Minnesota State University, Mankato

CORNERSTONE
MINNESOTA STATE UNIVERSITY MANKATO

Cornerstone: A Collection of Scholarly and Creative Works for Minnesota State University, Mankato

All Graduate Theses, Dissertations, and Other Capstone Projects

Graduate Theses, Dissertations, and Other Capstone Projects

2014

Effects of a Cross-Age Peer Tutoring Intervention on English Language Oral Reading Fluency in a Belizean Grade School

Marcia Ruth Sytsma
Minnesota State University - Mankato

Follow this and additional works at: https://cornerstone.lib.mnsu.edu/etds

 Part of the Educational Psychology Commons

Recommended Citation

Sytsma, M. R. (2014). Effects of a Cross-Age Peer Tutoring Intervention on English Language Oral Reading Fluency in a Belizean Grade School [Doctoral dissertation, Minnesota State University, Mankato]. Cornerstone: A Collection of Scholarly and Creative Works for Minnesota State University, Mankato. https://cornerstone.lib.mnsu.edu/etds/319/

This Dissertation is brought to you for free and open access by the Graduate Theses, Dissertations, and Other Capstone Projects at Cornerstone: A Collection of Scholarly and Creative Works for Minnesota State University, Mankato. It has been accepted for inclusion in All Graduate Theses, Dissertations, and Other Capstone Projects by an authorized administrator of Cornerstone: A Collection of Scholarly and Creative Works for Minnesota State University, Mankato.

Effects of a cross-age peer tutoring intervention on English language oral reading fluency in a Belizean grade school

by

Marcia Ruth Sytsma

A Dissertation Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Psychology

In

School Psychology

Minnesota State University, Mankato

Mankato, Minnesota

May 2014

Effects of a cross-age peer tutoring intervention on English language oral reading fluency in a Belizean grade school

Marcia Ruth Sytsma

This dissertation has been examined and approved by the following members of the student's committee.

_____
Carlos J. Panahon, Ph.D., Advisor

_____
Daniel D. Houlihan, Ph.D., Committee Member

_____
Lisa M. Perez, Ph.D., Committee Member

_____
Teresa L. Wallace, Ph.D., Committee Member

Dedication

Carrie and Maggie

Who are my inspiration

Greg

Whose love, support, and patience is complemented

by his reminders to not worry about problems I do not have

and

Monica

The best peer tutor a sister ever had

Acknowledgements

Every step of this journey has been possible and easier through the support and love from my extended family and friends. Every gesture, no matter how small, was a gift.

My family, both by birth and connected by friendship, allows me to go forth bravely. My beloved and brilliant daughters, Carrie and Maggie, and my best friend and partner Greg, have made sacrifices of time and companionship to support my dreams. You are the loves of my life. My sister, brothers, their children and spouses, and my parents have always held me to be my best and supported my every goal. I love you all so deeply, and I am so grateful you are in my life, now and always.

Thank you Sean and Liz for training tutors and assessing students, and for doing all the countless tasks that extended an intervention into research. Thank you Melissa Lenz, for fidelity checks and following up on my project during your own doctoral practicum. Particularly, thank you to all of the students and their families who agreed to participate. This project exists because of your willingness to take a chance to see if "Ms. Marcia's" ideas would help.

Thank you to the professors of the Doctoral Program in School Psychology at Minnesota State University, Mankato, who deserve my deepest gratitude for the training and vision they have given. I have infinite gratitude for my advisor, Dr. Carlos J. Panahon, for his cheerful yet persistent reminders to hold myself to the highest standards. Thank you Chip for reminding me of my goals whenever I hesitated or had doubts. I am

indebted also to Dr. Daniel Houlihan, from the first phone call, through the notorious oral "quizzes" in Behavioral Research Methods, to the time he spent reviewing graphs for just another doctoral candidate. I have never regretted the phone call or the decisions that followed. Thank you to my committee members, Dr. Lisa Perez and Dr. Teresa Wallace for your enthusiasm and expertise.

Jules, thank you for the glowing path. The way was clear because you went first and lit it so brightly. Jess, thank you for the reminder about the possibilities of peer tutoring. Here is the result. Natasha, thank you for your constant friendship and support in a million ways. RaeLynn and Cassie, you both jumped in at a moment's notice in multiple projects to ask how you could help. All the time. Thank you. My "Melissas," the two women who have been my closest research partners for the past two years. You simply blow me away with the sheer brilliance you wear so easily. Mel, I am so very grateful friendship is your superpower.

Thank you to the individuals who provided the training, support, protocols, and materials that made it possible to implement this program in Belize: James Wright of Intervention Central, Roland Good and Ruth Kaminski and their talented team at Dynamic Measurement Group, Inc., especially Katherine Bravo Aguayo, and the anonymous teachers and writers at ReadWorks. Thank you to Minnesota State University, Mankato for providing research and travel support, particularly the Department of Graduate Studies and Research. Finally, thank you to Lady Dixie Bowen, for your vision and support of the students in Belize.

Abstract of the Dissertation

Effects of a cross-age peer tutoring intervention on
English language oral reading fluency in a Belizean grade school

by

Marcia Ruth Sytsma

Doctor of Psychology in School Psychology
College of Graduate Studies and Research
Minnesota State University, Mankato, 2014
Carlos J. Panahon, Ph.D., Chair

A cross-age peer tutoring program was implemented in a small rural school in west central Belize, Central America. All students at the school were native Spanish speakers, and all general instruction was conducted in English. The program was devised to supplement existing reading and language arts instruction at all grade levels. Progress of both tutors and tutees was monitored on a weekly basis using DIBELS Next measures. Twenty-nine students started the program, with complete data available for twenty-one students: seven tutee-tutor pairs, and seven matched students who participated as controls. Three main hypotheses were tested with the present study. These hypotheses investigated the following criteria for effectiveness of this program: (a) statistically significant (reliable) results for students within tutoring pairs, (b) socially relevant changes (e.g., perceptions of improvement), and (c) maintenance of skills or transfer of skills. Progress monitoring results were mixed, but socially relevant outcomes were found for tutee progress on benchmark assessments. A survey of teachers at the school highlighted increased interest and motivation for reading and class participation attributed to the program.

Table of Contents

List of Tables

Chapter 1. Introduction

According to a report from the United States' National Reading Panel (National Institute of Child Health and Human Development [NICHD], 2000), reading instruction with an emphasis on early prerequisite skills is essential for student success. The panel evaluated research linking early literacy skills with successful reading and successful reading with academic success. According to the report, strong readers are more successful at automatically extracting meaning from written material (reading comprehension), through accurate and efficient reading of the material and applying what they have read to a specific context (NICHD, 2000).

The prerequisite skills that were identified as critical for successful reading are three skills related to connecting sounds to written text and using those concepts to decode and reassign sounds to specific words (phonemic awareness, alphabetic principle, and phonics; Kaminski & Good III, 2012; NICHD, 2000), automatically and efficiently using these skills to read words within text, and applying meaning to what has been read (NICHD, 2000). August and Shanahan (2006) describe these as word level and text level tasks. In addition to the word level texts of phonemic awareness, alphabetic principle and phonics, text-level tasks include accurate and fluent reading of text, vocabulary knowledge, and applying vocabulary knowledge efficiently to interpret written materials (August & Shanahan, 2006; Good et al., 2011a; Kaminski & Good, 2012), vocabulary development and reading comprehension (NICHD, 2000).

Within the National Reading Panel report, each of these skills was evaluated separately for their contribution to successful reading. That is, the research on educational

strategies to develop each skill was exhaustively reviewed by subcommittees within the panel. From these evaluative reviews, a set of best practices to address each skill were identified and formed the basis of the panel's recommendations for addressing reading problems (Good et al., 2011; Kaminski & Good, 2012; NICHD, 2000).

One of the points the National Reading Panel emphasized in this review was the critical role played by applying word-level skills to efficiently reading written material, a skill referred to as either reading fluency or fluency for short. Combined with vocabulary development and understanding of word meaning, this skill presents the last step prior to building reading comprehension. This skill was also distinct from earlier literacy tasks in that it requires using earlier skills in increasing both the accuracy and speed at which reading occurs. Therefore, the primary goal of instruction to build fluency is providing sufficient and appropriate practice opportunities (NICHD, 2000).

The Case for Peer Tutoring

One of the primary findings from the National Reading Panel report on reading fluency is that there are specific features of practice that are more effective in building fluency skills. One of the primary findings was that practice reading aloud (with or without feedback) resulted in significant gains in reading fluency (NICHD, 2000, p. 3-3). Any positive relationship found for studies showing effects of silently reading textual material could not be separated from the effects of reading comprehension on increasing fluency. That is, students who read well are more likely to enjoy reading on their own, and since most of the studies reviewed by the National Reading Panel regarding silent reading were correlational, it is difficult to separate out the effect of reading well on

practice, versus increasing and encouraging practice having an effect on increasing skills in reading comprehension (NICHD, 2000). A meta-analysis by Elbaum, Vaughn, Tejero Hughes, and Watson Moody (2000) found a weighted effect size of .41 (i.e., .41 standard deviations of improvement relative to students not receiving the intervention) for both group-based and individualized teacher-led practice as the National Reading Panel found for all oral reading practice effects on oral reading fluency (NICHD, 2000).

Multiple models for increasing oral reading practice have been evaluated since 2000. Subsequent meta-analyses and reviews and directs tests of these models have provided additional support for the NRP's general conclusions on the importance of opportunities to practice oral reading (e.g., Algozzine, Marr, Kavel, & Dugan, 2009; Ardoin, Williams, Klubnik, & McCall, 2009; Elbaum, Vaughn, Tejero Hughes, & Watson Moody, 2000; Fuchs & Fuchs, 2005; Ginsburg-Block, Rohrbeck, & Fantuzzo, 2006; Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003; Schreder, Hupp, Everett, & Krohn, 2012; Shapiro, 2011; Vostal & Lee, 2011).The findings from these studies support the importance of practice in general and features of such practice that lead to improvements in students' reading skills.

Ardoin, Williams, Klubnik, and McCall (2009) varied the amount of practice provided to students to identify the optimal level for developing fluency and maintaining performance levels. In their study, four students were asked to reread a target passage either three or six times, in an alternating treatments design. This means that each student was first asked to read a passage either three or six times and then was asked to read a separate passage under the other set of instructions (Ardoin et al., 2009). They found that

for these four participants, there was no difference between their accuracy and mean words correct per passage immediately after the practice, nor was there any difference between the amount of practice on a given passage and the students' ability to generalize their skills to reading a second passage. However, there was a significant practice effect on how well the students were able to maintain their performance levels in reading the practiced materials, with passages read six times showing a longer lasting effect. It is helpful to remember this finding when planning future tutoring session plans. The ultimate goal of any intervention is not only to show immediate improvement in skills, but to also demonstrate that these skills can be maintained.

The National Reading Panel report also identified parents and peers as potential sources for guided reading support (where appropriate reading is first modeled by the "guide" and then feedback is given during practice on the modeled reading) during oral reading practice. Each of these options has also been the focus of several programs and targeted research efforts. Parents who are proficient readers, with specific training on protocols, can provide such guided feedback, as demonstrated by Schreder, Hupp, Everett, and Krohn, (2012). In their study parents acted as coaches during a summer, and they were able to demonstrate long-term effects for two students with reading problems, while manipulating the specific type of oral reading practice (repeated reading, RR; or listening passage preview, LPP) in an alternating treatment design to determine practice type. Once the optimal practice type was identified, Schreder et al. used a multielement design to evaluate whether or not the addition of a reward improved performance to a greater degree than practice alone. Both types of RR and LPP provided opportunities to

practice reading, but the researchers adapted the task based on an analysis of students' pre-existing skills (Schreder et al., 2012). However, there are several situations where implementing either a parent-focused or teacher-led practice might not be feasible (Rohrbeck et al., 2003).

Much like adapting practice type to student needs was supported by Schreder et al. (2012), oral reading practice has been shown to yield greater gains in fluency and comprehension when the material is presented at a level corresponding to the students' reading proficiency (Marr, Algozzine, Nicholson, & Dugan, 2011; Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003; Shapiro, 2011). If a teacher has a classroom with multiple skill levels or grades, providing sufficient practice at an appropriate level for all students may warrant a more individualized approach (Fuchs & Fuchs, 2005).

One strategy that has been recommended for this purpose is having other students provide tutoring instead of teachers or parents (e.g., Algozzine et al., 2009; Allen, n.d.; Dufrene, Henington, & Townsend, 2006; Fuchs & Fuchs, 2005; Marr et al., 2011; Rohrbeck et al., 2003; Wendling & Mather, 2009; Wright, 2004; Yurick, Robinson, Cartledge, Lo, & Evans, 2006). These have varied between reciprocal strategies where students take turns acting as tutors (e.g., Dufrene, Hennington, & Townsend, 2006; Fuchs & Fuchs, 2005) and cross-age or cross-skills tutoring, where a more proficient reader models reading tasks and provides feedback and guided support for a less proficient reader (a tutee; Wright, 2004; Wright & Cleary, 2006). When students whose primary language is English are instructed to read in English, peer tutoring enhances learning in academic skills such as oral reading fluency and reading comprehension ((Fuchs &

Fuchs, 2005; Ginzburg-Block, Rohrbeck, & Fantuzzo, 2006; Ginsburg-Block, Rohrbeck, Fantuzzo, & Lavigne, 2006; NICHD, 2000; Rohrbeck et al., 2003; Wright & Cleary, 2006). Fuchs and Fuchs (2005), Ginsburg-Block, Rohrbeck, and Fantuzzo, (2006), and Ginsburg-Block, Rohrbeck, Fantuzzo, and Lavigne (2006) also endorse using peer tutoring to enhance social skills and improved cooperation between students engaged in reciprocal peer tutoring (when students take turns serving as tutors for each other). All of these studies and meta-analyses emphasize the effects of peer tutoring in a traditional, monolingual setting where tutors and tutees are proficient in the language of instruction.

Rohrbeck et al. (2003) evaluated several features of the research on peer tutoring and concluded that peer tutors (regardless of whether tutoring was implemented individually or in a class-wide intervention) were effective at improving tutee's fluency, as well as enhancing other academic skills among both tutors and tutees (e.g., Dufrene et al., 2010; Kourea, Cartledge, & Musti-Rao, 2007) and social interactions (Ginsburg-Block, Rohrbeck, & Fantuzzo, 2006). In 2006, Ginsburg-Block, Rohrbeck and Fantuzzo expanded their earlier meta-analysis on academic benefits of peer tutoring (e.g., Rohrbeck et al., 2003) to highlight the social benefits of peer tutoring. In this subsequent meta-analysis, they found smaller, yet still positive outcomes on the social interactions between tutors and tutees through the process of tutoring.

Rohrbeck et al. (2003) identified several specific characteristics of more effective peer tutoring programs. They found that effect sizes for reading skills were moderate overall but higher in studies that allowed student tutors to set goals and administer incentives, provided individualized outcome measures, or used group contingencies for

improved performance by tutees. Ginsburg-Block, Rohrbeck, Fantuzzo, and Lavigne (2006) and Wendling and Mather (2009) describe peer tutoring as one of the best interventions for supplementing standard teacher-directed instruction. Several models of peer tutoring have been proposed and generally evaluated for reading and other outcomes (Wendling & Mather, 2009). Besides the meta-analyses presented earlier, many studies have focused on small samples and single case implementations, which have also yielded "larger overall effect sizes," as stated by the NRP (NICHD, 2000). These effect sizes are adjusted during meta-analyses. These reviews support peer tutoring as an intervention to enhance student achievement in several academic areas, especially where students are being taught to read in their primary language within the United States.

A question emerges about how well these findings generalize to academic and social outcomes of peer tutoring among students who are learning to read English while they are learning English. As both the NRP report and a joint report by the Center for Applied Linguistics and the National Literacy Panel on Language-Minority Children and Youth National Panel (CAL/NLPLMCY) note, the linguistic and cultural diversity within schools is increasing (August & Shanahan, 2006; NICHD, 2000). Many of these students are being taught to read in English in spite of the fact that English is not their primary language (August & Shanahan, 2006).

Second Language English Literacy and Peer Tutoring

Although the National Reading Panel stated that its conclusions and recommendations were limited to students whose first language was English; a subsequent report was produced through the cooperative efforts of the Center for Applied

Linguistics and the National Literacy Panel on Language-Minority Children and Youth (CAL/NLPLMCY; August & Shanahan, 2006). This report also summarized research in countries other than the United States and research on literacy development in multiple languages (both first and second languages). These two reports addressed somewhat different issues yet found common themes. One consistent finding in the CAL/NLPLMCY report is that the same basic skills contribute to literacy in English, regardless of whether or not English is the student's first or second language (Lesaux & Geva, 2006). On average, for students who are being taught how to read, English literacy progresses from recognizing that words consist of sounds, progressing through fluency in reading words in context, and then deriving the meaning from those words (International Reading Association, 2013; Lesaux & Geva, 2006; NICHD, 2000). According to this model, good readers are able to automatically extract meaning from reading (i.e., are strong in reading comprehension). The NRP focused their efforts on defining the prerequisite, or foundational skills that contribute to proficient reading. The majority of the NRP report focused on the research to support this model along with summaries of research evaluating the most effective approaches for developing these skills.

Both panels emphasized the importance of providing supplemental interventions to students to remediate prerequisite skills (August & Shanahan, 2006; National Institute of Child Health and Human Development, 2000). Research summarized by Garcia, McKoon, and August (2006) and Lesaux and Geva (2006) within the CAL/NLPLMCY report highlights both the similarities and differences among students instructed in English as either a primary or secondary language. One of the conclusions drawn by

Garcia et al. is that most of the literacy skills among students learning to read and learning English simultaneously do not show differences between their results and those of students for whom English is their primary language. In other words, reading fluency scores in English for students whose primary language is Spanish are equally predictive of reading comprehension in English. However, the research summarized by Garcia et al. and generally described elsewhere in the report shows a significant mean difference in both reading comprehension and reading fluency scores depending on whether the reading instruction is provided in a student's primary or secondary language (August & Shanahan, 2006).

One of the basic skills emphasized by both panels was reading fluency, or the speed and accuracy of reading words embedded in meaningful text, described in more detail above. Although peer tutoring was not specifically evaluated within this report, general statements within the report include an observation that most literacy interventions for students learning English as a second language showed a lower effect size than similar studies among students for whom English is a primary language. However, a recent meta-analysis by Adesope, Lavin, Thompson, and Ungerleider (2011) found a similarly strong impact for peer tutoring interventions on the English oral reading fluency of students for whom Spanish is a primary language as Elbaum et al. (2000) found for teacher-led tutoring of similar students. This is also the same value found in the National Reading Panel meta-analysis for oral reading practice (e.g., weighted effect size of Hedge's g=0.41). This finding is surprising, given general statements within August and Shanahan (2006) that the task facing children whose primary language is not the

language in which they receive instruction is more complex and depends a great deal on the correspondence between their primary and secondary languages. For example, the correspondence between the phonological characteristics of letters in the English and Spanish alphabets may require more practice on these specific elements during phonological and decoding instruction (Genesee et al., 2006). Similarly, they are also translating from their own linguistic map of meaning to a new one, a task that Lesaux and Geva (2006) point out may require an additional step and thus is much more complicated than the task presented to a student for whom English is the primary language.

August and Shanahan (2006) and Ginsburg, Rohrbeck, Fantuzzo, and Lavigne (2006) caution that before embarking on implementing an intervention for students who are learning to read in a language other than their own, several other contextual features are important to consider. An important consideration is a careful analysis about the student's proficiency in the student's primary language (August & Shanahan, 2006) and the student's proficiency and/or literacy in other languages that are relevant within the community. August and Shanahan (2006) point out that it is possible for the official language and national language of a country to differ. They distinguish between a student's native language, the societal language of the student's community, the national and official languages (if an official language exists) of the country, and the language used for instruction. According to August (2006), official languages are mandated by law (p. 3; August, 2006) where they exist, and these mandates may not correspond to the prevailing language spoken in the community or country.

Several questions guide the design and structure of a process to choose, adapt, implement and evaluate a suitable intervention to be used in a particular school (Adesope et al., 2011; Rohrbeck et al., 2003). Meta-analyses may guide which features are likely to lead to stronger outcomes (e.g., Adesope et al., 2011; NICHD, 2000; Rohrbeck et al., 2003), but several investigators have highlighted the importance of acceptability, a finding that has led both national panels and other researchers to endorse gathering information about an intervention's acceptability as one outcome measure (August & Shanahan, 2006; NICHD, 2000). To guide the specific design for the present study, general recommendations from these panels and Adesope et al. (2011)'s and Rohrbeck et al. (2003)'s meta-analyses on key characteristics of effective studies are combined with contextual factors specific to the targeted school, provided in the next two sections of this introduction. The basic findings from Rohrbeck et al. (2003) are provided next.

Rohrbeck et al. (2003) concluded that these benefits from peer tutoring interventions are strongest for students who do not have as many non-school resources, such as students from a lower SES background, or students with a primary language other than English with parents who are not fluent enough in English to assist with feedback during oral reading practice. They surmised that positive benefits for students learning English may result from the informal and varied structure of language "inputs" provided by tutors who are proficient in English (Rohrbeck et al., 2003). This conclusion and rationale is shared by the joint panel on language minority literacy (August & Shanahan, 2006).

Several studies have evaluated peer tutoring specifically provided to students learning English as a second language in the United States, including the meta-analyses outlined above. Most of the studies listed below emphasized student perceptions of effectiveness, although several also mention quantitative summaries (Arquette, 2000; D. L. August, 1982; Calhoon, Al Otaiba, Greenberg, King, & Avalos, 2006; Calhoon, Al Otaiba, Cihak, King, & Avalos, 2007; Mooney, 2010; Pyron, 2007; S. B. Saenz, 2008; Serrano, 1987; Stryker, 1986; L. M. Sáenz, Fuchs, & Fuchs, 2005). The studies summarized below highlight some of the process issues and intervention characteristics that have been hypothesized to contribute to more effective peer tutoring interventions with students learning to read in English when their primary language is Spanish. Most of the research described below relied upon qualitative analyses of outcomes but are included here because one of the issues presented by the NRP and others (e.g., Wright & Cleary, 2006) is the importance of including perceptions by participants. The discussions from the qualitative studies may help guide the choices about the types of perceptions that are likely to be relevant for participants. However, because the proposed study is primarily focused on the impact of peer tutoring on measured outcomes, this section concludes with a review of quantitative outcome-focused studies.

Qualitative studies of tutoring. Almost all of the studies conducted with elementary students whose primary language is Spanish and who are learning both reading and English have demonstrated some positive outcomes. The studies listed below are restricted to studies where either spontaneous use of English or reading fluency or comprehension was a dependent variable. Although other studies have focused on the

effects of peer tutoring on vocabulary or word knowledge in English (e.g., Greenwood, 2001) and have demonstrated positive outcomes, vocabulary improvement was only included in the present review if it was combined with a reading or speaking fluency measure, as in Mooney (2010).

Arquette (2000) investigated the process quality and English reading comprehension during cross-age peer tutoring among students learning English. Students participating in the study had a range of tenure living in the United States from two to more than four years. Seventeen students enrolled in sixth grade were chosen to be tutors, and the journal entries and session audiotapes from a smaller sample of four students were chosen for more extensive qualitative analysis. These four students were chosen on the basis of three criteria: reading scores at least two levels below grade-level expectations, enrollment within the school for at least five years, and not receiving special education services. The primary dependent variables used to evaluate the effectiveness of the peer tutoring intervention: scores on an "informal reading assessment" (Arquette, 2000; p. 21), a qualitative analysis of taped and transcribed sessions and journal entries recorded by the students participating in the project. These tutors provided tutoring on a variety of tasks chosen and planned by the students. Tutors were allowed up to four hours for planning activities per week, and tutoring sessions were scheduled for a total of 70 minutes in two sessions per week, for a total of about four months of tutoring, or 30 total sessions (Arquette, 2000; p. 99). Fifteen of the tutors had improved scores in reading comprehension on the *informal* reading assessment after four months and most of the students reported positive responses to the tutoring process.

Pyron (2007) provided a qualitative description of the process of implementing a peer tutoring intervention for high school students who had previously failed the state language assessment. This study did not demonstrate that the peer tutoring was effective in improving testing outcomes for most of the students in the program. However, Pyron argued that the testing process was an inadequate measure of the program's overall effectiveness, particularly given the challenges of implementing a program with minimal support. This study was a rare attempt to evaluate peer tutoring for students learning English within a high school setting and Pyron's explanation of the contextual variables such as administration support for tutoring provided key insights for future intervention efforts.

August (1982) conducted two studies of incidental English instruction during a peer tutoring task that did not provide specific practice modeling and providing feedback to tutees among elementary students, all of whom were identified as Mexican-American. Twelve students spoke only English, while, the other fourteen students spoke Spanish as their first language and only limited English. The purpose of the study was to investigate the role of first language proficiency in learning a second language and the ancillary social benefits that can occur as a result of informal peer tutoring. Two studies took place over a thirteen week period, and included about a half-hour/day of tutoring, with several brief (2.5 to 10 minute) observation sessions and formal testing in both language proficiency (using the *Peabody Picture Vocabulary Test*, *Language Assessment Scale*, and *James Test of Language Dominance*) and cognitive abilities (*Raven Progressive Matrices*). The primary dependent variable, however, consisted of the observations of

spontaneous English usage of the students learning English (tutors in the first study, tutees in the second study) outside of tutoring sessions. In the first study, students learning English students were paired with tutees from the Early Childhood Program who only spoke English. In the second study, tutors who only spoke English were paired with students from the Early Childhood Program who only spoke Spanish. Initial results from each these studies did not show a significant increase in English spoken spontaneously during free play activities, but there was a significant increase in English words spoken thirteen weeks after the end of the intervention (August, 1982).(Mooney, 2010) provides some insights into the processes and features of tutors that may enhance a peer tutoring intervention with students whose primary language is Spanish. In this study, students provided extensive feedback about tutors and tutor characteristics that increased their motivation to participate in the tutoring. Tutors who did not provide enough positive feedback within a corrective feedback paradigm, or who expressed frustration with their tutees generated a negative reaction from tutees who asked to be reassigned. In this study, as in August (1980), tutee/tutor relationships were generally stronger and more spontaneous friendships emerged as a result of the tutoring process, in addition to increasing the accuracy of reading a specific children's book throughout the study, and increased accuracy on a vocabulary word list. One of the primary limitations of this study is that the reading fluency measure used was repeated readings of one children's book, so generalization of the results were not possible to other texts. Reading fluency and comprehension within this book improved, however, the study did not include a measure of how well the students' skills generalized to new reading material.

Quantitative studies. The following studies (concluding with a meta-analysis by Adesope et al., 2011) focused on how peer tutoring affected measured benefits in reading and English skills for students participating in the tutoring sessions. Serrano (1988) used thirty-two third through fifth-grade students proficient in English (17 of whom spoke English as a primary language) as tutors for forty-two students whose primary language was Spanish and who were not proficient in English in a cross-level (but same grade) peer tutoring intervention, where tutees were presented with pictures and coached in story sequencing by tutors. Cross-level status in this study was defined in terms of students' proficiency in English (as measured by a standardized oral proficiency assessment). The independent variable in this study was whether or not the tutor spoke English as a primary language or was bilingual in both Spanish and English. A pre-test of Spanish proficiency was administered to all students who were tutored to test for covariation with results from an English oral proficiency test. The *IDEA Oral Proficiency Test I (K-6)* or IPT (Ballard & tighe language assessment) was the primary dependent variable. All students who were tutored showed improvement in English oral proficiency, and there was no significant difference between groups based on the independent measures or covariate of Spanish literacy.

L. M. Sáenz, Fuchs, and Fuchs (2005) provided one of the most rigorous and largest investigations of the effectiveness of reciprocal peer tutoring in meeting the needs of students whose primary language is Spanish who were learning to read in English, although such an intervention is unlikely to be feasible within the current project. In this study, six different classrooms of students were randomly assigned either to participate in

the reciprocal peer tutoring effort or did not receive any supplemental reading practice. The researchers also matched students on the basis of whether they were diagnosed as having a learning disability, and by teacher and state assessment score categories (Low, Average, or High). Three different dependent variables were used to measure reading skill effects: number of words correctly read from a passage, the number of correctly answered questions about a passage, and the number of correctly added words in a passage where several words were randomly omitted. The first dependent measure was used to assess reading fluency, and the second and third dependent measures were used to assess reading comprehension. The only significant main effect of the intervention (difference between pretest and posttest scores) was found for the number of questions answered correctly for the passage. There was no significant interaction based on student's prior achievement or educational status. The overall effect size on this dependent measure was just over one standard deviation, such that students who participated in the peer tutoring showed more improvement in their ability to answer these questions correctly. In spite of non-significant main effects or statistically significant interactions based on student achievement status, the effect sizes across student categories ranged from .04 to 1.01. Both students and teachers indicated that they were very satisfied with the intervention. The NRP report (NICHD, 2000) recommended including teacher and student feedback in evaluating interventions. Another study of reciprocal peer tutoring by Calhoon, Otaiba, Cihak, King, and Avalos (2007) provided additional support for this approach. This study is viewed as one partially addressing the positive effects of reciprocal peer tutoring on the English reading skills of students

learning English, because although Calhoon et al. randomly assigned students to participate in the peer tutoring intervention, there was also disproportionate attrition among students who were learning English as a second language, compared to the attrition rate of students in the program whose primary language is English.

One issue in studies of reciprocal peer tutoring is that the effects on tutors' skills are confounded with the effects of peer tutoring on tutees' skills, since students serve both as tutors and as tutees. S. B. Saenz (2008) implemented a cross-age tutoring intervention as an intervention for seventeen third grade students learning English, whose primary language was Spanish. Students identified as tutees had all failed to make adequate progress in reading. In this study, students were either tutors or tutees. Students enrolled in fourth and fifth grade whose primary language was English served as the tutors. This study had a dual purpose: to evaluate how peer tutoring affects the reading skills of both tutors and tutees. Tutees scored significantly higher than comparable students on state reading assessments—with 100% of the students achieving proficiency by the end of the intervention. However, tutors did not improve significantly as a result of this intervention. Because tutors' skills were relatively strong initially, the limited improvement might reflect a ceiling effect. This emphasizes the importance of including reading measures where tutors' scores show substantial room for improvement.

Adesope et al. (2011) conducted a meta-analysis comparing peer-tutoring studies based on types of task, length and intensity of tutoring intervention, and other features of peer-assisted tutoring formats designed to address reading fluency and/or reading comprehension among students instructed in English who did not use English as their

primary language. Although their general findings held across studies on students with different first languages (i.e., other than Spanish), they are particularly relevant for students who use Spanish as their primary language. They found the largest gains in oral reading fluency when students were tutored by peers using a structured task that combined strategies from both languages or when structured, curriculum-based assessments were used to evaluate outcomes. Multimedia or computer-assisted materials were not as effective as the interaction between peers during the tutoring. In terms of the optimal time for the tutoring intervention, these authors concluded that although the most effective strategy was one of the longest in duration (about 160 hours total, or an entire year), the optimum interventions yielded similar results using a much shorter and more intense timeframe of 20-100 hours total, with the largest marginal effects found for programs that lasted no more than three months. This finding is consistent with the results for students whose primary language is English by Rohrbeck et al. (2003), where interventions implemented for a minimum of three sessions of 30 minutes per week were more effective than programs using shorter or fewer sessions implemented over a longer time frame.

These investigations into using peer tutoring highlight some advantages and concerns about effective peer tutoring interventions. Among the advantages are mixed but generally positive improvements for some students learning English as a second language, a positive appreciation (and potential benefits) for the tutors in a successful project, and potentially stronger relationships among students. At the same time, there is a greater chance of a negative result when peer tutoring has limited administration

support and when it is applied in lieu of core instruction rather than as a supplement to core instruction.

Features of Effective Peer Tutoring

These studies highlight some features that are likely to improve the impact of a peer tutoring intervention. As noted above, using a structured and consistent protocol and evaluating the fidelity of implementation has been associated with a greater positive effect on reading fluency in students learning to read English. Providing a direct link between features of the independent variable task and the dependent variable measure (e.g., practice on oral reading fluency and then measuring improvement on a standardized oral reading fluency measure versus practice on teaching a different task and then measuring improvement on a standardized oral reading fluency measure) not only makes logical sense, but it is more likely to directly affect improvement on the dependent measure (Adesope et al., 2011; L. M. Saénz et al., 2005). This means that tutors should have a structured and consistent task that is aligned with the type of dependent measure.

Curriculum based assessments are considered to be valid measures of similar skills in students who are learning English. However, the research suggests that cloze measures are likely to underestimate the performance of students learning English, while measures of fluency are less likely to show differences between students who have been instructed in English yet do not speak English as their primary language and students who speak English as their primary language. These measures have been proposed with the caveat that there may be mean differences even if they do not show statistical bias in prediction with other measures (Garcia et al., 2006). One option for using such

procedures is to evaluate the reading skills of tutors in a cross-age tutoring intervention, since these measures are still considered valid measures of reading comprehension in English among students whose primary language is not English (e.g., to avoid the potential effects of a ceiling effect on oral reading fluency, as found by S. B. Saenz, 2008).

Exclusionary criteria for tutors should include willingness to work with other students and to provide positive feedback. Relatively short (20-30 sessions, shorter than three months; optimal length is probably 12-13 weeks) but intensive (at least 3-5 sessions per week) tutoring programs are likely to yield the strongest results (Adesope et al., 2011). Although there is some limited support for cross-age versus reciprocal same-age tutoring (Scruggs & Osguthorpe, 1986), the present study proposes use of a cross-age peer tutoring intervention to isolate effects of tutoring on tutee versus tutor reading skills. Also, within the proposed setting (where all students have a range of reading skills and are instructed in multi-age classrooms), a cross-age peer tutoring program is a better fit with the needs and logistics within the school.

Extended repetitions of the same material are likely to have a minimal effect after three highly accurate readings (Ardoin, Williams, Klubnik, & McCall, 2009). Summaries and meta-analyses of the peer tutoring and peer tutoring for students learning English lead to a conclusion that while generally positive and robust to the effects of some issues, a successful peer intervention strategy should include a few key elements. Stable baseline trends for single-case evaluations and ideally, random assignment or matched comparison groups help improve the chances of finding positive results from an effective

intervention. Although in most cases, it does not seem to matter if tutoring occurs with consistent peer tutors or in a reciprocal format, when differences do arise between these options, the results lean towards support for cross-age tutors (e.g., Scruggs & Osguthorpe, 1986).

Some other features of peer tutoring outlined by these studies correspond to recommendations from Fuchs and Fuchs (2007), such as tailoring peer tutoring activities to the skill levels of tutees (e.g., reading comprehension strategies are less effective when students do not have adequate oral reading fluency, and oral reading fluency strategies are less effective when students are still developing word attack skills). As stated earlier, intensity of a peer tutoring intervention appears to facilitate positive outcomes (Ginsburg-Block, Rohrbeck and Fantuzzo, 2006), with shorter but more intense (measured in terms of time spent per session, e.g., 15 versus 30 minutes, and twice weekly sessions versus at least 3 times per week, Adesope et al., 2011; Rohrbeck et al., 2003) tutoring yielding more positive outcomes. This corresponds to results from both of the meta-analyses that included English language learners in peer tutoring efforts.

Several of these features are found in a protocol provided by Wright (2004), and evaluated in a large-scale cross-age intervention study by Wright and Cleary (2006). This study did not directly evaluate the effectiveness of the intervention on the reading fluency of students whose primary language is other than English. However, as earlier reviews have suggested (e.g., Adesope et al., 2011; August & Shanahan, 2006; Rohrbeck et al., 2003), it is reasonable to use this study as a guide for implementing the general protocol developed by Wright (2004).

Wright and Cleary (2006) implemented a cross-age tutoring protocol in a school district in the Northeast United States. Twenty-seven students (thirteen third graders and fourteen fourth graders) who were reading at or above the mid-second grade level (based on curriculum-based measures, using the protocol outlined in Wright, 2004) acted as tutors for twenty-seven tutees (fourteen second graders and thirteen third graders). Tutees were nominated by teachers as students who were not making adequate progress, but who would be expected to improve based on practice at or above the first grade level in a reading primer. Tutor-tutee matches were based on schedule compatibility and tutor reading skills at or above the level of their respective tutees. Tutoring sessions used one of two options outlined by Wright (2004): listening-while-reading (NICHD, 2000; Wright, 2004), a procedure where tutees read silently while a tutor models fluent reading, followed by the tutee practicing the same passage and receiving positive and corrective feedback from their tutor while they practice. Treatment integrity was measured by audiotaping 25% of the tutoring sessions and having independent observers code the tutors' behavior on an evaluation form. A four session tutor training protocol, and a three session site coordination training program (covering logistics, implementing the protocol, etc.) were conducted in the participating schools and with six site coordinators. Tutoring sessions were conducted for twenty minutes per session, two sessions per week, for a range of eight and a half to twenty-one weeks, for an average of nineteen weeks. This meant that treatment integrity was assessed every other week throughout the duration of the tutoring program. A relatively high level of treatment integrity (mean=90%; range=70-100%; Wright & Cleary, 2006) was reported. Although both tutors and tutees

showed some improvements in reading fluency, only the improvement for tutees was statistically significant. Students who were tutees had previously progressed in reading at a rate lower than their peers, and the tutoring program improved this to a rate of improvement equal to that of their peers. Unfortunately, the rate of reading improvement for tutors was less dramatic. This appears to be consistent to the results reported above for S. B. Saenz (2008) with students learning English who are tutored by students with stronger English proficiency. Wright and Cleary (2006) point out that one limitation of their study, relative to other peer tutoring efforts, was the lack of a measure of acceptability of the intervention. Wright and Cleary (2006) implemented the same protocol for 19 weeks, but as Adesope et al. (2011) and Rohrbeck et al. (2006) concluded, overall frequency of tutoring is likely to be more important, and in Wright and Cleary (2006), tutors did not meet for more than two sessions per week.

Tutors should be able to provide accurate and meaningful modeling and feedback for the targeted skills. Therefore, tutors should be performing at least at a level comparable to the tutees (even with a reciprocal strategy; Fuchs & Fuchs, 2005). For the proposed study, tutor feedback is a critical component of the intervention. Therefore, a major inclusionary criterion for tutors is that they should have oral reading fluency skills greater than tutees' instructional level (Wright, 2004; Wright & Cleary, 2006). Mooney (2010) and others (Pyron, 2007; Serrano, 1987) emphasize the importance of a peer tutor's relationship with the tutee as a critical component of the tutoring process. Therefore, positive, active encouragement of tutees needs to be maintained, and tutors who cannot achieve an adequate ratio of positive feedback are unlikely to motivate tutees.

This is particularly relevant within the context of the present study where there may be a tendency to emphasize the social aspects of literacy (Gordon, 1980).Tutees and tutors should only remain in the project if the tutees agree to continue and the tutors demonstrate that they are providing sufficient positive feedback. The tutees' willingness is inferred from attendance at voluntary tutoring sessions (see student assent script in Appendix A). Tutors are asked to provide positive feedback, which is measured via observations during tutoring. As part of that commitment, students in the project will need to minimize their use of languages other than English during tutoring sessions. These sessions are geared towards English proficiency as well as reading fluency practice, so incentives and other encouragement will emphasize the use of English. Most of this can be accomplished by maximizing the reading of passages during the tutoring sessions.

Ekstrand (2011) found a positive impact on reading fluency for four students in as few as three weeks of tutoring, with four fifteen minute sessions per week. The National Reading Panel's meta-analysis did not include such small samples of peer tutoring, but did indicate an average oral reading (including peer tutoring, reading to parents or teachers, oral group reading in classrooms, NICHD, 2000) of six weeks, with intervention lengths ranging from two to thirteen weeks. Adesope et al. (2011) as mentioned above, found optimal results for interventions of less than three months, with large effect sizes after only a few weeks. Rohrbeck et al. (2003) found that there was a slight but insignificant correlation between intervention duration and outcomes, but that one moderator of this relationship was the session length or frequency during the

intervention. Shorter, more intense interventions (e.g., with sessions at least 30 minutes long, with a minimum of 3 sessions per week, preferably 4) were more beneficial than lengthier, less intense interventions.

For the current study, a minimum of six weeks of tutoring is recommended for each intervention phase after baseline data are collected, or a minimum of five assessments of progress during the intervention.

Most studies of peer tutoring, whether with English proficient or English language learners have incorporated a range of designs, from quasi-experimental designs to carefully implemented alternating treatment or multiple baseline across participant designs. While a few studies also includes randomized assignment to treatment conditions (even with single case designs), attrition and other logistic concerns have compromised the rigor of these studies. The National Reading Panel reported that most studies of oral reading (as a general intervention) have fewer than a dozen participants. Both quasi-experimental and randomized control studies have demonstrated positive improvements in oral reading fluency for peer tutoring and other oral reading efforts.

The guidelines from the National Reading Panel (NICHD, 2000), the Center for Applied Linguistics/National Literacy Panel on Language-Minority Children and Youth (August & Shanahan, 2006) and others (Wright, 2004) emphasize the benefits of using direct measures of reading fluency such as curriculum-based measures during the peer tutoring intervention to measure ongoing tutee progress, with standardized assessments of related and target outcomes at the end of the study. A key feature of appropriate measures is whether scores on such assessments have similar meaning when used for students from

multiple backgrounds and cultures, as well as examining mean differences between students from different demographic, cultural, and linguistic backgrounds (Garcia, McKoon, & August, 2006). Within the larger review, Garcia et al. (2006) highlight several patterns of results for English language learners relative to English language proficient students. The least problematic measures for both monitoring progress and evaluating reading fluency are curriculum-based, or generalized passages that are similar in difficulty to passages used for tutoring students. Garcia et al. conclude that both researcher-developed and commercially available curriculum-based measures are appropriate for this purpose. For assessing reading comprehension, they conclude that in spite of consistent differences students speaking English as their primary language and students learning English as a second or third language in the level of results on so-called "cloze" tests (Garcia et al., 2006), where missing words from text passages are provided by students, such tests may provide the most appropriate index of reading comprehension skills, regardless of whether students are asked to recall words to include or choose among a set of words to complete the sentences.

Vocabulary has been hypothesized to moderate the relationship between oral reading fluency and reading comprehension (Garcia et al., 2006). According to Garcia et al., one of the most commonly used vocabulary assessment in literacy studies of Spanish-speaking English language learners is the *Peabody Picture Vocabulary Test*, a measure of receptive awareness of word meaning. Although several problems result in using this as a direct measure of language proficiency, this measure still assesses familiarity with English words, if used as an English vocabulary index only (Garcia et al., 2006). For

example, the item difficulty and item discrimination depend on a similar frequency of use within both the culture of assessment and the culture in which the items were developed. Pilot testing with an alternate form of the *PPVT-IV* within the school resulted in several students encouraging their friends to participate in testing. Although this is not a systematic evaluation of the attractiveness of this test, it provides some anecdotal support of the ease of use of the measure.

Several studies also include observations of student interactions, videotaped and coded sessions, and feedback questionnaires and interviews about the tutoring process (e.g., Johnson, 1980; Rohrbeck et al., 2003; S. B. Saenz, 2008). Adesope et al. (2011) also found that peer tutoring interventions that were implemented with stronger treatment fidelity (e.g., following a larger percentage of steps within the specific protocol for tutoring) resulted in larger effect sizes for the intervention. Therefore, monitoring intervention integrity is a critical component of this study.

Most of the content in standardized curriculum-based measures such as *DIBELS Next*© has been carefully reviewed and edited to increase relevance and connection to typical curricular materials used in other cultures, and to decrease the ambiguity of probes. While there are a few individual items on Nonsense Word Fluency that are actual words in Spanish, this is only likely to result in a slightly higher error rate (e.g., one or two points for NWF). Because a single NWF score is always interpreted in the context of repeated administrations (Good et al., 2011), this problem is likely to have a minimal impact. The content of DORF and DAZE passages is culturally specific, but written in a manner that is consistent with most basal readers, and the authors have incorporated

feedback to minimize this effect. One recommendation, based on research comparing average scores on these measures, is to consider U.S. benchmark norms provisionally until local norms can be established.

Curriculum-based assessments are advocated because they are efficient to administer, have strong reliability and validity evidence, even with non-native English speaking students (Albers, Kenyon, & Boals, 2009; Ardoin & Christ, 2009; Brice & Brice, 2009; Richardson, Hawken, & Kricher, 2012; Roberts, 2005; Yesil-Daglis, 2011; Young, 2009), however, several research studies have shown that specific subscore norms and benchmarks may be lower for students learning English in the U.S. (Brice & Brice, 2009; Grant, Gottardo, & Geva, 2012; Richardson et al., 2012; Wayman, McMaster, Saenz, & Watson, 2010; Young, 2009), and a logical extension of this research is that these benchmark goals may need to be adjusted for the present study, particularly while estimating instructional levels for the tutees. Standardized assessments are global measures of a more general skill being addressed during the intervention (e.g., for reading, this may include generalized vocabulary knowledge and reading comprehension) and provide an indication about how well skills practiced during the intervention generalize to a broader outcome goal. Because such assessments are more general, they are not as sensitive to smaller skill gains or issues and they cannot be repeatedly administered without compromising their validity and accuracy. There is also a danger with more global instruments of introducing cultural biases in items or format (with curriculum-based assessments, the more curricula and the assessments are aligned,

the less additional bias is provided by the content beyond what is given in the curriculum itself).

Specifically, both the NRP and the CAL/NLPLMCY (August & Shanahan, 2006; NICHD, 2000) recommend three types of outcome assessments: curriculum-based assessments of the target skill (collected throughout the intervention), pre and posttest measures of fluency, language skills such as word recognition, and reading comprehension, and some type of survey, questionnaire or qualitative measure of the tutors, tutees, and teachers' satisfaction with the intervention. As some of the earlier research summaries indicated, a range of standardized measures have been used with students learning English in the United States. Promising efforts are taking place with test development of language assessments in Belize.

Given that both reciprocal and cross-age peer tutoring have been demonstrated to provide a supplemental effect on oral reading fluency, especially for minority and students learning English within the United States, the question remains whether either type of peer tutoring would have a similar effect on reading fluency in the target school in Belize. This study expanded the knowledge base on cross-age peer tutoring into a novel setting, among a group of students with minimal English oral reading fluency skills. Also, although research has found less significant improvement on the oral reading fluency skills of peer tutors, another question this study was designed to address was whether it is possible that a cross-age peer tutoring program emphasizing oral reading practice may be effective at benefitting tutors' reading comprehension when their reading fluency is already at benchmark levels?

The major hypotheses for the present study were:

Hypothesis 1: Active participation in the peer tutoring program at the recommended levels of duration, frequency, and intensity will significantly improve the reading fluency of participating students, and the reading comprehension of tutors. Tutees will make the most growth in measures of reading fluency, followed by the growth of tutors and non-tutored students.

Hypothesis 2: Tutees, tutors, and teachers will perceive improved social interactions between tutees and tutors as a function of their participation in the program, possibly resulting from the positive feedback emphasis of the tutoring program.

Hypothesis 3: Consistent with the results of earlier studies, word use and oral reading fluency growth will be cumulative and sustainable after the completion of the oral reading intervention component of the proposed intervention, as determined by comparisons of benchmark results at the end of the study collected at least one week after the final progress monitoring assessment for all participants.

Rationale for Setting of Study

Belize provides an interesting opportunity to evaluate an English-based fluency intervention because it is considered a multilinguistic country (Belize Central Statistical Office, 2000; Belize Tourism Board, 2013, Hanson-Smith, 1988; Loskot, 2007) and the official language (English) is not the most commonly used language in daily life, nor is it the second language for many Belizean citizens (Belize Central Statiscal Office, 2000; Belize Ministry of Educatoin, 2007, 2008; Belize Tourism Board, 2013). This presents a more complex set of issues, due to the multiple languages (Belize Ministry of Education,

2007, 2008; Belize Tourism Board, 2013; Gordon, 1980; Hanson-Smith, 1988; Loskot, 2007; UNESCO-IBE, United Nations Educational Scientific and Cultural Organization International Bureau on Education, 2010, April). The Ministry of Education within Belize has specified that instruction is conducted in English (UNESCO-IBE, 2010, April), while recognizing that in many cases, this will need to be integrated with other languages. That is, proficiency in both English and Spanish is the goal, even though students are expected to pass English proficiency exams in particular (Ministry of Education, 2008; UNESCO-IBE, 2010, April). This is even though neither language is the prevailing societal language in Belize (Belize Tourism Board, 2013).

For example, in Belize, although the official language is English, most Belizeans speak a different language they call Kriol (Belize Tourism Board, 2013; or Creole, Belize Central Statistical Office, 2000). As the prevailing societal language in Belize, Creole may be considered a national language based on August and Shanahan's definition (August & Shanahan, 2006; p. 3). As Loskot (2007) points out, and as can be seen in several separate analyses of the most recently available census data, there are also individual societal languages within Belize, based on geographical region, community, or town (Belize Central Statistical Office, 2000).

Education within Belize is compulsory for all children between the ages of five and fourteen (UNESCO-IBE, 2010, April). A school year is 175 days; given the number of holidays within Belize, this means that students are expected to attend school from September until the end of June (UNESCO-IBE, 2010, April). Students attend classes for about five hours per day, and schools are broken into levels similar to the British system.

Although pre-school is a stated goal of the education system, most students do not attend school before age five, when they enter primary school (Belize Ministry of Education, 2007; UNESCO-IBE, 2010, April). Primary schools (similar to the United States model of Kindergarten through eighth grade) are broken into the following grades/levels: Kinders, Infants (I and II) and Standards 1-6 (corresponding to kindergarten, first and second grades, and grades 3-8 in the U.S. system; UNESCO-IBE, 2010, April). High school is referred to as Secondary, and Sixth Form corresponds to associate's level degrees in the U.S. Each transition requires examinations before continuing to the next level, corresponding to the British system on which it was based (UNESCO-IBE, 2010, April). Although the passing rate to go on to high school/secondary is very high, many students drop out before they are able to take this exam (UNESCO-IBE, 2010, April).

UNESCO-IBE reports that most schools receive some government funding, even if they are not considered typical public schools. Efforts are underway to improve the availability of early education and improve teacher certification. Curriculum for language arts within Belize incorporates many of the recommendations of the U.S. reports highlighted above (Belize Ministry of Education, 2007, 2008), emphasizing phonological awareness and phonics instruction and reading fluency as prerequisites to understanding written materials (Belize Ministry of Education, 2008). The curriculum guides also share the insight from reports such as August and Shanahan (2006) that transitioning towards literacy in a second language is facilitated by achieving a sufficient level of literacy in the student's first language (Belize Ministry of Education, 2007, 2008).

The specific research project used a protocol adapted from Wright (2004). It was a replication of the tutoring task used in Wright and Cleary (2006) and applied in a small primary school (K-8; UNESCO-IBE, 2010, April) in Belize. The setting characteristics provided below are supplemented by the statistical summaries for the three towns, villages or communities where students who attend the school live, using an interactive database from the 2000 Belize Census (Belize Central Statistical Office,). Even though this information is about twelve years old, comparing the information for the three communities served by the school shows a strikingly similar pattern to the contemporary population.

Chapter 2. Method

Setting

The study was conducted in a small primary school in a remote section of Belize, Central America. The school serves children from kindergarten through Standard VI, similar to kindergarten through grade 8 in the U.S. According to the most recent census information for the region, there was a total community population of 246 (Belize Central Statistical Office, 2000). In 2000, only two people in the area spoke any language other than English or Spanish, with the vast majority of the population speaking Spanish as a home language (238 out of 246; Belize Central Statistical Office, 2000). Therefore, it is reasonable to conclude that many of the students enrolled in the school do not speak English as their first language, and that a minority of them are likely to speak English "very well" or to have parents who speak English "very well," even though English is the official language of Belize. Belize is distinct among Central American countries because of this multi-literate and multicultural perspective (Loskot, 2007; Ministry of Education, 2007; 2008; UNESCO-IBE, 2010, April), particularly since the official and academic language of the country differs from the indigenous and commonly used languages of the surrounding areas (Gordon, 1980; Loskot, 2007). Several of the students at the school had either emigrated from other countries in Central America where Spanish was the official/predominant academic and social language, or had parents who primarily spoke and read in Spanish.

The structure of instruction at the school followed the guidelines for curriculum and instruction provided by the Belize Ministry of Education for primary students (see

Belize Ministry of Education, 2007, 2008) that focused on teacher-directed explanations and modeling of English phonetic rules and pronunciation in all grade levels (Belize Ministry of Education, 2007, 2008) and incorporated independent reading activities, where students were encouraged to read silently from a variety of books at their reading level. Although small group and individual instruction was encouraged by the Ministry, prior to implementing the peer tutoring program there was not a formal peer-tutoring intervention presented within the curricular guides, but neither were such options expressly prohibited. Within the school there was an existing "buddy reading" program, where older students read to the younger students so that the older students could practice their reading skills. Younger students were only asked to listen to the older students reading. Although this provided additional reading practice for the older student, it did not address the needs of most of the students whose oral reading proficiency was estimated to be between first and fifth grade reading levels, based on 2012-2013 benchmark assessments and teacher judgment.

Participants

Tutors. The initial tutor sample consisted of nine students from Standard III through Standard VI. Seven of the tutors were male, two were female. One tutor was in Standard III, two were in Standard IV, three were in Standard V, and three were in Standard VI.

Inclusionary and exclusionary criteria. Tutors were chosen to participate based on the following primary inclusionary criteria: (a) reading at or above Level 3 in *DIBELS Next Oral Reading Fluency (DORF)*, (b) reading accuracy of at least 96% on grade level

material, and (c) reading rate of at least 99 words per minute. The exclusionary criterion was not continuing to follow protocol after tutor training.

Wright (2004) and Wright and Cleary (2006) specified a minimum reading level of grade 2 in outlining the protocol used in this study. The principal at the school was concerned that tutors might not be able to master the peer tutor training or have sufficient vocabulary if criteria were strictly set at the second grade level. Therefore, the minimum standard for tutors was adapted to students meeting grade level benchmark levels for Standard I (corresponding to *DIBELS Next* Levels 3). This ensured that tutors who completed training were reading at least at one level above the mastery level of their tutees.

All of the recruited tutors successfully completed the tutor training and expressed enthusiasm about participating in the project.

Tutees. The tutee sample originally consisted of nine students in Standard I through Standard VI. Four of the tutees were male, and five were female. Two tutees each were in Standard I and II, and in Standard IV and V, and one was in Standard VI. None of the tutees were in Standard III.

Matched Controls. Nine students were selected to participate as control participants, or matched controls. Seven of these students were male, and two were female. One was in Standard I, two each were in Standard II and Standard III, three were in Standard IV, and one was in Standard VI. None of the controls were in Standard V.

Students needed to be able to read at least one line of a *DORF* Level 1 passage to be considered for participation as a tutee or matched control. One line of a *DORF* Level 1

passage corresponds to the minimum performance on *DORF* required for a valid score to be entered (Good et al., 2011a). Students who are not yet reading even at this level are likely to need very explicit instruction on phonics and word attack skills. Since peer tutoring is a fluency focused intervention, some researchers advocate setting a minimum standard of mastering decoding (Riley-Tillman, Burns, & Gibbons, 2013). This latter criterion corresponds to setting a minimum standard based on a student's instructional level. Tutees and matched controls were matched on the basis of test scores and instructional level relative to assigned grade level. Matched controls were chosen by random assignment from each matched pair. At least two of the matched pairs represented older students with reading skills substantially lower (e.g., 3 grade levels) than the students' assigned grades.

About two-thirds of the way through the study, one family moved out of the community. This resulted in a ripple effect across tutoring groups, because the family included a tutor for one pair and a matched control for another. The data from these pairs was excluded from the progress monitoring comparisons, yielding a final sample of 21 students in seven tutoring pairs, plus the tutees' matched controls. Analyses were based on the seven tutoring pairs with full participation.

Materials

Planning. The materials for planning and designing the implementation of this study consisted of introductory information provided in the training manual (Wright, 2004) retrieved from Intervention Central (www.interventioncentral.org, retrieved 11/20/2012). This included information about the importance of peer tutoring as a

supplement to classroom instruction, and an introduction to the importance of supporting peer tutoring within the school. Also, the roles of students, teachers, and program administrators were provided within this manual.

Recruitment. Recruitment materials consisted of parent consent letters briefly describing the project, available in both English and Spanish for parents of nominated students. Copies of these letters in both English and Spanish are provided in Appendix A. Student assent cards were distributed to every student whose parents provided consent, and the script in Appendix A was read individually to all students invited to participate. Assent cards were index cards with the students' names on one side and Yes/No on the other side. Students were asked by a bilingual teacher to provide assent, and requests were made in both Spanish and English.

Tutor Training. Tutor training materials were modified from those provided in Wright (2004) for each of the four training sessions. Each training session included an outline of the training curriculum, reminder posters used during training to illustrate points, worksheets and role-play practice sheets, tutor badges used throughout the training and tutoring program to identify tutors, and reward stickers for meeting training goals. The intervention fidelity was adapted from Wright (1994) to allow for non-visual recordings of peer tutoring sessions in the interest of preserving the privacy of tutoring pairs as much as possible. Fidelity focused on assessing whether or not tutors were able to consistently provide encouragement and support.

Tutoring Sessions. Each tutoring pair (tutor and tutee) had their own folder to use in each session. This folder was be kept in each teacher's room and passed out to tutors at

the beginning of each tutoring session. The folder contained the following materials: (a) instructions to remind the tutors to give compliments and a reminder of the listening while reading task, (b) a daily log sheet to record tutoring sessions, (c) a pencil or pen, and (d) instructions about the passages used for reading practice.

Practice passages were chosen from a website devoted to leveled passages designed for instructional and fluency practice at Read Works (www.readworks.org). The criteria used to select each passage was that the material covered in the passage was not specific to the culture or political structure of the U.S. A mixture of both narrative (fiction) and expository (nonfiction) passages were identified. A selection of passages at each grade level were chosen and school officials printed the passages. Final selection of practice reading materials was based on two criteria: (a) best fit with the interest of the tutee, and (b) fit with the tutee's instructional level. Rohrbeck et al. (2006) suggest that engaging students with high interest reading materials and providing choice is likely to facilitate improvement in reading skills. Each passage was at least one half page. Some passages included picture cues or photographs, and some passages were reused by pairs due to resource constraints and passage availability.

Assessment. Initial screening was based on benchmark assessments of student reading and pre-reading skills, using *DIBELS Next* benchmark assessment measures (Good et al., 2012). All students, including those in Standard V and Standard VI were assessed (corresponding to Grades 7 and 8 in the United States). Table 1 shows the level of *DIBELS Next* benchmark assessments each grade level completed.

Grade Level (Belize)	Grade Level (U.S.)	*DIBELS Next*© Level	Measures Included
Standard I	Third Grade	Grade 3	The set of benchmark assessments
Standard II	Fourth Grade	Grade 4	includes the same types of measures
Standard III	Fifth Grade	Grade 5	for Grade Level 3-6:
Standard IV	Sixth Grade	Grade 6	*DORF*: *WC*
Standard V	Seventh Grade	Grade 6	*DORF*: *Accuracy*
Standard VI	Eighth Grade	Grade 6	*DORF*: *Retell*

Table 1. Benchmark Assessments

Although the authors of *DIBELS Next* recommend calculating a Composite score (Good et al., 2011a) based on the indices from *DORF* combined with *DAZE*, the *DIBELS Next* cloze-formatted measure of reading comprehension, this strategy may present problems when the students are learning English. A review of literature on cloze-formatted based measures of reading comprehension with students who are learning English as a second or third language suggests that other indices (e.g., *DORF*) are likely to show a stronger match with the performance of students who speak English as their primary language. That is, the use of *DORF* without *DAZE* may be a more valid assessment. Although the ideal solution would be to develop local norms for establishing benchmarks for these measures, the current location had too few scores at each level of assessment for calculating local norms. Midyear benchmark results from 2012 showed a lower average score for all students than expected given the other *DIBELS Next* results. Researchers at the Dynamic Measurement Group, Inc. recommended using the

benchmark standards for *DORF Words Correct* and *DORF Accuracy* separately instead (personal communication; Katherine Bravo Aguayo, January 18, 2013).

Along with the assessments mentioned above, the teacher assigned to conduct progress monitoring sessions also kept a folder for the tutoring program that included the following materials: (a) a timer, (b) a progress monitoring record sheet for each student (tutors, tutees, and waitlist students) in the study (including the level of assessment to use for monitoring progress), (c) a clipboard, (d) a plastic overlay to use during assessment, (e) a dry erase marker, and (f) a pen to use for recording data on progress monitoring sheets.

The specific dependent measures for this study were: (a) *Words Correct per Minute (WCPM)*, the number of words correctly read as whole words in each passage in one minute, (b) *Accuracy*, the percentage of WCPM of total words attempted, and (c) *Retell Total*, the total number of words used by the student to recall the relevant details of the passage in one minute (Good et al., 2011b). Oral reading fluency was operationally defined as a joint examination of WCPM and Accuracy. Reading comprehension was defined by the Retell Total score.

When single passages were read, the calculation of accuracy was WCPM divided by Total Words. Inserted or added words were not counted in WCPM or Total Words, but omitted or skipped words were counted as part of the Total Words in the passage as errors. When three passages were read during the assessment, median counts are used for calculating Accuracy. That is, the median WCPM was divided by the median Total

Words and multiplied by 100 to yield the Accuracy index, even if the median Total Words and median WCPM were from different passages.

As shown in Table 1, the preliminary measures were based on the benchmark assessment battery *DIBELS Next*, published by the Dynamic Measurement Group, Inc. (Good et al., 2011a, b). This is the seventh edition of a set of measures designed to screen students on most of the prerequisite preliteracy and literacy skills identified by both the NRP and the CAL/NLPLMCY (August & Shanahan, 2006; NICHD, 2000). This set of assessments was chosen based on the availability of parallel forms across both benchmark and progress monitoring sets, enhanced specificity and difficulty levels, and both alternate forms and test-retest reliabilities. Research summarized in Good (2011b) concludes that alternate form reliability indices range from 0.68 for Retell and 0.88 for Accuracy to 0.95-0.98 for Words Correct Per Minute. Test-retest reliabilities for these same indices range from 0.57 for Retell among second graders (and 0.68 for third graders) to 0.84 or higher (Retell) for older students and reliabilities greater than 0.90 for other indices across all grade levels.

Validity of *DIBELS Next Oral Reading Fluency* has been assessed by examining the fit between the content and foundational literacy skills, along with evidence of both predictive and concurrent validity of the scores (as predictors of general achievement measures in similar domains). Each of the Indicators (the I in *DIBELS*) is linked to specific content areas from these skills. Furthermore, both the predictive and concurrent validity evidence of *DORF* scores against a variety of criteria such as the Group Reading Assessment and Diagnostic Evaluation (GRADE; Good et al., 2011b), a standard

achievement measure, have all been relatively strong predictors, ranging from 0.45 to 0.77.

Social validity survey. At the end of the tutoring program, teacher perceptions of the tutoring on student behavior and reading were assessed through a survey administered to all the current teachers in the school. A copy of the survey completed by teachers is provided in Appendix B.

Assessment Schedule

Three types of assessments were conducted with each of the students who participated in the project: (a) benchmark assessments at the beginning of the year and one week after the final progress monitoring assessment, (b) survey level assessments conducted just prior to the implementation of the tutoring program to establish instructional and progress monitoring levels for tutees and control participants, and (c) weekly assessments of progress in reading of all participants.

For most tutees and matched controls, the baseline was set with the survey assessment. For all tutors and the tutee and matched control for pair 8, baseline was established by the beginning of the year benchmark assessment. The survey level assessments and benchmark assessments were each administered in single testing sessions. Because of the observed variability across survey assessments, scores could be examined for trends, although these do not fully duplicate the characteristics of standard baseline assessments that occur over time. This constraint is the primary reason for using a matched pair design to evaluate the overall effectiveness of the intervention.

Procedures and Experimental Design

The first step in the study was recruiting and training tutors to participate. Once a cadre of tutors was recruited and trained in the tutoring process, the students eligible for being tutored were organized into a pool of eligible tutees, matched with other students in the tutee pool as described previously. Growth for all students (tutors, tutees, and monitored but not tutored students) was evaluated for improvement compared to U.S. normative growth in reading, operationally defined by benchmark goals provided by Good et al. (2011b). Repeated assessments of students were compared to expected change in each index from the beginning of the school year until the midyear benchmark.

Students were monitored for progress at the *highest* level at which they were reading with at least 90% accuracy. All students participating in the tutoring process were assessed weekly nine times, with a four week break for Winter Break. A matched control single case/repeated measures design was utilized to evaluate the effects of the cross-age peer tutoring intervention.

Teachers implementing the protocol worked with the principal at the school to go through the training and intervention protocol. The protocol for tutor training consisted of four one-hour training modules (Wright, 2004). The first module introduced peer tutoring and establishes behavioral expectations for all participants, as managed by tutors and teachers. The second module focused on teaching tutors how to provide positive feedback and support through compliments. The third module described two types of practice: "Paired Reading" and "Listening While Reading" (Wright, 2004; p. L3-1). Paired reading is simultaneous reading with error correction, with the tutee signaling when he or she

wants to read alone. Listening while reading consists of the tutor reading once while tutees follow along silently, followed by the tutor and tutee reading simultaneously, and finally by the tutee reading with error feedback.

Tutoring sessions were originally scheduled to occur 3 times/week, replacing the end of day recess or independent reading during these days. This schedule had to be adjusted based on student availability and absences, exam days and school holiday periods. The first tutoring sessions began on October 31, 2013 and continued until January 23, 2014. Between two and three tutoring sessions were conducted each week. When absences occurred, tutoring pairs made up sessions the following week. Table 2 provides a summary of the frequency of tutoring sessions by pairs throughout the study. As stated earlier and shown in Table 2, no tutoring occurred during the week ending November 22, 2013. As mentioned in an earlier section, one family moved partway through the study, which can be seen in the session log for Pair 4.

Week	Progress Assessment Date	Sessions per Tutoring Pair								
		Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Pair 7	Pair 8	Pair 9
1	11/8/2013	2	2	2	2	2	2	2	2	2
2	11/15/2013	3	3	3	3	3	3	3	2	3
3	11/22/2013	0	0	0	0	0	0	0	0	0
4	11/29/2013	3	3	3	3	3	3	3	2	3
5	12/6/2013	3	1	3	3	3	3	3	3	3
6	12/13/2013	1	2	2	2	1	2	2	2	2
7	1/10/2014	2	2	2	0	2	2	2	2	2
8	1/17/2014	2	2	2	0	2	2	2	2	2
9	1/24/2014	2	2	2	0	2	1	2	2	2

Table 2. Tutoring session frequency per week for each tutoring pair.

Table 3 summarizes the number of passage repetitions used in each tutoring session. The information for the incomplete pairs is excluded from this table. Three pairs repeated passages beyond the specified maximum of six repeated readings in a tutoring session. Although Pairs 2 and 9 only went over six repetitions once or twice, Pair 5 went over this maximum five times. The explanation provided by the school staff was that this occurred at Tutee 5's request because she wanted to master the passage before moving to another passage. Tutor 5 was instructed to break down the passages into smaller units (e.g., sentences) to facilitate Tutee 5's mastery of the passages. Also, some passages were repeated in later tutoring sessions, due to availability of passages that could be printed at the school.

Tutor Pair	Number of Sessions	Passage Repetition per Session				
		Mean	Standard Deviation	Median	Minimum	Maximum
1	18	4.61	0.70	4.5	4	6
2	17	4.24	1.30	4.0	3	7
5	18	5.06	1.92	4.0	3	9
6	18	4.00	0.77	4.0	3	6
7	17	4.00	0.71	4.0	3	6
8	17	4.13	0.50	4.0	3	5
9	19	3.63	1.12	3.0	3	7

Table 3. Descriptive statistics for passage repetition per tutoring session by tutoring pair.

After the end of the tutoring program, midyear (MOY) benchmark data were collected for all participants in the study. Novel passages from the unused progress monitoring passages replaced midyear benchmark assessments for students in Standard V and Standard VI who were assessed with Level 6 benchmark assessments the previous year.

Fidelity Assessment Results

One tutoring session per pair was audio recorded by an advanced doctoral student assigned to the school. These recordings were used to answer three questions about the interaction between the students: (a) did the tutor provide more positive than negative feedback (e.g., praise, inserting correct options versus statements such as "no" or "that's wrong)? Neutral statements (including prompting with correct responses) counted as positive under this definition. (b) Did the tutor read for a brief period of time before the tutee began reading? In this case, "brief" was defined as less than half of the tutoring session. Finally, (c) Did the tutor provide a correct option when no response was made within three seconds?

These criteria were adapted from guidelines in Wright (1994). Specifically, the ratio of five positive statements for every negative statement was removed, since precise counts of positive to negative comments could not be made (e.g., no negative corrective feedback was observed). Prompts or insertions (providing a correct option) were included as types of positive comments, because they replaced the option of telling a student that a word was incorrect.

Two raters listened to the audio recordings and answered these questions independently. The `doctoral student who recorded the sessions served as one of the two raters. The second rater was the author of the study. This resulted in 21 judgments across the seven pairs of students with complete participation. One set of ratings indicated that all three of the fidelity criteria were met across the seven pairs. The other researcher gave a single "no" rating to one session about whether or not the tutor provided positive

feedback versus strictly neutral feedback. As a result, agreement between the raters was over 95%, calculated using the following formula:

$$[(agreements)/(agreements + disagreements)] \times 100.$$

Chapter 3. Results

Overview

This section is organized by the timing or type of assessment sequence (e.g., benchmark assessments, survey level and progress monitoring assessments) within each perspective by pair or grade level, whichever was more appropriate for the data. For example, as stated in Chapter 2, benchmark assessments refer to the screening assessments that are used to evaluate global progress of all students across a school, and are usually used within a grade level.

Since progress monitoring scores are calculated based on single passages, when calculating baseline assessments the assessment used to establish baseline (either benchmark or survey assessment) was scored individually by passage instead of using the standard median scoring method outlined in Good et al. (2011b).

When examining trends within the baseline measures for tutees and matched controls, individual assessments are presented graphically. Evaluation of the progress monitoring data were based on changes from the baseline scores to the final progress monitoring score. All figures are in Appendices C through E.

Fluency

As stated earlier, reading fluency was operationally defined by both *DIBELS Next Oral Reading Fluency Words Correct per Minute (WCPM)* and *DIBELS Oral Reading Fluency Accuracy (Accuracy)*, which were examined as variables representing different aspects of fluency. Changes in fluency were evaluated by looking at progress monitoring charts relative to expectations and goals, per the guidelines in both the *DIBELS Next*

Assessment Manual and the *DIBELS Next Survey Manual* (Powell-Smith, Good, Kaminski, & Wallin, 2012). *Accuracy* results for each pair are presented first, based on a model of reading proficiency and learning outlined by Daly, Lentz, and Boyer (1996). In this model, accuracy is a component of both skill acquisition and a prerequisite skill for fluency, followed by the rate components of fluency that are operationally defined within the present study as *WCPM*.

Progress Monitoring Results

Tutees/Matched Controls. For each of the pairs, the progress of the tutee and the matched control's *Accuracy* and *Words Correct Per Minute* (*WCPM*) were compared to the baseline trends through percentage of non-overlapping data points (PND; Scruggs & Mastropieri, 2012). Percentage of non-overlapping data points (PND) is calculated by counting the number of scores during peer tutoring that exceeded the baseline maximum, and dividing that total by the number of assessments during the intervention. This proportion is converted into a percentage by multiplying by 100. As an estimate of effect size, the resulting percentage is compared to guidelines where 90% or higher is considered a Highly Effective intervention, 70-90% is considered Moderately Effective, 50-70% is Minimally Effective, and any intervention with less than 50% PND is judged to be Ineffective.

Practical significance of progress during tutoring was evaluated by comparing the median of each student's final three scores to a goal set for the student. Goal setting was based on two factors: whether or not a student was being monitored at grade level (or, in the case of Standard V and Standard VI students, whether or not the student was being

monitored using Level 6 materials), and the initial performance of the student. If a student was monitored at grade level, the goal was set at the midyear benchmark goal for the student's grade level. If the student was monitored below grade level, the goal was set at the end of year benchmark goal for the progress monitoring level.

Graphical representations of the individual data points are presented in Appendix C for each tutee-matched control pair for both Accuracy and WCPM that show more detail about how each pair of students responded on a weekly basis. Descriptions of these graphs, in the context of other indices, are provided for each of the seven pairs with complete participation for tutees, matched controls, and tutors. These seven pairs were Pair 1, Pair 2, and Pairs 5 through 9. Based on the survey assessment results, all students' reading progress was monitored at either the highest level where the student was reading with at least 90% accuracy, or one level below if the student was well below average on either Accuracy or WCPM, based on recommendations by .

Pair 1. Table 4 shows that both the tutee and the matched control in this pair had a Accuracy greater than the goal of 93% during baseline, and that both students' baseline median was just under the standard goal for WCPM for their shared instructional level. As a result, it was not difficult for the students to increase their median WCPM, because they had generally mastered the decoding requirements.

Statistic	Phase	Accuracy		WCPM	
		Tutee	Matched Control	Tutee	Matched Control
Max	Baseline	97.7%	97.8%	51	45
Mean	Baseline	94.5%	96.3%	44.67	43.67
	Intervention	92.3%	97.0%	52.00	50.11
Median	Baseline	93.0%	95.7%	43	44
	Intervention	94.8%	96.5%	52.5	49
	Final 3	96.6%	97.9%	54	55
Standard Deviation	Baseline	0.028	0.013	5.686	1.528
	Intervention	0.089	0.021	5.952	8.238
Which goal was met?		Standard	Standard	Standard	Standard
PND: Percent of Non-Overlapping Data		13%	44%	50%	67%
Standard Goal		93%		47	
Challenge Goal		98%		72	

Table 4. Pair 1 tutee and matched control Accuracy and WCPM progress.

In spite of meeting the practical consideration of achieving a specific reading goal, the direct effect of peer tutoring on Tutee 1's Indicator scores (Accuracy and WCPM) indicate a negligible effect based on guidelines for interpreting percentage of non-overlapping data (PND). A PND index of less than 70% is generally considered to reflect a minimal effect, and one of less than 50% indicates that an intervention is not considered effective.

The graphs of Accuracy and WCPM provided in Figures 1 and 2 in Appendix C illustrate the information provided in Table 4. Figure 1 also illustrates the variability that occurred in Tutee 1's performance during tutoring. Although Matched Control 1's initial reading accuracy was comparable to performance during the survey assessment, Tutee 1's accuracy in the first week was substantially lower than her median accuracy during the survey assessment. The data presented in Table 4 ignores the effect of the single

measure during week 1 of peer tutoring. After the second week of tutoring, Tutee 1 was able to recover and became more consistent in how accurately she read each assessment passage, even though her overall variability during tutoring was affected by her initially low accuracy.

Figure 2 shows the results for this pair of students on WCPM. The pattern of results in Figure 2 shows that Tutee 1 and Matched Control 1 had very similar results overall during the intervention, even though the control participant did not receive any tutoring. The median survey level performance of both students was only four words less than the standard goal, as stated above. Both students met the standard goal by the end of the intervention. They also showed a socially significant level of growth through the project, with a gain of two words per school week for Tutee 1 and just under two words per week gain for the Matched Control 1. Because both students improved at a similar rate, it is difficult to determine the actual impact of tutoring.

Pair 2. Pair 2 consisted of two students with similar patterns of Accuracy and WCPM who were monitored at different absolute levels, even though they were at the same assigned grade level. Although both students had survey level scores that established mastery at Level 1, Tutee 2's WCPM performance on Level 2 material was Well Below Benchmark levels. According to the guidelines, he was to be monitored using Level 1 progress monitoring materials. On Level 1 materials, his survey level WCPM was comparable to that of Matched Control 2. However, his Accuracy baseline was higher than Matched Control 2's Accuracy baseline. Both students were equally Accurate at the end of the intervention, and met the standard goals for growth in

Accuracy. Tutee 2 made substantially greater improvement in WCPM over the nine weeks of the intervention. Matched Control 2 also had more variable performance during progress monitoring.

Statistic	Phase	Accuracy		WCPM	
		Tutee	**Control**	**Tutee**	**Control**
Max	Baseline	97.4%	97.3%	76	71
Mean	Baseline	97.1%	92.3%	67.67	59.67
	Intervention	97.7%	93.9%	84.00	57.00
Median	Baseline	97.0%	94.4%	65	68
	Intervention	98.9%	94.9%	87	55
	Final 3	97.4%	97.4%	87	56
Standard Deviation	Baseline	0.003	0.064	7.371	17.098
	Intervention	0.028	0.034	11.619	17.875
Which goal was met?		Standard	Standard	Standard	Neither
PND: Percent of Non-Overlapping Data		67%	22%	78%	22%
Standard Goal		97%		87	
Challenge Goal		99%		100	

Table 5. Pair 2 tutee and matched control Accuracy and WCPM progress.

Figure 3 shows the variability of Accuracy rates across assessments. As shown in Table 5 and discussed above, Matched Control 2's performance was less consistent during the tutoring. Both students started with median scores that were relatively close to the goals for Accuracy.

Figure 4 shows that the survey assessment WCPM results were very similar for these two students, even though these results were based on passages one level apart. The score and performance paths of these two students diverged during tutoring, with Tutee 2 continuing to improve while Matched Control 2's growth was more modest overall. Tutee 2 also met the standard goal for WCPM, shown by comparing the median of his

final three assessments with the standard goal. The standard goal was set based on the benchmark for Tutee 2's instructional level, Level 1. Tutee 2 gained an average of three words per week in his WCPM score, or a total of 27 words over nine weeks. Matched Control 2 gained an average of about two words per week, for a total of 21 words over nine weeks. Although these results suggest that participation in peer tutoring resulted in the differential gain experienced by Tutee 2, these results may also reflect that practice at an easier level may have affected growth for these students.

Pair 5. Table 6 shows several interesting trends for this pair of students. Both students' improvement in Accuracy was considered moderately high, based on a PND index of 89%. Matched Control 5 also has a moderate effect on WCPM, while Tutee 5 has a negligible improvement on WCPM.

Figure 5 shows the growth in Accuracy. Although this growth appears small overall, because there is a relatively small variability in Accuracy, it is magnified relative to the other dependent measures. Figure 6 illustrates some of the patterns that may be reflected in the minimal PND result. Specifically, as Tutee 5's Accuracy improved, she appeared to read more slowly and carefully. This is likely to increase her Accuracy in a trade-off with reading rate. Both students showed a gain in median performance of about three percentage points.

Statistic	Phase	Accuracy		WCPM	
		Tutee	**Control**	**Tutee**	**Control**
Max	Baseline	94.2%	93.0%	74	66
Mean	Baseline	91.9%	92.2%	70.33	59.00
	Intervention	96.5%	96.0%	71.67	71.56
Median	Baseline	92.5%	92.6%	72	61
	Intervention	96.1%	96.5%	72	71
	Final 3	96.1%	97.3%	65	72
Standard Deviation	Baseline	0.027	0.010	4.726	8.185
	Intervention	0.018	0.018	12.884	7.683
Which goal was met?		Neither	Neither	Neither	Neither
PND: Percent of Non-Overlapping Data		89%	89%	22%	78%
Standard Goal		98%		120	
Challenge Goal		100%		130	

Table 6. Pair 5 tutee and matched control Accuracy and WCPM progress.

In spite of the Accuracy gains, this pair of students still did not meet the standard goals for their instructional level, although they are both getting closer. Figure 6 shows that after an initially promising start, neither of these two students continued to improve their reading pace after the third week of the tutoring project. A difference was noted on tutoring logs for the tutoring team for this group. Directions for the tutoring specified only repeating passages between three and six times, but the log entries for group 5 showed more than six repetitions of the training passages. When asked, the principal stated that the extra practice was at the request of Tutee 5, who wanted to master each passage before going to the next one. A decision to adapt the passage practice to subsets of the instructional passages before moving to the next passage may have inhibited Tutee 5's growth in WCPM, although she continued to make progress in accuracy. Tutee 5 had a net gain of 5 words over the nine weeks of tutoring, while the control participant's net

gain was one word over nine weeks. For Tutee 5, this is somewhat lower than the expected rate of growth compared to students in the U.S. who are in sixth grade (the highest grade equivalent norm group). However, Tutee 5 is fifteen years old, and it is well-established that typical increases in words gained per week decline with student age.

Pair 6. As Table 7 shows, both Tutee 6 and Matched Control 6 had substantial growth in Accuracy, based on the PND results (PND = 100%). This reflects that once the students began to improve their Accuracy, they were able to sustain this improvement throughout the peer tutoring. Tutee 6 also was able to make modest gains in his rate of reading, measured by the PND of 63%. In Figure 7, the gains in Accuracy were steady for both students. Figure 8 illustrates that most of Tutee 6's gains in WCPM occurred in the first three weeks of the intervention and declined to baseline levels after weeks four and five. The pattern of results suggests that for these students, Accuracy gains could not be accounted for solely by peer tutoring because both students showed improvement in Accuracy. Comparing Figure 7 and 8 shows that Accuracy improvement often accompanied a drop in WCPM, so the lag in fluency could be due to a renewed focus on accuracy.

Statistic	Phase	Accuracy		WCPM	
		Tutee	**Control**	**Tutee**	**Control**
Max	Baseline	93.1%	95.6%	67	87
Mean	Baseline	91.0%	94.1%	60.00	77.00
	Intervention	96.5%	98.5%	72.50	78.56
Median	Baseline	91.8%	95.1%	59	77
	Intervention	96.0%	98.5%	71	80
	Final 3	98.6%	98.5%	71	72
Standard Deviation	Baseline	0.026	0.021	6.557	10.000
	Intervention	0.020	0.013	10.941	7.143
Which goal was met?		Standard	Standard	Neither	Neither
PND: Percent of Non-Overlapping Data		100%	100%	63%	11%
Standard Goal		98%		120	
Challenge Goal		100%		130	

Table 7. Pair 6 tutee and matched control Accuracy and WCPM progress.

Pair 7. The results shown in Table 8 suggest differential improvement for Tutee 7 in both Accuracy and WCPM. Tutee 7 started out with slightly lower reading fluency compared to Matched Control 7, but by the end of the peer tutoring intervention, she had caught up with or overtaken Matched Control 7's earlier lead. Her gains in Accuracy show 100% PND, which is a substantial effect, particularly when compared to Matched Control 7's result of 56%, which is considered a small to nonexistent effect. Her improvement in WCPM, particularly compared to the matched control's weaker result. However, the median of her final three WCPM scores is lower than the median during the entire intervention.

Statistic	Phase	Accuracy		WCPM	
		Tutee	**Control**	**Tutee**	**Control**
Max	Baseline	95.8%	97.7%	70	86
Mean	Baseline	95.0%	97.3%	63.67	83.33
	Intervention	99.1%	98.0%	87.44	85.00
Median	Baseline	94.6%	97.6%	68	84
	Intervention	98.9%	98.7%	88	83
	Final 3	98.6%	98.7%	78	78
Standard Deviation	Baseline	0.007	0.006	9.292	3.055
	Intervention	0.008	0.019	11.001	10.296
Which goal was met?		Standard	Standard	Neither	Neither
PND: Percent of Non-Overlapping Data		100%	56%	89%	33%
Standard Goal		97%		103	
Challenge Goal		100%		115	

Table 8. Pair 7 tutee and matched control Accuracy and WCPM progress.

Figures 9 and 10 shed some light on the growth trajectory for these students. Figure 9 shows that Accuracy improved steadily throughout the intervention. Figure 10 shows that both students were steadily meeting a very challenging goal of reading improvement until after Winter Break, when both students' earlier improvement dropped. The growth after that drop in scores seemed to progress more slowly than earlier growth in reading.

Tutoring provided some supplemental benefit for Tutee 7 that allowed her to improve her reading fluency, but this benefit dropped after Winter Break. Both students had some skill regression over Winter Break, but Tutee 7 recovered by the end of tutoring. Note that Tutee's WCPM just before Winter Break was two words per minute less than the standard goal at the end of tutoring. Although Tutee 7 did not meet benchmark goals in reading based on assigned grade level, she had met mastery goals for

material at the previous grade level and was being tutored to catch up with benchmark goals for her assigned grade. The control participant in this pair was being monitored for progress at the same grade level, even though she was assigned to a higher grade level (Standard III or Grade 5) for core instruction, thus reflecting that she was two grade levels away from her instructional level.

Tutee 7's maximum WCPM performance during tutoring was encouraging and represents a growth of 20 words per minute over five weeks, or about four words a week. This is considerably greater than expected growth in reading fluency than expected in similarly aged peers in the U.S. (Riley-Tillman, Burns, & Gibbons, 2013). Even considering her net growth of 18 WCPM over the entire program, or a rate of two words per week of school is impressive, considering that average growth in the U.S. among sixth grade students is less than two words a week. The net effect for the matched control has her ending up almost where she began before tutoring started. This suggests that the effect of practicing the reading with her tutor provided some net benefit to Tutee 7.

Pair 8. This pair of students was the only one from Standard VI. Since they were in the highest grade level within the school, Tutee 8 had to be tutored by one of his classmates, who also was the fastest reader in their class. This particular tutor read extremely quickly, but also was very conscientious about providing consistent correct feedback. Tutee 8 made progress in Accuracy, even meeting a challenge goal of a median of 100% correct during the last three assessments. Matched Control 8 also improved his Accuracy, although he did not reach the levels of Tutee 8.

Statistic	Phase	Accuracy		WCPM	
		Tutee	**Control**	**Tutee**	**Control**
Max	Baseline	97.7%	97.8%	85	108
Mean	Baseline	97.7%	96.4%	84.00	92.67
	Intervention	98.4%	95.8%	79.22	79.00
Median	Baseline	97.7%	97.3%	84	89
	Intervention	98.9%	96.3%	80	78
	Final 3	100.0%	97.0%	80	96
Standard Deviation	Baseline	0.000	0.020	1.000	13.868
	Intervention	0.022	0.029	7.949	12.349
Which goal was met?		Challenge	Standard	Neither	Neither
PND: Percent of Non-Overlapping Data		0%	0%	0%	0%
Standard Goal		97%		109	
Challenge Goal		100%		120	

Table 9. Pair 8 tutee and matched control Accuracy and WCPM progress.

The trends in both Figures 11 and 12 show that Matched Control 8 had a more uneven and variable pattern throughout the program, in spite of reading more quickly than Tutee 8. This pattern was established during baseline, with a decreasing accuracy level across the three passages read in succession during the same session. Tutee 8 was more consistent and accurate. The primary reason they were not chosen as tutors was because Tutee 8 read significantly more slowly than benchmark goals, while the control participant was a substantially less accurate reader.

Pair 9. The results in Table 10 show equivalent patterns of improvement between the two students in both Accuracy and WCPM, although Matched Control 9 had substantially more variable performance on WCPM throughout the study. He had both the lowest score and the highest score during the intervention phase of the project. Both students had negligible improvement in Accuracy based on the PND results yet met the

Standard goal for Accuracy. They both showed moderate improvement overall in WCPM yet did not meet the goal for WCPM.

Statistic	Phase	Accuracy		WCPM	
		Tutee	Control	Tutee	Control
Max	Baseline	97.6%	97.6%	87	82
Mean	Baseline	94.9%	92.7%	84.33	72.33
	Intervention	96.5%	96.8%	95.00	91.78
Median	Baseline	95.5%	93.9%	85	77
	Intervention	98.0%	97.7%	97	93
	Final 3	98.0%	97.7%	97	93
Standard Deviation	Baseline	0.031	0.056	3.055	12.662
	Intervention	0.043	0.029	9.605	23.947
Which goal was met?		Standard	Standard	Neither	Neither
PND: Percent of Non-Overlapping Data		56%	56%	78%	78%
Standard Goal		97%		109	
Challenge Goal		100%		120	

Table 10. Pair 9 tutee and matched control Accuracy and WCPM progress.

Figures 13 and 14 illustrate some of the reasons why the results for Pair 9 had such paradoxes. A ceiling effect and high Accuracy score during the survey level assessments/baseline meant that there was only so far the students could improve beyond the baseline level, as shown in Figure 13. Figure 14 illustrates the difference between Tutee 9's steady and consistent growth and how this student's progress was just under the goal line until after Winter Break when her scores declined and then remained steady.

Based on the percentage of nonoverlapping data (PND) index across pairs of students, five out of fourteen (three out of seven tutees and two out of seven matched controls) had at least a moderate level of improvement in Accuracy. Six of the seven tutees and five out of the seven matched control participants met the standard goal for

Accuracy, based on the median score for the last three assessments during progress monitoring. One tutee met the challenge goal for Accuracy based on a median score of 100% during his last three assessments, while none of the matched controls met this level for Accuracy. Improvements in Accuracy occurred for tutees and matched controls in differe

nt pairs and monitored with passages across all levels of difficulty. A more consistent pattern emerged when baseline levels of accuracy were relatively low compared to goals or progress monitoring performance, indicating a possible ceiling effect arising from monitoring students below assigned grade level when grade level accuracy was less than 90% (measured through beginning of year benchmark assessments). This explanation has additional support from the large percentage of students (both tutees and matched controls) who met accelerated Accuracy goals (end of year versus midyear) for their progress monitoring level, yet did not meet these goals at the beginning of the intervention.

For WCPM, the results also show that a total of five students (three tutees and two matched controls) had at least moderate improvements over baseline maximums based on the PND index. A comparison of the median WCPM score on the last three assessments to the goals for tutees and matched controls shows that none of the students met the challenge goal for their level. Three students (two tutees and one matched control) met the standard (accelerated) goal for their progress monitoring level. All three of these students were monitored at the lowest level (Level 1) available for *DIBELS Next Oral Reading Fluency* (*DORF*).

The pattern of findings suggests that most of the improvement occurred for Accuracy and that while there were some instances of tutees improving somewhat more than their matched controls, there appear to be more consistencies based on pair and instructional level than based on tutee versus matched pair treatment effects. Visual inspection of the performance of control participants and tutees also shows a large amount of overlap between most tutees and matched controls, which means that even if statistically reliable, these results provide very limited support for a conclusion that the peer tutoring program had an incremental and meaningful effect on reading fluency of tutees. Tutee 7 was the only student remaining in the study whose improvement over baseline maximums (in spite of a post-Winter Break decline) showed the expected pattern of improvement compared to the progress of her matched control.

Tutors. Progress monitoring of the tutors' performance during the intervention was focused on both fluency (as defined above) and reading comprehension. Reading comprehension, as outlined in Chapter 2, was defined as both amount and quality of what a student could describe about the story he or she just read. Specifically, it was operationally defined as a combination of Retell Fluency (or Retell Total Score) and Retell Quality.

Accuracy. Tutors' Accuracy scores were compared to both beginning of year and midyear benchmark standards. Since tutors were chosen based on meeting beginning of year benchmark goals, it was not surprising that they all met these standards initially. Out of 63 possible assessments during peer tutoring (7 tutors × 9 weeks of assessments = 63), only 2 Accuracy scores were below 97%, and both occurred for the same tutor (Tutor 6).

This constitutes only about 3% of the total number of scores. Both scores were just under 97% (95.9% and 96.2%). The median Accuracy score for Tutor 6 was 98.3%, which is greater than the goal of 97%. As a result, the progress in tutor Accuracy scores was not analyzed.

Retell Quality. Retell Quality also was excluded from analyses due to relatively small variability in the ratings provided during assessments. Retell Quality is a four point rating based on the number of distinct details recalled from a story, and the way in which these details are organized during the retelling (1=fewer than two details, 2=two or more details, unordered, 3=two or more details, in a meaningful sequence, and 4=all of the features of a rating of 3 with the addition of a main idea). None of the tutors received a rating of 4 during progress monitoring. Table 11 shows the number of ratings received across tutors. As Table 11 indicates, only one tutor (Tutor 6) received any ratings of 1, and two of the tutors (Tutor 6 and Tutor 9) did not receive any ratings higher than 2. Tutor 9 had no variability in the ratings she received. Due to the limited variability of this index, it was excluded from further analysis.

Retell Quality Ratings	**Tutor**						
	T1	**T2**	**T5**	**T6**	**T7**	**T8**	**T9**
1	0	0	0	6	0	0	0
2	6	7	2	3	4	8	2
3	3	2	3	0	5	1	0

Table 11. Frequency of Retell Quality scores received by tutors during peer tutoring.

As a result, there were two indices of interest remaining for tutors: WCPM and Retell Fluency. Table 12 provides a summary of the results of tutors' WCPM scores, which served as the primary Indicator of changes in fluency during the peer tutoring

intervention. As the data summarized in Table 12 show, all of the tutors met the midyear benchmark goals for WCPM, indicated by comparing the median of the last three progress monitoring scores to the benchmark goal. Tutor 1 was the only tutor in a grade below Standard IV, and the Level 5 goals for WCPM are higher than the Level 6 goals, due to the increased difficulty of Level 6 passages. In spite of meeting the midyear goals, not all of the tutors showed improvement in WCPM during the tutoring. Specifically, only Tutor 1, Tutor 2, Tutor 5, and Tutor 8 had a noticeable improvement in their WCPM score during the tutoring intervention. Tutor 1, Tutor 5, and Tutor 8 increased their WCPM by 31 words (Tutor 1), 23 words (Tutor 5), and 30 words (Tutor 8) over the nine week period. This averages to over 2.5 words per week for Tutor 5 and over 3 words per week for Tutors 1 and 8. This is much greater progress than expected for students their age (Shapiro, 2008).

	Tutor						
	T1	T2	T5	T6	T7	T8	T9
Baseline Max	107	124	136	109	117	113	131
Baseline Mean	114.67	111.00	129.33	97.67	110.33	114.33	119.33
Intervention Mean	135.67	130.11	152.33	90.44	112.33	144.33	120.89
Baseline Median	115	106	126	99	114	113	117
Intervention Median	139	132	153	97	109	140	125
Median of Final 3	139	139	154	114	120	163	132
Midyear Benchmark Goal	120	109	109	109	109	109	109
PND	100%	78%	100%	22%	33%	100%	22%
Effect Size	High	Mod	High	None	None	High	None
Did student meet goal?	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 12. Tutors' progress in words correctly read during peer tutoring

Table 13 shows the tutors' progress in how many words they used in recalling the story they just read in each passage used during the progress monitoring assessments. As the data summarized in Table 13 show, all of the tutors except Tutor 6 were able to meet the midyear benchmark goal for Retell Total Score (or Retell Fluency). However, Tutor 6 made substantial gains in his Retell Total when compared to his performance during the baseline assessments. His baseline scores on this measure were extremely low. After being a tutor for nine weeks, his score was similar to the score Tutor 9 had at the beginning of the project. In contrast to Tutor 8's progress in WCPM, his Retell Total score was not noticeably higher than his baseline maximum, even though he was able to use eleven more words when he described the story during the end of the progress monitoring assessment.

	Tutor						
	T1	T2	T5	T6	T7	T8	T9
Baseline Max	28	33	47	2	30	37	15
Baseline Mean	39.67	34.00	47.67	1.00	31.67	33.33	15.67
Intervention Mean	42.67	41.11	54.56	12.56	41.78	36.89	33.89
Baseline Median	41	33	47	1	30	37	15
Intervention Median	47	37	55	13	39	35	34
Median of Final 3	47	37	60	13	39	48	34
Midyear Benchmark Goal	36	29	29	29	29	29	29
PND	89%	78%	89%	100%	100%	44%	89%
Effect Size	Mod	Mod	Mod	High	High	None	Mod
Did student meet goal?	Yes	Yes	Yes	No	Yes	Yes	Yes

Table 13. Tutors' progress in retell total score during peer tutoring

Almost all of the tutors were able to meet midyear benchmark goals for Accuracy, WCPM, and Retell Total score, shown by their progress monitoring assessment results. Since Accuracy was used as an inclusionary criterion for being a tutor, there was not sufficient variability in these scores to find any improvement during the intervention. Four out of seven tutors were able to show visible improvement in their WCPM after serving as tutors, and three of these students were able to increase their oral reading fluency a substantial amount when compared to their baseline scores. Similarly, almost all of the students were able to improve their ability to recall information from a novel passage after a single reading once they had participated as tutors during the project.

Benchmark Results

Tutees/Matched Controls. The tutees' and control participants' reading fluency was also assessed via benchmark assessments at the beginning of the year and between one and three weeks after the completion of the progress monitoring assessments. All students' benchmark assessments were conducted at a level corresponding to their assigned grade level, not their estimated instructional level. As a result, the benchmark assessments provided two opportunities to evaluate students' progress in reading skills. Since the benchmark assessments included three passages rather than the single passages used during progress monitoring, they are less subject to random fluctuations and have been shown to be more reliable (Good et al., 2011b). Also, assessing the students' skills relative to the grade/level benchmark assessments provides a direct assessment of whether or not growth in reading skills exceeded typical growth expectations compared to other students at similar age/grade levels, as recommended by the researchers who

developed *DIBELS Next* (Joshua Wallin, personal communication, March 26, 2014). No appropriate benchmark norms are available for students in the highest grades (Standard V/Grade 7 and Standard VI/Grade 8), therefore the results for these students are presented in a separate chart from the results for students in Standard IV/Grade 6. As a general rule, growth in reading fluency and reading comprehension is greatest when students are younger, with decreasing effects of reading interventions as students age (Shapiro, 2008).

The benchmark results for tutees and control participants are presented in Figures 15 to 26. The results are presented by student grade level. Since benchmark passages and goals are not available for seventh and eighth grade (Standards V and VI), students in these grades were compared to benchmark goals for grade 6.

Level 3. All three of the students who participated showed movement into a different benchmark category for Accuracy (Figure 15). This means that all three students' growth in Accuracy exceeded typical growth expectations compared to students their age/grade in a U.S.-based normative sample.

Figure 16 shows the benchmark assessment results and growth for students in Standard I on grade-level passages. In this grade, Tutee 2's pace of reading improved more than norm-based expectations, since he was able to move into the Below Benchmark (versus Well Below Benchmark) level by midyear. If this student sustained this rate of improvement, he is likely to reach the end of year benchmark targets. This would be the equivalent of making more than one grade level's worth of skill growth in a particular year of school. Tutee 1's rate of improvement was also faster than typical, but

her likelihood of reducing her risk of not meeting the end of year benchmark goal was still low. Matched Control 1's rate of progress was unlikely to improve her risk of meeting future benchmarks without additional supports or intensive instructional interventions.

Level 4. Figure 17 shows that Tutee 7, but not Matched Control 1 moved to a different benchmark category on Accuracy. Tutee 7's growth was closer to a typical pattern, since her beginning Accuracy benchmark score was close to meeting the benchmark goal (95.8% versus a beginning benchmark goal of 96%). However, her Accuracy score moved from a level where she was still at risk for not meeting the end of year Accuracy goal to a level where her reading Accuracy was well-established. Matched Control 1's Accuracy score at the midyear benchmark assessment was just above the cutoff for Well Below (Intensive Risk) based on a beginning benchmark goal, but he did not meet that goal until the midyear. If his Accuracy continues to improve at the same rate, his skills are likely to exceed expectations.

The rate of improvement for Tutee 7 on WCPM (Figure 18) indicates that she increased her likelihood of meeting end of year benchmark goals. Tutee 7 will have to increase her progress to meet these goals, even though she increased her reading pace to move into the Below Benchmark (Strategic) status.

Level 5. Figure 19 shows that Matched Control 7 was able to improve her Accuracy to the point where, by the midyear benchmark assessment, she was reading without making any errors. Her growth in Accuracy exceeded expectations for typical growth. Figure 20 shows that Matched Control 7 also improved her reading pace between

the beginning and midyear benchmark assessments. Her rate of improvement was greater than typically expected, as shown by the slope of the connection between her benchmark assessments relative to the slope of the benchmark goal differences. However, even with such an accelerated path, her WCPM on the midyear benchmark is still within the Well Below Benchmark/Intensive Support range.

Level 6. Almost all of the students in Standards IV through VI (Grades 6-8) improved their Accuracy more than expected when compared to sixth grade students of similar skill levels within the U.S. (Figures 21 to 23). All of the students included here had enough growth in Accuracy to move into a different benchmark category. Tutee 8's beginning of year benchmark assessment was very close to being At or Above benchmark, but his improvement to nearly perfect accuracy by midyear is still greater than typical. Tutee 5's improvement in Accuracy since the beginning of the year was very dramatic.

Figures 24 to 26 show how the oldest tutees and control participants WCPM benchmark scores changed from the beginning to midyear assessments. The first point is that expected growth rates at these grade levels are much lower than for younger students. This means that a rate of growth that seems somewhat low when compared to a younger student at the same instructional level may actually be better than expected when compared to grade-level improvement. Figure 24 shows how all of the sixth graders' WCPM except for Matched Control 6's were improving at a faster rate than expected. The WCPM rates for the students in Pair 9 (including Matched Control 9) moved from Well Below Benchmark to Below Benchmark status, thus increasing their chances for

reading At or Above Benchmark levels by the end of the year. During this time period, Tutee 9 improved her median WCPM more than Matched Control 9, even though their progress monitoring results did not show such a difference. The seventh graders' results suggest a faster rate of improvement than expected for both Tutee 5 and Tutee 6. However, both of these students' WCPM are still likely to be within the Well Below Benchmark range unless a more intensive level of instructional support is added. The eighth grade students' rates of improvement in WCPM illustrate some of the grade-related slowing down of WCPM growth. Matched Control 8 showed some improvement in his WCPM scores, but these were very similar to the expected growth rate. His performance also was less consistent during the progress monitoring assessments, which limits the conclusion that can logically be drawn from his midyear benchmark results. Tutee 8, a more consistently performing student, had a rate of improvement in WCPM scores that was better than expected. Since his overall performance was more consistent until Winter Break, the trend of his rate of improvement is expected to be an accurate estimate of his future performance.

Overall, the benchmark results show a consistent pattern of improvement in both Accuracy and WCPM among all of the tutees and control/progress monitored participants in the study, with tutees' progress exceeding expectations based on benchmarks, and control participants' progress exceeding expectations to a somewhat lesser extent. These results were also analyzed using a one-tailed paired t-test within each of these treatment groups.

Tutors. Tutors' progress based on benchmark assessments was considered to be the most relevant standard for evaluating their improvement. Since these students already met benchmark goals for Accuracy, their benchmark results are presented for WCPM and Retell Total Score, by tutor grade level.

Figures 27 to 30 show changes in tutors' WCPM benchmark results. These are presented separately for each grade level. All of the tutors had Accuracy scores that were At or Above Benchmark levels before peer tutoring began. Almost all tutors' scores ranged between 98% and 100%. Figures 27 to 30 show the changes in tutors' WCPM before and after peer tutoring. Most of the tutors' WCPM scores did not make a substantial shift. The exceptions to this trend are two sets of outlier scores: those from Tutor 5 and Tutor 8, and to a lesser extent, those from Tutor 1. The remaining tutors' performance on WCPM either stayed constant or declined slightly.

Paired t test results

Changes within groups. Table 14 shows the results of paired t tests conducted within each of the treatment groups (tutees, matched controls, and tutors). Tutees improved their WCPM an average of just over 18 words per minute, which is equivalent to an average weekly improvement of about a 1.33 words per week (with an estimate of about fifteen weeks of school between beginning and midyear benchmark assessments). They improved their Accuracy about seven percentage points during that same period, and improved their Retell Total score by almost nine and half words used in retelling the story. All of these results show a significant improvement between the beginning of the year and the midyear benchmark assessments.

Matched Control participants had an average improvement of about one word per week on WCPM using the same estimated timeframe as for the tutees. These students also improved their Accuracy scores on average between the beginning and midyear assessments, by about six and a quarter percentage points. Matched control participants also increased their median Retell Total score, on average, by just under eleven words.

The tutors had an average improvement of 23 words between beginning and midyear benchmark assessments. Tutors also show a significant increase in their total retell score.

There was significant improvement for all dependent variables in each of the groups, with the exception of no significant improvement in Accuracy for the tutors. Effect sizes represented by r^2range from 62.5% for tutors WCPM progress, to 83.5% for the matched controls improvement in retell. The modest results for tutors' WCPM reflect that three tutors' growth (Tutors 1, 5, and 8) accounted for most of the observed differences in benchmark assessment differences.

	Tutees	M_D	s_D	t	p	r^2
Tutees	WCPM	18.57	11.67	4.86	0.003	79.8%
	Accuracy	6.95%	4.21%	5.03	0.002	80.8%
	Retell Quantity	9.43	7.44	3.94	0.008	72.1%
Matched Controls	WCPM	15.29	10.32	4.55	0.004	77.5%
	Accuracy	6.25%	4.51%	4.27	0.005	75.2%
	Retell Quantity	10.71	5.91	5.51	0.002	83.5%
Tutors	WCPM	23.00	23.12	3.16	0.019	62.5%
	Accuracy	0.37%	0.92%	1.62	0.157	n/a
	Retell Quantity	16.57	10.74	4.73	0.003	78.8%

Table 14. Paired t test results for student skill improvement between beginning and midyear benchmark

Survey Results

Another aspect of the tutoring program was the effect of the program on tutor-tutee interactions in general, along with incidental socially relevant effects attributed to the program. Four of the five teachers at the school responded, for an 80% response rate. All of the teachers whose students participated in the program completed a questionnaire.

Changes in Behaviors between Tutors and Tutees. Out of the four teachers who responded, three stated that the behaviors between tutees and tutors were somewhat better than before peer tutoring began. The remaining teacher indicated that these had improved significantly. All of the respondents said this was true for most tutee-tutor pairs and that that there were no differences in behavior between tutors and the control participants.

Improvements in student behavior attributed to the peer tutoring program. Several improvements in student achievement, academic behavior, and social behavior were relayed by teachers. Some of the themes that emerged have focused on students' increased enjoyment of reading, exhibited by eagerly participating in tutoring, opting to stay in for recess to voluntarily continue reading, increased class participation when reading aloud, especially from students who previously struggled. Students are also more focused on reading during independent reading time.

Expectations for the program that did not occur. Two of the teachers mentioned that the program met all expectations, one teacher felt that Retell Total Score and Retell Quality could have been higher, and one teacher specified that the only expectation that was not met was a concern that tutees would not be motivated to

participate without external incentives. That teacher was pleasantly surprised to see tutees' motivation was not focused on earning external rewards for reading.

Alternative influences on outcomes from peer tutoring. There were several changes that co-occurred with the peer tutoring program. Teachers were asked to provide any alternative explanation for the change in reading skills. The teachers stated that the following interventions co-occurred during the study and could have affected students' progress in reading: (a) enforcing a "no Spanish" rule while at school, (b) adding high quality/high interest reading material to all classrooms.

Chapter 4. Discussion

The peer tutoring program was implemented within the school to facilitate and support student reading proficiency. Specifically, this project arose from teacher reports that more than half of the students within the school struggled with reading grade-level texts. Several of the students who were struggling appeared to be hesitant to attempt to read during group reading, a common strategy used within the school's curriculum.

The present study tested three main hypotheses. These hypotheses pertained to investigating the following criteria to determine the effectiveness of this peer tutoring program: (a) statistically significant (reliable) results for students within tutoring pairs, (b) socially relevant changes (e.g., perceptions of improvement), and (c) maintenance of skills or transfer of skills.

The first hypothesis stated that active participation in the peer tutoring program at the recommended levels of duration, frequency, and intensity will (a) significantly improve the reading fluency of participating students (tutees and tutors) relative to other students, and (b) the reading comprehension of tutors. Tutees will make the most growth in measures of reading fluency, followed by the growth of tutors and non-tutored students.

There were three sets of results that were relevant for the first part of this hypothesis: (a) comparisons of progress monitoring of Accuracy and WCPM for tutees and peers matched on relative instructional and progress monitoring levels, (b) comparisons between expected and obtained growth trajectories of Accuracy and WCPM

for tutees, tutors, and control participants, and (c) paired-comparison t-tests and effect size estimates for tutees and comparison peers.

Results presented support a general conclusion that Accuracy improved for both tutees and students who participated in progress monitoring. All but one pair of students in the matched pairs met the standard goals for accuracy based on the level at which they were monitored. As discussed in the Results, alternative explanations such difficulty differences in progress monitoring passages for Pair 2, or relatively greater improvement for matched controls when compared to tutees within the same pair indicate that the progress monitoring data were insufficient on their own to support a conclusion that tutees' Accuracy improved to a greater extent than their peers.

For WCPM, results from the progress monitoring support a conclusion that students at the lowest instructional level improved the number of words they read correctly, but were insufficient to support either a conclusion that students improved overall, or that tutees who did improve their WCPM during progress monitoring showed more of a benefit than similar students who were not tutored.

After examining the survey responses, it was evident all of the teachers perceived a significant and meaningful effect of the peer tutoring specifically on the general reading proficiency of tutees, particularly behaviors related to students' motivation and interest in pursuing independent reading, which in turn gave students more opportunities to practice reading skills outside of the tutoring program. More of these results will be discussed with respect to Hypotheses 2 and 3, below.

Results of the individual student benchmark assessments, when compared to the normative growth patterns by age or grade level (through grade 6) support a conclusion in this study that all of the students who participated in the progress monitoring and who had beginning benchmark scores for Accuracy that were Well Below Benchmark levels, were able to improve their Accuracy on grade-level passages. For most students, the degree of improvement means that their decoding skills are likely to be strong enough to increase the difficulty of instructional material, and to increase an emphasis on building automaticity of reading. These students' decoding skills have moved from emergent levels that required more of a decoding emphasis during instruction to a level where core instruction can focus on other aspects of fluency. Increased accuracy (particularly at levels > 93%; Burns, 2/19/2014; or above 90%; Cummings et al., 2011) reduces each student's dependence on a highly accurate peer or teacher and creates an opportunity for more independent reading for all of the students.

Social relevance of the results was determined by whether or not students were able to move into score bands associated with increased chances for meeting future benchmarks of adequate progress and teacher responses to the intervention. Teachers viewed tutees and tutors as making significant progress. The results from the progress monitoring assessments were mixed, but the benchmark assessment results suggest that improvement rates were somewhat greater for the tutees. Because most tutors were reading with perfect or nearly perfect Accuracy (as described earlier), there was less opportunity for tutors to show improved Accuracy after the tutoring project.

The results for WCPM for tutees, matched control participants who were only monitored for progress in reading fluency, and tutors compared to benchmark goals show that when tutees' rate of reading improved, it often led to moving into a score range associated with increased chances of meeting future benchmark goals, even if the tutee was still Below Benchmark goals at present. Monitored students who improved their WCPM, with one exception, were more likely to remain at risk for not meeting future benchmarks without more intensive reading fluency interventions. As a result, it is recommended that these students begin participating in paired reading, either with tutees as partners, or as part of an extension of the peer tutoring program, with the established tutors. Tutors' were less likely than tutees to show improvement in WCPM, even though some tutors had very dramatic improvements in WCPM. Most tutors' either remained within a few words of their beginning benchmark levels or improved at about the expected rate, based on benchmark goals.

The second part of Hypothesis 1 stated the primary effect for tutors would be an improvement in their reading comprehension skills. This effect was evaluated by looking at the benchmark results for tutors' reading fluency and comparing those results to the tutors' improvement in Retell Fluency compared to benchmark goals. Results shown in Figures 31 through 34 support the conclusion that the primary outcome for tutors was enhanced recall of passage details for a novel passage during benchmark testing. As a rough estimate of a student's ability to retain information through reading (a component of reading comprehension), the Retell Fluency results in Table 13 show that almost all

tutors experienced dramatic growth in this index relative to expected levels of improvement by grade level.

If an intact treatment group is defined as one where there was a tutee-tutor-peer control for the duration of the study, significant paired t-test results supported a conclusion that in general, tutees and peer controls in intact treatment groups showed significant improvement in their WCPM and Accuracy between the beginning and midyear assessment on grade level benchmark measures. All three groups had significant improvement in Retell Fluency, yet there were not significant differences between groups in the amount of improvement on any of the Indicators, which means that any apparent differences in growth are not substantial enough to directly link those effects to peer tutoring only.

The second hypothesis stated that tutees, tutors, and teachers will perceive improved social interactions between tutees and tutors as a function of their participation in the program, possibly resulting from the positive feedback emphasis of the tutoring program.

Tutees and tutors perceptions could not be directly assessed due to the time it took near the end of the study to complete the assessments. However, teachers who responded indicated that social interactions between tutors and tutees were somewhat better for most students, and significantly better for some of the student pairs. At the same time, teachers did not perceive any change in the social interactions between monitored students and tutors or tutees.

This finding provides limited but promising evidence that peer relations are positively affected by the strategies used within the peer tutoring paradigm. Although a direct link was not provided, both neutral and positive corrective feedback on reading were the primary differences between the structured peer tutoring protocol and less structured and more passive listening that occurred through the buddy reading program. The limited support is because teachers were surveyed about a program they like. Also, most of the teachers were informed that improved social relationships were a desired outcome for the intervention.

The third hypothesis stated that word use and oral reading fluency growth will be cumulative and sustainable after the completion of the oral reading intervention component of the proposed intervention. This hypothesis was intended to evaluate whether or not improvements specific to instructional passages would generalize to grade-level material and academic tasks within the school. Although reading fluency was the immediate goal of the program, the ultimate goal of peer tutoring was to accelerate students along a path towards reading proficiency. Word use and cumulative reading fluency skills were considered as options for assessing meaningful transfer of reading skills to improved academic progress and learning.

Although word use was not specifically measured, other, more academically relevant case information was provided through examples of improvement and unexpected outcomes described by the teachers in their open-ended responses to the survey. The most meaningful extensions of the tutoring program were given by three different summaries of student behavior primarily attributed to the effectiveness of

tutoring: (a) tutees' motivation to participate in tutoring, even without prizes, stickers, or other tangible rewards provided to tutors, (b) a reported increase in student participation in group and independent reading during class, and (c) reading and choosing to read during recess, especially nonfiction passages. Another support for the overall effectiveness or perceived effectiveness of the program was testimonials and repeated requests by the teachers at the school to extend the tutoring project beyond the designated tutees as soon as possible.

While it is compelling to attribute all positive outcomes within the school and for students' reading proficiency to the peer tutoring program, there are several considerations and threats to the internal and external validity of such a conclusion.

First, this program was implemented within the context of several other changes that occurred simultaneously with the timing of the peer tutoring program. A new principal and a teacher certified in primary education began teaching at the school in Fall 2012. This project was initiated in response to their observations about the reading skills of the students at the school, where both stated that they were frustrated by several students' lack of visible progress in reading, and limited participation in reading aloud in class.

Second, staff changes at the end of the 2012-2013 academic year and the hiring of two more teachers at the beginning of the 2013-2014 academic year (one of whom was replaced in November 2013), combined with reorganizing the teaching assignments in the school may also have affected the reading skills for the students. Third, one of the teachers also highlighted increased enforcement of a school policy of English-only

conversations as a possible influence on students' reading proficiency through her answer to the question about alternative influences on the causes of students' improvements in reading accuracy, fluency, and comprehension, where she stated: "I believe the no Spanish rule is helping kids improve their grammar and fluency as well."

Another potential confound was due to the popularity of the peer tutoring program. An enthusiastic adoption of structured peer tutoring within the school resulted in one teacher implementing peer tutoring in math and spelling throughout her classroom, as reported in the survey results and by the principal of the school. While it is possible that this program has covered material that is unrelated to the focus of peer tutoring for improving reading fluency, positive effects on either tutee or control participant's skills cannot be strictly attributed to the tutoring or to monitoring student progress on materials that are more closely aligned to students' instructional needs.

Finally, there were several constraints towards implementing the study as a fully controlled randomized trial to test the effectiveness of peer tutoring. The target school is very small and remote, thus limiting the options for finding matched control participants for tutees. The geographical isolation and available funding sources meant that most of the study had to be conducted via remote consultation with the principal and lead teacher using social media. However, these issues are not without precedent, and as a documentation of a field-based intervention, the robustness of the effects observed in this site are enhanced by their presence even with such limitations and constraints on the process.

Educational reforms are prevalent, and often invoke sweeping generalizations about systematic policies and philosophies about teaching and learning. At the heart of learning is sharing. The connections between learners or between tutees and tutors provide a conduit for this sharing. Even so, the cautions listed above restrict how much the present study may generalize to another school, another country, or even a new cohort of students within the target school. Questions still remain about how this program can be adapted and enhanced. There are also questions about isolating its influence within the context of other initiatives and programs within the school. Here are some of the questions that are still unanswered, and which invite a more extensive and rigorous testing in future efforts.

The effect of the peer tutoring was limited to the dependent variables in this study, which were limited in scope to specific reading skills. It would be worthwhile to test whether other prerequisite skills (e.g., rhyming and decoding) will be as enthusiastically endorsed by participants or lead to improved reading proficiency among students whose skills are not developed enough for a repeated reading intervention.

Comparing outcomes for matched peers is dependent on the variables used for matching. A decision was made to focus mainly on Accuracy and instructional level *relative* to assigned grade level. A review of the growth in benchmark assessments suggests that, once instructional fit is determined, matching peers to compare growth might be more effective if the matching is based on age or assigned grade level for this population.

This program's timing coincided with a new principal willing to try a variety of approaches, and whose training in organizational interventions was a key factor in implementing the program. As a pilot program that has the potential for being implemented in other schools in Belize, the influence of organizational supports is critical. It also would be worthwhile to investigate the relative effects of resources and fidelity monitoring that was beyond the scope of the present project.

Field research, and educational field research in particular, is marked by daily decisions in response to situational changes. This particular study has illustrated one case of implementing interventions in an action research paradigm that has changed students' perceptions and habits in reading. Educational standards are set according to essential, important, and enhancing characteristics. Applying a similar framework towards a taxonomy of intervention characteristics would help future interventions adapt protocols to best address the specific needs of their school and its students.

Summary

The present study was designed to assess whether or not an intervention designed to improve reading fluency skills in the U.S. could be adapted and implemented with fidelity in a developing country such as Belize. The general conclusion from this investigation is that even with limiting fidelity to a general emphasis on shared reading with feedback, it is possible to make dramatic and meaningful changes to students' prerequisites skills in reading fluency and reading comprehension. Using the basic framework provided by Wright (1994), combined with efforts to match instruction and progress monitoring to students' instructional needs, it is possible to assist and

supplement the effects of other change initiatives to improve students' access to learning through reading.

References

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2011). Pedagogical strategies for teaching literacy to ESL immigrant students: A meta-analysis. *British Journal of Educational Psychology, 81*(4), 629-653. doi: 10.1111/j.2044-8279.2010.02015.x

Albers, C. A., Kenyon, D. M., & Boals, T. J. (2009). Measures for determining English language proficiency and the resulting implications for instructional provision and intervention. *Assessment for Effective Intervention, 34*(2), 74-85. doi: 10.1177/1534508408314175

Algozzine, B., Marr, M. B., Kavel, R. L., & Dugan, K. K. (2009). Using peer coaches to build oral reading fluency. *Journal of Education for Students Placed at Risk, 14*(3), 256-270. doi: 10.1080/10824660903375735

Allen, C. A. (n.d.). *Using paired reading to increase fluency and peer cooperation - ReadWriteThink* Retrieved 1/18/2013 from http://www.readwritethink.org/professional-development/strategy-guides/using-paired-reading-increase-30952.html#research-basis

Ardoin, S. P., & Christ, T. J. (2009). Curriculum-based measurement of oral reading: Standard errors associated with progress monitoring outcomes from DIBELS, AIMSweb, and an experimental passage set. *School Psychology Review, 38*(2), 266-283.

Ardoin, S. P., Williams, J. C., Klubnik, C., & McCall, M. (2009). Three versus six rereadings of practice passages. *Journal of Applied Behavior Analysis, 42*(2), 375-380. doi: 10.1901/jaba.2009.42-375

Arquette, C. M. (2000). *Participation in an English language peer tutoring program: A case study of sixth-grade second language learners.* (Unpublished Ph.D.). New Mexico State University, United States -- New Mexico. (304617525)

August, D. (2006). Demographic overview. In D. August, & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth* (pp. 43-50). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

August, D. L. (1982). *The effects of peer tutoring on the second-language acquisition of hispanic elementary school children.* (Unpublished Ph.D.). Stanford University, United States -- California. (303234891)

August, D. L., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth*. London: Lawrence Erlbaum Associates.

Ballard & Tighe - language assessment Retrieved 3/10/2013, 2013, from http://www.ballard-tighe.com/products/la/oralEng/ipti.asp

Belize Central Statistical Office (2000). *Belize 2000 housing and population census tabulation system.* Retrieved 1/21/2013 from http://celade.eclac.org/cgibin/RpWebEngine.exe/PortalAction?&MODE=MAIN&BASE=CPVBLZ2000&MAIN=WebServerMain.inl

Belize Ministry of Education. (2007). *Belize national standards and curriculum web for language arts: Lower division.* Belize City, Belize: QADS, Belize Ministry of Education.

Belize Ministry of Education. (2008). *Belize national standards and curriculum web for language arts: Middle division.* Belize City, Belize: QADS, Belize Ministry of Education.

Belize Tourism Board (n.d.). *Language: Belize tourism board.* Retrieved 1/22/2013 from http://www.travelbelize.org/about-belize/language

Brice, R. G., & Brice, A. E. (2009). Investigation of phonemic awareness and phonic skills in Spanish-English bilingual and English-speaking kindergarten students. *Communication Disorders Quarterly*, *30*(4), 208-225. doi: 10.1177/1525740108327448

Calhoon, M. B., Al Otaiba, S., Cihak, D., King, A., & Avalos, A. (2007). Effects of a peer-mediated program on reading skill acquisition for two-way bilingual first-grade classrooms. *Learning Disability Quarterly*, *30*(3), 169-184. doi: 10.2307/30035562

Calhoon, M. B., Al Otaiba, S., Greenberg, D., King, A., & Avalos, A. (2006). Improving reading skills in predominantly Hispanic Title 1 first-grade classrooms: The promise of peer-assisted learning strategies. *Learning Disabilities Research & Practice, 21*(4), 261-272. doi: 10.1111/j.1540-5826.2006.00222.x

Daly, E. J., Lentz, F. E., & Boyer, J. (1996). The instructional hierarchy: A conceptual model for understanding the effective components of reading interventions. *School Psychology Quarterly*, *11*, 369-386.

Dufrene, B. A., Henington, C., & Townsend, A. E. (2006). Peer tutoring for reading fluency: Student implementation and effects on reading fluency. *Journal of Evidence-Based Practices for Schools, 7*(2), 118-137.

Dufrene, B. A., Reisener, C. D., Olmi, D. J., Zoder-Martell, K., McNutt, M. R., & Horn, D. R. (2010). Peer tutoring for reading fluency as a feasible and effective alternative in response to intervention systems. *Journal of Behavioral Education*, *19*(3), 239-256. doi: 10.1007/s10864-010-9111-8

Ekstrand, C. (2011). *The effects of peer tutoring on reading fluency.* (M.A.S.E., Caldwell College). ProQuest Dissertations and Theses (868528828).

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), 605-619. doi: 10.1037/0022-0663.92.4.605

Fuchs, D., & Fuchs, L. S. (2005). Peer-assisted learning strategies: Promoting word recognition, fluency, and reading comprehension in young children. *The Journal of Special Education*, 39(1), 34-44. doi: 10.1177/00224669050390010401

Garcia, G. E., McKoon, G., & August, D. (2006). Language and language assessment. In D. August, & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth* (pp. 597-630). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Genesee, F., Geva, E., Dressler, C., & Kamil, M. L. (2006). Synthesis: Cross-linguistic relationships. In D. August, & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth* (pp. 153-174). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Ginsburg-Block, M., Rohrbeck, C. A., & Fantuzzo, J. W. (2006). A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology*, *98*(4), 732-749. doi: 10.1037/0022-0663.98.4.732

Ginsburg-Block, M., Rohrbeck, C., Fantuzzo, J., & Lavigne, N. C. (2006). Peer-assisted learning strategies. *Children's needs III: Development, prevention, and intervention* (pp. 631-645) Washington, DC, US: National Association of School Psychologists.

Good, R. H., Kaminski, R. A., Cumming, K., Dufour-Martel, C., Petersen, K., Powell-Smith, K., Wallin, J. (2011a). *DIBELS® Next assessment manual*. Eugene, Oregon: Dynamic Measurement Group, Inc.

Good, R. H., Kaminski, R. A., Cumming, K., Dufour-Martel, C., Petersen, K., Powell-Smith, K., Wallin, J. (2011b). *DIBELS® Next technical manual*. Eugene, Oregon: Dynamic Measurement Group, Inc.

Gordon, M. E. (1980). *Attitudes and motivation of second language achievement: A study of primary school students learning English in Belize, Central America.* (Ph.D., University of Toronto). 42(1-), 141-142.

Grant, A., Gottardo, A., & Geva, E. (2012). Measures of reading comprehension: Do they measure different skills for children learning english as a second language? *Reading and Writing*, *25*(8), 1899-1928. doi: 10.1007/s11145-012-9370-y

Greenwood, C. R. (2001). Class wide peer tutoring learning management system. *Remedial & Special Education*, *22*(1), 34.

Hanson-Smith, E. (1988). Belize and Sri Lanka: Language planning in multilingual states. *Educational Research Quarterly, 12*(4), 23-31.

International Reading Association. (2013). English-language learners overview. Retrieved 1/18/2013 from http://www.reading.org/Resources/ResourcesByTopic/EnglishLearners/Overview.aspx

Johnson, D. M. (1980). *Peer tutoring, social interaction, and the acquisition of English as a second language by Spanish-speaking elementary school children.* (Unpublished Ph.D.). Stanford University, United States -- California. (303062070)

Kaminski, R. A., & Good III, R. H. (2012). *DIBELS Next® essential workshop.* Eugene, Oregon: Dynamic Measurement Group.

Klein, J. R., & Jimerson, S. R. (2005). Examining ethnic, gender, language, and socioeconomic bias in oral reading fluency scores among Caucasian and Hispanic students. *School Psychology Quarterly*, *20*(1), 23-50. doi: 10.1521/scpq.20.1.23.64196

Kourea, L., Cartledge, G., & Musti-Rao, S. (2007). Improving the reading skills of urban elementary students through total class peer tutoring. *Remedial and Special Education*, *28*(2), 95-107. doi: 10.1177/07419325070280020801

Lesaux, N. K., & Geva, E. (2006). Synthesis: Development of literacy in language-minority students. In D. August, & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth* (pp. 53-74). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Loskot, L. B. (2007). *Literacy practices and strategies of communication of the Mopan Maya of Belize: A rhetorical analysis.* (Unpublished Ph.D.). New Mexico State University, United States -- New Mexico. (304841100)

Marr, M. B., Algozzine, B., Nicholson, K., & Dugan, K. K. (2011). Building oral reading fluency with peer coaching. *Remedial and Special Education, 32*(3), 256-264. doi: 10.1177/0741932510362202

Mooney, C. (2010). *Effects of peer-tutoring on vocabulary recognition, fluency and interaction of low SES ELL students in a second grade classroom.* (Unpublished M.A.S.E.). Caldwell College, United States -- New Jersey. (305250198)

Muyskens, P., Betts, J., Lau, M. Y., & Marston, D. (2009). Predictive validity of curriculum-based measures in the reading assessment of students who are English language learners. *California School Psychologist, 14*, 11-21.

National Institute of Child Health and Human Development. (2000). *Report of the national reading panel, teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* (No. NIH Publication No. 00-4769). Washington, DC: United States Department of Health and Human Services.

Powell-Smith, K. A., Good, R. H., Kaminski, R. A., & Wallin, J. (2012). *DIBELS Next Survey Manual*. Longmont, CO: Cambium Learning Group.

Pyron, M. (2007). *"I hear you, but I don't understand you": The effects of peer tutoring for helping secondary ESL students achieve academic success.* (Unpublished Ph.D.). Louisiana State University and Agricultural & Mechanical College, United States -- Louisiana. (304836128)

Richardson, R. D., Hawken, L. S., & Kircher, J. (2012). Bias using maze to predict high-stakes test performance among Hispanic and Spanish-speaking students. *Assessment for Effective Intervention, 37*(3), 159-170. doi: 10.1177/1534508411430320

Riley-Tillman, T. C., Burns, M. K., & Gibbons, K. (2013). *RTI applications. Vol. 2: Assessment, analysis, and decision making*. New York, NY: Guilford.

Roberts, T. A. (2005). Articulation accuracy and vocabulary size contributions to phonemic awareness and word reading in English language learners. *Journal of Educational Psychology*, *97*(4), 601-616. doi: 10.1037/0022-0663.97.4.601

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, *95*(2), 240-257. doi: http://dx.doi.org.ezproxy.mnsu.edu/10.1037/0022-0663.95.2.240

Sáenz, L. M., Fuchs, L. S., & Fuchs, D. (2005). Peer-assisted learning strategies for English language learners with learning disabilities. *Exceptional Children*, *71*(3), 231-247.

Saenz, S. B. (2008). *The effects of a comprehensive tutor-assisted oral reading program on at-risk hispanic elementary students.* (Unpublished Ed.D.). University of Houston, US.

Schreder, S. J., Hupp, S. D. A., Everett, G. E., & Krohn, E. (2012). Targeting reading fluency through brief experimental analysis and parental intervention over the summer. *Journal of Applied School Psychology*, *28*(2), 200-220. doi: 10.1080/15377903.2012.670047

Scruggs, T. E., & Mastropieri, M. A. (2013). PND at 25: Past, present, and future trends in summarizing single-subject research. *Remedial and Special Education, 34*, 9-19. doi: 10.1177/0741932512440730

Scruggs, T. E., & Osguthorpe, R. T. (1986). Tutoring interventions within special education settings: A comparison of cross-age and peer tutoring. *Psychology in the Schools, 23*(2), 187-193. doi: 10.1002/1520-6807(198604)23:2<187::AID-PITS2310230212>3.0.CO;2-7

Serrano, C. J. (1987). *The effectiveness of cross-level peer involvement in the acquisition of English as a second language by Spanish-speaking migrant children.* (Unpublished Ph.D.). Florida State University, US.

Shapiro, E. S. (2011). *Academic skills problems: Direct assessment and intervention* (4th ed.). New York, NY, US: Guilford Press.

Shinn, M. R. (1989). Curriculum-based measurement: Assessing special children. New York, NY: Guilford.

Shinn, M. R. (2008). Best practices in using curriculum-based measurement in a problem-solving model. In J. Grimes and A. Thomas (Eds.), *Best practices in school psychology V (pp. 243-262).* Washington, DC: NASP.

Standley, L. (2006). *Cross-age peer-tutoring effects on the English literacy development and academic motivation of English language learners identified with, and referred for, mild and moderate disabilities.* (Unpublished Ph.D.). The University of New Mexico, United States -- New Mexico. (305295761)

Stryker, S. B. (1986). *Analysis of a communicatively oriented ESL program utilizing native-speaking peer tutors.* ProQuest Information & Learning. ProQuest Information & Learning, 46(8-), 2218.

United Nations Educational Scientific and Cultural Organization International Bureau on Education. (2010, April). *World data on education: Belize.* (No. IBE/2010/CP/WDE/BH).UNESCO-IBE.

Vostal, B. R., & Lee, D. L. (2011). Behavioral momentum during a continuous reading task: An exploratory study. *Journal of Behavioral Education*, *20*(3), 163-181. doi: 10.1007/s10864-011-9129-6

Wayman, M. M., McMaster, K. L., Sáenz, L. M., & Watson, J. A. (2010). Using curriculum-based measurement to monitor secondary English language learners' responsiveness to peer-mediated reading instruction. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, *26*(4), 308-332. doi: 10.1080/10573569.2010.500260

Wendling, B. J., & Mather, N. (2009). *Essentials of evidence-based academic interventions.* Hoboken, NJ: John Wiley & Sons.

Wright, J. (2004). *Kids as reading helpers: A peer tutor training manual | intervention central.* Retrieved 11/20/2012 from http://www.interventioncentral.org/academic-interventions/reading-fluency/kids-reading-helpers-peer-tutor-training-manual

Wright, J., & Cleary, K. S. (2006). Kids in the tutor seat: Building schools' capacity to help struggling readers through a cross-age peer-tutoring program. *Psychology in the Schools*, *43*(1), 99-107. doi: 10.1002pits.20133

Yesil-Dagli, U. (2011). Predicting ELL students' beginning first grade English oral reading fluency from initial kindergarten vocabulary, letter naming, and phonological awareness skills. *Early Childhood Research Quarterly*, *26*(1), 15-29. doi: 10.1016/j.ecresq.2010.06.001

Young, M. C. (2009). *A comparative study of English language learners and non-English language learners' in a school district in Delaware; performance in national standardized achievement tests of reading assessment: Dynamic Indicators of Basic Early Literacy Skills (DIBELS).* (Unpublished Ed.D.). Wilmington University, US.

Yurick, A. L., Robinson, P. D., Cartledge, G., Lo, Y., & Evans, T. L. (2006). Using peer-mediated repeated readings as a fluency-building activity for urban learners. *Education & Treatment of Children*, *29*(3), 469-506.

Appendix A: Parental Consent and Student Assent Materials

Parent Consent Form—Tutee Participation

[Date]

Dear Parent:

My name is Marcia Sytsma and I am a graduate student in the School Psychology program at Minnesota State University, Mankato. In cooperation with the teachers at the Casey Community School, I would like to have your child, _____, participate in a peer tutoring program in reading. This tutoring program gives children at the school extra opportunities to practice their reading skills. Your child's teacher selected your child for this tutoring program. Participation in this project is voluntary and you may choose to have your child participate or not. Below is a description of what the project is about. If after reading this description you have any questions, feel free to contact me 001-612-978-1367 at marcia.sytsma@mnsu.edu.

If you agree to let your child participate in this tutoring program, your child will meet individually with an older student from the Casey Community School who is trained as a reading tutor. During tutoring sessions, your child will read aloud from books as the tutor listens, will have the tutor correct any reading mistakes, and will receive praise from the tutor for trying his or her best.

- These tutoring sessions will take place 3-4 times per week during school hours and will be supervised by adults.
- Each session will last about 20-30 minutes.
- Tutoring sessions will be scheduled so that your child does not miss any classwork.
- The goal of these tutoring sessions is to help your child to become a more skilled and confident reader.

We need your permission in order to give your child this extra reading tutoring. If you would like your child to participate, please read through the form on the next page and sign your name at the bottom of that form and return it to your child's teacher.

Sincerely,

Marcia R. Sytsma, MA
Doctoral Candidate in School Psychology

Carlos J. Panahon, Ph.D.
Assistant Professor

Parent Permission for Student to receive Peer Tutoring

I am the legal guardian of_____ . I consent for her or him to participate in a research project on peer tutoring at the Casey Community School. I understand that Marcia Sytsma, MA is conducting this research under the supervision of Carlos Panahon, Ph.D. and Daniel Houlihan, Ph.D. from the Psychology Department at Minnesota State University Mankato (MSU). I understand that participation in this study includes the following commitment for my child and me:

1) Read and sign this consent form.

2) If you agree to let your child participate in this tutoring program, your child will meet individually with another student who is trained as a reading tutor. During tutoring, your child will read aloud from books as the tutor listens, will have the tutor correct any reading mistakes, and will receive praise from the tutor for trying his or her best.

I understand that my child will participate in the following procedures:

- These tutoring sessions will take place 3-4 times per week during school hours and will be supervised by adults.
- Each session will last about 20-30 minutes.
- Tutoring sessions will be scheduled so that my child does not miss any classwork.
- The goal of these tutoring sessions is to help my child to become a more skilled and confident reader.

I understand that I can contact Ms. Sytsma, Mr. Sean Houlihan (Principal at Casey Community School) at 604 2139 or Dr. Panahon at 001-507-389-2815 or carlos.panahon@mnsu.edu about any concerns I have about this project. I understand that I also may contact the MSU Institutional Review Board Administrator, Dr. Barry Ries, at 001-507-389-2321 with any questions about research with human participants at MSU.

Confidentiality

All information obtained in this project will be kept private by the staff of this research project. All information will be stored in a locked file cabinet at Minnesota State University, Mankato. It can be viewed only by authorized research staff members. I understand that no information about my child will be released and no names will be recorded other than the consent forms. By law, the only times when information will not be kept confidential is if my child or I state that we are in imminent danger of harming ourselves or others, or in suspected cases of child abuse.

Risks and Benefits

I understand that the risks of participating in this study are no more than those in normal school life. I understand that my child will not receive any direct compensation. I understand that I can request a copy of the study's results, which would be mailed to me after the end of the study. I understand that participating in this study may help the

researchers better understand whether or not peer tutoring improves reading skills for both my child and the other students participating in the project.

Right to Refuse Participation

I understand that participation in this project is voluntary and my child has the right to stop at any time. My child can choose to inform the school staff, Ms. Sytsma, Dr. Panahon, or Dr. Houlihan if he or she wishes to discontinue participating in the project. My child can stop participating by saying he or she does not want to be in the study any more. My decision whether to allow my child to participate, or my child's decision whether to participate will not affect our relationships with Minnesota State University, Mankato or the Casey Community School.

Signed: _____ Date: _____

With my signature, I state that I am at least 18 years of age and I have received a copy of the consent form to keep.

WRITTEN CONSENT DOCUMENT FOR PARENTS OF PARTICIPANTS WHO DO NOT READ ENGLISH (will be translated into the Parents' primary language)

Consent to Participate in Research

You are being asked to give your consent to your child's participation in a research study.

If you agree to let your child participate in this tutoring program, your child will meet individually with another student who is trained as a reading tutor. During tutoring, your child will read aloud from books as the tutor listens, will have the tutor correct any reading mistakes, and will receive praise from the tutor for trying his or her best.

- These tutoring sessions will take place 3-4 times per week during school hours and will be supervised by adults.
- Each session will last about 20-30 minutes.
- Tutoring sessions will be scheduled so that your child does not miss any classwork.
- The goal of these tutoring sessions is to help your child to become a more skilled and confident reader.

All information obtained in this project will be kept private and information will be stored in a locked file cabinet at Minnesota State University, Mankato. It can be viewed only by authorized research staff members. No information about your child will be released and no names will be recorded other than the consent forms.

The risks of participating in this study are no more than those in normal school life. Your child will not receive any direct compensation.

Participation in this project is voluntary and your child has the right to stop at any time by saying he or she does not want to be in the study any more. Your decision whether to allow your child to participate, or your child's decision whether to participate will not affect your relationships with Minnesota State University, Mankato or the Casey Community School.

If you agree to allow your child to participate, you will be given a signed copy of this document and a written summary of the research.

You may contact Marcia Sytsma at 001-612-978-1367 any time you have questions about the research.

You may contact Dr. Barry Ries at 001-507-389-2321 if you have questions about your child's rights as a research participant.

Signing this document means that the research study, including the above information, has been described to you orally, and that you voluntarily agree to allow your child to participate.

Signed: _____ Date: _____
(parent)

Signed: _____ Date: _____
(witness)

Student Assent

ASSENT FORM

Student's Name _____

You are being asked to be part of a research project that will help your teachers learn about how to help students at the school become better readers. We have talked to the teachers about reading, and all of you have met with Ms. Liz to practice some reading activities. Now, we are asking some students to work with other students and practice reading out loud, together.

We are working with Ms. Marcia to see if this program works for Casey Community School. You will either be asked to work with another student and practice your reading. Some of you will be helping other students with reading, and others will be asked to work with a student helper to practice your reading.

Those of you who are being asked to help other students will go through a class to learn how to coach other students. During this training, you will practice with the teachers and other peer tutors to make sure you follow instructions carefully and to teach you some ideas about how to help other students.

After you are trained in tutoring, you and your tutoring partner will meet 4 days a week at the end of the day for 20-30 minutes to practice reading. During some of these times, you may either be observed or the session will be recorded so Ms. Marcia and Mr. Sean can make sure you are following directions on how to work with your partner. Every week, Ms. Liz will ask you and your partner to show her how well you are reading.

Your parents and teacher have said that it is okay for you to be out of your classroom for about 20 minutes each day to be part of this project. If you decide that you do not want to be a tutor or tutoring partner, just let me know and I will take you back to your class. You do not have to be part of this project. If you want to be part of the project, write the word "Yes" on the card that we gave each of you. If you do not want to be part of the program, just write the word "No" on the card we gave each of you. I will come and collect all the cards and tomorrow we will begin the tutoring training.

Parent Consent Form—Tutor Participation

[Date]

Dear Parent:

My name is Marcia Sytsma and I am a graduate student in the School Psychology program at Minnesota State University, Mankato. In cooperation with the teachers at the Casey Community School, I would like to have your child, _____, participate in a peer tutoring program in reading. This tutoring program gives children at the Casey Community School extra opportunities to practice their reading skills. Your child has demonstrated the necessary reading skills to be a successful peer tutor and help other students with their reading. Participation in this project is voluntary and you may choose to have your child participate or not. Below is a description of what the project is about. If after reading this description you have any questions, feel free to contact me 001-612-978-1367 at marcia.sytsma@mnsu.edu.

If you agree to let your child participate in this tutoring program, you child will work individually as a reading tutor with a other students at the school, listening to the other student read aloud, correcting the student's reading mistakes, and offering the other reader praise and encouragement.

- Before the tutoring starts, staff at the school will train your child to be a reading tutor.
- Supervised tutoring sessions will take place 3-4 times per week during recess periods.
- Each session will last about 20-30 minutes.
- Tutoring sessions will be scheduled so that your child does not miss any classwork.

We need your permission in order to allow your child to be a reading tutor. If you would like your child to participate, please read through the form on the next page and sign your name at the bottom of that form and return it to your child's teacher.

Sincerely,

Marcia R. Sytsma, MA
Doctoral Candidate in School Psychology

Carlos J. Panahon, Ph.D.
Assistant Professor

Parent Permission for Student to Serve as Peer Tutor

I am the legal guardian of_____ . I consent for her or him to participate in a research project on peer tutoring at the Casey Community School. I understand that Marcia Sytsma, MA is conducting this research under the supervision of Carlos Panahon, Ph.D. and Daniel Houlihan, Ph.D. from the Psychology Department at Minnesota State University Mankato (MSU). I understand that participation in this study includes the following commitment for my child and me:

1) Read and sign this consent form.

2) My child will participate as a tutor, and will work individually as a reading tutor with other students at the school, listening to these students read aloud, correcting the reading mistakes, and offering the other student praise and encouragement.

I understand that my child will participate in the following procedures:

- Before the tutoring starts, staff at the school will train my child to be a reading tutor.
- Supervised tutoring sessions will take place 3-4 times per week during recess periods.
- Each session will last about 20-30 minutes.
- Tutoring sessions will be scheduled so that my child does not miss any classwork.

I understand that I can contact Ms. Sytsma, Mr. Sean Houlihan (Principal at Casey Community School) at 604 2139 or Dr. Panahon at 001-507-389-2815 or carlos.panahon@mnsu.edu about any concerns I have about this project. I understand that I also may contact the MSU Institutional Review Board Administrator, Dr. Barry Ries, at 001-507-389-2321 with any questions about research with human participants at MSU.

Confidentiality

All information obtained in this project will be kept private by the staff of this research project. All information will be stored in a locked file cabinet at Minnesota State University, Mankato. It can be viewed only by authorized research staff members. I understand that no information about my child will be released and no names will be recorded other than the consent forms. By law, the only times when information will not be kept confidential is if my child or I state that we are in imminent danger of harming ourselves or others, or in suspected cases of child abuse.

Risks and Benefits

I understand that the risks of participating in this study are no more than those in normal school life. I understand that my child will not receive any direct compensation. I understand that I can request a copy of the study's results, which would be mailed to me after the end of the study. I understand that participating in this study may help the researchers better understand whether or not peer tutoring improves reading skills for

both my child and the other students participating in the project.

Right to Refuse Participation

I understand that participation in this project is voluntary and my child has the right to stop at any time. My child can choose to inform the school staff, Ms. Sytsma, Dr. Panahon, or Dr. Houlihan if he or she wishes to discontinue participating in the project. My child can stop participating by saying he or she does not want to be in the study any more. My decision whether to allow my child to participate, or my child's decision whether to participate will not affect our relationships with Minnesota State University, Mankato or the Casey Community School

Signed: _____ Date: _____

With my signature, I state that I am at least 18 years of age and I have received a copy of the consent form to keep.

WRITTEN CONSENT DOCUMENT FOR PARENTS OF PARTICIPANTS WHO DO NOT READ ENGLISH (will be translated into the Parents' primary language)

Consent to Participate in Research

You are being asked to give your consent to your child's participation in a research study.

If you agree to let your child participate in this tutoring program, you child will work individually as a reading tutor with a other students at the school, listening to the other student read aloud, correcting the student's reading mistakes, and offering the other reader praise and encouragement.

- Before the tutoring starts, staff at the school will train your child to be a reading tutor.
- Supervised tutoring sessions will take place 3-4 times per week during recess periods.
- Each session will last about 20-30 minutes.
- Tutoring sessions will be scheduled so that your child does not miss any classwork.

All information obtained in this project will be kept private and information will be stored in a locked file cabinet at Minnesota State University, Mankato. It can be viewed only by authorized research staff members. No information about your child will be released and no names will be recorded other than the consent forms.

The risks of participating in this study are no more than those in normal school life. Your child will not receive any direct compensation.

Participation in this project is voluntary and your child has the right to stop at any time by saying he or she does not want to be in the study any more. Your decision whether to allow your child to participate, or your child's decision whether to participate will not affect your relationships with Minnesota State University, Mankato or the Casey Community School.

If you agree to allow your child to participate, you will be given a signed copy of this document and a written summary of the research.

You may contact Marcia Sytsma at 001-612-978-1367 any time you have questions about the research.

You may contact Dr. Barry Ries at 001-507-389-2321 if you have questions about your child's rights as a research participant.

Signing this document means that the research study, including the above information, has been described to you orally, and that you voluntarily agree to allow your child to participate.

Signed: _____ Date: _____
(parent)

Signed: _____ Date: _____
(witness)

Appendix B: Teacher Survey

Peer Tutoring Evaluation

1. How have the behaviors between Tutors and their Tutees changes since the tutoring project started?
    a. Significantly worsened
    b. Somewhat worse than before
    c. No change; pretty much the same
    d. Somewhat better than before
    e. Significantly improved
2. Thinking about the answer to question 1, would you say this is true for Some, Most, or All of the tutoring pairs?
    a. Some
    b. Most
    c. All
3. If “some” or “most” of the peer tutoring pairs changed in this way, how were the behaviors between the rest of the tutoring pairs?
    a. Worse than some or most
    b. Better than some or most

4. How has the peer tutoring program affected the behaviors between students who were not tutored?
    a. Significantly worsened
    b. Somewhat worse than before
    c. No change; pretty much the same
    d. Somewhat better than before
    e. Significantly improved
5. What are at least two improvements in Casey Community School that you think are due to Peer Tutoring?
6. What expectations did you have for the tutoring project that you did not see?
7. Besides peer tutoring, what other changes in the school do you think may be helping students in Standards I-VI improve their reading accuracy, fluency, and comprehension?

Appendix C: Tutee and Matched Control Progress Monitoring

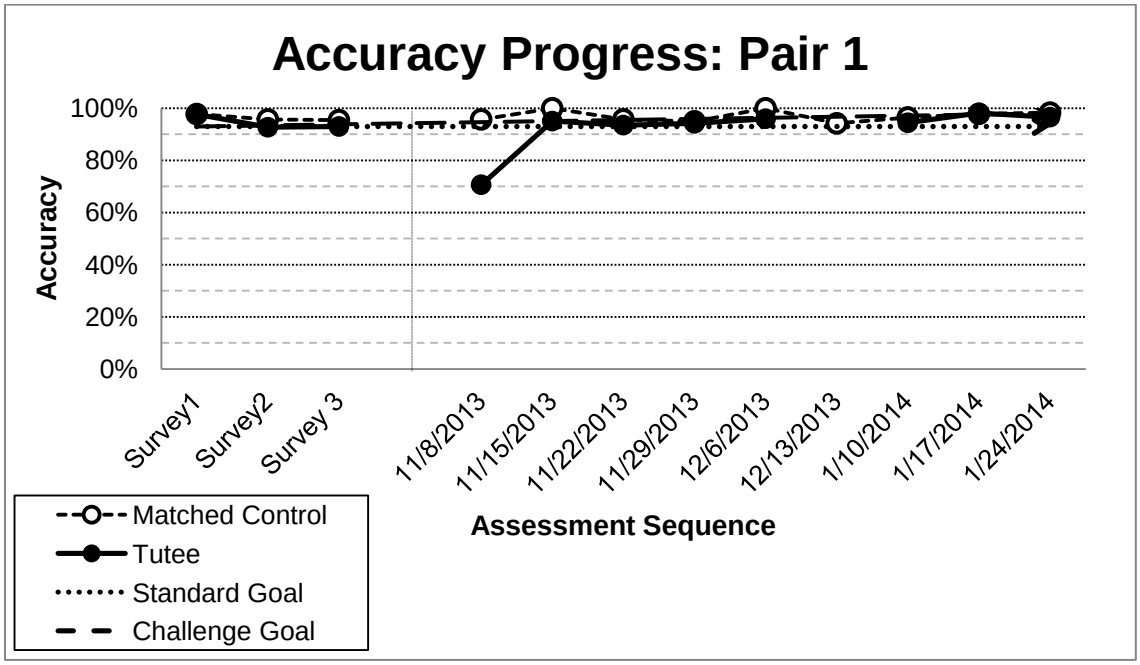

Figure 1. Accuracy progress for tutee and matched control participant in pair 1 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

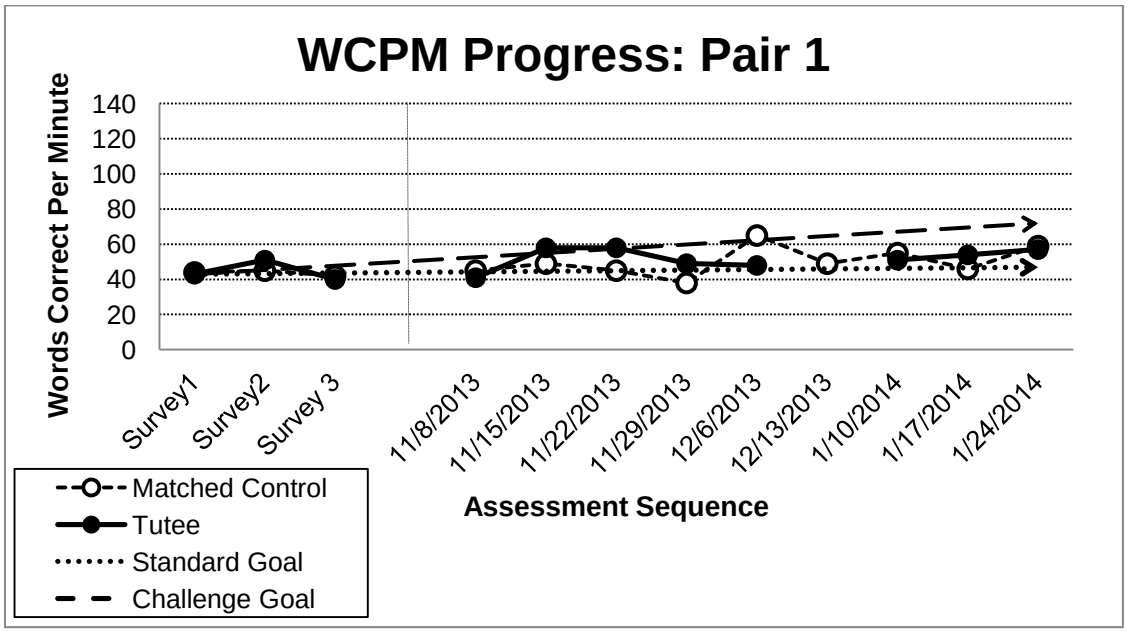

Figure 2. Progress in words correctly read for tutee and matched control participant in pair 1 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

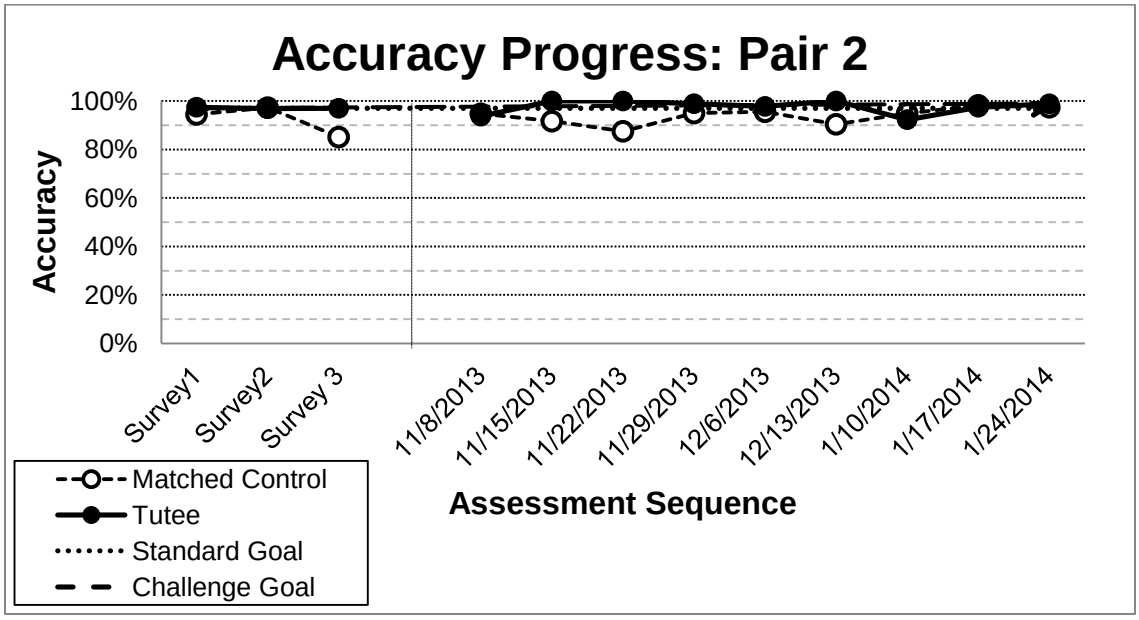

Figure 3. Accuracy progress for tutee and matched control participant in pair 2 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

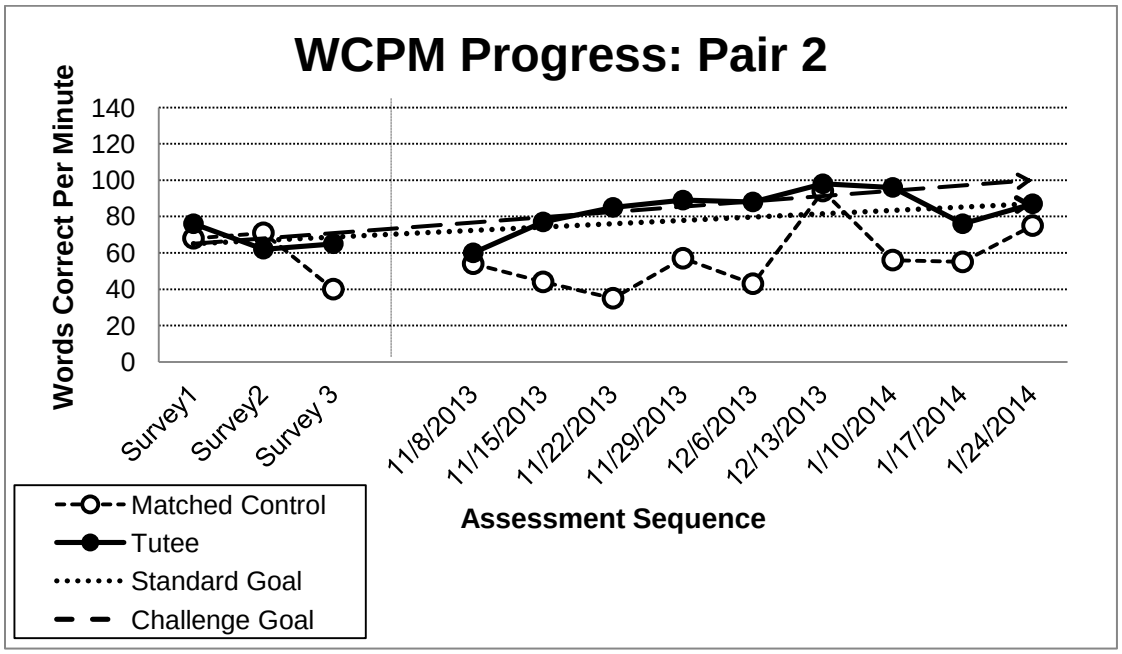

Figure 4. Progress in words correctly read for tutee and matched control participant in pair 2 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

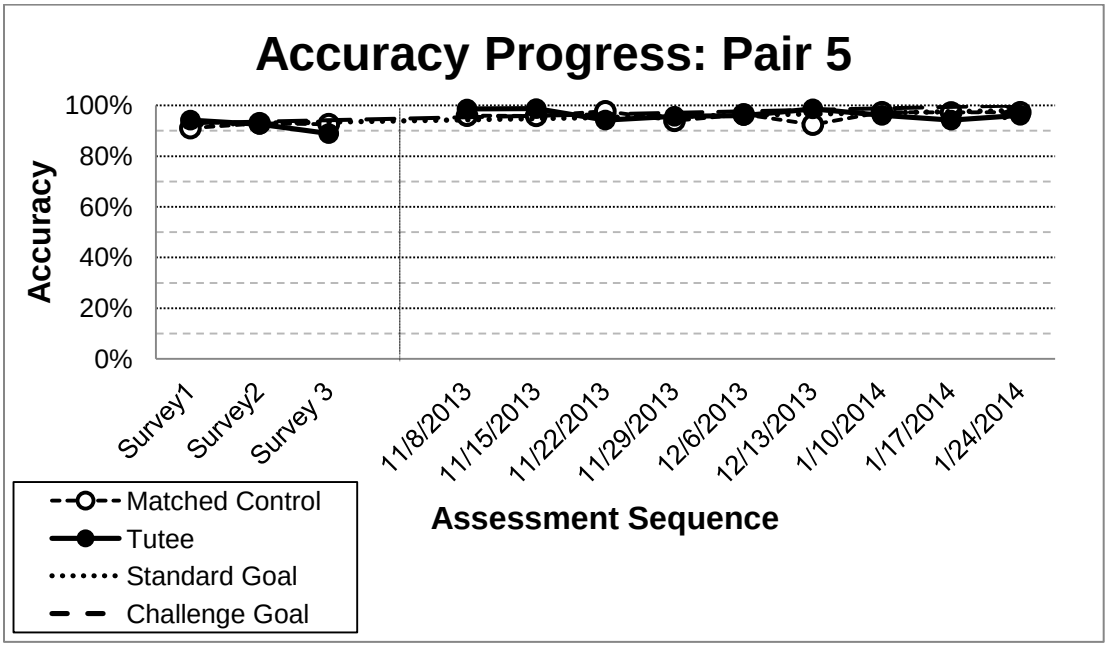

Figure 5. Accuracy progress for tutee and matched control participant in pair 5 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

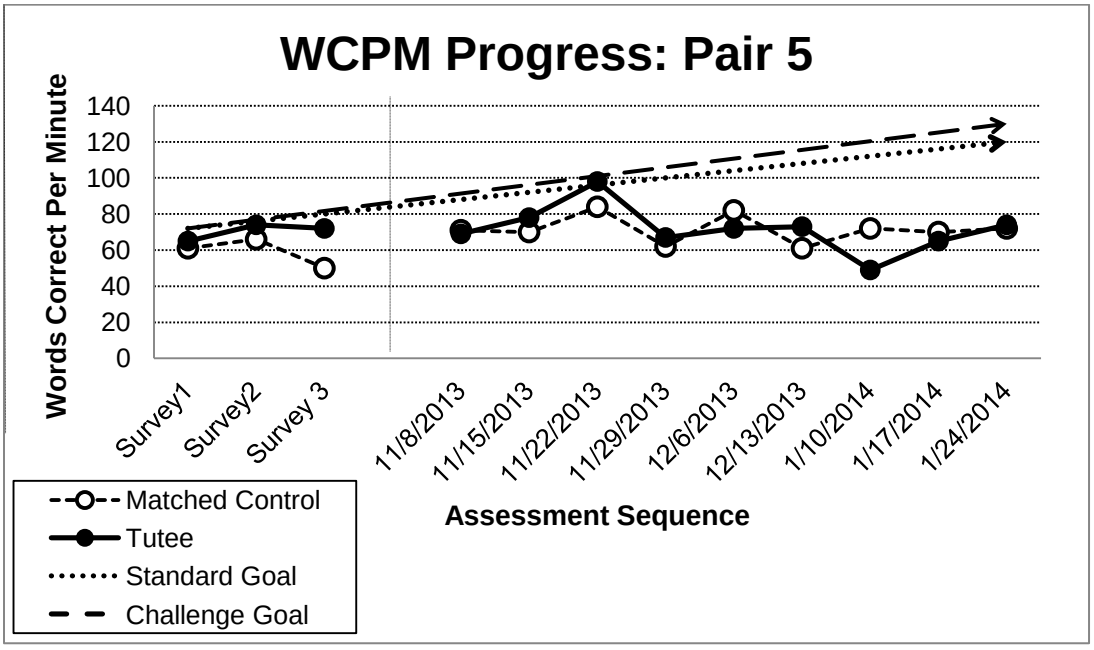

Figure 6. Progress in number of words read correctly for tutee and matched control participant in pair 5 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

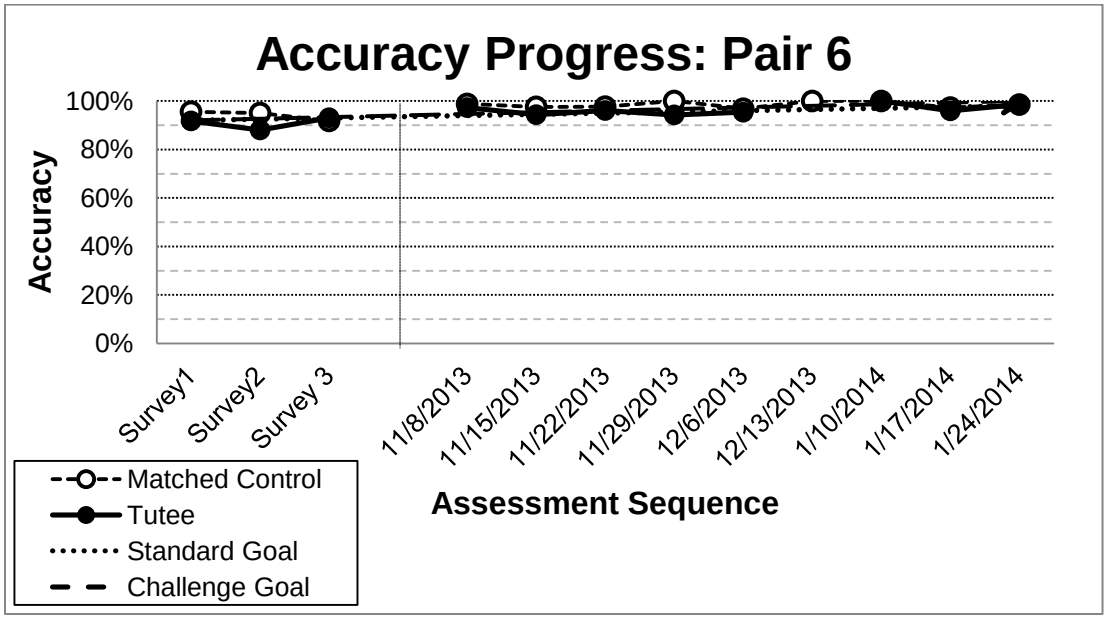

Figure 7. Accuracy progress for tutee and matched control participants in pair 6 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

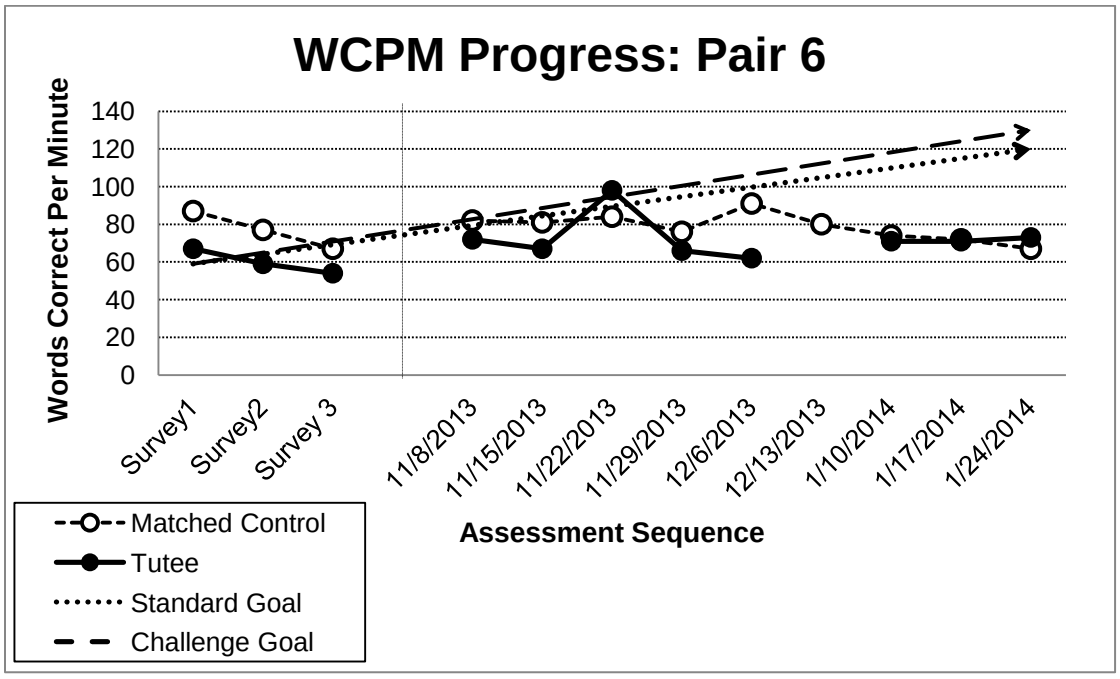

Figure 8. Progress in words correctly read for tutee and matched control participant in pair 6 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

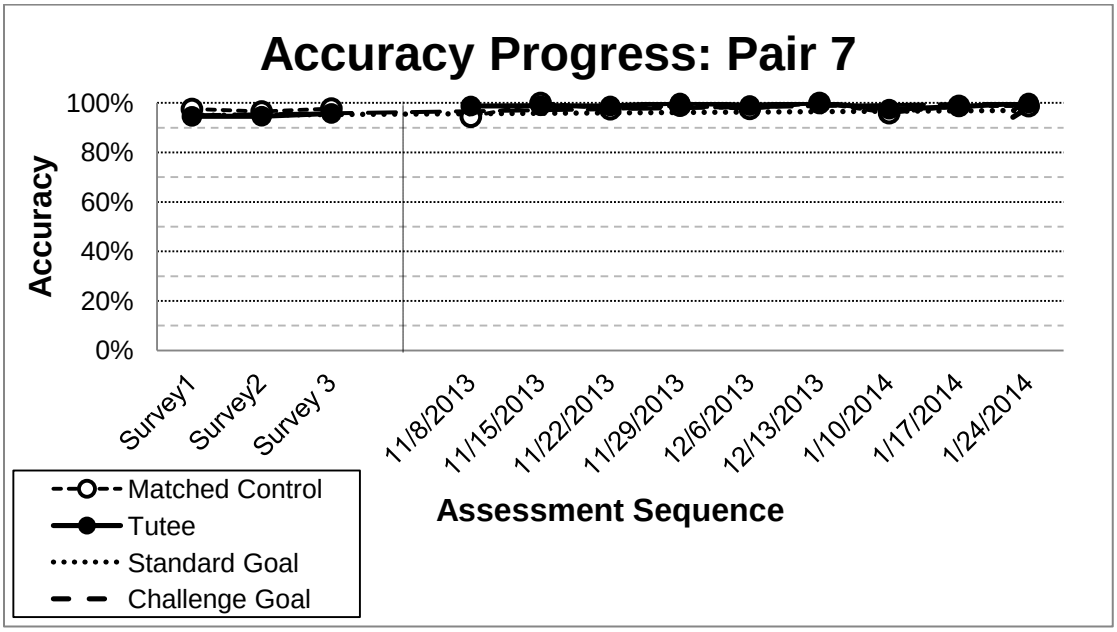

Figure 9. Accuracy progress for tutee and matched control participants in pair 7 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

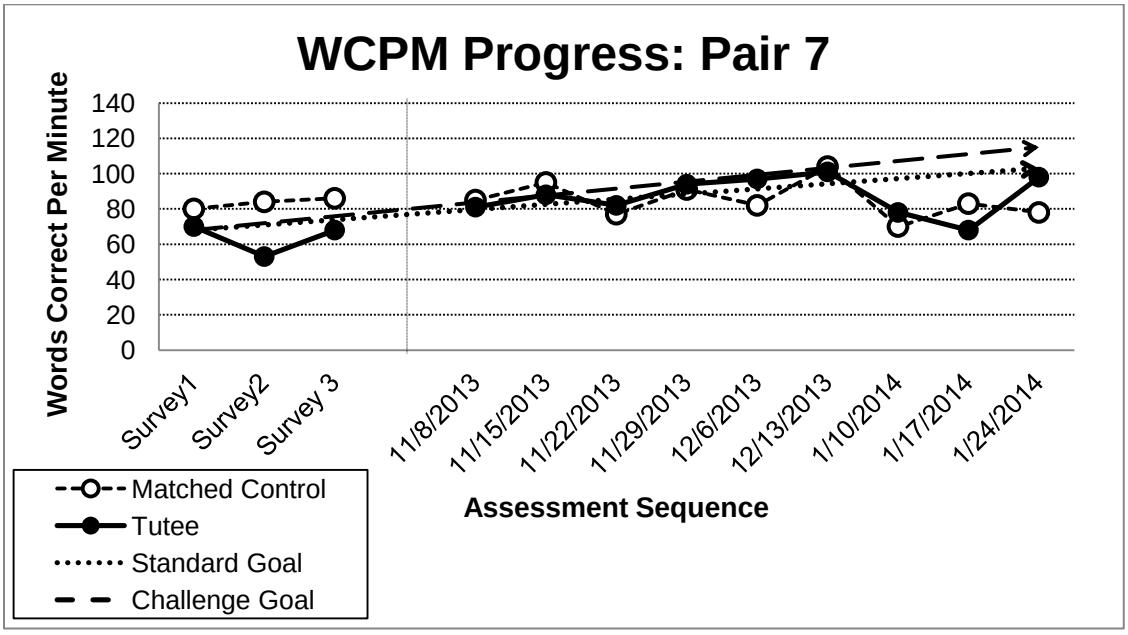

Figure 10. Progress in words correctly read for tutee and matched control participant in pair 7 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

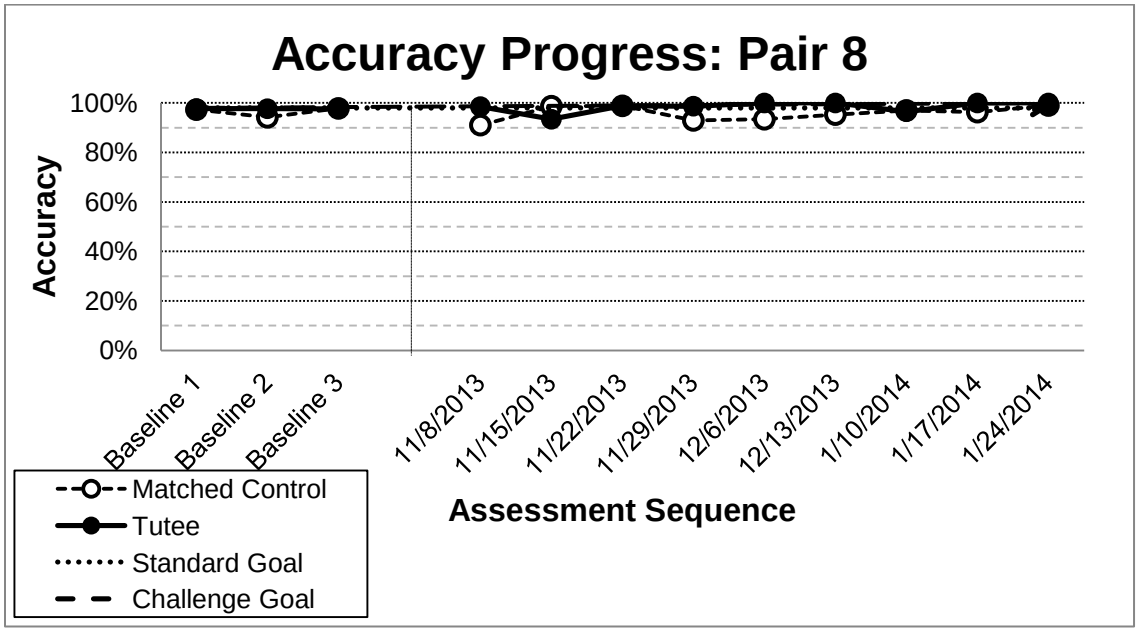

Figure 11. Accuracy progress for tutee and matched control participants in pair 8 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

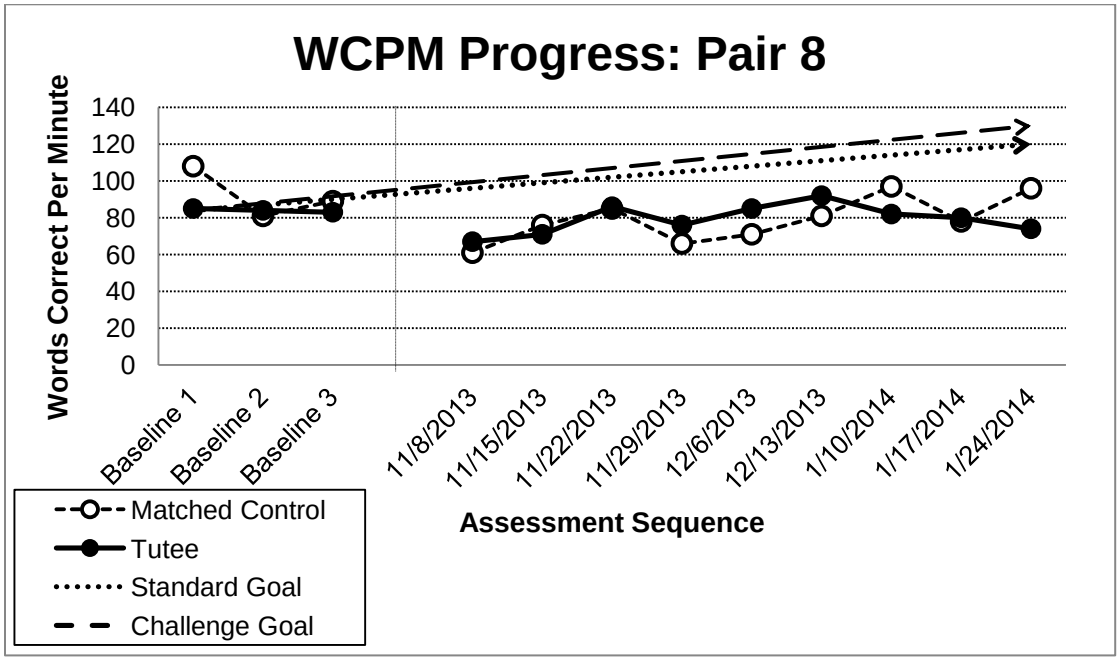

Figure 12. Progress in words correctly read for tutee and matched control participant in pair 8 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

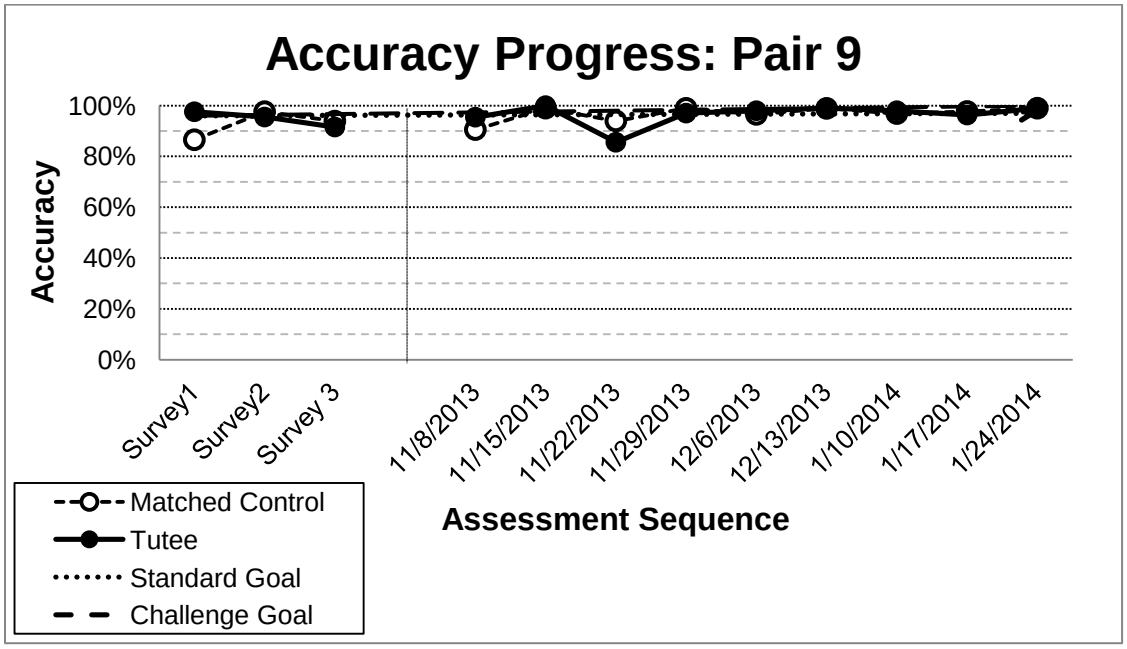

Figure 13. Accuracy progress for tutee and matched control participants in pair 9 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

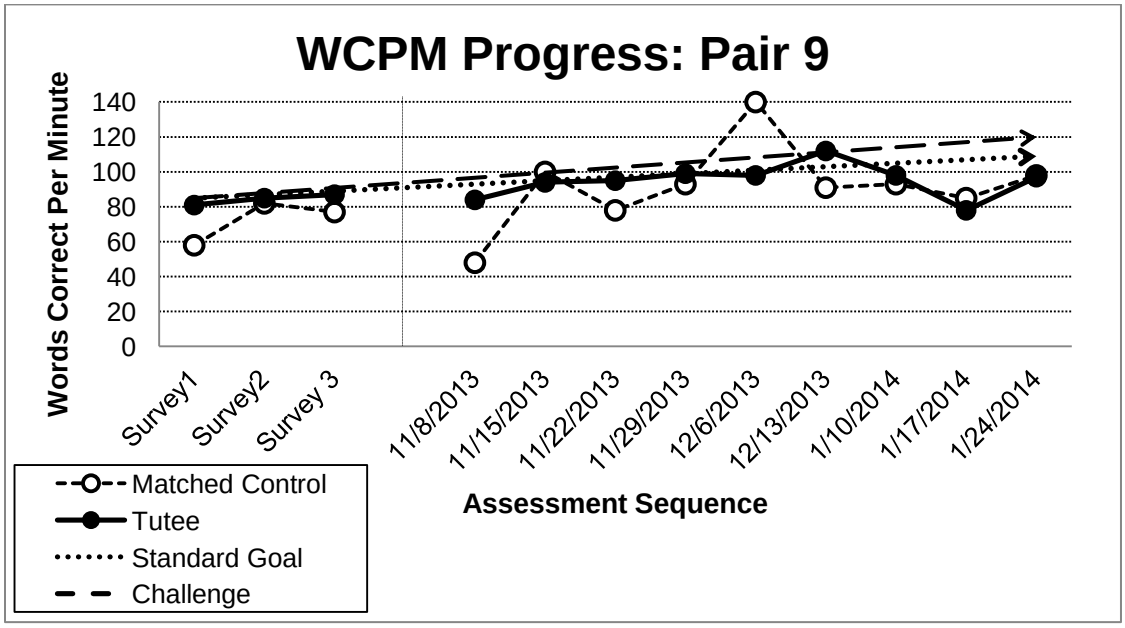

Figure 14. Progress in words correctly read for tutee and matched control participant in pair 9 during peer tutoring. Survey assessment was three measures in a single setting; progress was monitored at the end of each school week for nine weeks.

Appendix D: Benchmark Results for Tutees and Matched Controls

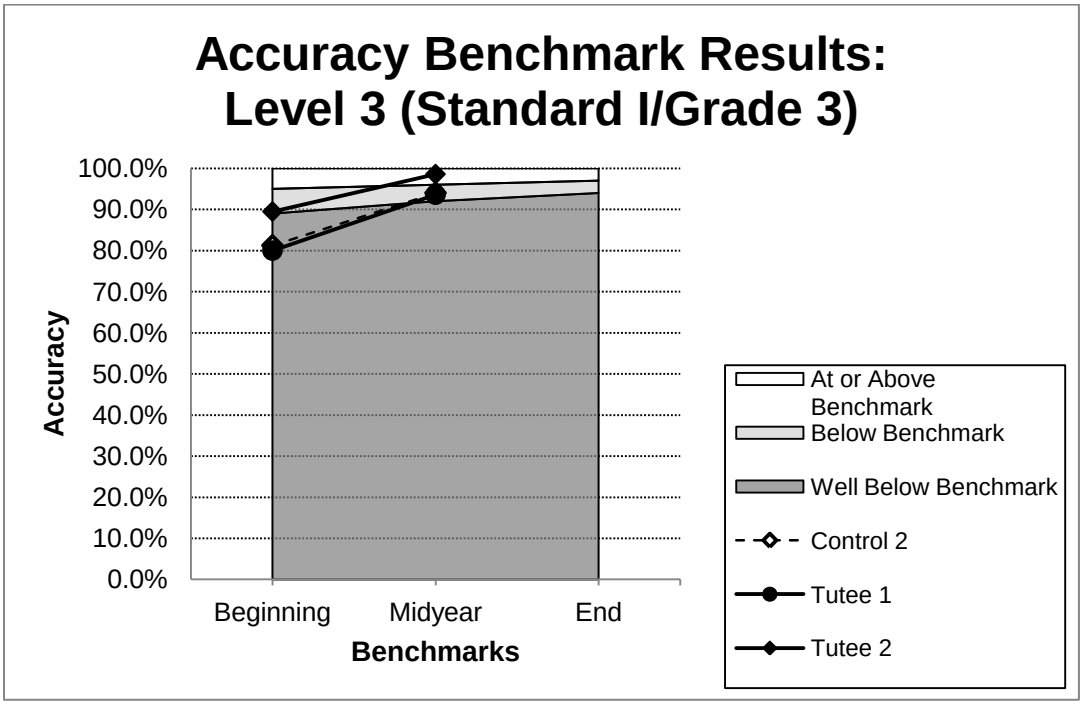

Figure 15. Growth in accuracy from beginning to midyear benchmark *DIBELS Oral Reading Fluency* compared to benchmark goals for students in Standard I.

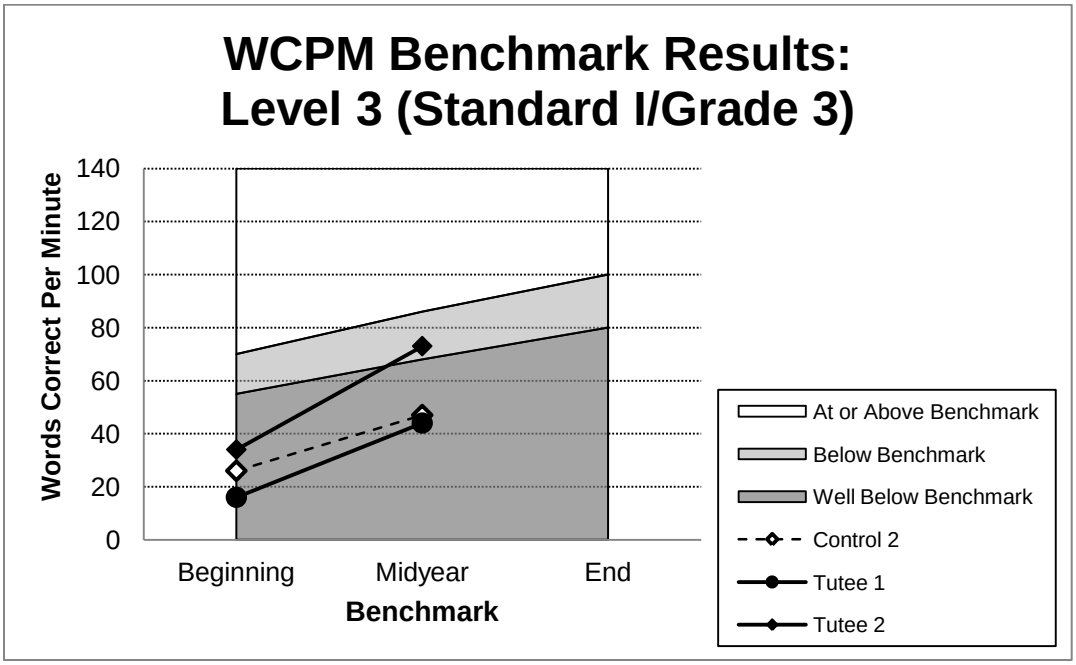

Figure 16. Growth in words correctly read per minute from beginning to midyear benchmark *DIBELS Oral Reading Fluency* compared to benchmark goals for students in Standard I.

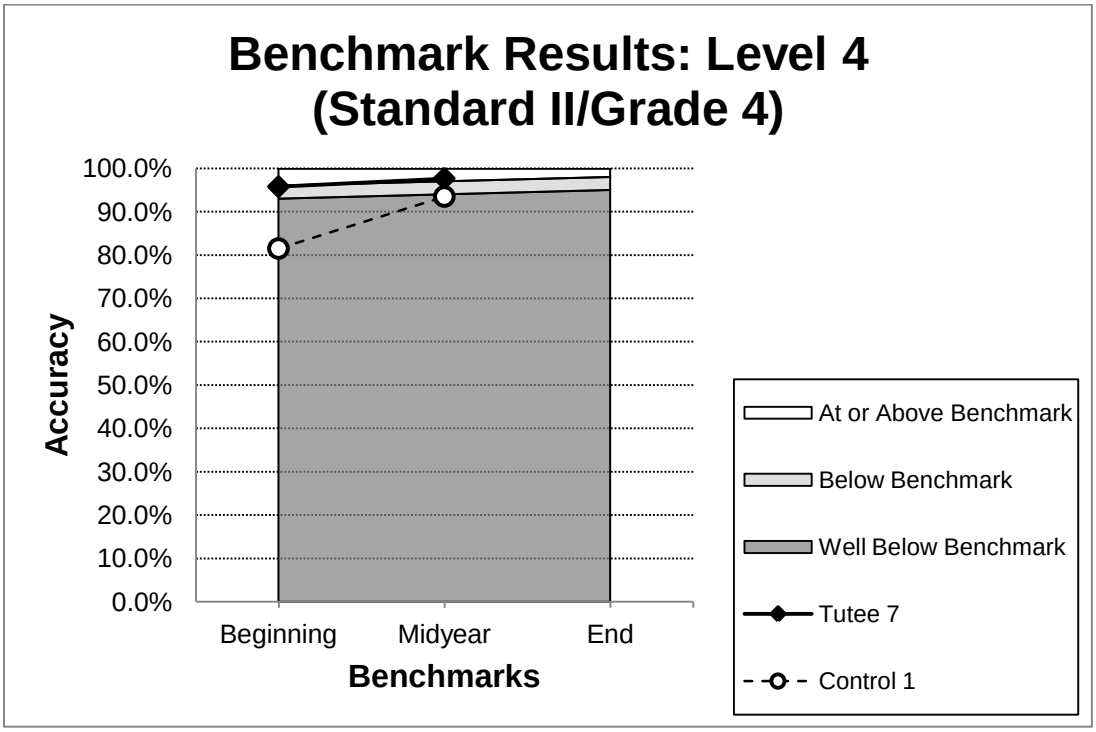

Figure 17. Growth in accuracy from beginning to midyear benchmark *DIBELS Oral Reading Fluency* compared to benchmark goals for students in Standard II.

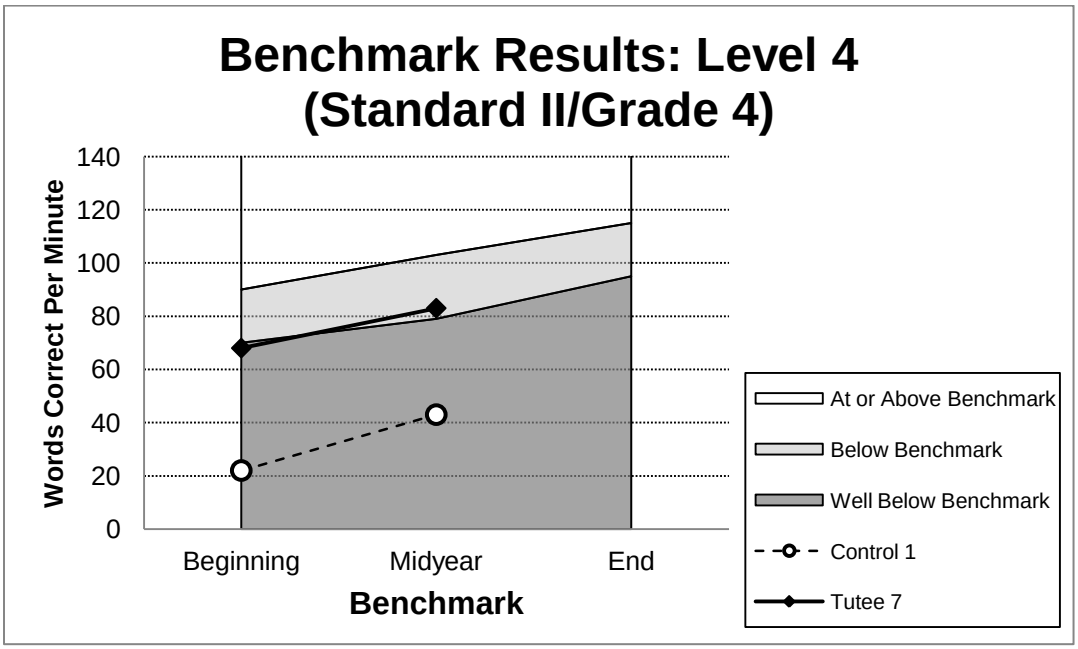

Figure 18. Growth in words correctly read per minute from beginning to midyear benchmark *DIBELS Oral Reading Fluency* compared to benchmark goals for students in Standard II.

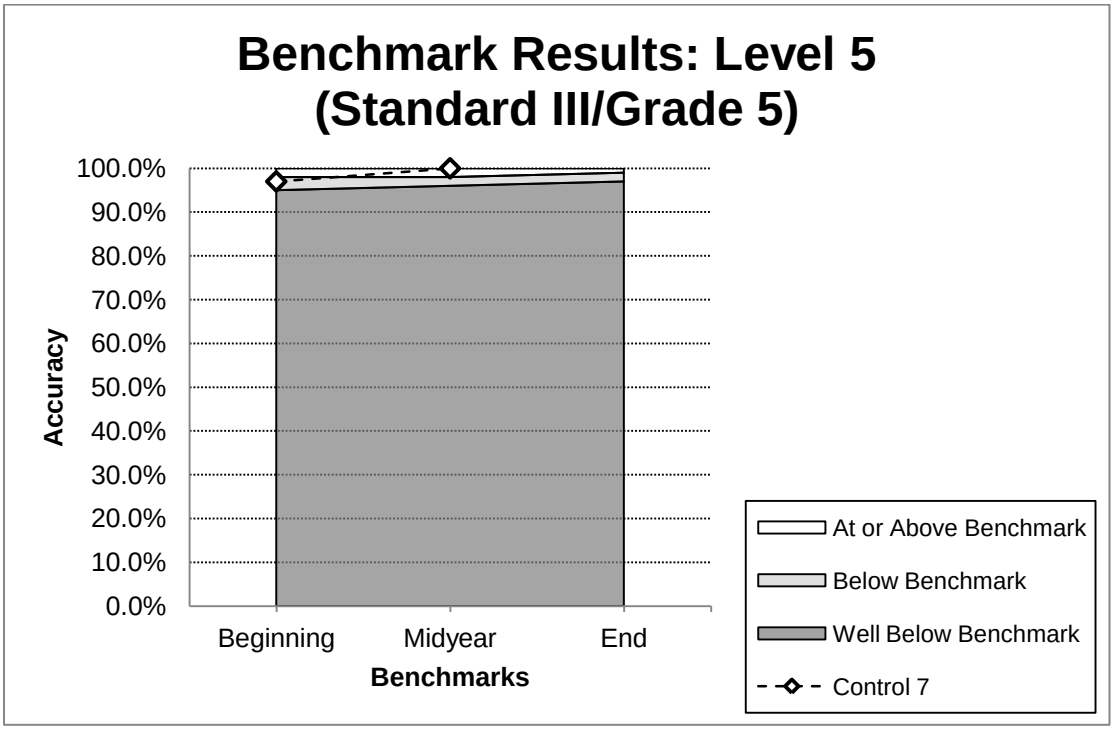

Figure 19. Growth in accuracy from beginning to midyear benchmark *DIBELS Oral Reading Fluency* compared to benchmark goals for student in Standard III.

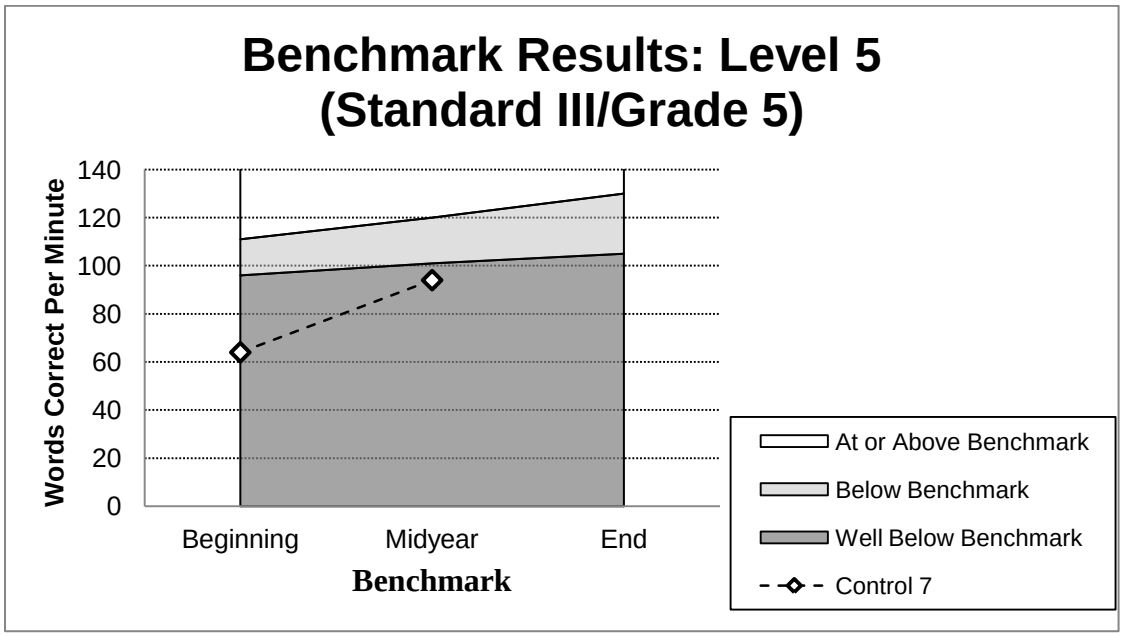

Figure 20. Growth in words correctly read per minute from beginning to midyear benchmark *DIBELS Oral Reading Fluency* compared to benchmark goals for student in Standard III.

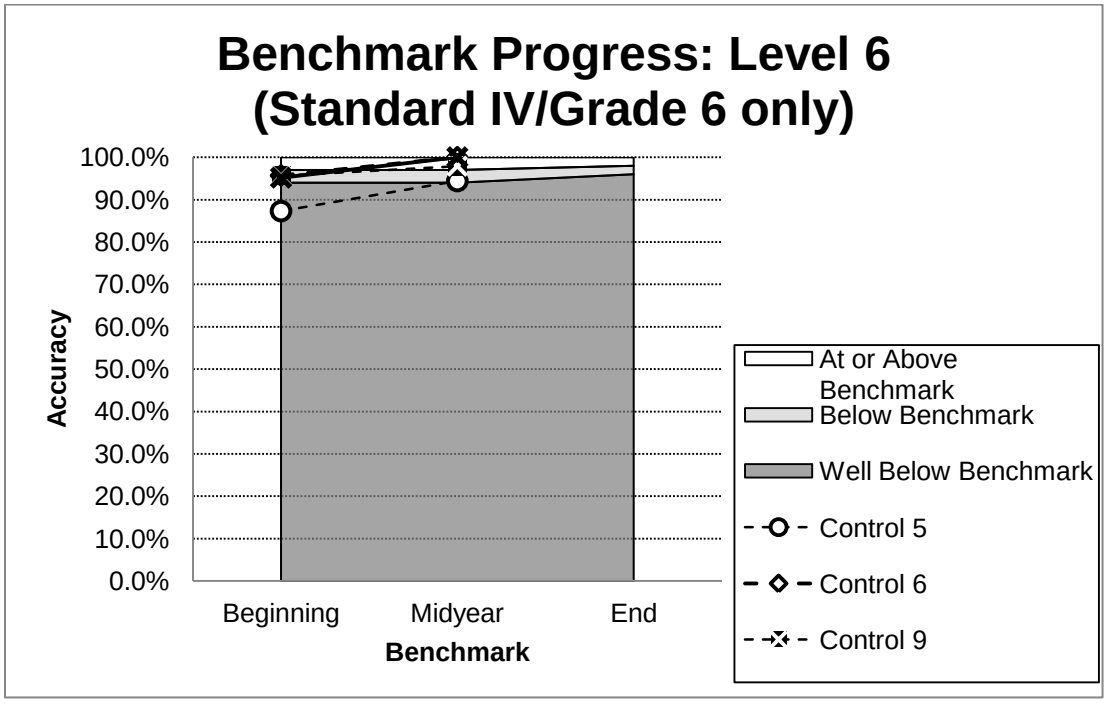

Figure 21. Growth in accuracy from beginning to midyear benchmark *DIBELS Oral Reading Fluency* compared to benchmark goals for students in Standard IV.

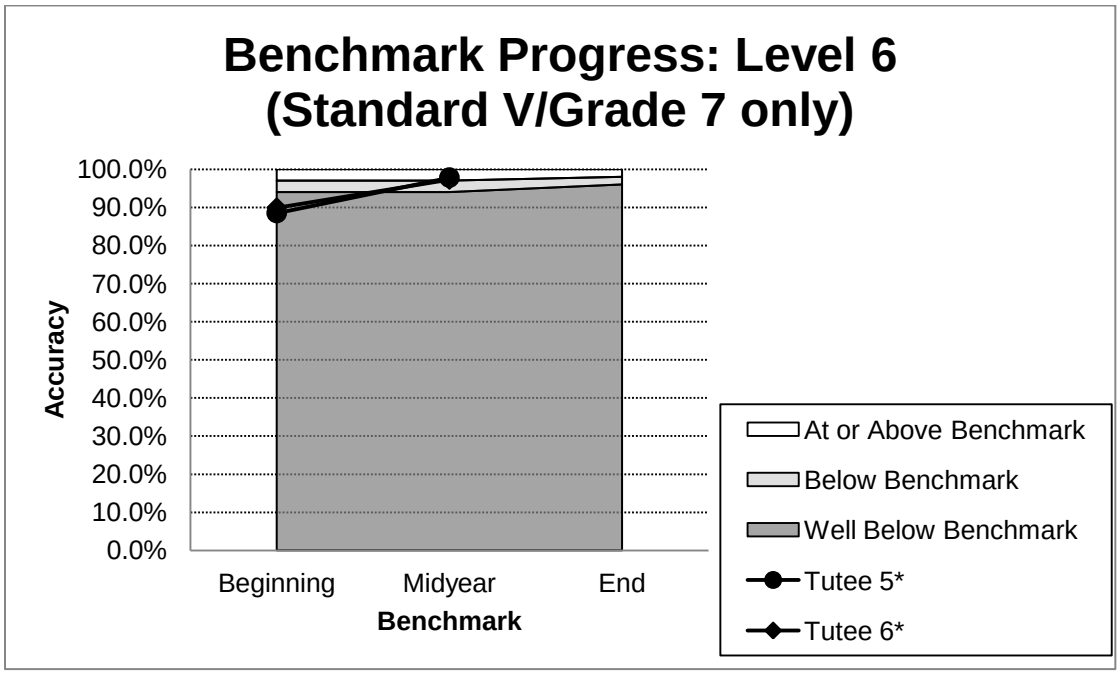

Figure 22. Growth in accuracy from beginning to midyear benchmark *DIBELS Oral Reading Fluency* compared to benchmark goals for students in Standard V.

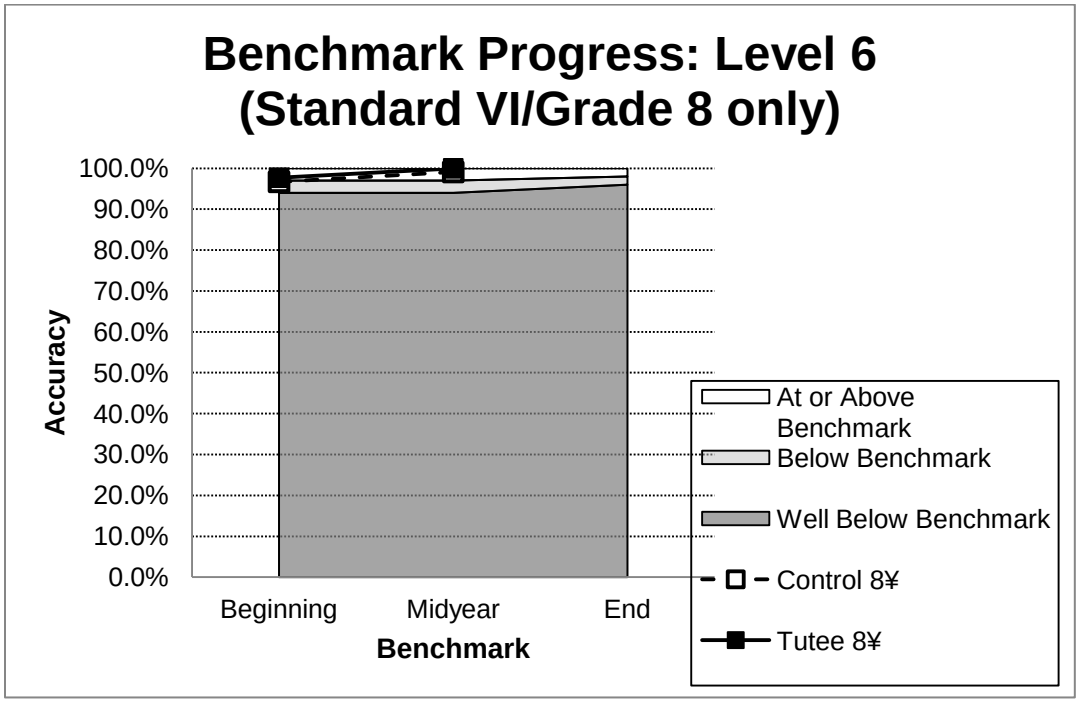

Figure 23. Growth in accuracy from beginning to midyear benchmark *DIBELS Oral Reading Fluency* compared to benchmark goals for students in Standard VI.

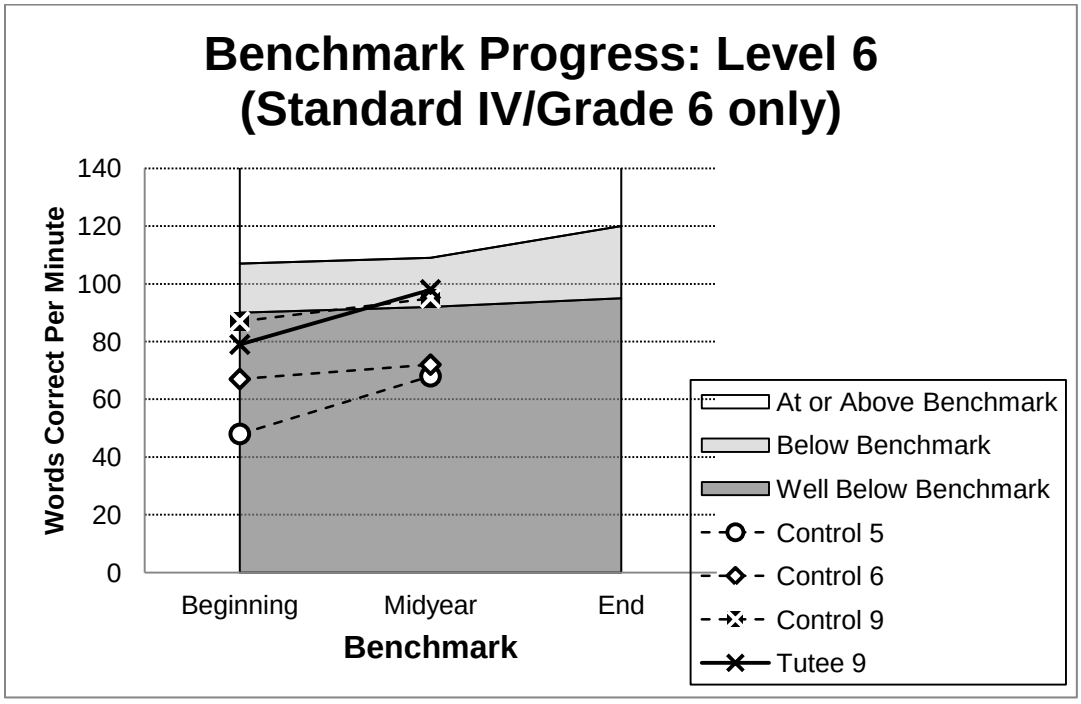

Figure 24. Growth in words correctly read per minute from beginning to midyear benchmark *DIBELS Oral Reading Fluency* compared to benchmark goals for students in Standard IV.

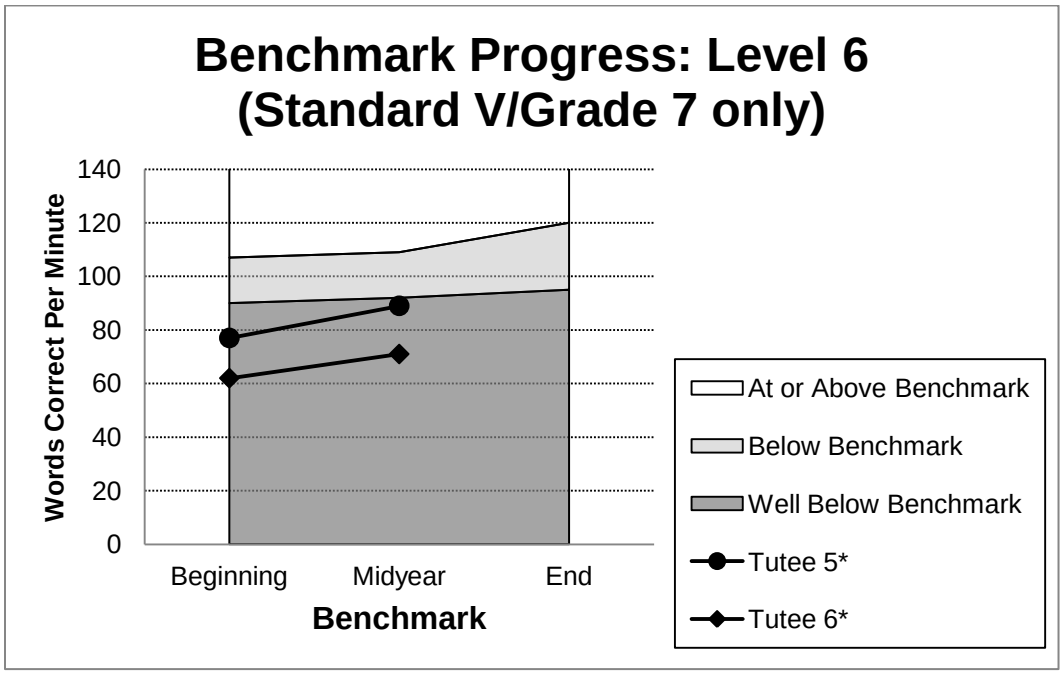

Figure 25. Growth in words correctly read per minute from beginning to midyear benchmark *DIBELS Oral Reading Fluency* compared to benchmark goals for students in Standard V.

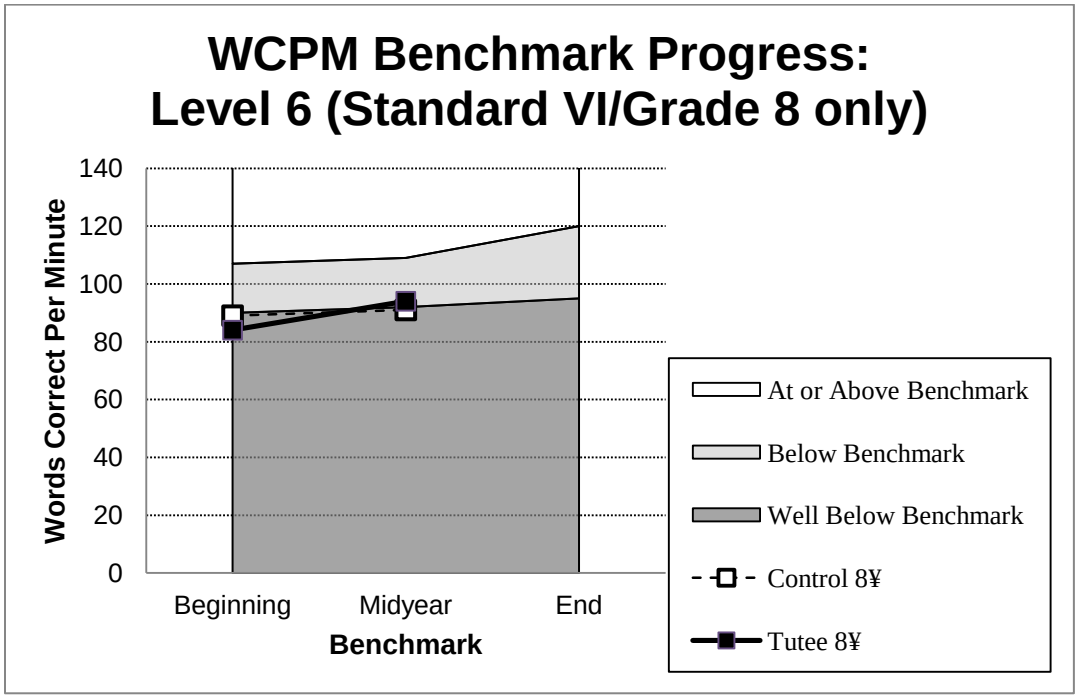

Figure 26. Growth in words correctly read per minute from beginning to midyear benchmark *DIBELS Oral Reading Fluency* compared to benchmark goals for students in Standard VI.

Appendix E: Benchmark Results for Tutors

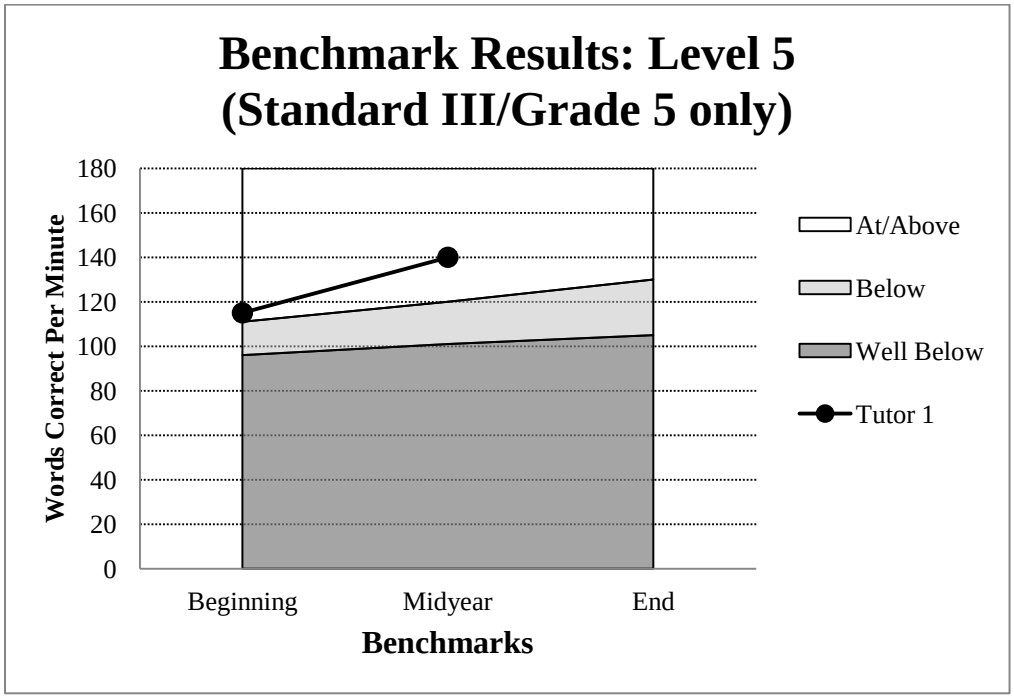

Figure 27. Tutor WCPM benchmark results for Standard III (Grade 5)

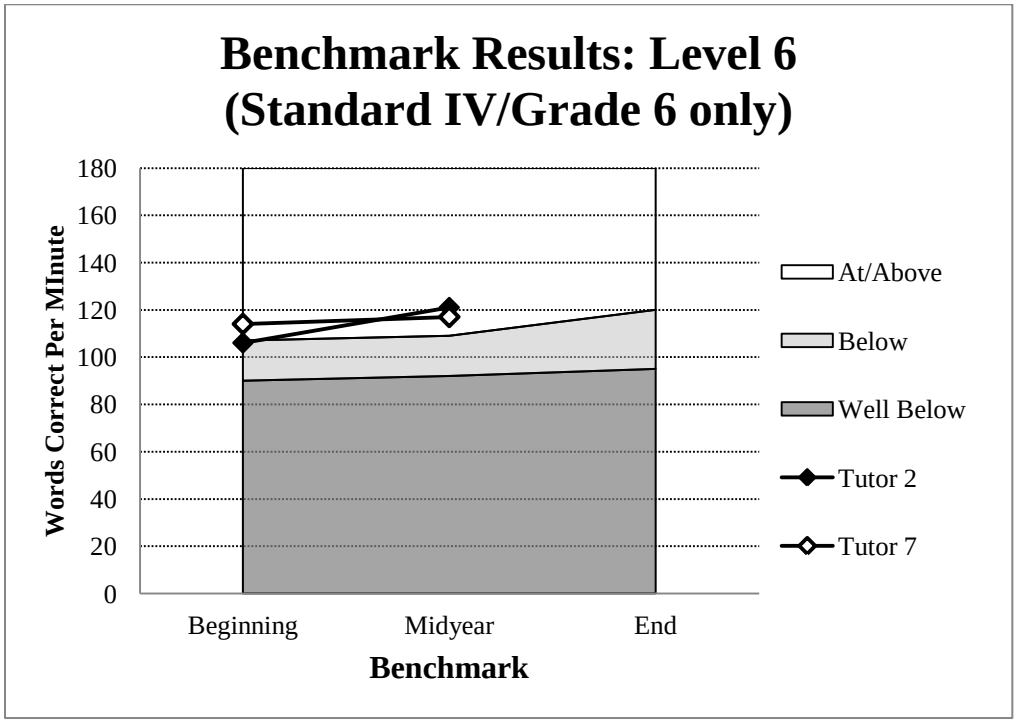

Figure 28. Tutor WCPM benchmark results for Standard IV (Grade 6)

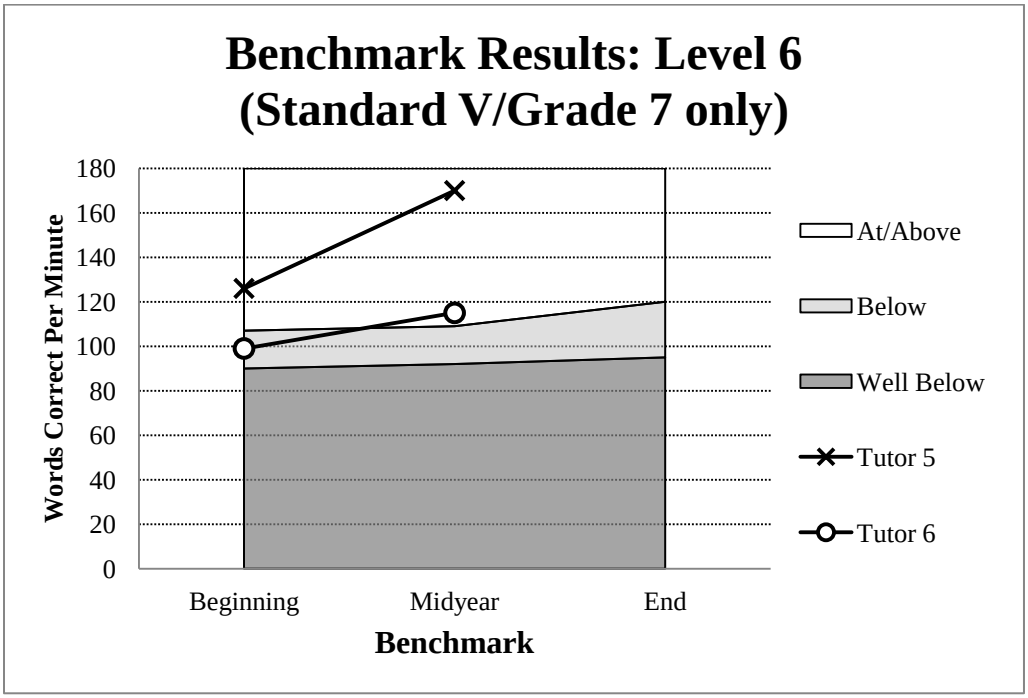

Figure 29. Tutor WCPM benchmark results for Standard V (Grade 7)

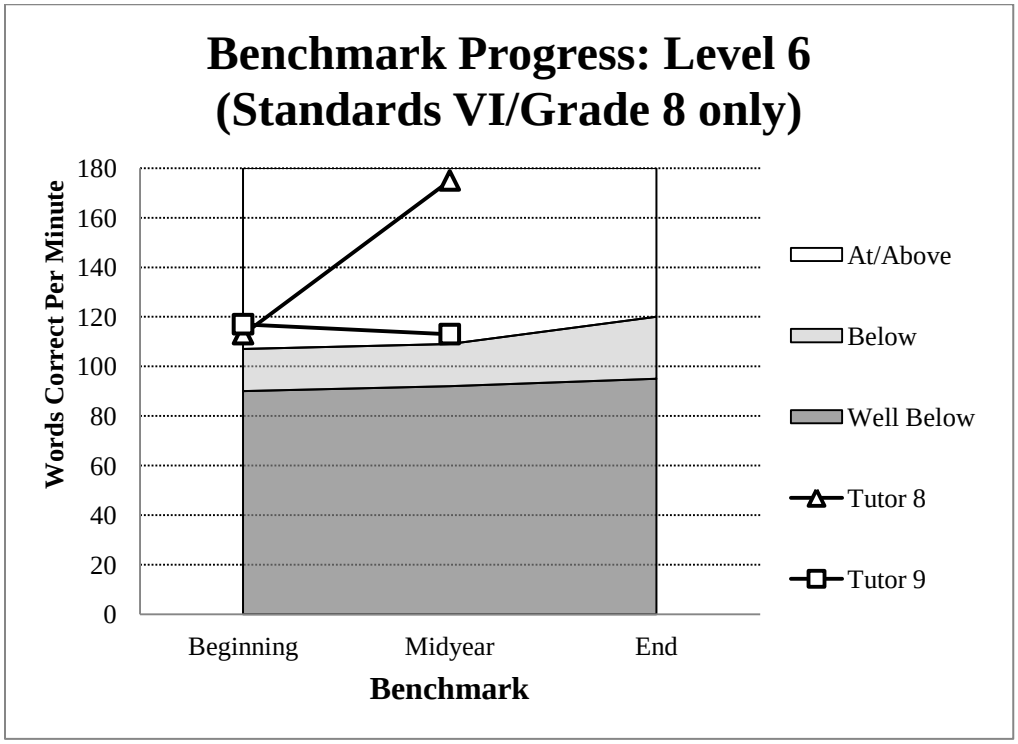

Figure 30. Tutor WCPM benchmark results for Standard VI (Grade 8)

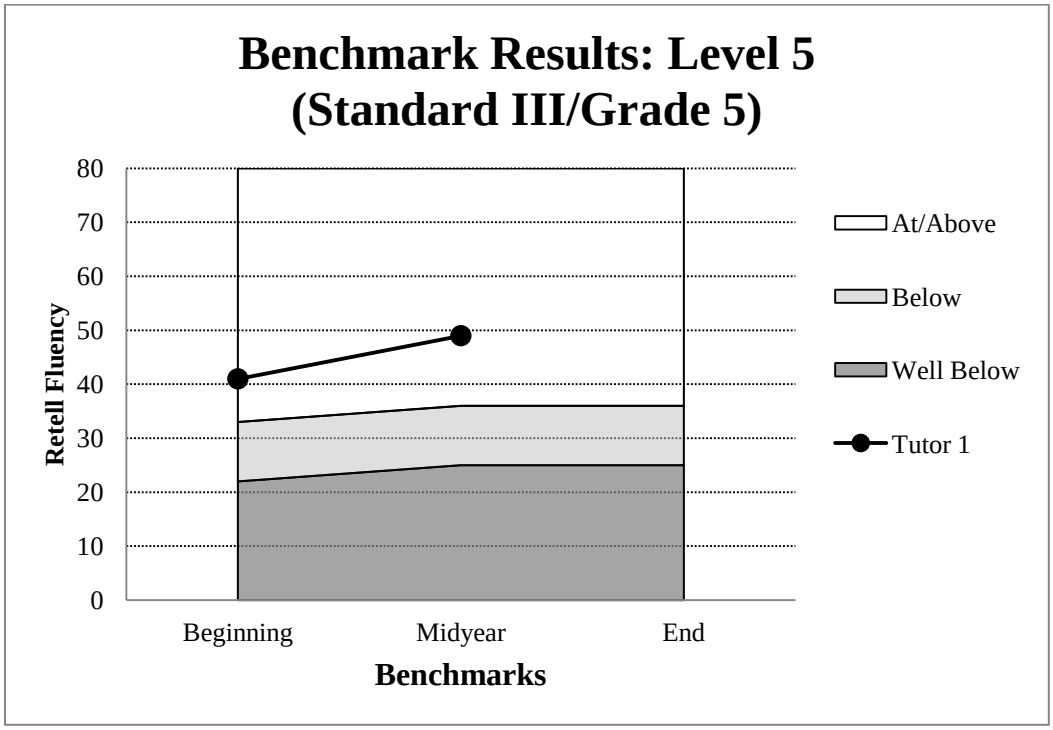

Figure 31. Tutor retell fluency benchmark results for Standard III (Grade 5)

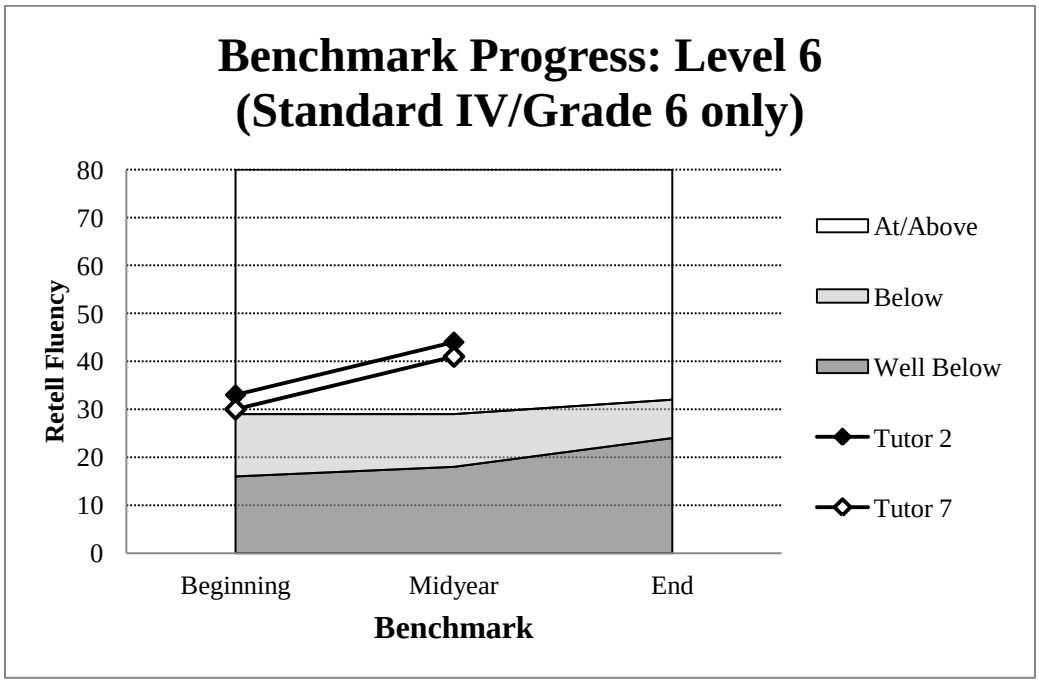

Figure 32. Tutor retell fluency benchmark results for Standard IV (Grade 6)

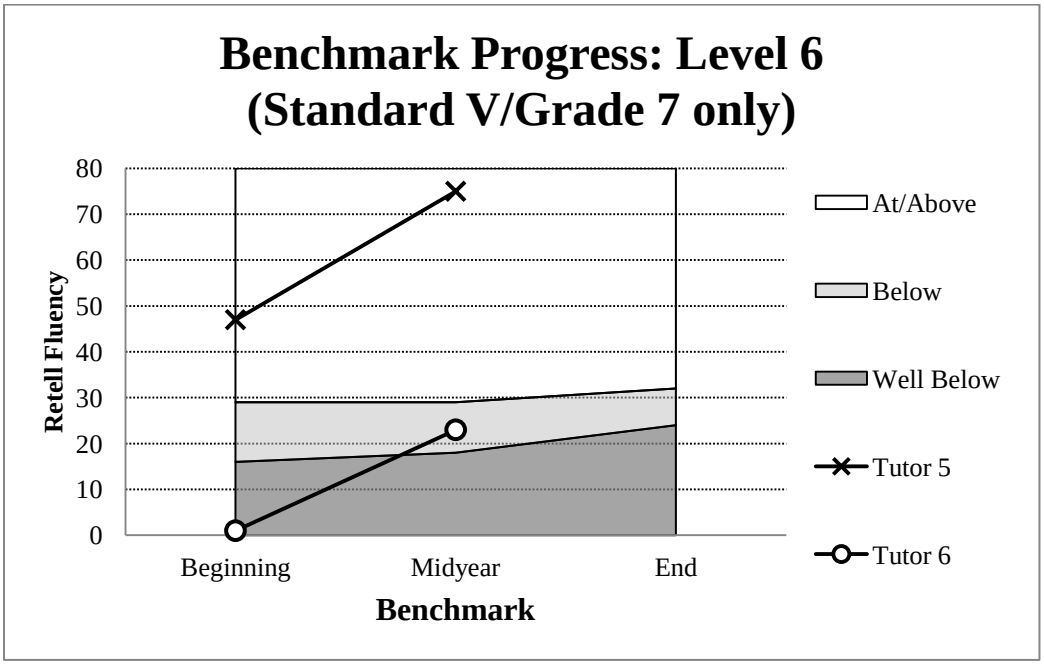

Figure 33. Tutor retell fluency benchmark results for Standard V (Grade 7)

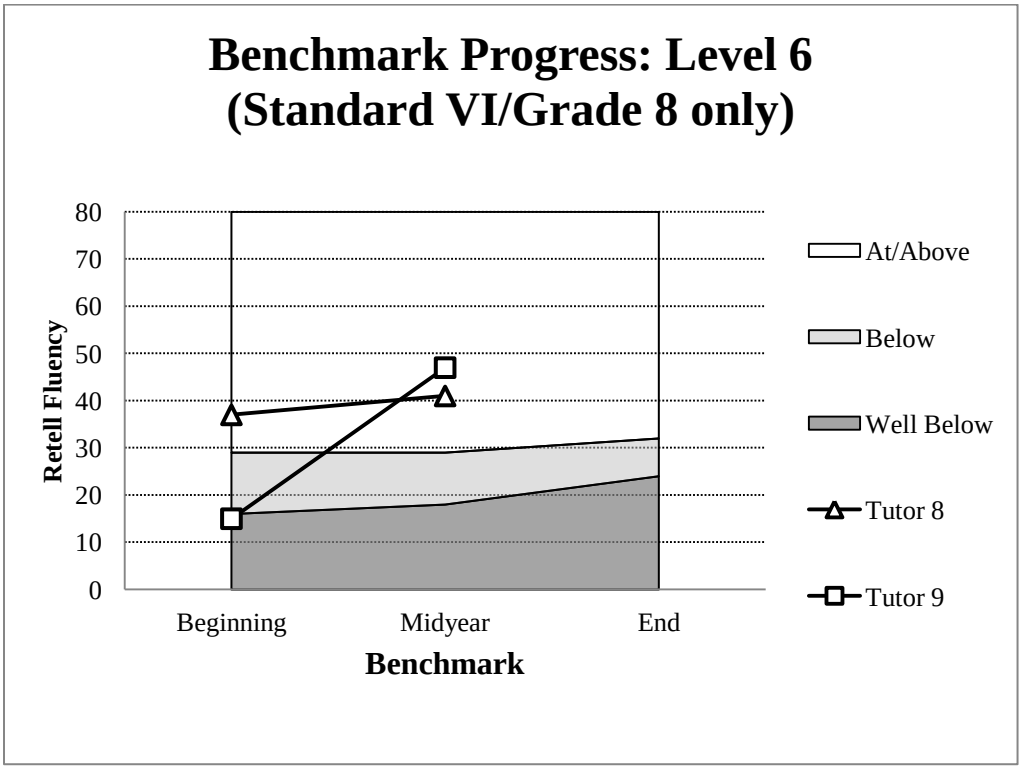

Figure 34. Tutor retell fluency benchmark results for Standard VI (Grade 8)