setting with uncertified

tutors was Al Otaiba and Lake (2007), which focused on preservice teachers in a practicum course. Through a qualitative analysis of tutors' weekly written reflections and final reports, Al Otaiba and Lake (2007) determined that tutors were beginning to use data from the assessments they were administering to make judgments about their own efficacy and to adjust their instruction accordingly, albeit in a preliminary way. It is through these adjustments that instruction can be improved in accordance with the QAIT model (Wasik & Slavin, 1990). Overall, the message is clear that, ideally, tutors should be problem-solvers, adapting in response to input from their students.

Data-Based Individualization

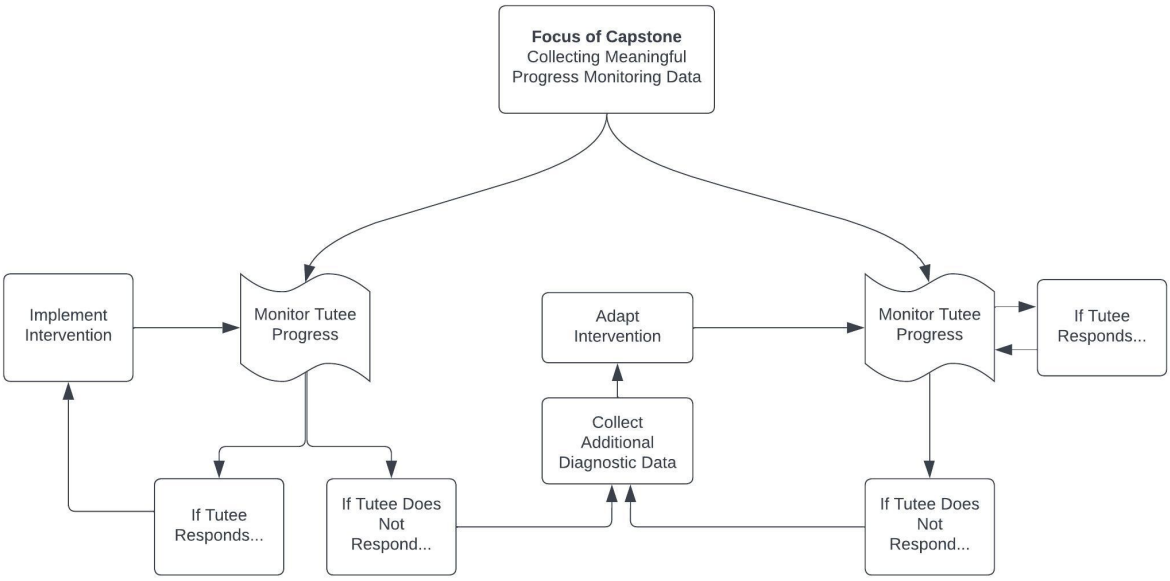
One evidence-based process for systematically gathering and using this input to inform instruction is data-based individualization (DBI). The foundational principles of DBI relate to the idea that educators should commit to a problem-solving approach for students experiencing academic difficulty. These principles emerged in the 1970s in response to critical needs identified by teacher-educator and researcher Stanley Deno and colleagues (Deno, 2016). Deno's aim was to build a process that established goals for students and that systematically evaluated whether the students were making progress toward those goals. Essentially, the process mirrors single-case study research in that the effect of a given instructional program on a student is treated as a hypothesis, and data is collected to test it along with any subsequent modifications to that program. As he explains, the process necessitates a way to test "vital signs" of educational development, by which he means progress monitoring measures that are technically adequate and logistically feasible to use (Deno, 2016, p. 12). Critically, these vital signs must correlate with overall achievement in a specific domain; for example, rate and accuracy of oral reading fluency (ORF) measures should provide a window into reading proficiency more generally

(Busch & Reschly, 2007; Wayman et al., 2007). Further, practitioners must be well-trained to interpret this data to draw valid conclusions about the impact of the instructional program (Austin & Filderman, 2020; Filderman et al., 2018).

As operationalized in schools, DBI generally occurs within the context of a Response to Intervention model, where teachers provide instructional support for students across "tiers" of increasing intensity (Filderman & Gesel, 2022). Figure 2.2 (NCII, 2013) illustrates the steps of the DBI process and notes the focus of the current capstone in context.

Figure 2.2

Focus of Capstone within the Context of DBI (Adapted from NCII, 2013)



Teachers implement an intensive intervention and monitor progress toward a goal to determine the efficacy of the intervention. If the intervention appears to be effective, they stay the course. However, if not, they administer diagnostic assessments to determine the student's specific instructional needs. Equipped with that data, they modify the intervention to better meet the

student's needs. Subsequently, they engage in a cyclical process, continuing to monitor progress and to collect diagnostic data and modify instruction as needed (Austin & Filderman, 2020).

Through the DBI process, teachers make decisions both across and within RTI tiers; this is to say, they make determinations about when to provide students with a more intensive tier of instruction as well as when to adapt instruction within a tier when it is not having the desired impact (Ball & Christ, 2012; Filderman & Gesel, 2022). Generally, to support an individual student, stakeholders come together as a team, employing a protocol to discuss instructional needs, analyze the data available, and determine appropriate next steps to support that student's growth.

In contrast, settings using uncertified tutors, like DRC, are challenged because they lack the infrastructure for this type of team-based approach, and the uncertified tutors' toolbox for adapting instruction is more limited than that of a certified teacher. However, working within those limitations to engage in more comprehensive and structured DBI practices can enable these programs to improve instructional outcomes for students experiencing reading difficulty through providing more responsive instruction.

Evidence Base for DBI

Data-based individualization is generally touted as a "best practice" for supporting students in a Response to Intervention framework, but there are relatively few empirical studies validating DBI as a whole. Because DBI is embedded within systems, and because it has so many nuances and moving parts, isolating and quantifying its impact is difficult. Despite these challenges, evidence is emerging that DBI can improve student achievement outcomes. Although correlations in studies of DBI do not generally meet the threshold to be considered "strong" evidence of impact, it must be noted that the students who comprise these study populations have

generally been experiencing significant academic difficulties. For these students, who have not responded to typical instructional practices, any progress is hard-won. Therefore, in interpreting data, the relative impact of DBI practices on these students must be considered (Filderman et al., 2018).

One synthesis and meta-analysis by Filderman et al. (2018) quantified the impact of DBI, demonstrating a modest effect size on student progress. The researchers conducted a systematic search of peer-reviewed research that incorporated DBI ranging from 1975 to 2017 to identify 15 high-quality studies examining DBI in reading for students in Grades K-12. Nine of the studies compared outcomes for students under intervention/DBI conditions with outcomes under "business as usual" (BAU) conditions, and the researchers found a weighted mean effect of 0.24 (95% confidence interval) in favor of DBI. However, in those studies, the intervention was markedly different from the BAU treatment, so it is difficult to determine how much of the difference can be attributed to DBI alone.

In contrast, the other six studies identified by Filderman et al. (2018) involved students participating in parallel reading interventions, some with and some without a DBI component. In these studies, DBI as a variable was more isolated, so the results speak more directly to the contribution of DBI alone. For these studies, the weighted mean effect was 0.27, again with a 95% confidence interval. These effect sizes, though positive, are fairly modest, and the researchers themselves state that "the results of this synthesis reveal that claiming DBI to be an evidence-based practice is tenuous at best" (p. 185). However, it is important to note that, for students who have experienced persistent reading difficulties, like many of the students in these study populations, "moving the needle" can be difficult; this is especially true for the secondary-age students represented in some of the reviewed studies. For students who have not responded

to typical instructional practices, effect sizes must be interpreted relative to their practical significance, and even an effect size of 0.2 may translate into meaningful gains (Filderman et al., 2018).

Having identified a paucity of research on DBI, Filderman and Toste (2021) devised a randomized controlled study demonstrating a significant positive impact of DBI with upper elementary students experiencing word-reading difficulties. Their purpose was to isolate and examine the impact of DBI within a multisyllabic word reading intervention for 88 students in Grades 4 and 5 who were at risk for reading disability, comparing results across three conditions. Some students participated in a researcher-designed intervention, while others received what constituted "business as usual" (BAU) instruction in their school (which ranged from independent reading/homework time to small-group intervention with a program like Leveled Literacy Intervention, which has a comprehension focus). All students receiving the multisyllabic word reading intervention experienced some initial customization of the program based on an assessment given prior to the intervention. For a subset of these students, however, the intervention was adjusted at the five-week mark in accordance with a standardized DBI procedure.

The impact of the intervention and of the DBI component varied across measures. Although students across all treatment conditions performed comparably on a standardized measure of word identification, students who participated in the intervention significantly outperformed their BAU peers on a researcher-designed mastery measure aimed at assessing what had been specifically taught. The most noteworthy result, however, was evident on a standardized measure of nonsense word decoding, on which students who received the intervention and whose instruction was intensified based on DBI procedures significantly

outperformed the BAU condition (effect size = 0.87). In contrast, the intervention-only group did not improve on this measure to a significant degree as compared to BAU, demonstrating that the DBI component was a critical addition to the intervention program. Although further study is needed to quantify the impact of DBI across skills, age groups, and contexts, the existing evidence indicates that it is a promising practice for increasing student achievement in reading.

Using CBMs for Monitoring Progress in a DBI Framework

To implement a DBI framework, practitioners must select and utilize tools to collect data about student progress and engage in a systematic process of analysis. The following section of this literature review examines curriculum-based measures (CBMs), their advantages and disadvantages, and considerations for effective use within a DBI framework for students with needs at the word level, which is to say students who benefit from instruction focused on phonemic awareness, word knowledge to support decoding and spelling, and fluency (reading rate, accuracy, and expression), rather than comprehension. This capstone is focused on a code-based intervention for students with word-level reading difficulties. Therefore, this literature review focuses on progress monitoring tools that are most relevant for students with these instructional needs.

What are CBMs?

CBMs are administered using a highly standardized protocol with an eye toward efficiency (Austin & Filderman, 2020; Deno, 1985). Generally, they are designed to capture the number of correct responses a student can provide in one minute. In the case of reading CBMs, this could involve tasks such as reading nonsense words, or reading real words in isolation or in context.

Although CBMs may measure a narrow skill, the skills were selected so that overall achievement in a domain can map onto them with the goal of "assessing proficiency in a global, longitudinal manner" (Busch & Reschly, 2007, p. 224). The same materials and metrics are used across a period of time—perhaps months, perhaps an academic year—to track progress toward long-term goals. CBMs of nonsense word reading fluency, word reading fluency, and oral reading fluency are examples of skills that provide a window into overall reading achievement, albeit to different degrees for different ages of children. For this reason, CBMs are a type of general outcomes measure, or GOM. The teacher and others on a team can track a student's progress over time as well as compare their performance to grade-level benchmarks (Stecker et al., 2005; Wayman et al., 2007).

CBMs and DBI

A vast number of studies have addressed and established the technical adequacy of CBMs in and of themselves (see e.g., Goffreda & DiPerna, 2010; Goffreda et al., 2009). In contrast, the evidence base for using CBMs for progress monitoring within a DBI framework, though emerging, is relatively slim, which has serious implications for their application in educational contexts. Most of the studies of CBMs are about a single point in time with the purpose of screening students or identifying those at risk for reading difficulties (Goffreda & DiPerna, 2010; Jenkins et al., 2007; Reschly et al., 2009; Wayman et al., 2007). In fact, Filderman et al. (2018) were surprised to note that, of the 15 DBI studies included in their meta-analysis, only three used CBMs to collect data and guide decision making. So, although CBMs are generally recommended for DBI, more empirical research is needed to disambiguate how they might contribute to effective practice.

As recently as 2013, in a literature review, Ardoin et al. examined over 100 studies about the use of oral reading fluency (ORF) CBMs for progress monitoring and decision making, and, although they found no shortage of recommended best practices based in expert opinion, the empirical evidence behind those recommendations was more preliminary. More recently, according to Van Norman et al. (2018), contradictions abound in the emerging evidence base regarding the best methods for summarizing student performance, coping with variability of scores, and establishing the minimum duration and number of data points necessary to draw meaningful conclusions. Along with this, a 2018 meta-analysis by Jung et al. examined 14 studies and concluded that DBI using CBMs had a moderate positive impact on student outcomes (0.37 effect size), but they explained that a variety of factors influenced efficacy, including but not limited to the specific CBM tasks selected, the frequency of data collection, and the support provided to teachers. In the absence of research-based consensus, Ardoin et al. (2013) cautioned that, for individual students, "effective treatments might be terminated prematurely and ineffective treatments might be continued longer than necessary," undermining their progress (p. 2). For students who are most at risk for long-term reading challenges, stakes are high, and scholars and practitioners must take care to avoid generalizations and make nuanced recommendations that are sensitive to context.

Choosing a Measure for Progress Monitoring

Development of and research into CBMs that are technically adequate has flourished over the last decade or so, and resources such as the "Academic Progress Monitoring Tools Chart" from the National Center for Intensive Intervention provide guidance to practitioners who want to find out if a particular measure is reliable and valid for its intended use (Austin & Filderman,

2020). Of course, the tool must also be available for practitioners to use. DRC used this guidance in selecting an appropriate and available set of measures for use with tutees.

Guidance on selecting a CBM encourages educators to select the measure that (a) represents an area of risk for the student and (b) aligns with the intervention a student is receiving (Hosp et al., 2016). In general, tasks of nonsense word reading fluency (NWF) and oral reading fluency (ORF) are most sensitive to growth over time (University of Oregon, 2018). However, students who need to build word recognition but who have strong decoding skills might benefit from a word reading fluency (WRF) measure (University of Oregon, 2018). Because all measures are not equally well suited for all students, the following guidance can help educators determine which measures to use when and with whom. Specifically, this guidance informed how DRC weighed three viable CBM options while accounting for the specific limitations in their context.

Nonsense Word Reading Fluency

Because ORF may not be sensitive enough to show growth for beginning readers (Fuchs et al., 2004; Hosp et al., 2016), educators should consider incorporating measures of nonsense word reading fluency or word reading fluency for our first-grade students, at least during the fall semester. One reason to consider nonsense word reading fluency is that beginning readers are solidifying their knowledge of the alphabetic principle, meaning that they are working to master letter-sound correspondences and to use those correspondences to sound out words. A nonsense word fluency task foregrounds those particular skills and enables a teacher to monitor progress in this critical area. For students in Kindergarten and first grade, nonsense word reading fluency has demonstrated predictive validity for later reading skills in late first and second grade (Catts et al., 2009; Clemens et al., 2001; Cummings et al., 2011; Kim et al., 2010). In a large-scale 2010 study

involving over 3500 first graders, Fien et al. found that gains in nonsense word reading fluency across an academic year were highly predictive of end-of-year fluency and comprehension performance. This held true for nearly all students, regardless of their initial decoding skills; the only exception was for students who performed in the top 5% on the beginning-of-year assessment. Further, the researchers found that, for the students who demonstrated the most significant initial decoding needs, the earlier the gains in NWF, the more strongly predictive of later reading outcomes. Taken together, the data indicate that, for all but the strongest decoders, improvements in decoding during first grade will bolster achievement in fluency and comprehension across the year. Further, the earlier in the year those gains take place, the better. In part, this is likely because when students start decoding words proficiently, they become more efficient at adding words to their bank of automatically recognized words, bolstering their fluency and freeing cognitive resources for comprehension.

Word Reading Fluency

Another option for educators to consider would be a CBM of word reading fluency (WRF), which, in a comparison with NWF, was found to be even more predictive of later reading success (Clemens et al., 2014). As with NWF, a strong evidence base attests to the predictive power of WRF for later reading outcomes (Clemens et al., 2011; Compton et al., 2006; Compton et al., 2010; Fuchs et al., 2004; Speece et al., 2011; Zumeta et al., 2012). In a 2014 study, Clemens and colleagues compared these measures of NWF and WRF with a group of 80 first graders and found that the slope for word reading fluency better predicted skills in oral reading fluency, real and nonsense word reading, and basic comprehension than did the slope for nonsense word reading fluency. This makes sense because WRF is a more "downstream" measure of growth than NWF.

Although NWF provides a window into a student's automaticity with letter-sound correspondences, application to actual word reading may be obscured because task demands are different. Words on the NWF lists are all phonetically regular, while, for WRF, the target words incorporate a wider range of features with varying degrees of regularity. Although the ability to decode quickly and accurately supports the ability to recognize words on the WRF task, decoding ability alone is not sufficient for success. Rather, WRF represents a more distal skill in which word knowledge has been consolidated and leveraged to support word recognition. For students who have made good progress with their decoding but still experience undue difficulty in connected text, particularly in Kindergarten through third grade, WRF may be an appropriate measure with which to monitor progress.

Oral Reading Fluency

A measure of ORF is recommended for capturing growth over time, and it aligns with the instructional needs of many students receiving reading intervention (Hosp et al., 2016; University of Oregon, 2018). CBMs of oral reading fluency (ORF) have been shown to correlate with other reading proficiency measures, including standardized test scores (Fuchs et al., 1988; Marston, 1989; Reschly et al., 2009; Wayman et al., 2007). Notably, some research has indicated this correlation weakens as students age and comprehension demands increase, especially with expository text (Cho et al., 2018; Espin & Foegen, 1996; Hosp & Fuchs, 2005; Shinn et al., 1992). This is logical because reading rate and accuracy are constrained skills, meaning that there is a limit to the rate and accuracy with which someone can read. As students encounter more complex text, assuming they have reached a particular threshold of fluency success, their comprehension of language is much more of a determining factor for their comprehension. On the other hand, in their 2009 meta-analysis, Reschly et al. identified a strong correlation between

ORF and other standardized measures of reading achievement that did not differ in magnitude across the grade levels included (first through sixth), running counter to their hypothesis. All told, despite the potential for waning utility among older students, ORF is an appropriate choice for progress monitoring in first through eighth grade because of the preponderance of evidence of a strong correlation with reading comprehension at that grade level. Additionally, students receiving intervention services generally have needs in fluency, so it is unlikely that the ceiling effect would be a limiting factor. In summary, different progress monitoring measures have advantages and disadvantages with different student populations, and educators must take care to select measures that align with students' instructional needs as well as the intervention they are receiving. In the local context of DRC, ORF is the only measure that (a) covers the full grade range of students served and (b) aligns with the instructional focuses of word-based skills and fluency.

Obtaining High-Quality Data from ORF CBMs

Since the 2013 publication of Ardoin et al.'s literature review, researchers have explored how ORF CBMs can be used meaningfully in a DBI context, mostly by using simulation studies to investigate the accuracy of advice regarding decision rules (see e.g., Van Norman & Christ, 2016; Van Norman et al., 2018). Although an in-depth treatment of decision rules is beyond the range of this literature review, it is important to note that, in DBI, the purpose of collecting progress monitoring data is to provide the basis for a team to come together and use decision rules to determine if a student's instructional program is supporting sufficient progress or if an adjustment is necessary. Therefore, practitioners need to know how many data points, and over what duration, are needed to support sound decision making. They also need to ensure as much

as possible that differences in scores are due to true differences in reading proficiency rather than some other source of variability.

Frequency and Duration of CBM Administration

One major point in favor of CBMs is that they are quick to administer, so they can be administered often, enabling a teacher to obtain frequent "temperature checks" about an aspect of a child's reading. That said, any time spent on assessment is time taken away from instruction, and teachers have been known to balk at the time needed to administer even one-minute probes (Begeny et al., 2011; Mellard et al., 2009; Rowe et al., 2014). Ultimately, the goal of a recommended administration schedule is to collect sufficient data to draw valid conclusions while minimizing demand on teachers and students to collect that data.

Recommendations on frequency and duration vary across sources. Guidance from some sources encourages progress monitoring every two weeks for students in Kindergarten through Third Grade, and not more than once per week (University of Oregon, 2018). For students in fourth grade and higher, they recommend administering an ORF CBM every second or third week to account for the slower growth typically observed as students advance through the grades. However, in their 2018 meta-analysis that found moderate support for DBI, Filderman et al. noted that the majority of studies involved data collection once per week, or more frequently. Generally, this aligns with other research-based recommendations for collecting data between one and three times per week (Christ et al., 2013; Van Norman & Christ, 2016). Because these recommendations represent a range of frequencies, educators must balance practicality with ensuring valid data, and a weekly administration schedule seems to best strike that balance (Filderman et al., 2018).

To further complicate the matter, the frequency of data collection has implications for how quickly a response to instruction can be determined, because the determination depends on obtaining a number of data points. If the time between those data points is protracted, the determination is delayed, as is an instructional response. In general, eight or more data points have been recommended across studies (Christ, 2006; Christ et al., 2012; Van Norman & Christ, 2016) to obtain the most accurate sense of a student's growth based on modeling. However, if collecting bi-weekly data as recommended by some sources (e.g., University of Oregon, 2018), the elapsed time period would be so long as to be impractical. To balance the competing priorities of data collection frequency and quantity, examiners can create an end-of-year goal, graphing an aimline based on the student's baseline performance, and plotting points for each administration to see where they fall relative to the aimline. Based on the number of points above or below the aimline, the student's response to the instruction can be determined (Hosp et al, 2016; University of Oregon, 2018). Educators engaging in progress monitoring must work within the parameters of their context, which may demand some compromise as they interpret data to fairly reflect student growth. For example, in the context of DRC, a semester-long session of nine weeks may not provide the opportunity to obtain a sufficient number of data points; however, across an academic year, weekly CBM administrations could yield insights relative to an end-of-year goal.

Minimizing Variability

Despite their advantages, one major critique of ORF CBMs in particular has been the high degree of variability that can be evident in a student's performance across multiple passages, even on passage sets carefully created to be as equivalent as possible (Cummings et al., 2013; Hintzie & Christ, 2004; Poncy et al., 2005). However, some researchers have found that

when ORF CBMs are investigated solely with their intended audience, students reading well below benchmark expectations (as opposed to a more generalized study population), variability decreases (O'Keeffe et al., 2017; Tindal et al., 2016). Minimizing variability is important for DBI because outlying data points muddy the waters for decision making. If the goal is to draw valid conclusions about student progress from the minimum number of data points, it is critical that each of those data points truly represents what it purports to measure (Nelson et al., 2017; Van Norman & Christ, 2016).

Bundock et al. (2018) have identified three potential sources of variability to be aware of in hopes of minimizing it: passage factors, student factors, and setting factors. The first, passage factors, has diminished in prominence because of the work done by publishers to develop equivalent sets through the use of readability formulae, statistical methods, and field testing. Including both narrative and expository text introduces some potential for variability (Briggs, 2011; O'Keeffe et al., 2017), but omitting one type or the other would lessen the applicability of the measure to overall reading achievement. Therefore, Bundock et al. (2018) recommend focusing on the second two sources of variability, student factors and setting factors. In terms of student factors, Bundock et al. (2018) emphasize that, as long as the measures are used with the target population, students experiencing reading difficulty, variability is a minimal concern (O'Keeffe et al., 2017; Tindal et al., 2016).

According to Bundock et al. (2018), the most significant source of variability is setting factors, or factors related to where the measures are administered, who administers them, and to what extent the administrators follow standard protocol for administration and scoring. In the context of DRC, anecdotal observations align with Bundock et al's (2018) conclusion about setting factors; therefore, this source of variability forms the basis for the current investigation. A

2013 study by Reed and Sturges found that, even among trained assessors, 8% of administrations had uncorrectable errors (such as a timing error or providing guidance beyond the standard directions) and 91% had errors in scoring. Further, the researchers found that fidelity to administration protocols diminished over time if refresher training was not provided, potentially undermining the utility of the data for decision making.

Training Recommendations

To minimize variability and obtain valid data, training and support is crucial for anyone who will administer a CBM. Bundock et al. (2018) offer a series of recommendations to improve assessment fidelity with CBMs. Directly related to training, they recommend that all individuals be trained in administration procedures and that a review of training should be implemented periodically to prevent drift from standardization and variability due to setting factors. The training itself should include procedures for administration and scoring along with opportunities to practice and receive feedback, employing a "gradual release" format, thus building the examiners' procedural knowledge (Bundock et al., 2018; Reed & Sturges, 2013). Such training should enhance the examiners' adherence to the standard administration procedures. Further, Reed and Sturges (2013) state that the training in their study explained the "theoretical basis" of CBMs, which aligns with an educative training lens (p. 265); although they do not use the term "quality," they emphasize that examiner beliefs may play a role in assessment fidelity, and those beliefs are likely to be impacted by the educative aspects of the training they receive. As an instrument for evaluating fidelity, both initially and on an ongoing basis, Bundock et al. (2018) recommend the use of checklists, either developed by the CBM publisher or by personnel in the local context.

In alignment with guidance from Bundock et al. (2018) and Reed and Sturges (2013), DRC has worked to minimize variability as much as possible. Regarding passage factors, DRC has selected a set of passages for progress monitoring that have been extensively vetted and determined to be of equivalent difficulty by grade level (University of Oregon, 2018). In terms of student factors, DRC is using this set of passages with the target population: students, primarily in the elementary grades, who are experiencing reading difficulty. Additionally, DRC follows guidance for off-grade progress monitoring as recommended by the developers of the CBM to better capture progress for students who have substantial difficulty accessing grade-level text. Therefore, student factors are unlikely to represent a significant source of variability at DRC (Bundock et al., 2018; O'Keeffe et al., 2017; Tindal et al., 2016). Having accounted for the passage and student factors, DRC is focused on minimizing the impact of setting factors. DRC has used the guidance from Bundock et al. (2018) and Reed and Sturges (2013) to create the tutor training process as described in Chapter 1, which involves a learning module to build both procedural and educative knowledge as well as opportunities for feedback in a gradual release sequence of observations and debriefs. Through the current investigation, I hope to gain information about how DRC can refine their training process to better meet the needs of tutors in their context and support them in their assessment fidelity, both in terms of adherence and quality.

Summary

In summary, one-to-one tutoring programs show promise in addressing the literacy needs of school-age children. The tutorial setting enables instruction that targets a student's individual needs with appropriate pacing and scaffolding. Even when the tutors are not certified teachers, they can be successful given proper training and supervision, particularly when delivering code-

focused instruction using highly structured lesson formats that enable responsiveness to students' changing instructional needs.

Although few studies examine the practice of DBI within a tutorial setting, DBI provides a framework for systematizing the responsiveness that makes tutoring successful. High-quality data collection is an essential component of DBI, and CBMs are a valuable tool provided that appropriate measures are chosen for the context and that they are administered with fidelity. Setting factors have the most significant impact on assessment fidelity, and even when examiners are thoroughly trained, errors in administration and scoring undermine results. Employing a gradual release model for training alongside fidelity checklists can ameliorate some of the threats to assessment fidelity of CBMs.

Chapter 3: Methods

In this chapter, I begin by briefly reviewing the purpose of the study and the research questions guiding the investigation. Next, I discuss my selection of a case study approach and its appropriateness for exploring these research questions. Following this, I describe the research context, participants, data sources, analysis procedures, and ethical considerations that shaped the design of the inquiry.

Purpose and Research Questions

As noted previously, challenges inherent to administering CBMs with fidelity undermine the collection of high-quality data on tutees' progress. This phenomenon has been observed globally as well as in the local context of Douglass Reading Center (DRC) (Bundock et al., 2018; Derr-Minneci & Shapiro, 1992; Reed & Sturges, 2013). Through a case-study approach, I investigated practices for training tutors to administer CBMs and the impact of that training on their adherence to standard procedures and the quality of their administration. As seen in Chapter 5, my findings formed the basis of recommendations to stakeholders regarding potential modifications to the training with attention to procedural and educative elements of the training. These recommendations are especially timely given increasing demands among stakeholders for data that accurately reflects tutees' progress (L. Hayes, personal communication, November 2022). The research questions guiding the inquiry are:

1. What challenges do DRC tutors experience when administering and scoring CBMs?
2. In what ways do a DRC tutor's knowledge, skills, and beliefs impact their administration and scoring of CBMs?
3. In what ways can DRC improve current CBM training practices to build DRC tutors' capacity for administering and scoring CBMs?

Methodology

Qualitative Case Study Research Design

Because most data collection for progress monitoring occurs in school settings, DRC is a distinct case, with many contextual factors at play. I have therefore selected a qualitative case study approach for my inquiry because it is suited to in-depth exploration of a phenomenon as it is enacted within a specific setting (Yin, 2003). My choice of approach stems from a pragmatist paradigm, with its focus on "the actions, situations, and consequences of inquiry" (Creswell, 2007, p. 22). In my case, the intended consequence is to learn about building tutor capacity for data collection practices. The pragmatist paradigm prioritizes trying to find "what works" by investigating a contextualized problem through whatever methods of data collection best match it (Creswell, 2007). A case study approach is ideal for addressing my problem of practice because it emphasizes "how" and "why" questions and enables the study of a phenomenon in its natural context (Denzin, 1989). Because I serve as DRC's Lead Clinician, I have internally guided interest in this particular case, my inquiry is an intrinsic case study, and conclusions drawn are addressed to the specific site, DRC, even though it is possible to situate findings in the broader global context of challenges in high-quality data collection (Hays & Singh, 2011; Stake, 1995).

Additionally, my study draws from the tradition of action research, which emphasizes systematic inquiry for the sake of generating change (Efron & Ravid, 2019). As a participant researcher, I employed my unique knowledge of my context to generate knowledge about preparing tutors to engage in effective data-collection practices. The fruits of the inquiry led to recommendations for subsequent actions, as seen in Chapter 5; because of my ongoing engagement with the study context, I will have the opportunity to apply this knowledge and

engage in future cycles of generating and investigating new research questions to impact practices in developing tutor capacity for CBM use at DRC.

Methods

This study was conducted during the spring semester of 2023 in the context of an after-school tutoring program run by DRC, which is based at a university in the mid-Atlantic region. Within this context, I examined aspects of tutor training that impacted the tutors' capacity to administer CBMs with fidelity and obtain high-quality data about their tutees' progress.

Researcher Access

As Lead Clinician at DRC, I have an existing relationship with DRC leadership, and I am familiar with the goals of the organization. Because of my intimate involvement, I, with the support of the DRC Director, identified the problem of practice as an impediment to those shared goals and determined it to be worthy of investigation (L. Hayes, personal communication, August 2022). While many of the tutors were unaffiliated with the Teacher Education program at the university, some were, and I worked with program leadership to obtain permission to recruit them as participants. By coordinating with the Teacher Education office, I ensured that participation in my study did not present an undue burden to any tutors who are pre-service teachers, as they are frequently recruited for studies.

Study Context and Participants

DRC has been an established presence in the larger community for decades, serving children through intervention and assessment work. DRC offers an after-school tutoring program where approximately 50 tutees in Grade 1 through Grade 8 participate in one-on-one reading lessons with a tutor. Each tutee meets with their tutor twice each week for approximately 50 minutes, with a total of 18 sessions across the semester, or 36 across the academic year. The

program offers both in-person and virtual options, and, although most of the tutees come from the counties surrounding the university, some virtual tutees log on from various other regions of the country. That said, this investigation focuses on in-person tutorials because research and guidance around progress monitoring and CBM administration is situated in face-to-face settings.

Because the focus of my inquiry is the impact of training on tutors' knowledge base and behaviors, my participants were the population of tutors working in DRC's after-school tutoring program in the spring of 2023. At that time, 10 tutors were working with tutees in person at DRC. As part of their regular training, all generated artifacts (i.e., scoring checks and observations), and all consented to the use of their artifacts for the purposes of this study. Additionally, nine of the tutors participated in semi-structured interviews about the training process. Because I interviewed all tutors who consented to that part of the study, a wide range of demographic characteristics were represented among my interview respondents, increasing "the likelihood that findings will reflect differences or different perspectives" (Creswell, 2007, p. 126). Characteristics of interest included class year; program of study; previous experience administering CBMs; and previous experience administering other standardized assessments. I anticipated that these characteristics would impact how a tutor participated in the training and how they subsequently administered the CBMs. Of note, three of the tutors (Leilani, Philonese, and Janice) had participated in a previous version of ORF CBM training through DRC that was revised prior to the current study. Information about each participant interviewed is included below and organized into Table 3.1:

Table 3.1

Participant Information: Name, Class Year, Program of Study, Previous Assessment Experience

Participant Pseudonym	Undergraduate or Graduate	Program of Study	Number of Semesters with DRC	Previous CBM Experience?	Previous Standardized Assessment Experience?
Wanda	Graduate	SLP	2	No	Yes
Kamaria	Graduate	SLP	2	No	Yes
Teri	Graduate	SLP	2	No	Yes
Renee	Undergraduate (4th Year)	Elementary Education	2	Yes	Yes
Leilani	Undergraduate (4th Year)	History/Spanish	3	No	No
Philonese	Undergraduate (4th Year)	Sociology/African-American Studies	4	Yes	No
Janice	Undergraduate (3rd Year)	History/Women, Gender & Sexuality	4	Yes	No
Pam	Undergraduate (2nd Year)	Elementary Education (Accepted)	2	No	No
Sharon	Undergraduate (2nd Year)	Elementary Education (Accepted)	2	No	No

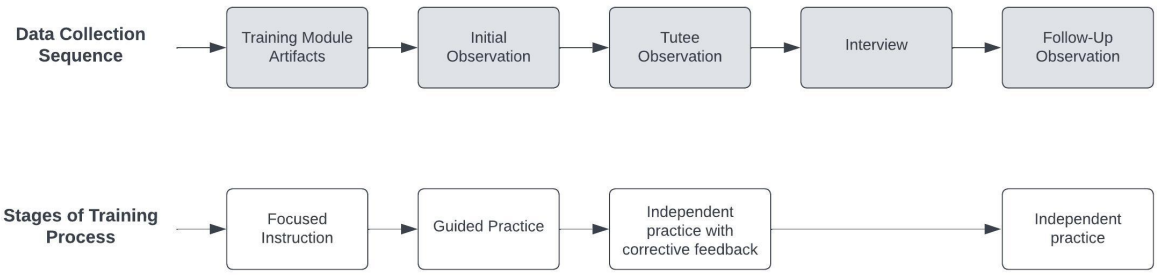
Data Sources and Collection

In order to address my research questions, I collected data from multiple sources. Key artifacts included training module artifacts; observations incorporating a fidelity checklist and field notes; and semi-structured interviews conducted with all consenting participants. Data collection occurred within the context of the overall training process. Figure 3.1 shows how the

sequence of data collection corresponded with the stages of the training process, beginning with focused instruction and culminating in independent practice.

Figure 3.1

Data Collection Sequence and Stages of Training Process



Training Process Overview

Representing the focused instruction stage, the training process began with a virtual, asynchronous module that tutors completed through Canvas, the university's online learning management system. To start, the tutors answered a series of baseline questions to prompt some initial reflection on the topic of monitoring student progress. They watched a brief video that had an educative focus, explaining the rationale of progress monitoring and the selection of a CBM as the tool. The next section of the training had a procedural focus; tutors read the guidelines for CBM administration and watched a video example of a sample administration. They were provided with a list of fidelity criteria to guide their viewing, and they were also instructed to practice scoring the administration and to compare their scoring sheet to an example. Finally, they listened to a recording of a tutee reading a passage, which they scored; these scores were submitted for the scoring check to complete the online module. The baseline question responses and the scoring check responses comprise the training module artifacts.

From there, moving into the guided and independent practice stages, I observed the tutors three times as they administered a CBM: an initial observation where they administered a CBM to me; a tutee observation, where they administered a CBM to a tutee; and a follow-up observation, where they administered a CBM to a tutee after approximately six weeks had elapsed. After each observation, I provided corrective feedback to the tutors and allowed them the opportunity to ask questions as needed as well as to provide formative feedback to me.

Training Module Artifacts: Baseline Questions

In January, to prepare for the focused instruction phase of training, all tutors responded to several open-ended questions via a quiz in the Canvas learning management system. The purpose of these questions was to gain information about tutors' initial beliefs about and attitudes toward progress monitoring. The baseline questions served as a foundation for the educative aspects of tutor training, because they primed tutors to consider the purpose of monitoring tutees' progress and how that practice might influence current tutoring practices. In developing the baseline questions, I considered the "gradual release" framework, beginning with a question designed to prime tutors to the topic of progress monitoring before leading them to consider the potential impact of progress monitoring in their investment in tutoring and their tutoring practices.

Training Module Artifacts: Scoring Checks

In order to learn whether tutors could accurately score a CBM after completing a training module, all tutors working at DRC were required to submit a scoring check at the end of the module. This data source stems from the procedural aspects of the training and relates to the tutors' adherence; the scoring checks determined whether the tutors could adhere to the standard procedure for scoring a CBM. The tutors listened to an audio example of a child reading a passage, and they scored that passage reading, obtaining scores of rate and accuracy. The tutors

then entered their rate and accuracy scores into a Canvas quiz and submitted it. The scoring checks took place in mid to late January.

Initial Observations

After completing the online training, moving into the guided practice phase of training, tutors administered the CBM to me one time so that I could observe fidelity criteria that pertained to administration in addition to scoring. The goal of these observations was to track adherence to fidelity characteristics while also describing the quality of the administration. I observed this initial administration using a fidelity checklist and field note template (See Appendix E). This document was developed through consideration of both procedural and educative training aspects with the purpose of collecting data both on tutor adherence and quality. The items on the checklist were adapted from widely-adopted procedures for CBM administration (e.g., University of Oregon, 2018) and address aspects of adherence such as timekeeping, reading standardized directions, and scoring. For the checklist items, I noted "pass" or "needs practice" for each item and recorded any additional details under "Commentary." For example, if the tutor started the timer before the first word of the passage is read, I recorded "needs practice" for the criterion "Starts timer when student says the first word" and noted "started before first word was read" under "Commentary." To address administration quality, the template also provided space to record notes on tutor affect (e.g., tone and body language) that may be pertinent to the tutors' comfort with the administration and the degree to which they "bought into" the process of using CBMs for progress monitoring.

I took notes on my laptop using my field note template. I also audio recorded each initial administration so that I could engage in peer debriefing to confirm my impressions of administration fidelity; I asked a peer to watch the recordings and discuss my checklist and notes

with me, which built the credibility of my findings (Hays & Singh, 2011). Following the initial administration, I held a five-minute debrief with the tutor to review the fidelity checklist with them and answer any of their questions, which may pertain to their procedural or educative learning. For any areas marked "needs practice," I provided specific corrective feedback and/or modeling to reinforce the tutors' procedural learning. Immediately after the debrief, I copy/pasted my field notes into a new document where I added any additional contextual notes and then created a brief written summary of the observation and my interpretations as an analytic memo. Crafting this memo enabled me to consolidate data from the checklist, field notes, and tutor debrief; gather ideas about what to look for in future observations, both with this tutor and with other tutors; and hone my protocol for tutor interviews later in the data collection process. These initial observations took place in late January.

Tutee Observations (Initial Tutee Observation and Follow-Up Observation)

Following the initial observation with me, I used the same checklist and observation template as the tutor administered the CBM with a tutee as independent practice. I completed two of these tutee observations; one was early in the semester and the other took place four weeks later to account for the "drift" from standardized protocols that may occur even after initial fidelity has been achieved (O'Donnell, 2008; Reed & Sturges, 2013). In this way, I collected data on whether tutors could bring to bear their knowledge of scoring and CBM administration in the context of working with their tutee as well as the degree to which they could maintain fidelity over time. As with the initial administration, I recorded each tutee administration to confirm scoring, completed the fidelity checklist and observation template, debriefed with the tutor, and crafted an analytic memo as described for the practice

administration. The first tutee observations took place in mid February, while the second set took place in mid-April.

Interviews

After collecting artifacts through the scoring checks and observations, I conducted semi-structured interviews with all consenting tutors so that I could better contextualize the artifacts. I selected a semi-structured interview format because it would enable me to direct the conversation to topics of interest while also allowing latitude to pursue lines of questioning that vary based on previous answers or the participant's background (Hays & Singh, 2011). In my interview protocol, my topics and questions were based on my conceptual framework as well as literature related to tutor training and assessment fidelity. Topic areas included questions related to the helpfulness of the training to support procedural and educative learning; adherence (e.g., tutors' confidence that they are administering and scoring the CBMs in alignment with standard protocol); and quality (e.g., the value that tutors perceive in using CBMs). For instance, to learn about tutors' perceptions of their own adherence after completing the training, I asked, "How confident do you feel that you can administer and score the CBMs according to the standard procedure?" As another example, to address the educative aspects of training, I asked the question, "How did the training contribute, or not contribute, to your investment in administering CBMs?" From this question, I hoped to learn about any potential impact on assessment quality. The complete interview protocol is available in Appendix F.

I conducted nine interviews and include a table below (Table 3.2) showing the interview dates and durations. Interviews ranged in length from 15 to 32 minutes, lasting an average of 25 minutes each:

Table 3.2

Interview Dates and Durations

Participant	Interview Date	Interview Duration
Janice	April 3	17 minutes
Kamaria	April 3	32 minutes
Leilani	March 29	24 minutes
Pam	March 29	20 minutes
Philonese	April 3	33 minutes
Renee	April 3	30 minutes
Sharon	April 3	24 minutes
Teri	April 4	31 minutes
Wanda	March 27	15 minutes

In preparation for the interviews, I referred back to my analytic memos from the tutors' observations before and after the interviews, as previous data from specific participants at times influenced the interview questions I chose to ask and my interpretations of their interview responses. During the interview, I took notes on the interview protocol to help determine what follow-up questions were appropriate and to record a brief analysis. Immediately following the interview, I took reflective notes to synthesize my interpretations of the interview and identify my areas of potential bias (Hatch, 2002). I used Zoom to record and transcribe each interview. Upon review of the recording and transcription, I ensured accuracy and added any analytical notes from the interview protocol.

Summary of data sources and collection. To address my research questions, I incorporated various data sources, including artifacts (baseline questions, scoring checks, observations) and interviews. Table 3.3 indicates which data sources I used to answer each question:

Table 3.3

Data Sources Used to Address Research Questions

Research Question	Baseline Questions	Scoring Checks	Observations	Interviews
What challenges do DRC tutors experience when administering and scoring CBMs?	x	x	x	x
In what ways do a DRC tutor's knowledge, skills, and beliefs impact their administration and scoring of CBMs?	x		x	x
In what ways can DRC improve current CBM training practices to build DRC tutors' capacity for administering and scoring CBMs?	x		x	x

Data Analysis

The goal of my study is to better understand the process of building tutor capacity for collecting meaningful data on tutees' progress at DRC. With this goal in mind, I developed my data analysis strategy, which involved making a detailed description of the case, analyzing multiple sources of data to look for patterns across data types, and developing naturalistic

generalizations based on the data to inform recommendations to the site (Creswell, 2007). In qualitative research, data analysis is integrated with data collection and with report writing, so the process might be conceptualized as a spiral that moves "in analytic circles rather than using a fixed linear approach" (Creswell, 2007, p. 150). Overall, I followed these steps: develop a system of data management, read all of the data, begin coding the data, identify themes, and craft a narrative interpreting the themes (Creswell, 2017).

Because I analyzed data from artifacts as well as data from interviews, I engaged in document analysis, using the documents in several ways as suggested by Bowen (2009). First, I used the information contained in my participants' responses to baseline questions and scoring checks to shape subsequent questions that I asked and situations I observed subsequently. For example, if a participant's response was inaccurate during a scoring check, I would observe them in the initial administration to determine the source of the error. Additionally, I used the documents to track change and development; for example, I interpreted participants' interview responses in the context of their responses to the baseline questions to see how their investment in progress monitoring shifted across the course of the training. Excerpts and quotations from participants' responses to the baseline questions and scoring checks were incorporated into my analysis as I applied codes and identified themes. By incorporating participant responses from training module artifacts, I triangulated my findings, seeking "convergence and corroboration" across data sources (Bowen, 2009 p. 28). This process served to minimize the impact of potential biases, building credibility and enriching the conclusions I drew.

To aid in the data management process, each participant in my study was assigned a pseudonym, and all artifacts and interview protocols were labeled with this pseudonym. For each participant, I created a folder in UVA's Box, which is a password-protected file storage that

requires multi-factor authentication. There, I stored artifacts (responses to baseline questions; results of scoring checks; observation protocols with field notes), interview protocols and transcripts, and relevant analytic memos. As I began the process of coding the data, I also stored Microsoft Word documents with tables for each participant in these folders.

To begin coding the data, I developed drafts of two *a priori* codebooks, one focused on adherence and one focused on quality. I divided the codebooks in this way because the adherence-focused codes were more closed-ended (e.g., the tutor either read the directions verbatim or they did not) in contrast to the quality-focused codes, which were more open-ended. I developed the codebooks by considering my conceptual framework, my research questions, my findings from the literature, and my observation and interview protocols (see Appendices G and H). For example, the adherence-focused codes were selected to align with fidelity in terms of adherence to the standard administration protocol for the CBM; the quality-focused codes aligned more with the quality aspect of fidelity, such as expressing an interest in a tutee's progress, or alluding to the rationale for engaging in progress monitoring. I read through all of the data and applied initial codes to the artifacts, interview transcripts, and memos. As I implemented this process, additional codes emerged, and I incorporated those into my codebook as well.

To help me identify themes, I recorded my thoughts and questions throughout the coding process, and I also wrote a more comprehensive memo at the end of each coding session where I began integrating my ideas. I returned to my research questions throughout this process to help shape my initial findings as they emerged.

Trustworthiness

In designing my study, I attempted to maximize trustworthiness, with attention to credibility, dependability, and confirmability of my findings (Shenton, 2004). I triangulated key findings by incorporating multiple sources and types of data. Through my range of artifacts, field notes, and interviews, I was able to address my research questions from multiple angles, gaining a more complete understanding of what my participants took from their training experiences. By collecting data over several weeks, I ensured that my data represented a more complete picture of a contextualized process, not just one moment in time. To further build trustworthiness, I used a peer reviewer to support the data collection and analysis process by providing feedback on my field notes, codes, and conclusions drawn from the data. Finally, I ensured that my descriptions of my context and procedures were specific and thorough enough that others can accurately gauge the applicability of my findings to their own contexts.

Ethical Considerations

To ensure that I maintain high ethical standards for my study, I followed guidelines from the University of Virginia and its Institutional Review Board. Because I work directly with the tutors, I took care in how I recruited and consented my participants. I asked my peer reviewer to email the tutors (see Appendix B) to explain the purpose of the study and obtain their consent to become participants. The text of the email stated that participation in the study was fully voluntary and would have no bearing on their placement at DRC. It also explained the nature of the study's activities, noting that, although artifacts were collected from all tutors as part of their work with DRC, their artifacts would be excluded from the study should they chose not to participate. Further, they were free to decline the interview request if they so chose, as consent is an ongoing process.

While the anticipated risk to participating in the study was low, I notified participants that, although I could not guarantee their anonymity, I would take measures to maintain confidentiality. For example, all participants were assigned a pseudonym to be used to label all artifacts and interview transcripts and used in the final report. I also housed all documents in the UVA Box platform, which is protected behind a two-step login procedure.

Researcher Reflexivity and Role

In all research design, it is important to recognize the ways and the extent to which the researcher's involvement with the subject matter may impact their choices in planning, execution, and data analysis and interpretation. This is especially true for qualitative studies, where the researcher is often more intimately involved with the subject and context (Hays & Singh, 2011). I have worked at DRC for the last seven years, and I am highly invested in improving reading outcomes for the children who participate in tutoring. As a reading specialist and doctoral student in reading education, I have some credibility due to my background and credentials. My personal experiences as a K-12 reading specialist and at DRC, along with the research base, have convinced me that data-based individualization is critical for providing tutees with the instruction they need in a timely manner. In this way, my personal goals and the organizational goals of DRC are essentially one and the same. The design of this study, and my interpretation of the data, were informed by the beliefs I have developed through my professional experience.

Further, in my work at DRC, I have spoken with parents and families who are concerned about whether their child is making appropriate progress toward meeting grade-level benchmarks in reading. My project was driven by my desire to collect high-quality data from which to draw

these kinds of conclusions. In the past, when I have tried working with tutors to administer CBMs, I have noted inconsistencies in administration and challenges in maintaining routines over the course of the semester. This experience shaped the content and design of the tutor training as well as how I developed my study and interpreted my data.

With its multiple data sources, my study's design helped me to work against some of my assumptions that could potentially have undermined the credibility of my findings. In analyzing my data, I foregrounded my problem of practice, conceptual framework, and the relevant literature to ground my interpretations. Further, I made use of a peer reviewer, or peer debriefer, throughout the research process (Hays & Singh, 2011). This peer reviewer has expertise in case study research as well as in reading education, so she had insightful feedback to share about the study's design and my process of data analysis. Because she is unaffiliated with DRC, she provided a more objective lens and helped me note when my assumptions or biases might have interfered with the conclusions I drew from my data.

Role of Researcher

Since the tutors and I already had an established relationship, I took care that our interactions did not exacerbate our perceived differences. In general, the tutors had experience with observations and feedback, but they may have misunderstood the use of the fidelity checks, seeing them as purely evaluative, especially because they knew I was looking specifically for any errors in administration they might make. They had the potential to feel self-conscious in the interview setting, which brought a level of formality not usually present in the context of their workplace. To help them feel as comfortable as possible, I shared information about my study so that they could anticipate the types of questions I would be asking them, what I was looking for in my observations, and so forth. When they had this information in advance, they gained a

better sense of how I was approaching our interactions, which was in the spirit of improving outcomes for our shared tutees. Creating a dynamic of mutual goals helped me establish reciprocity.

Chapter Summary

In this chapter, I explained and justified the design of my study, a qualitative case study. I then described my site and participants. Next, I detailed my data sources and procedures for analysis. Finally, I outlined my strategies for building trustworthiness, ethical considerations for my study, and how I perceive my role as a researcher. My next chapter will discuss my key findings and supporting themes.

Chapter 4: Findings

This capstone study was designed to investigate how tutors at Douglass Reading Center (DRC) use curriculum-based measures, or CBMs, to collect data on their tutees' progress. High-quality data is foundational to instituting DBI practices, which is an organizational goal of DRC. However, the DRC tutors, who are paid volunteers, experience challenges in administering CBMs with fidelity to established procedures, which undermines the utility of the data they collect. To learn more about this organizational challenge, the following research questions guided the current inquiry:

1. What challenges do DRC tutors experience when administering and scoring CBMs?
2. In what ways do a DRC tutor's knowledge, skills, and beliefs impact their administration and scoring of CBMs?
3. In what ways can DRC improve current CBM training practices to build DRC tutors' capacity for administering and scoring CBMs?

To investigate these questions, I conducted a case study of DRC and its ten in-person tutors. The tutors represented varied levels of tutoring experience and courses of study. I collected data across twelve weeks, beginning with training module artifacts (i.e., baseline questions and scoring checks) in January at the beginning of the training process. In February, I conducted two observations of tutors as they administered the CBMs; the first of these was the initial observation in the context of a 1:1 practice session with me, while the second was an administration with their tutee, the tutee observation. After each observation, I provided feedback and addressed any questions or concerns raised by the tutors. In these observations, I noted aspects of adherence to the standard administration procedure along with the tutors' affect. Additionally, I recorded the tutors' questions and comments before and after the two

observations. Then, in late March and early April, I conducted interviews with all tutors who consented and elicited responses pertaining to comfort with and investment in administering CBMs as well as any feedback on the training process. Finally, I conducted a follow-up observation in April, again noting aspects of adherence to the standard administration procedures. Table 4.1 summarizes the data collection process and timeline:

Table 4.1

Data Collection Process and Timeline

Stage of Data Collection Sequence	When Collected	Context Collected	Data Yielded
Training Module Artifacts	January/Early February	Embedded in training module in Canvas	Responses to baseline questions and scoring check
Initial Observation	Late January/Early February	Tutor administers a CBM to the researcher, who provides feedback	Observation checklists and field notes, tutor questions and comments
Tutee Observation	Mid to Late February	Tutor administers a CBM to a tutee, researcher provides feedback	Observation checklists and field notes, tutor questions and comments
Interview	Late March/Early April	Researcher interviews tutor	Responses to interview questions
Follow-Up Observation	Mid April	Tutor administers a CBM to a tutee	Observation checklists and field notes

In this chapter, I address the research questions by discussing three findings about challenges to assessment fidelity, tutors' investment in using CBMs, and the impact of the training process on their investment. I use direct quotations from the baseline questions and

interviews as well as from the interactions following the observations. I also describe general trends in assessment fidelity that I observed across the twelve weeks as tutors progressed through the training process. In my analysis, I relied most heavily on direct quotations from tutors across the training process, and I used the fidelity checklists as a tool to contextualize and interpret these quotations. This was because, as I discovered, reasons tutors deviated from standard procedure were often individualized and specific to a given administration. With the training module scoring check, for example, the scores tutors reported seemed sufficiently haphazard that I questioned my use of an audio recording in the training, and, therefore, the utility of those scores for my analysis. Similarly, instances of tutors backsliding, or scoring "needs practice" on a criterion met in a previous observation, demanded individual consideration based on the particular tutor and administration instance. Ultimately, the various types of data that I collected address the research questions in an integrated way, enabling me to identify themes, and, ultimately, the findings shared in this chapter:

- Finding 1: Challenges in achieving assessment fidelity often stem from tutors' care for their tutees' comfort and confidence.
    - Theme 1: Tutors as Guides
    - Theme 2: Tutors Minimizing Tutee Stress
    - Theme 3: Administration Errors that Stem from Care for Tutees
- Finding 2: While tutors were already invested in their tutees' success, the training process heightened their investment in measuring success through CBMs.
    - Theme 1: Tutor Investment in Tutee Success
    - Theme 2: Impact of Training on Tutor Assessment Priorities
    - Theme 3: Investment Does Not Guarantee Fidelity on Any One Administration

- Finding 3: Many tutors made relevant connections across the data they collected, the instruction they provided, and next steps.
    - Theme 1: Tutor Perceptions of Standardized Measures and Instructional Goals
    - Theme 2: Tutor Perceptions of Variability
    - Theme 3: Tutor Ideas about and Attempts at Responsive Instruction

These findings are specific to the in-person tutors at DRC (n=10 for artifacts and observations, n=9 for interviews) and are not generalizable to other tutoring centers or groups of paid volunteer tutors. However, by identifying commonalities and tensions, I lay a foundation for recommendations to DRC as seen in Chapter 5.

Finding 1: Challenges in achieving assessment fidelity often stem from tutors' care for their tutees' comfort and confidence.

Almost without exception, the DRC tutors indicated that they valued the relationships they built with their tutees over the semester, or even across the academic year. They aimed to create a tutoring environment that fostered their tutees' comfort and confidence. Because the tutors recognized that the CBM passages could be challenging for tutees, many tended to sacrifice assessment fidelity in favor of easing tutees' discomfort, whether that discomfort was observed or anticipated. The most common administration errors by far related to modifying task directions (which should be read verbatim) and neglecting to provide a word when a tutee hesitated while reading.

In this section, I explain how tutors perceive their role and how they describe their relationships with their tutees. Next, I explain how these relationships influence how tutors frame the CBM task for their tutees. Finally, I trace the most common administration errors back to those relationships. I present this overall finding through analysis of several themes, including

tutors as guides, tutors minimizing tutee stress, and administration errors that stem from care for tutees. This finding addresses Research Question 1 because it articulates challenges to administration fidelity that tutors face because of their tutee relationships. This finding also sheds light on Research Question 2 because it reveals the variety of ways tutors cope with these challenges based on their previous experiences that shape their knowledge, skills, and beliefs. Indirectly, it also addresses Research Question 3 by illuminating specific aspects of fidelity that should be better emphasized in the tutor training.

Tutors as Guides

To contextualize trends in assessment fidelity at DRC, it was necessary to explore how tutors conceptualized their role. The role of the tutor provided the foundation for the one-to-one relationships that they built with their tutees, and these relationships had implications for the types of administration errors tutors tended to make and why.

Contrasting with the Role of "Teacher"

When explicitly asked about their role in the context of an interview, nearly all tutors indicated that they saw their main job as supporting their students, more like guides than like teachers. In fact, several tutors drew an explicit contrast between their supporting role as a tutor and the role of a teacher. For instance, Kamaria emphasized, "I definitely don't see my role as a teacher or something, it's not like I'm teaching our kids something for the first time. It's more supporting what they're already learning hopefully in school, and strengthening their strategies" (Interview, April 3, 2023). Similarly, Philonese expressed, "Often, what we're working on they've already worked on in school, so my role is not to be a teacher but to help them practice what they're doing and just to guide them to the right answer" (Interview, April 3, 2023). Wanda, too, stated that she sees her role as a tutor is to "provide extra support for them. I feel

like I'm not really teaching them new things, but just aiding them in what they already know and building their confidence in reading" (Interview, March 27, 2023). These tutors were reluctant to take on the title of "teacher," perhaps because, to them, that title embodied authority that they would feel presumptuous to claim. Rather, the tutors wanted to provide support and extra practice in hopes of building tutees' confidence.

Building Confidence

Like Wanda in the quote above, several tutors expressed that they saw their role as one of confidence-building. Leilani, who particularly emphasized this, was an experienced tutor, having worked with DRC for three years. In her baseline questions, she expressed that, to her, "confidence is the most meaningful measure of progress" (Training Module Artifacts, February 1, 2023). She described confidence as students feeling empowered to use the tools she had given them, especially when they employed them without prompting. She wrote, "If I can instill the confidence to trust in their abilities and use their tools, then I have faith that a student will be able to progress." In her interview, Leilani echoed these sentiments, explaining that, over time, the 1:1 relationship with her students enabled her to build their confidence by allowing room for mistakes, stating:

> So, with my new student, [student name], she seems hesitant to make a mistake and so sometimes she won't even try if she thinks she's going to get it wrong. But with students I've had longer like [student name], he's okay with trying and making mistakes, because he knows I'll correct him, and I'll help him learn, and give him those tools for next time when he comes to a similar word that he can't understand or has a hard time with (Interview, March 29, 2023).

The idea of creating a "safe space" and minimizing "pressure" for tutees was apparent in other tutors' responses as well (e.g., Teri, Interview, April 4, 2023; Sharon, Interview, April 3, 2023). In general, the tutors' understanding of their own role as guides for their tutees aligned with DRC's vision of a successful tutor-tutee relationship and set the stage for positive interactions in the context of instruction.

Tutors Minimizing Tutee Stress

Since tutors prioritized their supportive relationships with their tutees and even, in some cases, perceived assessments as a source of stress, many tutors expressed empathy for their tutees in the context of CBM administration. Two tutors, Sharon and Pam, anticipated tutee anxiety even during the initial observation, before their first attempt to administer the CBM to a tutee. Notably, both tutors made connections to their own testing anxieties. Pam stated, "I just remember when I was in school being like… stressed when we would do things like this, like a test" (Initial Observation, January 30, 2023). Sharon commented, "I hate testing" and added, "I've never tested like this before!" indicating that, while administering a CBM was foreign to her experience, she connected it to her own anxieties about being assessed. In fact, she appeared to be nervous during the practice administration (Initial Observation, February 1, 2023).

Introducing the CBM Task to the Tutee

It seemed that Sharon's affective response to administering the CBM, even in practice, led her to ask about an aspect of administration that was outside the scope of the asynchronous training modules– how to introduce the CBM task to the tutee for the first time. Embedded in her comments about her nervousness, she asked whether she should discuss the task with her tutee first so that it wouldn't feel like "throwing a test at him" (Initial Observation, February 1, 2023). This empathetic consideration of the impact CBMs might have on tutees was apparent for many

of the tutors. One other tutor, Kamaria, commented on this during the initial observation ("I want to try to frame it for the kid like– no pressure on you!" (Initial Observation, February 1, 2023). These comments indicated that tutors were cognizant of the potential for their tutees to experience some performance anxiety with the CBMs, and, even before an initial administration, they were thinking about how to temper that impact. Because tutors value their relationships so highly, they took care to ensure that the "testing" stance did not undermine their tutees' confidence.

Striking a Balance

By the interview stage, having administered the CBMs several times, many of the tutors discussed the challenges of striking a balance between presenting the CBM as a serious task while avoiding placing undue stress on the student. Renee commented that CBM administration by the tutee's own tutor, rather than someone unknown to the tutee, could help alleviate stress and capitalize on the relationship built between tutor and tutee; as she said, it shows the tutee that "it's not something to feel extremely anxious about" (Tutee Observation, February 22, 2023). On the other hand, if the tutee was *too* comfortable with their tutor, they may not have been inclined to take the CBM seriously. In her interview, Teri said that she told her tutee, "We're not grading you. We just want to make sure that you're getting enough help and support" (Interview, April 4, 2023). However, of that same tutee, she said, "I have to really hone in and be like, this is serious, you need to sit in the chair, and you are reading this, you know, because sometimes they'll read it in a funny voice" (Interview, April 4, 2023). Observationally, this tutee brought a great deal of energy to his tutorials, and engaging him in instruction required a great deal of creativity and accommodation on Teri's part. However, the CBM administration did not allow this kind of leeway, so it stood in contrast with how instruction generally worked in the lesson.

This desire to balance tutee comfort with the demands of CBM administration was evident for other tutors as well. Kamaria described her struggles:

> I'm trying to keep [the tutee] motivated and going along without invalidating [the CBM] or going too much off the standard or taking away too much time. Like trying to find the balance of letting him do his thing– but also, if he did all of his own things, then it would be crazy (Interview, April 3, 2023).

In a similar vein, Leilani stated that her biggest challenge in administering CBMs is finding "the sweet spot where the student knows to take it seriously, but also not to feel too stressed… like, the performance matters, but also, I'm not testing you, it's not the end of the world type of testing" (Interview, March 29, 2023). In these comments, it was apparent that the tutors cared about their tutees' emotional response to assessment and were grappling with how to achieve the goals of the assessment while honoring their relationships with their tutees.

Administration Errors that Stem from Care for Tutees

Because tutors were so invested in their relationships with their tutees, and because aspects of CBM administration seemed to the tutors to undermine those relationships, several elements of administration fidelity appeared, time and again, to be challenging for the tutors. In the practice session, the most prevalent error was modifying the directions of the CBM, which are expected to be read verbatim. Generally, tutors modified these directions in an effort to help the hypothetical tutee feel more comfortable. When provided with feedback that the directions were to be read verbatim, they generally met this fidelity criterion in subsequent observations. When administering CBMs with tutees, the most common administration error was failing to provide the correct word after three seconds when the tutee hesitates. Tutor comments during the interviews revealed that this error, too, stemmed from tutors trying not to undermine their tutee's

confidence. Tutors' prior experiences seemed to impact the degree to which their concerns about using CBMs with their tutees went on to impact their administration fidelity. Table 4.2 summarizes the following section, highlighting four tutors with different backgrounds and the impact those various backgrounds had on assessment fidelity:

Table 4.2

Tutor Background and Impact on Fidelity

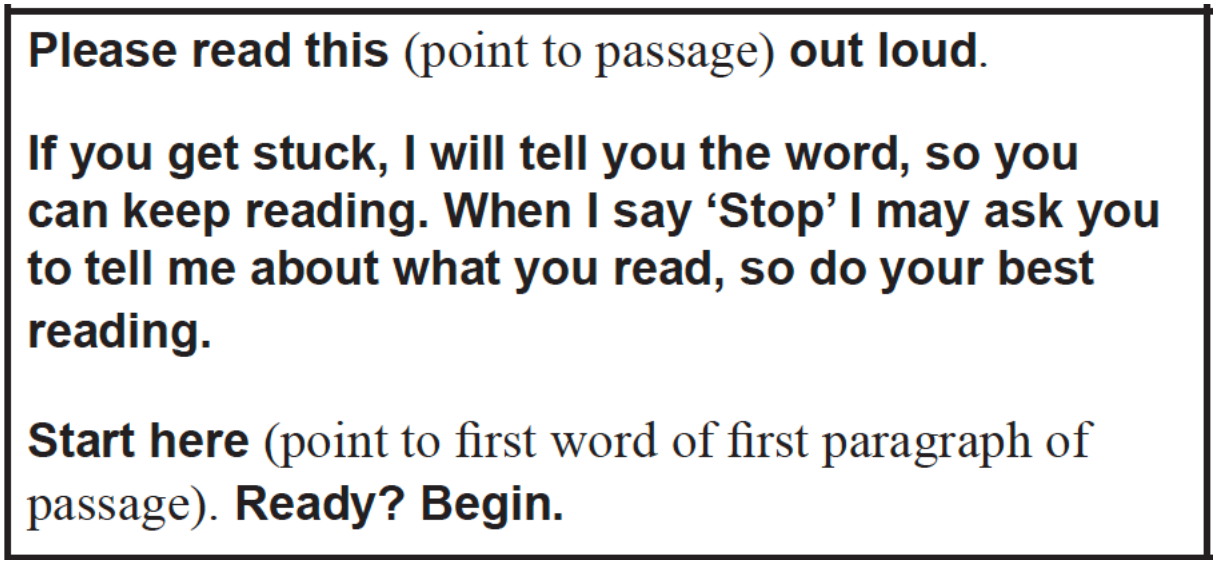
Fidelity Criterion	Tutor	Prior Experience with Standardized Assessments?	Did the Tutor Consistently Meet this Fidelity Criterion?	Tutor Comment Regarding Administration Challenge
Task Directions	Philonese	No	No (error on Initial Observation, January 30, 2023)	I noticed that, like, **my word choice could definitely affect the way kids choose to read it**. Because I remember I initially said, like, 'Take your time,' but no, you don't want them to take their time, because you want them to meet a time goal. **But you also don't want to pressure them**.
	Renee	Yes	Yes	The kids would ask questions or just not understand my directions. So, then I would just have to, like, **stick to the script** and make sure I wasn't adding anything to it… **even if it's something slight, it might shift everything**. So **that's hard for me, because I love talking to the kids.**
Feedback During Reading	Sharon	No	No (error on Tutee Observation, February 15, 2023; error Follow-Up Observation, April 12, 2023)	I do feel bad for the kids because if they don't say the word within three seconds, **I have to tell them and mark it wrong, and I hate doing that** because I know if I gave them another second or two, they'd be able to get it. I think they notice it, especially [student name], because she'll say, oh, I was going to get that, so then **I just feel really bad in that moment**, because **I want to give them the time to process it, but I can't**… I just feel like it hurts their confidence to read.
	Wanda	Yes	Yes	**It's hard for me letting them struggle** for those three seconds, or if I see them self-correct not to be like, 'Oh good.' Generally, if my student does something good, I want to reward them, **but for CBMs it's not so much about learning as demonstrating what are their current skills.**

Changes to Task Directions

As a point of reference, the standardized directions for the CBM read as follows in Figure 4.1:

Figure 4.1

Standard CBM Task Directions



Several features of these directions are noteworthy. For one thing, they are explicit in telling the tutee exactly what is expected of them on the task. Also, they let the tutee know what they can expect from the tutor in terms of feedback– providing a word for them if they get stuck so that they can keep moving through the passage. Finally, by telling the tutee that they may be asked about what they read, the tutor sends the message that reading is about making meaning, not just saying the words on the page; the implication to the tutee is that their "best reading" is not about reading quickly, but about engaging with the text.

If a tutor modifies these directions, a tutee may experience confusion about how to proceed or what to expect. The tutee's interpretation of expectations could shape how they

approach the task, and, therefore, impact their score (Christ et al., 2013). In the initial observation, some tutors made relatively minor modifications that, nonetheless, could undermine trust in the data collected. For instance, Sharon incorporated extra wording into her directions: "Today you're going to read me just a story and just read as well as you can" (Initial Observation, February 1, 2023). Her use of the word "just" implies that she was trying to downplay the task for the hypothetical tutee, seemingly in an attempt to alleviate anticipated anxiety.

Other changes that tutors made to the directions were more substantial. In her initial observation, Philonese modified the task directions to say, "Today we're going to be practicing some reading, I just want you to take your time and read it– okay, whenever you feel ready. No pressure, you got this. Go ahead" (Initial Observation, January 30, 2023). By telling the hypothetical tutee to "take their time," Philonese could have impacted the rate at which a tutee read the passage, undermining the quality of the data collected. When we discussed this at the end of the initial observation, Philonese thanked me for bringing it to her attention, expressing that she was unaware she had even said it. In subsequent observations, she delivered the instructions verbatim, and she reflected on her growth in her interview, saying, "I noticed that, like, my word choice could definitely affect the way kids choose to read it. Because I remember I initially said, like, 'Take your time,' but no, you don't want them to take their time, because you want them to meet a time goal. But you also don't want to pressure them" (Interview, April 3, 2023). In this comment, Philonese made an explicit connection between her modification of the directions and her concerns about her tutees' anxiety about testing. Through the training process, she came to recognize the potential impact of those changes, which likely contributed to her fidelity on subsequent observations.

Similarly, from the experience of administering CBMs over time, Renee came to recognize the potential impact of deviating from the standard directions, even if those deviations were in the spirit of making her tutees more comfortable. She stated that the hardest aspect of administration for her was that "the kids would ask questions or just not understand my directions. So, then I would just have to, like, stick to the script and make sure I wasn't adding anything to it… even if it's something slight, it might shift everything. So that's hard for me, because I love talking to the kids" (Tutee Observation, February 22, 2023). In this statement, Renee articulates the challenges of maintaining fidelity with task directions when administering CBMs, especially as a tutor who values her relationships with tutees.

Tutor Background and Task Directions. Notably, tutor background appeared to impact assessment fidelity on task directions most visibly in the initial observation. Both Sharon and Philonese, who did not read the directions verbatim, are tutors who do not have a background in teaching or clinical practice. On the other hand, Renee, Wanda, Teri, and Kamaria, all of whom have relevant experience in teaching or in the SLP graduate program, met this fidelity criterion on the initial observation and maintained it throughout the study. Sharon and Philonese, after receiving feedback following the initial observation, met the criterion in subsequent observations.

In the quotes shared from Philonese and Renee above, both tutors expressed similar ideas about the tension they experienced trying to maintain assessment fidelity while honoring their relationships with their tutees. Even though both tutors were tempted to elaborate on the directions to make their tutees feel more comfortable, Renee had internalized from her previous clinical experiences that, for an assessment like a CBM, directions are to be read verbatim. Therefore, Renee met this criterion from the beginning of the training process, while Philonese

did not. This contrast serves as an example of how a tutor's knowledge, skills, and beliefs might impact their adherence to standard administration procedures.

Feedback During Reading

In observations of tutors working with actual tutees, the most common error observed was improper feedback provided while the tutee was reading their passage. While a few tutors tended to jump in too quickly to provide a word when their tutee hesitated, more tutors waited longer than the required three seconds, or did not offer the word at all. For some, this tendency persisted even into the follow-up observation.

In their interviews, several tutors raised this feedback as an area of challenge for them in learning to administer CBMs correctly, and, in a few cases, tutors expressed feelings of guilt about providing the word for the tutee. Teri, an SLP graduate student, commented on her tutee's reaction to feedback while reading a CBM passage, saying that he specifically asked her to stop reading the words for him because "he is insecure and doesn't like to be corrected" (Interview, April 4, 2023). Despite her observations about her tutee, Teri provided feedback correctly during her tutee observation as well as her follow-up observation.

Sharon, a tutor who experienced more difficulty providing this feedback, expressed more of her own affective response to engaging in the standard correction procedure:

> I do feel bad for the kids because if they don't say the word within three seconds, I have to tell them and mark it wrong, and I hate doing that because I know if I gave them another second or two, they'd be able to get it. I think they notice it, especially [student name], because she'll say, oh, I was going to get that, so then I just feel really bad in that moment, because I want to give them the time to process it, but I can't… I just feel like it hurts their confidence to read (Interview, April 3, 2023).

In this quote, Sharon shows how the procedures for CBM administration are in tension with her desire to support her tutees and build their confidence. In one of her baseline responses, Sharon had written, "I enjoy tutoring because I like interacting with children and guiding them with tools and tips from lessons they are already learning at school" (Training Module Artifacts, January 22, 2023). The attitudes expressed in her baseline quote likely connect to her strong affective response in the interview, and her sense of herself as a tutor led to challenges with assessment fidelity.

Tutor Background and Feedback During Reading. As with task directions, tutor background impacted fidelity on this criterion. However, errors on this criterion proved harder to correct over time. For tutors like Teri, who provided feedback correctly from the beginning, fidelity was generally maintained. However, Sharon allowed her tutee to hesitate on a word for more than three seconds, both in the tutee observation and the follow-up observation. Of note, Sharon was recently admitted into a teacher education program, so she is a pre-service teacher who has not yet taken any education coursework. Perhaps, through that coursework, she will develop an understanding of assessment that better enables her to negotiate the tension that she feels.

In contrast to Sharon, Wanda is an SLP graduate student, and, although she too expressed a feeling of guilt, her comments conveyed an understanding of the role of assessment in the tutorial context. She stated, "It's hard for me letting them struggle for those three seconds, or if I see them self-correct not to be like, 'Oh good.' Generally, if my student does something good I want to reward them, but for CBMs it's not so much about learning as demonstrating what are their current skills" (Interview, March 27, 2023). Because of her coursework and clinical experience, Wanda is likely more accustomed to reconciling her dual roles of instructor

and assessor. Because of her care for her tutees, she found assessment administration to be "hard" at times, but she was one of the few tutors who administered the CBM with fidelity beginning with the initial observation and maintained it through all subsequent observations. The contrast between Wanda and Sharon illustrates how a tutor's background can impact assessment fidelity.

Summary of Finding 1

Tutors at DRC prioritize the relationships built with their tutees, and, at times, they see these relationships as being at odds with administering CBMs with fidelity. Tutors conceive of themselves as guides for their tutees, and they strive to create a learning environment that builds tutee confidence. In their view, the use of CBMs has the potential to undermine that confidence and to engender anxiety in their tutees. Tutors attempt to strike a balance with their tutees, trying to convey that the CBMs should be taken seriously but are not meant as a source of pressure for tutees. At times, tutors' attempts to strike that balance result in assessment errors, particularly in how they convey task directions and provide corrective feedback while tutees are reading. While errors with task directions were more easily corrected by the tutors, feedback challenges persisted for a few, even through the follow-up observations. Tutors' varied backgrounds likely contributed to some of the differences observed in assessment fidelity.

Finding 2: While tutors were already invested in their tutees' success, the training process heightened their investment in measuring success through CBM data.

Because tutors work with tutees consistently across a full semester, or longer in many cases (e.g., Teri, Wanda, Kamaria, Leilani, and Pam have all been with at least one of their tutees for the academic year) they feel a responsibility to help their tutees succeed. At times, they even question their own efficacy as tutors if they perceive a lack of success in their tutee, as they

define success. Through the training process, tutors became more invested in CBMs as a measure of tutee success for two reasons. First, the in-depth nature of the training made it seem important; second, the educative aspects of the training conveyed why CBMs were valuable. For some tutors, this understanding of the value directly connected to greater assessment fidelity, but not for everyone. Two tutors, Sharon and Janice, minimized the importance of CBMs, but they demonstrated assessment fidelity nevertheless, at least in terms of adherence. Other tutors expressed investment in using CBMs at DRC but made one-off administration errors that impacted their adherence. This group included tutors such as Renee and Kamaria, who had prior experience with standardized assessments and who expressed high levels of investment. Though these trends may seem to indicate that assessment fidelity is unrelated to tutor investment or tutor background, in the long run, the invested tutors may exhibit better quality, because they think deeply about how to handle these one-off errors. This finding mainly addresses Research Question 2 through considering how tutors' different beliefs about the utility of CBMs impact their assessment fidelity. Additionally, it relates to Research Question 3 by (a) highlighting the efficacy of the educative aspects of the training in its current iteration, and (b) suggesting that future iterations might consider differentiating training based on tutors' varied backgrounds.

Tutor Investment in Tutee Success

Because the tutoring format at DRC enables tutors to build sustained, one-to-one relationships with their tutees, many of the tutors expressed investment in their tutees' reading success, and, at times, they couple their sense of worth as a tutor to that success. One tutor, Pam, particularly conveyed this sentiment with consistency across all stages of the training process, and, by the end, she came to see the CBM data as one potential measure of success. As she wrote in her baseline questions, "I have become invested in [my students], and I think it would make

me feel proud if they improved" (Training Module Artifacts, January 23, 2023). After her practice session, she expressed that she was "curious" to see how her students would do on the measures, especially one particular student with whom she had been working across multiple semesters of tutoring. Then, in her interview, she again referred to her sustained relationship with her tutees and indicated that, because of that relationship, she felt a "personal responsibility to look at the [CBM] scores and act on them, or at least take notice of them" (Interview, March 29, 2023). Later in the interview, Pam stated some specific ways in which she interpreted and took action in response to her tutee's data. Overall, although Pam was invested in her tutee's progress from the beginning, administering CBMs furthered her investment and encouraged her to start thinking about what she might do instructionally to support her tutee.

Tutors Questioning their Efficacy

Notably, Pam, along with several other tutors, also seemed to become invested in the CBM data to a degree that made them question their own efficacy as tutors. As Pam discussed the efficacy with whom she had been working across the full academic year, she shared, "I was looking at her scores, and they just kept going down every week, and I was like, I don't know if it's me, I don't know what we're doing" (Interview, March 29, 2023). Similarly, in her interview, Philonese compared her tutee's scores to those of his twin brother, who works with a different tutor. She noted that both tutees were improving in their accuracy, but that her tutee's rate was building more slowly in comparison to that of his brother. This prompted her to ask, "So, as a tutor, what am I not doing?" (Interview, April 3, 2023). Of her tutees' CBM data, Leilani, too, stated:

> Sometimes I'm curious if it reflects if I am doing a good job with them and if I'm doing a good job with getting the information and the concepts to them. Because I know

> sometimes when I log the data and their accuracy or word count has gone down sometimes, I think, "Oh my gosh, have I failed the student?" (Interview, March 29, 2023).

All tutors who expressed feelings of failing their tutees were undergraduates who did not have previous experience in any other clinical or school contexts. In contrast, tutors who are SLP graduate students (e.g., Wanda, Teri, and Kamaria) or the pre-service teacher who had completed student teaching (i.e., Renee) discussed tutee progress without the connotation of somehow being at fault for a tutee's CBM score. Overall, the majority of tutors seemed invested in their tutee's success and looked to the data they collected as a measure of success. For less-experienced tutors, there was also a tendency to question their own worth as a tutor if, in their interpretation, the data did not show adequate progress.

Impact of Training on Tutor Assessment Priorities

For many tutors, the training process impacted the importance they placed on CBMs as a measure of tutee progress. This was the case for two reasons: (a) the in-depth nature of the training conveyed its importance and (b) the training provided more information about the rationale behind CBM administration than it had in prior iterations. As an encapsulation of those reasons, in her interview, Renee observed that she had "seen it in the [other] tutors, that they now give [CBM administration] a lot of importance" to a degree that she had not observed in the past, and she credited that shift to the training's greater emphasis on "why it's necessary" (Interview, April 3, 2023). Table 4.3 features key quotes from tutors regarding the educative aspects of the tutor training and those tutors' administration fidelity on the follow-up administration, which represents the "independent practice" stage of the training process, reflecting overall learning:

Table 4.3

Tutor Quotes about Educative Aspects of Training and their Administration Fidelity

Tutor	Fidelity on Follow-Up Administration?	Key Quote
Renee	Yes	I think [the educative aspect of training] is important so people can understand why [progress monitoring is] necessary, and I've seen it in the tutors now that they **give it a lot of importance**, to really see the growth, and they're aware of that.
Pam	Yes	[The training] made [giving CBMs] at least feel like it was something I **needed to take seriously** and be very accurate with... [the educative aspect of training] helped **put into perspective why it's important**. Maybe if I didn't have that information, I wouldn't take it as seriously.
Kamaria	Yes	Psychologically, for the tutor, [the training] is **giving them more responsibility**, involving them more, and it makes them feel more professional. The training is not just saying, "You have to start doing this," but it's explaining **why I want to buy into that and use it**… it's **involving us more actively** and treating us more like real tutors
Philonese	Yes	I think last year I knew this was something we did, I just thought it was part of a procedure. I think this semester, **[the training] made me think [the CBM] was for the students**, for the kids. It wasn't just information [to] track but it was also **to help us adjust so we could help the students.**

The shift that Renee noted was borne out in several tutors' interview responses when they were asked directly about the impact of training on their comfort with and investment in CBM administration. Pam's impression was that the training made her "feel like it was something I needed to take seriously and be very accurate with." When asked whether she found the educative aspects of the training to be helpful, she responded that having the rationale "helped put into perspective why it's important. Maybe if I didn't have that information, I wouldn't take

it as seriously" (Interview, March 29, 2023). Similarly, Leilani stated that the training conveyed to her a sense of importance and that "feeling qualified that I can give something made me feel like, 'Ok, this is something that I can give and that will be important to give'... you feel really prepared, it's real material" (Interview, March 29, 2023). On a follow-up question, she, like Pam, indicated that the educative elements of the training process helped engender that significance for her. Although these tutors did not independently connect the sense of importance to knowing the rationale, they affirmed that they do see those things as related.

Educative Training Elements Lead to Investment

For other tutors, the link between the importance of CBMs and the educative elements of the training process came more organically, because they made the connection without being asked about the training directly. One of these tutors was Kamaria, who, in her interview, made an explicit connection between the educative aspects of the training and her mindset about assessment fidelity. "I want to make sure I'm ready," she said, "because a difference of one error can make a big difference." The training impacted her, she said, because:

> Psychologically, for the tutor, [the training] is giving them more responsibility, involving them more, and it makes them feel more professional. The training is not just saying, "You have to start doing this," but it's explaining why I want to buy into that and use it… it's involving us more actively and treating us more like real tutors (Interview, April 3, 2023).

Interestingly, earlier in her interview, Kamaria had been careful to state that she did not see herself as "a teacher." Yet, the comments she offered later in the interview conveyed a sense of deep responsibility on her part for administering the CBMs with fidelity. For her, the educative aspects of the training process were what helped create that sense of responsibility.

Another tutor, Philonese, provides an additional example of how the educative aspects of training directly impacted her assessment fidelity. She explicitly conveyed how the revised training, with its emphasis on why CBMs are important, changed her perspective. In her initial observation, she made several administration errors, and, when we discussed these, she credited the training process with identifying these errors, stating, "I do appreciate how much effort you're putting into [the training], I can see how you're trying to refine our practices… having this meeting is helpful, I appreciate you clarifying that I've been doing it wrong" (Initial Observation, January 30, 2023). Then, in her interview, she connected the shifts in her administration practices to the rationale shared in the training process, saying:

> I think last year I knew this was something we did, I just thought it was part of a procedure. I think this semester, [the training] made me think [the CBM] was for the students, for the kids. It wasn't just information [to] track but it was also to help us adjust so we could help the students (Interview, April 3, 2023).

For Philonese, the procedural aspects of the training process alone were not sufficient. To attain higher standards of adherence, she needed the educative aspects of the training, because she needed to understand more about the purpose of CBM administration.

Educative Training Not Impactful for All Tutors

As a caveat to this finding, there were two tutors for whom the educative aspects of training did not have a discernable impact on their adherence or quality. In interview responses, both Sharon and Janice explicitly contrasted their roles as tutors with what it means to be a teacher, and both also explained that they were administering CBMs in their tutorials simply because it is the expectation at DRC, not because they see value in the practice. When asked about the impact of the training on her investment in CBM administration, Sharon replied, "I

know this sounds really bad, but you told me to administer the test to the students and you gave me the reason why and I went along with it" (Interview, April 3, 2023). I then asked her if having that reason why influenced her investment, to which she replied, "To be honest, I don't think it really made a difference. You know, I'm not a teacher or anything, so if you tell me to do it, I'm like, 'Ok.' But if I was a teacher, I think knowing the rationale would be very helpful" (Interview, April 3, 2023). The phrases with which Sharon introduced these two responses convey that she was being frank in the interview, and she seemed apologetic. Despite feeling like she was "supposed to" be more invested in administering CBMs, her honest responses revealed that the educative aspects of training did not lead to greater administration quality for all tutors.

Like Sharon, Janice's interview responses undermined the connection between educative aspects of training and administration adherence and quality. Previously in her interview, Janice had stated that her tutee's performance on CBMs would not impact her instruction in any way; based on that comment, she went on to say, "Perhaps it's not for me, but maybe other tutors can use this [CBM data] in helping their kid" (Interview, April 3, 2023). When asked whether the training impacted her investment in administering CBMs, she said that it did not; rather, she said, "because I had the training, I know how to do it… so I might as well do it and do it well" (Interview, April 3, 2023). She continued, "I'm [administering the CBMs] because we're supposed to. I wouldn't do it if we weren't supposed to. Like, doing the training wouldn't make me want to do it if we didn't have to (Interview, April 3, 2023). Given this lack of investment, one might expect Janice's adherence to standard procedures to suffer; however, the opposite was true. Janice was one of very few tutors to reach adherence on her practice administration and maintain it through all subsequent administrations, and she did not ask any questions. Therefore, her case serves as a counter-example to this overall finding.

Investment Does Not Guarantee Fidelity on Any One Administration

Just as a lack of tutor investment in CBM administration does not necessarily undermine their fidelity, evidence shows that, in some cases, investment is not a guarantee of fidelity, either. Furthermore, achieving fidelity is not necessarily a linear process. While some tutors made errors on the practice administration or initial observation that they later corrected, others made different errors each time, or made errors in their follow-up observation that they had not made earlier in the process. For these tutors, these instances of "backsliding" did not seem to be a function of their degree of investment, indicating that adherence and investment do not always directly correlate.

To illustrate this point, three tutors whose paths were not always linear were Teri, Kamaria, and Renee. Importantly, through their responses to interview questions, all of these tutors conveyed a high degree of investment in administering CBMs correctly, and they all had previous experience with standardized assessments as SLP graduate students or as teacher candidates. However, various situations that arose during their administrations seemed to throw them off track, generally because the training conditions did not reflect a wide enough array of circumstances to sufficiently prepare them. For example, for Teri and her spirited student, each stage of the training process seemed to present a new challenge whereby, even when she achieved a "pass" on fidelity criteria she had previously missed, different errors took their place. Ultimately, on the follow-up observation, she met all fidelity criteria, but, despite her desire to get things right, her learning process was haphazard.

Unlike Teri, Renee demonstrated assessment fidelity from the beginning of the training process. However, on her follow-up observation, her tutee skipped a line while reading the passage. Prior to her CBM training through DRC, Renee had assisted with a research study

where she had been extensively trained to administer CBMs, and she had ample practice administering them with students in various schools to collect data for the study. Despite all her training and prior experience, Renee had never encountered a student skipping a line, and she did not know the correct way to react in the moment. Therefore, she serves as an additional example that adherence can be compromised despite a tutor's investment.

As a final example, Kamaria expressed a high level of investment in administering CBMs correctly, and she did so from the beginning on every observation. However, she approached me about an administration that I did not observe and explained that she did not feel it was representative of what her tutee could do. As we discussed it together, she explained that her tutee was having an unusual emotional reaction to the challenges of the task. In response, Kamaria had provided feedback for him that was outside the standard administration procedure. She asked if the proper course of action would be not to report that day's score, and I agreed.

This instance raised some questions about the relationship between investment and fidelity that I had not previously considered. Although Kamaria's adherence to procedure in that administration was undermined, she acted to preserve her relationship with her tutee in that moment. Perhaps, by letting go of administration fidelity on one occasion, she prevented an adverse reaction from her tutee that could have undermined all future administrations. Furthermore, rather than simply reporting the score, she chose to ask about it, indicating that she was invested in abiding by proper procedures. In her interview, Kamaria spoke about the training leading her to feel "professional" and "responsible," and it may have been those sensibilities that prompted her to inquire about the right course of action, rather than simply reporting the score or estimating what it "should have" been, as a less-invested tutor may have done. I also reflected on Renee, who engaged with me at length as we discussed what to do when a tutee skips a line.

Perhaps a less invested tutor would have simply accepted the correction without considering why the procedure was what it was; that tutor may have gone on in future administrations to redirect the tutee inappropriately or to score based on how the tutee "would have" done had they not skipped the line. While tutor investment does not guarantee fidelity on any particular administration, it does encourage tutors to ask questions and seek clarification that could improve their administration fidelity over time.

Summary of Finding 2

Nearly all of the tutors expressed a desire to see their tutees improve in reading. When they did not see what they considered to be sufficient progress, several tutors attributed the lack of progress to their own actions, questioning their own efficacy as tutors and wondering what they might do differently to better support their tutees. For many tutors, undergoing the CBM training process increased their investment in using CBMs as a measure of tutee progress. This change resulted from how the training conveyed the high expectations for CBM administration as well as the rationale for administering them– in other words, the change resulted from the educative aspects of the training. For some tutors, this increased investment led directly to improvements in assessment fidelity, both in terms of adherence and quality. That being said, for other tutors, the educative aspects of training did not have an appreciable impact on fidelity, at least in terms of adherence. Quality is more difficult to observe and is perhaps more evident in the long run. Even for tutors who report a high level of investment in proper CBM administration, any one administration occasion may be compromised in terms of adherence; however, a tutor's reactions to singular circumstances that might arise across multiple administrations is an indicator of quality.

Finding 3: Many tutors made relevant connections across the data they collected, the instruction they provided, and potential next steps.

Some tutors made a natural connection between the CBM data and the instructional focus of their tutorials; therefore, they view CBMs as a meaningful measure of tutee progress. Tutors who did not see the relationship between their instruction and the CBM data tended to be less invested in CBMs, and they did not see the data as meaningful to them and their instructional practices. As noted in Finding 2, less investment does not necessarily equal less fidelity, at least not in terms of adherence to standard procedures. Several tutors noted variability among passages and among student factors, but, with few exceptions, these questions about variability did not correlate directly with lowered tutor investment. In fact, some tutors saw the data as being sufficiently meaningful that they made, or intend to make, changes in response to the data. With the right support, they could adjust their instruction in meaningful ways. Helping the tutors feel empowered to use data thoughtfully could also reinforce for them the importance of their data collection role in a DBI context. This finding addresses Research Question 2 because the tutors' different perspectives impacted their valuation of objective measures of progress as well as the degree of trust placed in CBM data. It also relates to Research Question 3 by suggesting directions for better supporting tutors in interpreting data to yield more responsive instruction.

Tutor Perceptions of Standardized Measures and Instructional Goals

As discussed in Chapter 2, CBMs are intended to be a general outcome measure, not tied to a particular curriculum; rather, they align broadly with instructional goals in a given domain. For some tutors, the connection between the oral reading fluency CBM and the instructional goals of the lesson was readily apparent. Pam, when asked in the baseline questions to suggest a measure of progress that would be meaningful to her, suggested a "semi-standard fluency

passage"; before she even began learning about CBMs, she recognized that an instructional goal of the lessons is to build tutees' rate, accuracy, and expression in connected text. Later, in her interview, she made this connection overtly when she said, "[CBMs] are helpful because it's an easy, fast, standardized way of checking their fluency, which is a main focus of what we're teaching" (Interview, March 29, 2023). Similarly, at baseline, Wanda indicated that a measure of fluency would be a meaningful way to gauge tutee growth in alignment with instructional goals. She wrote, "Having a steady increase in accuracy and speed in reading is definitely one of the main goals of working with these kids. Being able to see that progress in a physical way [as through a standardized measure] would be great motivation for both the tutor and the tutee" (Training Module Artifacts, January 22, 2023). For tutors like Pam and Wanda, who intuited the relationship between the content of the lessons and a measure of fluency, and who prioritized some degree of standardization from the beginning, CBMs were a natural choice for monitoring tutee's progress.

A Few Tutors Saw CBM Data as Disconnected from Instruction

For other tutors, the relationship between the lesson content and the CBMs was more opaque. Sharon and Janice, who are both undergraduate students without prior experience with standardized assessments, demonstrated through questions and comments that the CBM data did not seem connected to instruction. At the conclusion of her interview, Sharon asked, "Do the lessons that we teach the children, is that in partner with the CBM? Like a certain skill that we're teaching them?" (Interview, April 3, 2023). By asking that question, Sharon indicated that she did not see how the CBM was related to the instructional goals of the lessons. Seeing the CBMs as disconnected from instruction seems likely to undermine her investment in administering the CBMs with fidelity and aligns with her impressions at baseline, when she wrote, "I don't think

having an objective measure would impact my investment (Training Module Artifacts, January 22, 2023). Even before the training process began, Sharon was disinclined to attend to an objective measure of progress, and the training process had little impact in terms of her seeing the relevance of CBMs to her tutoring.

Janice, too, seemed to see the CBM as disconnected from her instruction. This was the case at baseline, when she indicated that she did not see the value of an objective measure of progress because "it's generally easy to see them make progress (make less mistakes or reading words more fluidly is usually how 'progress' manifests)… I don't think an objective measure would really affect anything" (Training Module Artifacts, February 3, 2023). In her interview, too, she conveyed that the data the CBMs yielded was immaterial when compared to the observations she made in the course of instruction:

> How well they do on the CBM doesn't always translate to how well they do in reading… I don't know if the information from the CBM is helpful because [during the lesson] I'm getting real data, like in real time, whether they're improving or not. And you don't really see that with the CBM… so I guess it would be nice to have the [CBM] data available, but it's not necessary, because I can see how they're progressing in other ways (Interview, April 3, 2023).

For Janice, the "real" data came from what she observed during the lessons, and, as an experienced tutor, she was used to adjusting her instruction in response to that data. When Janice considered the idea of progress, she thought of it in terms of tutees making progress against themselves– for example, recognizing a vowel pattern they did not recognize the day before, or spelling a word with a particular ending that had previously been too challenging. Janice has generally shown herself to be an observant tutor, and she likely noticed these specific

improvements. However, other types of progress she suggested, such as reading words "more fluidly," or "making less mistakes," are reliant only on tutor perception and, therefore, more subject to bias. Janice may perceive growth in those areas, but, without an objective measure, it would be difficult to know for sure.

Contrasting Lesson Content with Grade-Level Material

Some of the tutors recognized that another potential issue with assessing progress strictly through observation, as Janice did, is that they cannot get a sense of whether a tutee is improving relative to grade-level performance. Several of the tutors drew a contrast between the CBMs and the lesson content, emphasizing the difficulty of a grade-level passage relative to the reading their tutees practiced in the context of the tutorials. Kamaria, in particular, attempted to reconcile this contrast throughout the training process. In reflecting on her observation with her tutee, she remarked on how his reading of the CBM was slower than she would have expected given his success in the lessons. However, she connected back to the training and the fact that the CBM is representative of grade-level text, in contrast to the lesson content that is highly scaffolded. She said, "[In the lessons], I know we've been doing a lot of the same words and stuff, so he gets really good at those words… so I was like, oh, this [CBM] is a lot slower…but this is real life" (Tutee Observation, February 15, 2023). After making this initial realization, Kamaria continued to reflect on the benefits of using a CBM to supplement the observations she made in the course of her instruction. In her interview, she encapsulated that when she said:

> Of course, you can use informal ways to gauge like how is his accuracy going, or like how fast are we doing the lessons, or like parent reports saying, "Oh yes, he's getting better." But it's not standardized. I feel like you need something like the CBM that doesn't change much from week to week. They're all going to be equally different so it's

> a good way to tell without having the structure and support from the lessons, which always build up very systematically. In our lessons we use a lot of the same words, which is great, but it's not mirroring what he's getting in school and how supported he would be reading things in school. So, I feel like the CBM is more representative of how far his reading skills are going to get him as they are and how that's changing with the intervention that we're doing (Interview, April 3, 2023).

This quote demonstrates how Kamaria honed her understanding of the relationship between CBMs and observational data across several weeks of collecting progress monitoring data. When she first administered a CBM with her student and realized the demands of grade-level text, her affect seemed resigned. She expressed how "bad" she felt for her tutee but seemed resolved to make the CBMs work despite the challenges. However, once she had administered a handful of passages with her tutee, she seemed to appreciate having a sense of what he faced in school, and she ended on an optimistic note, implying that she expected him to improve relative to grade-level standards as a result of their work together. For Kamaria, the contrast between the CBM and lesson content was initially jarring; however, by working through the process, she successfully connected her instruction to anticipated improvement on a standardized measure, and she perceived significant value in making that connection.

Tutor Perceptions of Variability

In her above quote, Kamaria implied that she viewed all of the grade-level CBM passages as equivalent, but she had previously wondered about that. After her tutee observation, she asked:

> Was the one we did last week the same grade level? Because I think the words, he just maybe knew a lot better, because he was reading through it faster. This time, he was just,

like, everywhere, so he wasn't focused. But when he did focus, he was able to get through a bunch of words at once (Tutee Observation, February 15, 2023).

Through these questions and comments, Kamaria revealed that she was taking into consideration two factors that could impact her tutee's performance on the CBM independent of instruction: passage-level factors (i.e., the difficulty of the passage) and student factors (i.e., her tutee's level of focus on that particular day). In wondering about those two factors, Kamaria was not alone– in fact, several tutors raised these factors as potential explanations for what they perceived as variability in their tutees' CBM performance. With a few exceptions, this perceived variability did not seem to undermine investment in CBM administration or trust in the data collected; however, given the number of tutors who raised questions, it is relevant to this overall finding.

Passage-Level Factors

Along with Kamaria, two additional tutors, Pam and Janice, questioned the equivalence of the passages. In her interview, Pam mused, "I was wondering if the passages go up in level or if they're all the same level… [student name] has actually said, like, I don't know what's going on in this passage, and I'm just, like, it's a little abstract, yeah" (Interview, March 29, 2023). In this example, Pam was not questioning the passage equivalence in terms of readability, but in terms of its conceptual underpinnings. Without knowing which specific passages engendered this reaction for Pam and her tutee, it is difficult to explain what they were sensing. However, of note, this anecdote illuminates two themes previously discussed in this chapter. For one, Pam's response indicated that she responded to her tutee with empathy to preserve their relationship, turning what could have been a moment of frustration for the tutee into a moment of bonding. Secondly, based on Pam's other interview responses, the question of variability did not appear to undermine her confidence in the CBM data of her investment in administering them.

Janice, in contrast, raised her perception of passage variability more than once in her interview, both when asked about disadvantages of CBM use at DRC and when asked for feedback to improve the tutor training process. In her initial response, she stated, “Sometimes I think the difficulty levels of each passage are, like, different across the weeks… but it's not like, going from easy to hard. It's just like kind of random, so I don't know if that's like the best method [for measuring progress]” (Interview, April 3, 2023). Later in the interview, she reiterated that the training process could be improved by “making it so it’s not so random, the difficulty of the passages. Like either make it linear or make them equal in difficulty” (Interview, April 3, 2023). In Janice’s statements, she did not draw a direct connection between her perception of passage variability and her relatively low investment in CBMs as a measure of tutee progress; however, it stands to reason that, if she perceived the passages as “random,” those factors may have been correlated.

Tutee Factors

Along with passage-level factors, several tutors raised the issue of tutee factors accounting for perceived variability in tutee CBM scores. Some of these tutors noted differences on individual days. For example, after Renee’s tutee skipped a line during the tutee observation, she commented that this was unusual for her tutee, saying, “Sometimes, a kid might have one of those days where they’re extra jittery, and they’ll skip over a line” (Tutee Observation, February 22, 2023). Along those same lines, Pam referred to her tutee as having a “bad day, and she didn’t want to read, so I thought that was affecting her” (Interview, March 29, 2023). Similarly, Sharon noted that, for one of her tutees, scores appeared pretty consistent, but, for another, “You can tell when there’s a day she didn’t go to school that day” because her scores seemed to be lower than normal (Interview, April 3, 2023). For other tutors, the perceived variability seemed to be more

of a pervasive issue. For instance, in her interview, Kamaria reiterated her remark from the tutee observation, stating, “This is just for my specific student, but I wonder how accurate [the CBM data] is” (Interview, April 3, 2023). Leilani elaborated on this idea in her interview, saying:

> [Student name], because he’s more comfortable with me now, he’ll like either read slow if he doesn’t feel like reading it because Leilani is giving him and assessment today, or he’ll read super fast because Leilani is giving an assessment and he wants to show that he can be super fast. So, I feel like sometimes his performance fluctuates. He doesn't read to read, but instead he reads to perform (Interview, March 29, 2023).

This quote illustrates a few interesting things. In Leilani’s tutee observation and follow-up observation, the tutee’s tendency toward “performance” was not apparent to the observer; therefore, concerns about fidelity were not noted. However, by Leilani’s account, administration quality may have been compromised at times, even if Leilani maintained adherence by following all the administration steps correctly. In general, Leilani’s responses illustrated a high degree of investment in administering CBMs with fidelity and in the utility of CBMs, so it was noteworthy that she did not approach a supervisor for support. Leilani’s case further illustrates the theme that tutor investment does not necessarily yield administration fidelity.

Tutor Ideas about and Attempts at Responsive Instruction

As previously discussed, the focus of this study is not DBI itself; rather, it is to learn more about how DRC tutors collect data as a prerequisite for the ultimate goal of engaging in DBI practices. Although DBI practices were outside the scope of the investigation, several tutors expressed curiosity about how DRC intended to use the data they were collecting. For instance, at the conclusion of her interview, Leilani asked what would happen with the “final results.” After I responded with an overview of the DBI process and DRC’s organizational goals, she

replied, “I think knowing that makes me interested and want to give the assessment more. Now, I’m interested to see if my teaching is working by this assessment (Interview, March 29, 2023). Moreover, several tutors alluded to ways in which they adjusted their tutoring practices in response to the progress monitoring data, or ways they might in the future. By asking these questions and by looking for ways to be responsive to CBM data, tutors implied that they saw that data as meaningful. Therefore, they would likely be more invested in collecting it with adherence and quality, even though this was not necessarily apparent in the observations completed in this study.

Some tutors offered general comments about the importance of being responsive to CBM data. For example, Wanda stated that she wanted to “make adjustments to [her] teaching technique based off of what their progress is looking like,” but she did not offer specifics (Training Module Artifacts, January 22, 2023). Similarly, in her interview, Renee said she would like to start “to try and see how we can shift something or move something…it’s not just, we’re going to go through the lessons, if we see no growth. It should be intentional” (Interview, April 3, 2023). More specifically, Kamaria provided examples of adjustments she might make that were well within the purview of a paid volunteer tutor, including adjusting the pace within a given lesson; adjusting overall pacing across lessons; providing additional examples of concepts; and teaching to mastery. All of these ideas would be appropriate for a paid volunteer tutor to discuss with a supervisor or even implement independently in response to data trends they might observe.

Some tutors responded to the CBM data in ways that demonstrated a misunderstanding of their purpose. In the absence of specific examples guidance about use of CBM data, this was understandable, but it indicated a need for some additional support or training. For instance,

Sharon tried to treat the CBM data more like a running record, where she would analyze specific errors. She explained, “If I notice in the passage that she’s struggling with something different, like a phonics concept, I’ll try to do that with her, like we practice ‘AI’ and ‘EA’ words that aren’t in the lessons. When I see her read on those tests, I can see what she’s struggling with” (Interview, April 3, 2023). Using CBMs in this way is not in keeping with their purpose, and deviating from the intervention scope and sequence could be detrimental. That being said, Sharon’s impulse to identify her tutee’s instructional needs and make changes accordingly should be supported, perhaps through explicit guidance on using the mastery measures built into the curriculum to yield more actionable data for a paid volunteer tutor.

Summary of Finding Three

For a few tutors, the connection between CBM data and the instructional focus of the tutorials was tenuous; however, most tutors perceived that connection, and, therefore, the utility of CBMs as a measure of tutee progress. Tutors identified several sources of perceived variability, including passage factors and student factors. Though some had questions about variability, and even expressed concerns, many nonetheless indicated that they have tried to interpret the CBM data and even adjust instruction in response to it. These attempts at responsiveness demonstrate that most tutors perceive the CBM data as meaningful, which would likely lead to improvements in administration quality over the long term.

Chapter Summary

In this chapter, I integrated data from a range of sources, including training artifacts, observations, and interviews. I then presented the key findings and themes synthesized through my thematic analysis of the data and related to my research questions. I explained that, because tutors value their relationships with their tutees, two aspects of administration fidelity, task

directions and feedback while reading, proved especially challenging for them (i.e., Finding 1). I discussed the impact of tutor training, particularly the educative aspects of training, on tutor investment in CBM administration and, ultimately, on their administration fidelity (i.e., Finding 2). Finally, I conveyed how many of the tutors made connections between the CBM data and the lesson content and how they sought ways to use the data to support responsive instruction (i.e., Finding 3). In Chapter 5, I share recommendations for DRC about adjustments to the training content and process that could better support the tutors' administration fidelity and instructional responsiveness.

Chapter 5: Recommendations

At Douglass Reading Center (DRC), an ongoing organizational challenge has been supporting tutors, who are paid volunteers, in collecting high-quality progress monitoring data through CBMs. Obtaining this data would further the organization's goals of engaging in DBI practices at DRC. However, previous attempts at implementing CBMs have revealed difficulties with assessment fidelity, both in terms of adherence to standardized procedures and in terms of administration quality.

To better understand this phenomenon in context, I conducted a case study of DRC and its in-person tutors. Through various data sources, I sought to identify challenges tutors face when learning to administer CBMs as well as the ways in which their varied backgrounds might impact their administration fidelity. My case study was guided by a conceptual framework built upon andragogical principles (Alford, 2013; Knowles, 1984). By considering procedural and educative aspects of tutor training and how they impact assessment fidelity for DRC tutors, I collected data across the stages of the training process, which followed a gradual release model. In January through April of 2023, I collected data in the form of training module artifacts (i.e., baseline questions and scoring checks); checklists and tutor quotes from an initial observation and a tutee observation; tutor quotes from interviews; and checklists from a follow-up administration. I then integrated data across these sources, synthesizing findings to address the following research questions:

1. What challenges do DRC tutors experience when administering and scoring CBMs?
2. In what ways do a DRC tutor's knowledge, skills, and beliefs impact their administration and scoring of CBMs?

3. In what ways can DRC improve current CBM training practices to build DRC tutors' capacity for administering and scoring CBMs?

The purpose of this chapter is to articulate recommendations to DRC for future implementation of CBMs as part of the DBI process. These recommendations stem both from my findings and from the literature on CBM administration and assessment fidelity. Because the recommendations are so tailored to the context, their basis in the literature is sometimes indirect. For example, I did not find research that recommends the particular modifications to the training module as outlined in Recommendation 1. However, the suggested modifications align with the broader principles at work in my conceptual framework of supporting assessment fidelity through a training process grounded in a gradual release model. Ultimately, the goal of these recommendations is to support DRC staff in improving tutors' data collection practices in terms of fidelity and quality.

When designing this capstone, I expected that my recommendations would relate directly and exclusively to CBM administration. However, as I collected data and spoke with tutors, I determined that supporting tutors in linking CBMs to instruction also impacts fidelity. This is because tutors who related the CBM data to their instruction found the CBMs to be more purposeful. Understanding the purpose of CBMs, in turn, yielded greater investment and, therefore, assessment fidelity in terms of quality. For this reason, I include a recommendation about data interpretation and instructional responsiveness, even though, if taken at face value, that recommendation may appear to be beyond the scope of the inquiry. My ultimate recommendations to DRC include:

- Recommendation 1: Keep the current training process in place with modifications to the training module that address specific fidelity criteria that proved challenging for tutors.

    - Modification 1: Revise training module language related to task directions and feedback.
    - Modification 2: Include more practice opportunities in the training module.
    - Modification 3: Provide specific guidance around introducing the task to a tutee for the first time.
- Recommendation 2: Address tutors' concerns about variability and alignment of CBMs to their instruction.
    - Action Step 1: Explore alternative passage sets.
    - Action Step 2: Investigate alternative measures, such as Nonsense Word Fluency.
    - Action Step 3: Provide specific opportunities for tutors to discuss tutee factors impacting variability.
- Recommendation 3: Support tutors as they explore data interpretation and instructional responsiveness.
    - Action Step 1: Institute tutee progress meetings with Lead Clinician.
    - Action Step 2: Create a system to track data from mastery measures along with CBM data.
    - Action Step 3: Create a system to help tutors visualize the data relative to goals.

Recommendation 1: Keep the current training process in place with modifications to the training module that address specific fidelity criteria that proved challenging for tutors.

Findings from this study suggest that the current training process is working well in terms of assessment fidelity and tutor satisfaction. From this inquiry, I received feedback to answer questions I had about the training as Lead Clinician at DRC. As discussed in Chapter 1, DRC

experienced a rapid scaling up of services that exerts pressure on training paid volunteer tutors who arrive with varied backgrounds. Therefore, in designing the tutor training for CBM administration, I have attempted to provide key information, feedback, and sufficient practice opportunities for tutors while keeping their time commitment to a minimum. While trying to strike this balance, I have been curious about the perceived helpfulness to tutors of two elements that extend training time: (a) the educative aspects of the training and (b) the use of a multi-stage gradual release model. In embarking on this study, I was eager to obtain the tutors' perspectives on these elements, wondering if the time invested would yield improved assessment fidelity.

Through my observations of and conversations with tutors, it seems that tutors feel supported, rather than burdened, by both of these training elements. As illustrated in Finding 2 from Chapter 4, in support of the educative aspects of training, tutors expressed that having the rationale for CBM use at DRC enhanced their investment in administering the CBMs with fidelity. Further, several tutors specifically credited the series of gradual-release observations with their confidence and comfort with CBM administration, as demonstrated by Finding 2. Although the educative material built into the module added to its length, and the initial observation necessitated an appointment with me outside of their regular work hours the time investment was outweighed by the benefits of these training elements.

While the training process was successful overall, I recommend that several adjustments be made to the training module to enhance the process. Some of these adjustments were explicitly suggested by tutors, while I extrapolated others to address some of the more pervasive challenges to assessment fidelity. The modifications I recommend are to (a) revise training module language related to task directions and feedback; (b) include more practice opportunities

in the training module; and (c) Provide specific guidance around introducing the task to a tutee for the first time.

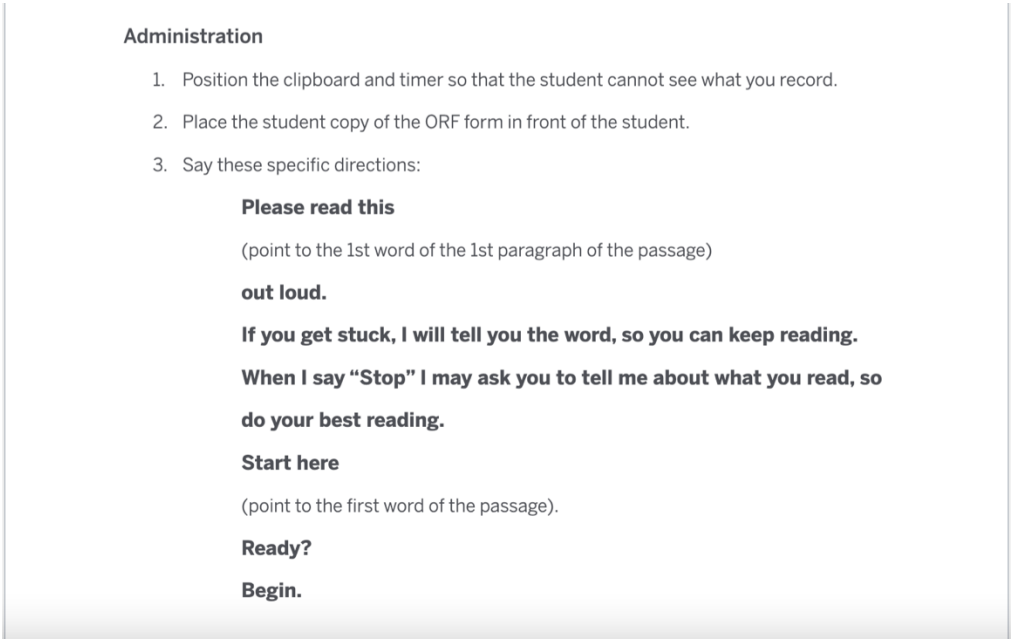
Modification 1: Revise training module language related to task directions and feedback

This modification addresses the two fidelity criteria that proved most challenging for tutors to fulfill: reading task directions verbatim and providing timely corrective feedback for tutees. As established in Finding 1, errors on these criteria were the most pervasive, and, although tutors were more likely to correct errors with task directions after receiving feedback during the initial observation, errors with corrective feedback persisted. Likely, this was the case because tutors value their relationships with their tutees and want them to feel confident while reading. Therefore, they want to frame the task in a "tutee-friendly" way, and they are reluctant to interrupt a tutee who is working to attack an unfamiliar word. Unfortunately, despite the tutors' good intentions, these deviations from the standard procedure can undermine the validity of the data collected.

Knowing that these errors are most common, and knowing *why* they are the most common, adjustments should be made to language in the module to address these areas explicitly and to try to prevent the errors in the first place. Because lengthening the module text should be avoided when necessary (Sachs et al., 2013), I suggest annotating the administration and scoring manual pages embedded in the module. That way, the manual will be more personalized to the tutors' areas of need, and tutors will directly see the context for applying the suggestions. For example, Figure 5.1 shows the manual page as it stands now:

Figure 5.1

Current Training Module Guidance on Task Directions (University of Oregon, 2018)



Administration

1. Position the clipboard and timer so that the student cannot see what you record.
2. Place the student copy of the ORF form in front of the student.
3. Say these specific directions:

Please read this

(point to the 1st word of the 1st paragraph of the passage)

out loud.

If you get stuck, I will tell you the word, so you can keep reading.

When I say "Stop" I may ask you to tell me about what you read, so do your best reading.

Start here

(point to the first word of the passage).

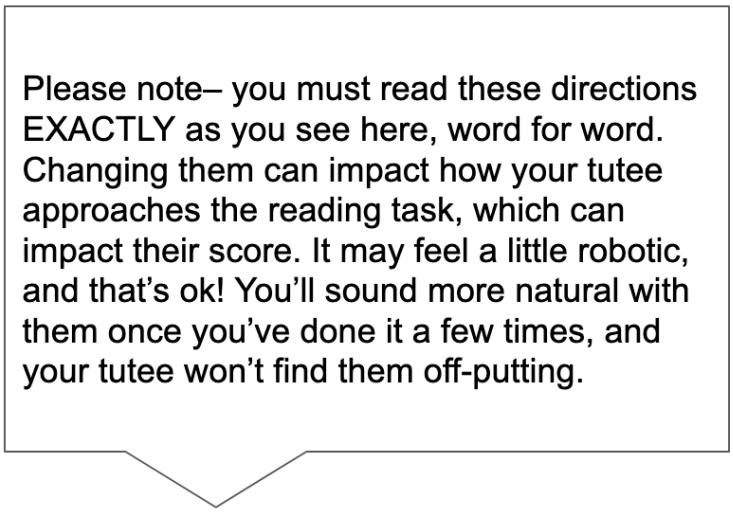
Ready?

Begin.

Figure 5.2 suggests language for a callout to be added to that same manual page by Step 3, "Say these specific directions," which, in the current training, did not provide sufficient guidance to the tutors to ensure fidelity. The callout language is crafted to emphasize the importance of reading directions verbatim; provide the rationale (Christ et al., 2013); and anticipate and address tutor concerns about impact on their tutee. In other words, the language involves an educative element intended to enhance both adherence and quality in CBM administration:

Figure 5.2

Callout to Be Added to the Task Directions Guidance Page



By investigating opportunities to annotate current resources, rather than adding to the length of the module, the training can continue to balance conveying key information to tutors while respecting their time and making clear to them the context where the suggestions are relevant. These suggested modifications will support tutors' administration fidelity within the limitations of the DRC contest.

Modification 2: Include more practice opportunities in the training module

As Finding 2 demonstrates, tutor investment in CBM use is not sufficient to ensure administration fidelity. Further, achieving fidelity is not a linear process, as seen in tutors "backsliding," or meeting a criterion in one administration but going on to make an error in a subsequent administration. In nearly all cases, backsliding occurred when a tutor encountered a circumstance that was new to them. For example, if a tutor incorrectly scored a tutee self-correction as an error, it was likely because the tutee had never self-corrected on prior administrations, and the tutor did not have a chance to receive feedback on how to handle that

situation. In other words, tutors did not generally meet criteria and later miss them; rather, the practice opportunities they were given did not cover the wide array of circumstances that might arise when working with tutees.

To address this piece of Finding 2, I recommend that the training module be expanded to include a wider variety of example recordings coupled with brief narrative in the form of Frequently Asked Questions. In the current module, the recording examples feature readers whose performance approximates grade-level standards, and the errors made are not sufficiently varied to demonstrate a range of possibilities including self-correcting, rereading words or phrases, hesitating, skipping lines, or losing focus on the task. DRC staff should intentionally build a library of example recordings that better reflect the variety of errors students can make. These examples should be contextualized for the tutors with answers to questions that address common scenarios, such as the ones in Figure 5.3:

Figure 5.3

Frequently Asked Questions to Address in the Training Module

- What should I do if my tutee makes a mistake?
- What should I do if my tutee is stuck on a word?
- What should I do if my tutee skips a line?
- What should I do if my tutee loses focus while reading?
- What should I do if my tutee does not seem to put forth their best effort?
- What should I do if my tutee seems anxious?
- What should I do if I make a mistake giving the CBM?

By including guidance about these topics, the training can better support administration fidelity, both in terms of adherence and in terms of quality. Generally, the questions earlier on the list

address issues of adherence, as they pertain to maintaining the standard administration procedure. The latter questions, however, relate more to quality. These aspects of administration are more subjective, but they are important to tutors given their care for their tutees; addressing them would likely heighten the tutors' investment in administering the CBMs as well as their trust in the data they collect.

Modification 3: Provide specific guidance around introducing the task to a tutee for the first time

As seen in Finding 1, the tutors are invested in their relationships with their tutees, and they do not wish to undermine their confidence. However, the tutors rightly perceive that a CBM passage often poses a significant challenge to tutees, and some tutors see this as potentially detrimental to the relationship. In its current iteration, the training module does not include a framework or script for how tutors should frame the CBM for their tutees. Finding 1 demonstrates that tutors thought deeply about how to do this, and providing explicit guidance would help support them better. As seen in Figure 5.4, I recommend formatting this guidance in two parts: a list of key points to convey, and a sample script. This format is differentiated for the tutors' varied backgrounds. Experienced tutors, or tutors with prior clinical experience, may want the freedom to share the information with their tutee more flexibly, while tutors who need an additional scaffold can use the sample script:

Figure 5.4

Guidance Around Introducing the CBM Task for the First Time

Key Ideas to Convey to the Tutee:

- Let them know you're starting a new weekly routine.
- Convey that the purpose is to make sure our teaching is effective– it's about us, not about the tutee!
- Tell them what you will be doing so they know what to expect– emphasize that they are only reading for one minute.
- Contrast the assessment with the reading you do in the instructional context so they know it will feel different.
- Let them know that it might feel difficult, but they only need to do their best reading
- Give them the opportunity to ask questions.

Sample Script:
"Starting this week, I will be asking you to read a passage out loud to me every week. We are doing this so that I can make sure I'm teaching you correctly to help you grow as a reader. As you read, I will be taking notes and keeping track of the time, because you will only read the passage for one minute. The passage might feel challenging to you. It will feel different from the other reading we do in our lessons, where we have practiced sounding out the words. And that is totally ok! Your only job is to do your best reading for one minute. It'll go by so quickly, and then we'll get back to our regular lesson! What questions do you have for me before we try it?"

Giving the tutors sample language they could use with a tutee also legitimizes some of the concerns the tutors have about using CBMs. For example, as explained in Finding 3, some tutors grappled with the contrast between material represented in the CBMs and the content of their lessons with their tutees, because the lesson content is carefully controlled while the ORF passage represents more "authentic" text. By suggesting sample language that acknowledges this contrast, the training shows the tutors that the tension they feel is an expected reaction. The advice to discuss it openly with the tutee ensures that all parties—DRC leadership, tutors, and tutees— share a common understanding about the relationship between the CBMs and the lesson content. With this common understanding established, tutors should take greater ownership over the process of administering CBMs, leading to improved fidelity.

Recommendation 2: Address tutors' concerns about variability and alignment of CBMs to their instruction.

Finding 3 discusses some questions and concerns tutors had regarding the alignment of CBMs to the lesson content as well as variability they perceive in the CBM scoring. Several tutors commented on, or asked questions about, what they interpreted as differences within the passage sets that purport to be of equivalent difficulty. They also wondered about tutee behaviors that would impact their scoring unevenly. Additionally, some tutors noted that the CBMs felt disconnected from their day-to-day instruction.

As stated in Chapter 1, the choice of an ORF CBM was deliberate as a starting point for data collection at DRC precisely because of its alignment with the instructional goals of the tutorials. In terms of passage factors impacting variability, Chapter 2 explains that these should not be at play in the DRC context, because (a) the passage sets have been validated as equivalent and (b) DRC tutees are the target population (i.e., students experiencing reading difficulty) (Briggs et al., 2011; O'Keeffe et al., 2017; Tindal et al., 2016). However, what matters for fidelity is the tutors' *perception* of the CBMs, because, if they question their validity or relevance to the lesson content, their administration quality could suffer. Therefore, I recommend that these disconnects be addressed directly through the following action steps: (a) explore alternative passage sets; (b) investigate alternative measures, such as Nonsense Word Fluency; and (c) provide specific opportunities for tutors to discuss tutee factors impacting variability.

Action Step 1: Explore alternative passage sets

As Bundock et al. (2018) explain, passage factors are not generally seen as a source of variability because many published passage sets have been validated through a rigorous process. Even so, as noted in Finding 3, several tutors commented on perceived differences in the

passages; for example, when asked how DRC could improve CBM administration procedures, Janice identified perceived passage differences as the primary problem to address. Notably, Janice also expressed a low degree of investment in CBM administration, which is a logical consequence of perceiving the passages as haphazard. Generally, if tutors perceive variability in the passages, it could undermine their trust in the CBM, and, therefore, their investment in administering it with adherence and quality.

The passage set used at DRC includes both narrative and expository passages, which, according to Bundock et al. (2018), is the only relevant passage factor that can introduce some degree of variability (Briggs, 2011; O'Keeffe et al., 2017). In her interview, Pam mentioned that her tutee found some of the passages to be "a little abstract," and it is possible that the narrative/expository difference may contribute to that perception (Interview, March 29, 2023). In general, CBM guidance is to include both so as to better represent general reading outcomes (Bundock et al., 2018). However, if passage sets that mix genres undermine tutors' confidence in the measure, they may not be as well suited to the DRC context, where they will be used by paid volunteer tutors. Therefore, I recommend that DRC explore alternatives, taking guidance from a recommended resource such as the "Academic Progress Monitoring Tools Chart" created by the National Center on Intensive Intervention (Austin & Filderman, 2020).

Action Step 2: Investigate alternative measures, such as Nonsense Word Fluency

Finding 3 shows that, although many tutors made the connection between the ORF CBM and the day-to-day lesson content, a few did not grasp the relationship. If tutors perceive an assessment as unrelated to instruction, their investment, and, therefore, assessment fidelity, would likely be impacted, especially over the long term. As explained in Chapter 2, DRC

selected the ORF CBM in accordance with publisher guidance because it aligns with the lesson instructional goals of building code-based skills and fluency (University of Oregon, 2018).

While an ORF measure was a good starting point for DRC to begin formal data collection, other CBM tasks, such as Nonsense Word Fluency (NWF) and Word Reading Fluency, also relate to DRC's lesson content (see Chapter 2 for additional detail). For some students, NWF in particular would be a better fit than ORF (Fuchs et al., 2004; Hosp et al., 2016; University of Oregon, 2018). For example, for a student in the early stages of developing their letter/sound knowledge and phonological awareness, the NWF measure would be more sensitive, capturing growth in a way that ORF does not. If tutors were to see tighter alignment between the assessment and their instruction, and if they were to see their tutee making growth based on the CBM data, they would likely see the CBMs are more purposeful and, therefore, administer them with improved fidelity.

Action Step 3: Provide specific opportunities for tutors to discuss tutee factors impacting variability

Along with passage factors, as seen in Finding 3, several tutors expressed concerns about tutee factors that might contribute to variability. Some of the training modifications suggested in Recommendation 1 can help address this by framing the task more deliberately. Additionally, if a tutor feels confident in their CBM administration, they are more apt to see the data they collect as meaningful. Finding 3 noted that there were some tutors who felt comfortable with the CBM procedures in the abstract but felt unsure about their ability to implement those procedures with a specific tutee. This inspired feelings of nervousness in the tutors. If left unaddressed, these tutor concerns could contribute to a sense that the CBM data is too unstable to be meaningful, and fidelity would then suffer.

To address this, I recommend dedicated points in the tutor training process where the Lead Clinician asks tutors directly if they anticipate or if they have experienced any tutee-related challenges to maintain standardized administration procedures. The particular points I suggest are at the end of the initial observation and the end of the follow-up observation. By asking tutors about this, the Lead Clinician brings greater transparency to the training process and avoids the implication that tutee factors are assumed to well in hand. The tutor has a safe space to express doubts and to troubleshoot with support as needed. Addressing these concerns up front also preserves fidelity because the tutor will not feel pressured to pretend that their administrations are all going smoothly; rather, they will ask for help, ensuring that the data collected is ultimately useful because assessment fidelity was maintained.

Recommendation 3: Support tutors as they explore data interpretation and instructional responsiveness.

As noted in Chapter 2, tutor responsiveness is a key factor in achieving positive outcomes in the tutorial environment (Allor & McCathren, 2004; Morris, 2022; Wasik & Slavin, 1990). To be clear, this responsiveness is distinct from DBI, which is comprised of a series of formalized steps. Rather, responsiveness is about tutors using the tools they have at their disposal to adapt instruction in response to data, both formal (such as data from CBMs) and observational (such as observing that a tutee consistently confuses their short vowel sounds when spelling words). As evidenced by Finding 3, many DRC tutors intuit that responsiveness is important, and they seek ways to better meet their tutees' instructional needs. However, as paid volunteer tutors, the tutors at DRC do not have sufficient knowledge to adapt instruction as a teacher or reading specialist might; this was evident when, for example, Sharon deviated from the program scope and sequence based on patterns she noticed in her tutee's CBM reading (this example is detailed in

Finding 3). With improved guidance, the tutors could engage in responsive instruction more appropriately through maneuvers such as adjusting pacing, providing additional examples, and ensuring mastery of key concepts. When tutors understand the relationship between the CBM data and their instruction, they are more likely to administer the CBMs with fidelity.

To capitalize on tutor assets and address challenges identified in Finding 3, I recommend a three-part approach to supporting data interpretation and use for tutors. The action steps of this approach are (a) institute tutee progress meetings with Lead Clinician; (b) create a system to track data from mastery measures along with CBM data; and (c) create a system for helping tutors visualize their data relative to goals.

Action Step 1: Institute tutee progress meetings with Lead Clinician

In Finding 2, I summarized the ways in which some tutors, especially tutors with less experience, tended to view CBM data as a reflection on their efficacy as tutors, even expressing feelings of guilt if they perceived a lack of sufficient growth in their tutees. In reality, a tutor's work with a tutee is just a drop in the bucket of that tutee's overall instructional program, a mere two hours out of their week. Further, there are plenty of reasons a tutee might not appear to be making growth on a CBM, many of which are not in the tutor's control—and, in fact, that tutee may actually be making growth, but this is difficult to see from a list of scores alone. Rather than letting the tutor carry these misplaced feelings of guilt, DRC's Lead Clinician should institute periodic progress meetings to help tutors interpret data appropriately, both relative to their tutees and relative to themselves and their success as tutors.

Along with supporting tutors in interpreting data, these meetings would allow the Lead Clinician to advise the tutors around more responsive instruction. As discussed in Chapter 2, coaching is a key element of successful tutoring, especially with paid volunteer tutors, and these

meetings would provide the opportunity to discuss the relationship between tutees' data and instruction (Al Otaiba & Lake, 2007; Morris, 2006; Samson et al., 2015). The Lead Clinician could help tutors consider the CBM data and mastery data (see Action Step 2) alongside their observations from lessons, such as:

- Pacing within lessons (Does the tutee seem to need more processing time or more frequent breaks?)
- Pacing across lessons (Does this content seem like a review for the tutee? Does the tutee need the optional "reteach" lessons built into the program?)
- Tutee placement (Does it seem like the tutee is in the right spot within the scope and sequence?)
- Common tutee errors (Are there any letter sounds or high frequency words the tutee frequently confuses?)
- Confusion of key concepts (Is the tutee experiencing difficulty with a particular concept, such as inflected endings, in their decoding and/or spelling? Or, do they have difficulty applying a strategy to decoding and/or spelling?)
- Need for instructional scaffolds (Does the tutee consistently need the most intensive of the suggested scaffolds built into the instructional activities?)

For tutors who are less experienced, more support may be required to engage in these conversations. This could involve the Lead Clinician observing a tutorial and leading the conversation, modeling for the tutor the kinds of things they should be observing. Having worked together to identify patterns based on data and observation, the tutor and Lead Clinician can discuss possible next steps, resulting in more responsive instruction. Creating an explicit link

for tutors between their data and instruction should lead to greater investment in administering CBMs with fidelity.

Action Step 2: Create a system to track data from mastery measures along with CBM data

As evidenced in Finding 2, tutors grew more invested in administering CBMs with fidelity because the in-depth nature of the training conveyed their importance. Further, as seen in Finding 3, many tutors had begun to think about how to interpret their CBM data and engage in more responsive instruction. Because they are paid volunteer tutors without an in-depth understanding of CBMs and their purpose, some tutors, like Sharon, used them in a way that was not in keeping with best practice. Sharon noted specific phonics patterns in the CBM that posed a challenge for her tutee, which is an admirable impulse. However, a preferred source of that information would be the mastery measures built into the lessons, since those align with what her tutee has been taught (See Chapter 1 for information on mastery measures at DRC).

While CBMs are designed to capture progress toward general outcomes, mastery measures take a narrower focus, measuring proficiency with content directly targeted by an intervention or program (Austin & Filderman, 2020; Begeny et al., 2015; Fuchs & Deno, 1991; Wayman et al., 2007). The use of mastery measures alone could give a false impression about what a student has learned because they have a narrower scope than a general outcome measure (VanDerHeyden & Burns, 2018). However, the information they provide is distinctly relevant and important when considering student progress, and it complements information gleaned through CBMs (VanDerHeyden & Burns, 2018).

An advantage of a mastery measure is that it can take place within the context of instruction, and, therefore, what tutors are already doing. In this way, it could seem more directly relevant to them, especially in terms of providing data that informs their instruction (Rowe et al.,

2014; Stecker et al., 2008). Additionally, because the focus of mastery measures is on short-term objectives, they may be more sensitive to tutee growth, enabling a quicker gauge of how the tutee is responding to instruction (Filderman & Toste, 2021).

By underemphasizing mastery measures while focusing on CBMs, DRC's Lead Clinician may have inadvertently conveyed that CBM data takes precedence, rather than that CBMs and mastery measures can be used in tandem to consider a tutee's progress in the lesson sequence as well as their progress relative to grade-level standards. To rectify this misconception, I suggest creating a system of tracking mastery measure data alongside CBM data. Currently, tutors complete the mastery measure with their tutees and use rough guidance to decide whether to review current material or move on to a new unit, but these scores are not recorded anywhere. Asking tutors to record the mastery measure data would convey a sense of purpose to that data. Then, in the progress meetings described in Action Step 1, the Lead Clinician could support the tutor in interpreting both types of data and identifying instructional practices in keeping with their distinct purposes. When tutors understand the relationship between these two data sources, they will be less likely to misalign the CBM data with its purpose. With enhanced clarity of purpose, the tutors should administer CBMs with greater quality.

Action Step 3: Create a system for helping tutors visualize the data relative to goals

In their interviews, several tutors with previous clinical experience stated that, in other contexts, they have found it helpful to have a visual representation of data relative to goals. Additionally, research indicates that graphs can be helpful to novice teachers when they receive proper support to interpret them (Nelson et al., 2017). Therefore, I would recommend that DRC devise a method to support tutors in graphing and interpreting data. Prior to this inquiry, the Lead Clinician did not understand the degree of investment tutors have in the CBM data they are

collecting. As seen in Finding 2, this level of investment likely has not always existed, but was heightened by the educative aspects added to the training module. Nonetheless, now that the tutors' investment is made clear, the Lead Clinician should devise a system for data reporting that generates a graph, showing a trend line relative to a goal (Hosp et al, 2016; University of Oregon, 2018). This would provide transparency and support for the tutor, enabling them to have a realistic goal in mind for their tutee and seeing that progress is not about any one data point, but an overall trend. It could also lay a foundation for future goal-setting conversations with tutees, which could increase tutee motivation (Al-Bataineh et al., 2019; Cabral-Márquez, 2015). By getting a more representative sense of their tutees' progress, the tutors should feel more motivated and invested in the data they are collecting, and, therefore, collect it with greater fidelity.

Limitations

Regarding these recommendations, as informed by the study's findings, several limitations should be noted. For one, these recommendations are highly specific to the context of DRC's in-person tutoring services. Therefore, they are not generalizable to DRC's virtual tutors, or to any other university-based tutoring center. Secondly, while advantageous in some ways, my position as DRC's Lead Clinician was also a liability to trustworthiness in this study. As outlined in Chapter 3, I employed strategies to counteract this, including triangulation of data sources, collecting data over a period of several months, and use of a peer reviewer. I also carefully considered the ways in which my investment in the tutor training process could lead to erroneous assumptions and biases, and I worked to counteract these when possible through ongoing awareness and reflection. Additionally, while my relationship with my participants enabled insight, it could also undermine trustworthiness. Although I am not the tutors' manager, we have

a working relationship that could have impacted the data I collected. For example, tutors may have behaved differently during observations because they knew I was observing them. Further, their interview responses may have reflected what they thought I wanted to hear, rather than their true opinions and feelings. This concern is somewhat mitigated by some tutors' frank interview responses, at times framed with language such as, "I know this sounds really bad, but…" (Interview with Sharon, April 3, 2023). Even so, my level of involvement with the site and the tutors represents a limitation of this investigation.

Summary

In Chapter 5, I presented a series of recommendations to DRC that are grounded in my conceptual framework and relevant literature as well as in the findings of my inquiry. These recommendations are specific to the context and provide guidance on addressing challenges to assessment fidelity and adapting tutor training to meet those challenges. My hope is that these recommendations yield the short-term outcomes of enhanced administration fidelity (in terms of both adherence and quality) and, therefore, trustworthy data that DRC staff can use to consider the efficacy of instruction, both for individual tutees and for their program as a whole. Ultimately, this inquiry represents the first step in enabling DRC's organizational goals of implementing a DBI framework to improve reading achievement for tutees and to better communicate tutee progress with university and community stakeholders.

References

Al-Bataineh, A. T., Brenwall, L., Stalter, K., & York, J. (2019). Student growth through goal setting. *International Journal of Learning and Teaching*, *11*(4), 147-161.

Alford, B. (2013). Adult learning theory. In B. Irby, G. H. Brown, R. Lara-Aiecio, & S. A. Jackson (Eds.), *The handbook of educational theories* (pp. 125-131). IAP.

Allor, J., & McCathren, R. (2004). The efficacy of an early literacy tutoring program implemented by college students. *Learning Disabilities Research & Practice*, *19*(2), 116-129.

Al Otaiba, S., & Lake, V. E. (2007). Preparing special educators to teach reading and use curriculum-based assessments. *Reading and Writing*, *20*(6), 591-617.

Ardoin, S. P., Christ, T. J., Morena, L. S., Cormier, D. C., & Klingbeil, D. A. (2013). A systematic review and summarization of the recommendations and research surrounding curriculum-based measurement of oral reading fluency (CBM-R) decision rules. *Journal of School Psychology*, *51*(1), 1-18.

Austin, C. R., & Filderman, M. J. (2020). Selecting and designing measurements to track the reading progress of students with disabilities. *Intervention in School and Clinic*, *56*(1), 13-21.

Austin, C. R., Vaughn, S., & McClelland, A. M. (2017). Intensive reading interventions for inadequate responders in grades K–3: A synthesis. *Learning Disability Quarterly*, *40*(4), 191-210.

Baker, S., Gersten, R., & Keating, T. (2000). When less may be more: A 2-year longitudinal evaluation of a volunteer tutoring program requiring minimal training. *Reading Research Quarterly*, *35*(4), 494-519.

Ball, C. R., & Christ, T. J. (2012). Supporting valid decision making: Uses and misuses of assessment data within the context of RTI. *Psychology in the Schools*, *49*(3), 231-244.

Bambera, L., & Kern, L. (2005). Individualized supports for students with problem behaviors.

Barnes, N., Brighton, C. M., Fives, H., Meyers, C., & Moon, T. R. (2022). Where's the data to support educators' data use for instructional practice?. *Theory Into Practice*, *61*(3), 277-287.

Bazeley, P. (2020). Qualitative data analysis: Practical strategies. *Qualitative Data Analysis*, 1-584.

Begeny, J. C., Krouse, H. E., Brown, K. G., & Mann, C. M. (2011). Teacher judgments of students' reading abilities across a continuum of rating methods and achievement measures. *School Psychology Review*, *40*(1), 23-38.

Begeny, J. C., Whitehouse, M. H., Methe, S. A., Codding, R. S., Stage, S. A., & Nuepert, S. (2015). Do intervention-embedded assessment procedures successfully measure student growth in reading?. *Psychology in the Schools*, *52*(6), 578-593.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational researcher*, *13*(6), 4-16.

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative research journal*, *9*(2), 27-40.

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., ... & Fernandez, M. (2009). How we design feasibility studies. *American journal of preventive medicine*, *36*(5), 452-457.

Briggs, R. N. (2011). *Investigating variability in student performance on DIBELS Oral Reading Fluency third grade progress monitoring probes: Possible contributing factors* (Doctoral dissertation, University of Oregon).

Bundock, K., O'Keeffe, B. V., Stokes, K., & Kladis, K. (2018). Strategies for minimizing variability in progress monitoring of oral reading fluency. *TEACHING Exceptional Children*, *50*(5), 273-281.

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*(4), 223-230.

Cabral-Márquez, C. (2015). Motivating readers: Helping students set and attain personal reading goals. *The Reading Teacher*, *68*(6), 464-472.

Case, L., Speece, D. L., Silverman, R., Ritchey, K. D., Schatschneider, C., Cooper, D. H., ... & Jacobs, D. (2010). Validation of a supplemental reading intervention for first-grade children. *Journal of Learning Disabilities*, *43*(5), 402-417.

Catts, H. W., Petscher, Y., Schatschneider, C., Bridges, M., & Mendoza, K. (2009). Floor effects associated with universal screening and their impact on the early identification of reading disabilities. *Journal of Learning Disabilities*, *42*(2), 163-176.

Century, J., Rudnick, M., & Freeman, C. (2010). A framework for measuring fidelity of implementation: A foundation for shared language and accumulation of knowledge. *American journal of evaluation*, *31*(2), 199-218.

Cho, E., Capin, P., Roberts, G., & Vaughn, S. (2018). Examining predictive validity of oral reading fluency slope in upper elementary grades using quantile regression. *Journal of Learning Disabilities*, *51*(6), 565-577.

Christ, T. J., White, M. J., Ardoin, S. P., & Eckert, T. L. (2013). Curriculum based measurement of reading: Consistency and validity across best, fastest, and question reading conditions. *School Psychology Review*, *42*(4), 415-436.

Clemens, N. H., Shapiro, E. S., & Thoemmes, F. (2011). Improving the efficacy of first grade reading screening: An investigation of word identification fluency with other early literacy indicators. *School Psychology Quarterly*, *26*(3), 231.

Clemens, N. H., Shapiro, E. S., Wu, J. Y., Taylor, A. B., & Caskie, G. L. (2014). Monitoring early first-grade reading progress: A comparison of two measures. *Journal of Learning Disabilities*, *47*(3), 254-270.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (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*(2), 394.

Cook, B. G., Tankersley, M., & Landrum, T. J. (2009). Determining evidence-based practices in special education. *Exceptional children*, *75*(3), 365-383.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

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*(2), 284-295.

Cummings, K. D., Park, Y., & Schaper, H. A. (2013). Form effects on DIBELS Next oral reading fluency progress-monitoring passages. *Assessment for Effective Intervention*, *38*(2), 91-104.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*(3), 219-232.

Deno, S. L. (2016). Data-based decision-making. In *Handbook of response to intervention* (pp. 9-28). Springer, Boston, MA.

Denton, C. A., Nimon, K., Mathes, P. G., Swanson, E. A., Kethley, C., Kurz, T. B., & Shih, M. (2010). Effectiveness of a supplemental early reading intervention scaled up in multiple schools. *Exceptional Children*, *76*(4), 394-416.

Denton, C. A., Beegle, R., Solari, E. J., & Madsen, K. (2016). Reading RULES Grade 1: A Tier 1 and Tier 2 intervention in word study, text reading, vocabulary, and comprehension. *Houston, TX: The Children's Learning Institute*.

Denzin, N. K. (1989). *Interpretive biography*. Newbury Park, CA: Sage.

Derr-Minneci, T. F., & Shapiro, E. S. (1992). Validating curriculum-based measurement in reading from a behavioral perspective. *School Psychology Quarterly*, *7*(1), 2.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational researcher*, *38*(3), 181-199.

DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical education*, *40*(4), 314-321.

Elbaum, B., Vaughn, S., Tejero Hughes, M., & Watson Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading

failure? A meta-analysis of the intervention research. *Journal of educational psychology*, *92*(4), 605.

Efron, S. E., & Ravid, R. (2019). *Action research in education: A practical guide*. Guilford Publications.

Espin, C. A., & Foegen, A. (1996). Validity of general outcome measures for predicting secondary students' performance on content-area tasks. *Exceptional Children*, *62*(6), 497-514.

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 to reading outcomes for beginning readers. *School Psychology Review*, *39*(4), 631-653.

Feinberg, A. B., & Shapiro, E. S. (2009). Teacher accuracy: An examination of teacher-based judgments of students' reading with differing achievement levels. *The Journal of Educational Research*, *102*(6), 453-462.

Filderman, M. J., & Gesel, S. A. (2022). Data Teams: A Collaborative Approach to Intensifying Intervention Using Student Data. *TEACHING Exceptional Children*, 00400599221096753.

Filderman, M. J., & Toste, J. R. (2021). Effects of Varying Levels of Data Use to Intensify a Multisyllabic Word Reading Intervention for Upper Elementary Students with or At Risk for Reading Disabilities. *Journal of Learning Disabilities*, 00222194211048405.

Filderman, M. J., Toste, J. R., Didion, L. A., Peng, P., & Clemens, N. H. (2018). Data-based decision making in reading interventions: A synthesis and meta-analysis of the effects for struggling readers. *The Journal of Special Education*, *52*(3), 174-187.

Fives, A., Kearns, N., Devaney, C., Canavan, J., Russell, D., Lyons, R., ... & O'Brien, A. (2013). A one-to-one programme for at-risk readers delivered by older adult volunteers. *Review of Education*, *1*(3), 254-280.

Freire, P. (1970). *Pedagogy of the oppressed*. New York, NY: Continuum.

Fuchs, D., Compton, D. L., Fuchs, L. S., Bryant, J., & Davis, G. N. (2008). Making "secondary intervention" work in a three-tier responsiveness-to-intervention model: Findings from the first-grade longitudinal reading study of the National Research Center on Learning Disabilities. *Reading and Writing*, *21*(4), 413-436.

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional children*, *57*(6), 488-500.

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children*, *71*(1), 7-21.

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education*, *9*(2), 20-28.

Gersten, R., Haymond, K., Newman-Gonchar, R., Dimino, J., & Jayanthi, M. (2020). Meta-analysis of the impact of reading interventions for students in the primary grades. *Journal of Research on Educational Effectiveness*, *13*(2), 401-427.

Goffreda, C. T., & DiPerna, J. C. (2010). An empirical review of psychometric evidence for the Dynamic Indicators of Basic Early Literacy Skills. *School Psychology Review*, *39*(3), 463-483.

Goffreda, C. T., DiPerna, J. C., & Pedersen, J. A. (2009). Preventive screening for early readers: Predictive validity of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS). *Psychology in the Schools*, *46*(6), 539-552.

Hagen, N. A., Biondo, P. D., Brasher, P. M., & Stiles, C. R. (2011). Formal feasibility studies in palliative care: why they are important and how to conduct them. *Journal of pain and symptom management*, *42*(2), 278-289.

Hatch, J. A. (2002). *Doing qualitative research in education settings*. Suny Press.

Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing fidelity with flexibility and fit: What do we really know about fidelity of implementation in schools?. *Exceptional Children*, *79*(2), 181-193.

Hays, D. G., & Singh, A. A. (2011). *Qualitative inquiry in clinical and educational settings*. Guilford Press.

Hintze, J. M., & Christ, T. J. (2004). An examination of variability as a function of passage variance in CBM progress monitoring. *School Psychology Review*, *33*(2), 204-217.

Hosp, M. K., & Fuchs, L. S. (2005). Using CBM as an indicator of decoding, word reading, and comprehension: Do the relations change with grade? *School Psychology Review*, *34*(1), 9-26.

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2016). *The ABCs of CBM: A practical guide to curriculum-based measurement*. Guilford Publications.

Jacob, R., Armstrong, C., Bowden, A. B., & Pan, Y. (2016). Leveraging volunteers: An experimental evaluation of a tutoring program for struggling readers. *Journal of Research on Educational Effectiveness*, *9*(sup1), 67-92.

Jarvis, P. (1992). Leaders of adult and continuing education should come from outside the field. In M. W. Galbraith & B. R. Sisco (Eds.), *Confronting controversies in challenging times: A call for action* (pp. 53-58). San Francisco, CA: Jossey-Bass.

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at-risk readers in a response to intervention framework. *School Psychology Review*, *36*(4), 582-600.

Jones, C. J., Christian, M., & Rice, A. (2016). The Results of a Randomized Control Trial Evaluation of the SPARK Literacy Program. *Society for Research on Educational Effectiveness*.

Jung, P. G., McMaster, K. L., Kunkel, A. K., Shin, J., & Stecker, P. M. (2018). Effects of data-based individualization for students with intensive learning needs: A meta-analysis. *Learning Disabilities Research & Practice*, *33*(3), 144-155.

Kim, Y. S., Petscher, Y., Schatschneider, C., & Foorman, B. (2010). Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? *Journal of Educational Psychology*, *102*(3), 652.

Kirschner, P. A., & Hendrick, C. (2020). *How learning happens: Seminal works in educational psychology and what they mean in practice*. Routledge.

Knowles, M. S. (1984). *The adult learner: A neglected species* (3rd ed.). Houston, TX: Gulf.

Knudson, R. S. (1980). An alternative approach to the andragogy/pedagogy issue. *Lifelong Learning: The Adult Years*, *3*(8), 8-10.

La Pelle, N. (2004). Simplifying qualitative data analysis using general purpose software tools. *Field methods*, *16*(1), 85-108.

Lee, Y. S., Morrow-Howell, N., Jonson-Reid, M., & McCrary, S. (2012). The effect of the Experience Corps® program on student reading outcomes. *Education and Urban Society*, *44*(1), 97-118.

Lemire, C., Rousseau, M., & Dionne, C. (2022). A Comparison of Fidelity Implementation Frameworks Used in the Field of Early Intervention. *American Journal of Evaluation*, 10982140211008978.

Lemons, C. J., Kearns, D. M., & Davidson, K. A. (2014). Data-based individualization in reading: Intensifying interventions for students with significant reading disabilities. *Teaching Exceptional Children*, *46*(4), 20-29.

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

Mellard, D. F., McKnight, M., & Woods, K. (2009). Response to intervention screening and progress-monitoring practices in 41 local schools. *Learning Disabilities Research & Practice*, *24*(4), 186-195.

Merriam, S. B. (2001). Andragogy and self-directed learning: Pillars of adult learning theory. *New directions for adult and continuing education*, *2001*(89), 3-14.

Mertens, D. M., & Wilson, A. T. (2018). *Program evaluation theory and practice*. Guilford Publications.

Miller, S., & Connolly, P. (2013). A randomized controlled trial evaluation of time to read, a volunteer tutoring program for 8-to 9-year-olds. *Educational Evaluation and Policy Analysis*, *35*(1), 23-37.

Morris, D. (2022). The Case for Tutoring Struggling Readers in the Primary Grades. *Reading & Writing Quarterly*, 1-16.

Morris, D. (2006). Using noncertified tutors to work with at-risk readers: An evidence-based model. *The Elementary School Journal*, *106*(4), 351-362.

Morris, D., Trathen, W., Perney, J., Gill, T., Schlagal, R., Ward, D., & Frye, E. M. (2017). Three DIBELS tasks vs. three informal reading/spelling tasks: A comparison of predictive validity. *Reading Psychology*, *38*(3), 289-320.

National Center for Education Statistics (2022). *2022 Grade 4 assessment report card: Summary data tables for national and state average scores and achievement level results.* Washington, DC: US Department of Education.

National Center for Intensive Intervention. (2013, March). *Data-based individualization: A framework for intensive intervention*. https://intensiveintervention.org/sites/default/files/DBI_Framework.pdf

Neitzel, A. J., Lake, C., Pellegrini, M., & Slavin, R. E. (2022). A synthesis of quantitative research on programs for struggling readers in elementary schools. *Reading Research Quarterly*, *57*(1), 149-179.

Nelson, P. M., Van Norman, E. R., & Christ, T. J. (2017). Visual analysis among novices: Training and trend lines as graphic aids. *Contemporary School Psychology*, *21*(2), 93-102

Nichols, J. D., Kim, I. H., & Nichols, G. W. (2020). The effect of parent and community volunteerism on early literacy development. *Educational Review*, *72*(4), 411-426.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of educational research*, *78*(1), 33-84.

O'Keeffe, B. V., Bundock, K., Kladis, K. L., Yan, R., & Nelson, K. (2017). Variability in DIBELS Next progress monitoring measures for students at risk for reading difficulties. *Remedial and Special Education*, *38*(5), 272-283.

Orsmond, G. I., & Cohn, E. S. (2015). The distinctive features of a feasibility study: objectives and guiding questions. *OTJR: occupation, participation and health*, *35*(3), 169-177.

Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary educational psychology*, *8*(3), 317-344.

Pearson, P. D., McVee, M. B., & Shanahan, L. E. (2019). In the Beginning: The Historical and Conceptual Genesis of the Gradual Release of Responsibility1. In *The gradual release of responsibility in literacy research and practice* (Vol. 10, pp. 1-21). Emerald Publishing Limited.

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical psychology: Science and practice*, *12*(4), 365.

Poncy, B. C., Skinner, C. H., & Axtell, P. K. (2005). An investigation of the reliability and standard error of measurement of words read correctly per minute using curriculum-based measurement. *Journal of Psychoeducational Assessment*, *23*(4), 326-338.

Rachal, J. (1983). The andragogy-pedagogy debate: Another voice in the fray. *Lifelong learning: The adult years*, *6*(9), 14-15.

Reed, D. K., & Sturges, K. M. (2013). An examination of assessment fidelity in the administration and interpretation of reading tests. *Remedial and Special Education*, *34*(5), 259-268.

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*(6), 427-469.

Ritter, G. W., Barnett, J. H., Denny, G. S., & Albin, G. R. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: A meta-analysis. *Review of Educational Research*, *79*(1), 3-38.

Rogers, E. M., Singhal, A., & Quinlan, M. M. (2014). Diffusion of innovations. In *An integrated approach to communication theory and research* (pp. 432-448). Routledge.

Rowe, S. S., Witmer, S., Cook, E., & Dacruz, K. (2014). Teachers' attitudes about using curriculum-based measurement in reading (CBM-R) for universal screening and progress monitoring. *Journal of Applied School Psychology*, *30*(4), 305-337.

Sachs, D. E., Langan, K. A., Leatherman, C. C., & Walters, J. L. (2013). Assessing the effectiveness of online instruction literacy tutorials for millennial undergraduates. College & Undergraduate Libraries, 20(3–4), 327–351. doi:10.1080/10691316.2013.829365

Samson, J. F., Hines, S. J., & Li, K. (2015). Effective use of paraprofessionals as early intervention reading tutors in grades K-3. *Mentoring & Tutoring: Partnership in Learning*, *23*(2), 164-177.

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of learning disabilities*, *48*(4), 369-390.

Schwartz, R. M. (2005). Literacy Learning of At-Risk First-Grade Students in the Reading Recovery Early Intervention. *Journal of Educational Psychology*, *97*(2), 257.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for information*, *22*(2), 63-75.

Shinn, M. R., Knutson, N., Good III, R. H., Tilly III, W. D., & Collins, V. L. (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review*, *21*(3), 459-479.

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, *6*(1), 1-26.

Solari, E. J., Denton, C. A., Petscher, Y., & Haring, C. (2018). Examining the effects and feasibility of a teacher-implemented Tier 1 and Tier 2 intervention in word reading, fluency, and comprehension. *Journal of Research on Educational Effectiveness*, *11*(2), 163-191.

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*(4), 585-607.

Stake, R. (1995). *The art of case study research*. Thousand Oaks, CA: Sage publications.

Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools*, *42*(8), 795-819.

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*(2), 48-58.

St. Martin, K., Vaughn, S., Troia, G., Fien, H., & Coyne, M. (2020). *Intensifying literacy instruction: Essential practices.* Lansing, MI: MiMTSS Technical Assistance Center, Michigan Department of Education.

Tindal, G., Nese, J. F., Farley, D., Saven, J. L., & Elliott, S. N. (2016). Documenting reading achievement and growth for students taking alternate assessments. *Exceptional Children*, *82*(3), 321-336.

University of Oregon. (2018). *8th Edition of Dynamic Indicators of Basic Early Literacy Skills (DIBELS®): Administration and Scoring Guide.* https://dibels.uoregon.edu/sites/dibels1.uoregon.edu/files/2021-05/dibels_8_admin_and_scoring_guide_05_2020.pdf

University of Virginia School of Education and Human Development. (2021). *Examining the impact of COVID-19 on the identification of at-risk students: Fall 2021 literacy screening findings.* *https://literacy.virginia.edu/sites/g/files/jsddwu1006/files/2022-04/PALS_StateReport_Fall_2021.pdf*

Vadasy, P. F., Sanders, E. A., Peyton, J. A., & Jenkins, J. R. (2002). Timing and intensity of tutoring: A closer look at the conditions for effective early literacy tutoring. *Learning Disabilities Research & Practice*, *17*(4), 227-241.

Vadasy, P. F., Sanders, E. A., & Peyton, J. A. (2006). Paraeducator-supplemented instruction in structural analysis with text reading practice for second and third graders at risk for reading problems. *Remedial and Special Education*, *27*(6), 365-378.

VanDerHeyden, A. M., & Burns, M. K. (2018). Improving decision making in school psychology: Making a difference in the lives of students, not just a prediction about their lives. *School Psychology Review*, *47*(4), 385-395.

Van Norman, E. R., & Christ, T. J. (2016). Curriculum-based measurement of reading: Accuracy of recommendations from three-point decision rules. *School Psychology Review*, *45*(3), 296-309.

Van Norman, E. R., Nelson, P. M., & Parker, D. C. (2018). Curriculum-based measurement of reading decision rules: Strategies to improve the accuracy of treatment recommendations. *School Psychology Review*, *47*(4), 333-344.

Vaughn, Sharon, Jeanne Wanzek, Christy S. Murray, Nancy Scammacca, Sylvia Linan-Thompson, and Althea L. Woodruff. "Response to early reading intervention examining higher and lower responders." *Exceptional Children* 75, no. 2 (2009): 165-183.

Villiger, C., Hauri, S., Tettenborn, A., Hartmann, E., Näpflin, C., Hugener, I., & Niggli, A. (2019). Effectiveness of an extracurricular program for struggling readers: A comparative study with parent tutors and volunteer tutors. *Learning and Instruction*, *60*, 54-65.

Wanzek, J., Al Otaiba, S., & McMaster, K. L. (2019). Intensive Reading Interventions for the Elementary Grades. The Guilford Series on Intensive Instruction. *Guilford Press*.

Wanzek, J., & Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *School Psychology Review*, *36*(4), 541-561.

Wasik, B. A., & Slavin, R. E. (1990). Preventing early reading failure with one-to-one tutoring: A best-evidence synthesis.

Wayman, M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education*, *41*(2), 85-120.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Child Psychology & Psychiatry & Allied Disciplines*.

Yin, R. K. (2003). *Case study research: Design and methods*. (Applied Social Methods Research Series, Vol. 5). Beverly Hills, CA: Sage publications.

Zumeta, R. O., Compton, D. L., & Fuchs, L. S. (2012). Using word identification fluency to monitor first-grade reading development. *Exceptional Children*, *78*(2), 201-220.

Appendix A

Key Terms as Defined in the Capstone Project

Adherence: Following the standard procedures of an intervention (Lemire et al., 2022); in the context of this capstone, tutors accurately performing each step of administering and scoring a CBM

Assessment Fidelity: The extent to which assessments are performed and scored as intended (Reed & Sturges, 2013)

Curriculum-Based Measure: A set of standardized procedures used to obtain valid and reliable data about student progress toward proficiency in a specific domain (NCII, 2013)

Data-Based Decision Making (DBDM): A systematic method for using assessment data to determine when and how to intensify an intervention (NCII, 2013)

Data-Based Individualization (DBI): Data-based decision making as applied to an individual student (NCII, 2013)

Educative Learning: Training that conveys what the user needs to know to implement an intervention with fidelity (Century et al., 2010); in the context of this capstone, aspects of the training that foster tutors' tutors' knowledge base about the rationale for administering CBMs

Implementation Fidelity: Delivering an intervention exactly the way it was designed and in the way it was implemented during research studies that validated its effectiveness (NCII, 2013)

Paid Volunteer Tutor: A tutor such as an AmeriCorps member or tutor whose time was donated by employers **(**Neitzel et al., 2022); Tutors at DRC are best classified as paid volunteer tutors

Procedural Learning: Training that conveys what to do to implement an intervention with fidelity (Century et al., 2010); in the context of this capstone, aspects of the training that foster

tutors' knowledge of the basic steps to administer a CBM

Progress Monitoring: A form of assessment in which student learning is evaluated at regular intervals to gauge responsiveness to instruction (NCII, 2013)

Quality: The enthusiasm with which someone implements a program and their perception of its effectiveness (Lemire et al., 2022)

Appendix B

Recruitment Email Text

Dear Tutor,

My name is Emma Pearson, and I am helping Lauren Hauser conduct a research study in the McGuffey tutoring program around using Curriculum-Based Measures (CBMs) to collect data demonstrating student progress. CBMs are brief measures of a student's reading, such as reading a passage aloud for one minute. I am asking for your voluntary participation in this research study. The information you share for the study will be used to help us understand how to support you as you collect data to monitor your students' progress.

Participation in the study will not impact any aspect of your work with McGuffey.

Participation in the study includes a collection of data from two sources:

1.) Training Data—This data has already been collected. Training Data includes data from the CBM training module you completed in Canvas as well as data from your observations with Lauren.
2.) Interviews—Lauren will conduct an interview with you after your training process that is completed as part of your job with MRC. It will take place in person and will last no longer than 30 minutes. The interview will be audio recorded and transcribed. Once transcribed, the audio recording will be deleted.

You have the option to share your Training Data only, or to share your Training Data AND participate in the interview. Sharing your data and participating in the interview is voluntary, optional, and not part of your job responsibilities.

Data included in the study will be anonymized, meaning your name will not be associated with any of your information. No data produced from this research will identify you in any way.

The anticipated risk to participating in this study is low. If you are under 18, please notify me that you cannot participate.

If you interested in participating in this study, or if you have any questions or concerns about the study, please contact:

Emma Pearson, Sub-Investigator
Telephone: 434-365-4582
Email: ecp4ey@virginia.edu

Lauren Hauser, Principal Investigator
Reading Education and Outreach Specialist
University of Virginia School of Education and Human Development

Department of Curriculum, Instruction, and Special Education
PO Box 400273
Charlottesville, VA 22904
Telephone: 434-924-8151
Email: leh8mu@virginia.edu

To obtain more information about the study, ask questions about the research procedures, express concerns about your participation, or report illness, injury or other problems, please contact:

Tonya R. Moon, Ph.D.,
Chair, Institutional Review Board for the Social and Behavioral Sciences
One Morton Dr Suite 500
University of Virginia, P.O. Box 800392
Charlottesville, VA 22908-0392
Telephone: (434) 924-5999
Email: irbsbshelp@virginia.edu
Website: www.virginia.edu/vpr/irb/sbs
Website for Research Participants: https://research.virginia.edu/research-participants

Appendix C

Consent Form

Informed Consent Agreement

Please read this consent agreement carefully before you decide to participate in the study.

Purpose of the research study: The purpose of the study is to learn how McGuffey Reading Center (MRC) can support tutors in collecting high-quality data using curriculum-based measures (CBMs). By collecting this progress monitoring data, you can provide valuable information to MRC staff to inform potential instructional modifications and recommendations to families. Through this study, we aim to identify challenges to administering CBMs and to gain your feedback to improve our tutor training process. Please note that your ability to tutor is not dependent upon your participation in the study.

What you will do in the study: As part of your job tutoring at MRC, you have already completed a CBM training module in Canvas as well as observations with MRC staff. If you consent to participate in the study, you are giving your permission for the study team to use data from the module and tutoring observations in the study (referred to as Training Data).

Additionally, you may choose to participate in an interview. You may skip any question that you do not wish to answer or stop the interview at any time. The interview will take place in person and will last no longer than 30 minutes. The interview will be audio recorded and transcribed. Once transcribed, the audio recording will be deleted.

Time required: If you consent to the interview, the study will require about 30 minutes of your time beyond your regular tutoring responsibilities.

Risks: The anticipated risks for this study are low. Data collected for the study will be anonymized, meaning your name will not be associated with any of your information. No data produced from this research will identify you in any way.
Benefits: There are no direct benefits to you for participating in this research study. The study may help us understand how to better support tutors, and, ultimately, to improve instructional outcomes for tutees.

Confidentiality: The information that you give in the study will be handled confidentially. Your information will be assigned a code number. The list connecting your name to this code will be kept in a secure file. When the study is completed and the data have been analyzed, this list will be destroyed. Your name will not be used in any report.

Voluntary participation: Your participation in the study is completely voluntary. Your decision to participate will have no effect on any aspect of your work with MRC.

Right to withdraw from the study: You have the right to withdraw from the study at any time without penalty.

How to withdraw from the study: If you want to withdraw from the study, please notify Lauren Hauser (434-924-8151; leh8mu@virginia.edu). In this case, your Training Data will be excluded from the study, and, if you have already provided an interview, that will be excluded as well. If you decide during the interview that you wish to withdraw, tell the interviewer to stop the interview. Withdrawing does not impact any aspect of your work with MRC.

Payment: You will receive no payment for participating in the study.

Using data beyond this study: Your data will not be used beyond the study. The data you provide in this study will be retained in a secure manner by the researcher for three years and then destroyed.

If you have questions about the study, contact:
Lauren Hauser, Principal Investigator
Reading Education and Outreach Specialist
University of Virginia School of Education and Human Development
Department of Curriculum, Instruction, and Special Education
PO Box 400273
Charlottesville, VA 22904
Telephone: 434-924-8151
Email: leh8mu@virginia.edu

Latisha Hayes, Faculty Advisor
Associate Professor
University of Virginia School of Education and Human Development
Department of Curriculum, Instruction, and Special Education
PO Box 400273
Charlottesville, VA 22904
Telephone: 434-982-2021
Email: llh6e@virginia.edu

To obtain more information about the study, ask questions about the research procedures, express concerns about your participation, or report illness, injury or other problems, please contact:
Tonya R. Moon, Ph.D.
Chair, Institutional Review Board for the Social and Behavioral Sciences
One Morton Dr Suite 400
University of Virginia, P.O. Box 800392
Charlottesville, VA 22908-0392
Telephone: (434) 924-5999
Email: irbsbshelp@virginia.edu
Website: https://research.virginia.edu/irb-sbs
Website for Research Participants: https://research.virginia.edu/research-participants
UVA IRB-SBS # 5671

Agreement:

Please indicate your consent to participate in the study by initialing one or both statements below.

_____ I consent to participate in the study as described above by sharing my Training Data.

_____ Additionally, I consent to participate in the interview as described above.

Print Name: _____ Date: _____

Signature: _____

You will receive a copy of this form for your records.

Appendix D

Baseline Questions

Instructions: Before you start the module on Progress Monitoring, please respond to these short answer questions. I appreciate your honest responses-- there are no right or wrong answers here!

The questions are phrased to be about just one student-- if you have worked with more than one student, please answer generally based on your experiences across students.

1. Do you ever wonder whether your student is making progress? Why or why not?
2. What do you think would be a meaningful measure of your student's progress, and why do you think that?
3. To what extent would having an objective measure of your student's progress impact your *investment* in tutoring, and why?
4. To what extent would having an objective measure of your student's progress impact what you actually *do* in tutoring, and why?

Appendix E

Field Notes Template

Date:	Start Time:	End Time:	Elapsed Time:
Tutor Name:		Passage:	
Phase of Administration	Criteria	Pass/Needs Practice	Commentary
Setup	Holds clipboard and timer so student can't see what is recorded		
	Places student copy in front of student		
Passage Intro	Performs standardized directions verbatim		
	Starts timer when student says the first word		
During Reading	Follows along and marks the scoring book as the student responds		
	Administers acceptable prompts correctly, if appropriate		

	Applies the discontinue rule correctly and when appropriate		
After Reading	At the end of 60 seconds, puts a bracket (]) after the last word read and says **"Stop"**.		
Scoring	Applies scoring rules consistently and correctly		
	Accurately determines and records the number of words read correctly. Score is within two points of the expert examiner.		
Notes on Tutor Affect (tone, body language):			
Questions Tutors Ask Before:			
Questions Tutors Ask After:			
My Questions or Interpretations:			
Reflective Summary:			

Appendix F

Interview Protocol

Participant: _____
Date: _____
Start Time: _____
End Time: _____

Consent

- Thank you for agreeing to participate in this interview and this research project. This interview should take around 30 minutes. As you know, the goal of this interview is to learn more about how tutors at DRC collect data so we can track our tutees' progress.
- Before we start, I wanted to let you know that you can end the interview at any time. Just let me know if you feel uncomfortable or would like to stop for any reason.
- I will be audio recording this interview so that I can refer back to it later. Do you have any concerns about being recorded? If you change your mind at any point, just let me know. After I have transcribed the interview I will delete the recording.
- Do you have any questions before we begin?

Background

- What is your class year and program of study?
    - Have you ever taken any courses related to education? What was the nature of those courses?
- Before coming to DRC, what were some of your prior experiences working with children?
- How long have you been tutoring with DRC?
- How would you describe your experience of tutoring?
- What do you see as your role as a tutor in this context?

CBMs

- Before this spring, did you have previous experience administering CBMs, either at DRC or in a different context?
    - If yes, how did the training process you just completed change your understanding CBMs?

Adherence

- How would you describe your comfort level with administering the CBMs?
    - scoring them?
- How confident do you feel that you can administer and score the CBMs according to the standard procedure?
- What elements of the CBM administration procedure did you find most challenging?

Quality

- What is your understanding of why we are using CBMs at DRC?
- What do you see as the major advantages and disadvantages for us in using CBMs?

- In your opinion, is it worthwhile for us to use CBMs? Why or why not?
    - Are CBMs a meaningful way to gauge progress?
- To what extent are you curious about your tutee's CBM data?
- Can you imagine taking any sort of action in response to your tutee's CBM data?

Training

- What elements of the training did you find to be most helpful?
    - What elements did you find to be unhelpful?
    - Did you find it helpful to have info about the rationale for giving CBMs?
        - Why or why not?
- How did the training contribute, or not contribute, to your comfort with administering CBMs?
- How did the training contribute, or not contribute, to your investment in administering CBMs?
- What suggestions do you have for how we might improve the training?

Appendix G

Excerpt from Adherence-Focused Codebook

a priori codes		
Category	**Code**	**Definition**
1. Proficient in Administering on Initial Administration: Use these codes when tutor demonstrates proficiency on a criterion during the initial administration	1.01	Holds clipboard and timer so tutee can't see what is recorded
	1.02	Places tutee copy in front of tutee
	1.03	Performs standardized directions verbatim
2. Needs Practice in Administering on Practice Administration: Use these codes when tutor receives a score of NP on a criterion during the practice administration	2.01	Does not hold clipboard and timer so tutee can't see what is recorded
	2.02	Does not place tutee copy in front of tutee
	2.03	Does not perform standardized directions verbatim
emergent codes		
11. Nature of Error: Use these codes when tutor receives a score of NP on a criterion during any administration	11.01	Starts timer before tutee reads the first word
	11.02	Counts insertions as errors
	11.03	Adds words to standardized directions
	11.04	Omits words from standardized directions
	11.05	Does not prompt within 3 seconds when tutee hesitates

Appendix H

Excerpt from Quality-Focused Codebook

a priori codes			
Category	**Code**	**Definition**	**Example**
Tutor Expresses Concerns about CBMs	Stressful	Tutor expresses concern that the CBM is a source of stress for their tutee.	I just remember when I was in school being like… stressed when we would do things like this.
	Off-Day	Tutor expresses concern that their tutee's performance on a given day does not reflect their ability.	Sometimes, a kid might have one of those days where they're extra jittery, and they'll skip over a line.
Tutor Refers to Impact of Training on Investment in Assessment Fidelity	Responsibility	Tutor expresses that the training conveyed that they should take responsibility for CBM administration fidelity.	I feel like it was something I needed to take seriously and be very accurate with.
	Educative	Tutor refers to an educative aspect of the training specifically as increasing investment.	The training is not just saying, "You have to start doing this," but it's explaining why I want to buy into that and use it.
emergent codes			
Tutor Expresses Concerns about CBMs	Unrelated	Tutor indicates that they see a disconnect between CBM and instructional goals.	How well they do on the CBM doesn't always translate to how well they do in reading.
	Variability (Passage)	Tutor indicates that they do not see passage sets as equivalent.	Sometimes I think the difficulty levels of each passage are, like, different across the weeks… but it's not like, going from easy to hard. It's just like kind of random, so I don't know if that's like the best method [for measuring progress]
Tutor Refers to Impact of Training on	Confidence	Tutor expresses that the training process helped them feel	[The training made me feel] qualified...you feel really prepared.

Investment in Assessment Fidelity		confident they can administer CBMs with fidelity.	
	No Impact	Tutor expresses that the training did not increase their investment in administering CBMs.	Doing the training wouldn't make me want to [administer CBMs] if we didn't have to.