J Behav Educ (2010) 19:239–256
DOI 10.1007/s10864-010-9111-8

ORIGINAL PAPER

Peer Tutoring for Reading Fluency as a Feasible and Effective Alternative in Response to Intervention Systems

Brad A. Dufrene · Carmen D. Reisener · D. Joe Olmi · Kimberly Zoder-Martell · Marlena R. McNutt · Dana R. Horn

Published online: 14 August 2010

Abstract Peer tutoring is an evidence-based procedure for improving academic performance for a variety of skill areas. The current study evaluated the feasibility and impact of a peer tutoring package for reading fluency with 4 middle school students receiving Tier II remedial supports. This study used a multiple baseline design across participants to evaluate impact of the peer tutoring procedure on students' oral reading rate on instructional passages. Results indicated that students' oral reading rate on instructional probes increased following implementation of the peer tutoring procedure. Moreover, peer tutors implemented most steps of the procedure with a high degree of integrity. Results are discussed in terms of contributions to the peer tutoring and Response to Intervention literatures, as well as application to applied practice.

Keywords Peer tutoring · Treatment integrity · Response to Intervention · Curriculum-based measurement

Introduction

After many years of unreliable, invalid, an largely ineffectual practices related to identifying and serving students with learning disabilities, Ysseldyke (2005) asked whether those practices were "as good as it gets" (p. 125). Fortunately, recent federal legislation (IDEIA 2004; NCLB 2001) has paved the way for a paradigm shift with regard to identification, decision making, and service provision for

B. A. Dufrene (✉) · D. J. Olmi · K. Zoder-Martell · M. R. McNutt · D. R. Horn
The University of Southern Mississippi, 118 College Drive #5025, Hattiesburg, MS 39406, USA
e-mail: brad.dufrene@usm.edu

C. D. Reisener
Department of Counseling and Educational Psychology, 175 President's Circle, 508 Allen Hall, Box 9727, Mississippi State, MS 39762, USA

students with, or suspected of, having learning disabilities. Response to Intervention (RtI) has emerged as a leading service delivery model for providing a continuum of supports to meet the academic and behavioral needs of all students.

RtI systems routinely include three tiers of service delivery (Marston 2005). Tier I includes scientifically based instructional procedures for all students. Additionally, Tier I includes universal screening of all students to identify those who may be at-risk for failure. Tier II includes supplemental instruction for students identified as at-risk for failure during universal screening or through other performance measures. Additionally, in accordance with NCLB and IDEIA, Tier II includes progress monitoring of students receiving supplemental instruction. Progress monitoring allows school personnel to make data-based decisions regarding students' response to instruction. Tier III includes intensive intervention for students displaying poor performance and who fail to respond to supplemental instruction. Like Tier II, Tier III includes progress monitoring that allows for data-based decision making regarding students' response to instruction.

As RtI has evolved, differing approaches to supplemental and intensive interventions have emerged. Some researchers and policy makers have advocated for a standard protocol (SP) approach to intervention (Fuchs et al. 2003), which uses a standardized criterion to determine at-risk status. Then, students at each tier are exposed to evidence-based interventions with only minimal analysis of their skill deficits. SP interventions provide comprehensive instruction for a given skill area. For example, all first-grade students identified as at-risk for reading failure, regardless of their specific skill deficit, might receive a Tier II supplemental instruction package that targets each of five critical reading areas (i.e., phonics, phonemic awareness, fluency, vocabulary, comprehension) in a series of lessons delivered over an extended period of time. The advantages of this approach include (a) reduced decision making on the part of school personnel, (b) training interventionists to conduct a small number of interventions correctly and to assess the accuracy of implementation, (c) large numbers of students can participate in generally effective treatment protocols, and (d) potentially greater quality control. The SP approach includes scripted intervention protocols to ensure greater integrity of instruction. One major limitation of this approach, however, is inflexibility in terms of individualizing interventions based on prior analysis of instructional or environmental conditions and skill deficits (Fuchs et al. 2003).

Recently, SP approaches to service delivery have received increasing attention in the research literature. Most of the research is in the area of reading; however, research in mathematics is emerging. O'Connor et al. (2005) evaluated the effects of an SP approach to reading intervention within an RtI system for a cohort of students in kindergarten through third grade. Participants included approximately 100 students per grade with nearly 45% of students receiving free or reduced-price lunch and the majority of students of European-American decent. Tier II reading instruction included an SP approach implemented in small groups three times per week, while Tier III included an SP approach implemented daily to individual students or in groups of two. Results from this evaluation indicated that 89% of students receiving Tier II supports did not require additional intensive intervention. Additionally, only 8% of students who received Tier III intensive interventions

received a special education ruling. Moreover, students who received three-tiered service delivery experienced moderate to large gains in decoding, word identification, fluency, and comprehension compared to students in the control group. Additional research by O'Connor et al. (2005) found similarly promising evidence for the SP approach as students who experienced the three-tiered system evidenced overall improvements in reading, improved reading for at-risk students, and decreases in identification of students with learning disabilities in reading. Across both studies, the authors provided limited information regarding implementation of the SP approach and level of support required for program implementation.

As use of RtI systems has increased, researchers have focused on instructional and intervention procedures designed to improve oral reading fluency (i.e., reading rate, accuracy, and prosody). Reading fluency is an important aspect of overall reading performance, as fluency adequately predicts reading comprehension (Deno et al. 1982; Hintze et al. 2002; Pinnell et al. 1995). Numerous intervention procedures including Listening Passage Preview (LPP) and Repeated Reading (RR) have been demonstrated effective for improving students' reading accuracy, fluency, and comprehension (e.g., Freeland et al. 2000; Skinner et al. 1997; Therrien 2004). LPP includes a model reading a passage aloud to a student as the student follows along reading silently (Skinner et al.). RR procedures involve the student reading and rereading a passage across multiple trials in order to develop fluency for that passage (Sindelar et al. 2002). Several studies have effectively combined LPP and RR components resulting in greater improvements in fluency than either intervention component alone (Daly et al. 1999, 2002; Dufrene and Warzak 2007). Moreover, the National Reading Panel's report on fluency (National Institute of Child Health and Human Development 2000) stated that procedures such as LPP and RR may substantially improve word recognition, reading fluency, and comprehension for a variety of students across several grades.

Despite demonstrated effectiveness, there are at least two concerns regarding use of LPP and RR in RtI systems. First, there is concern that sustained implementation of individualized reading interventions like LPP and RR may not be likely because of limited personnel resources (Klubnik and Ardoin 2010). Second, while there is ample evidence demonstrating the effects of reading fluency interventions on instructed passages, less evidence is available demonstrating their impact on untrained passages. Below is an expanded discussion of these issues.

Noell et al. (2005) indicated that sustained intervention implementation is effortful, requires substantial resources and may compete with other tasks for which teachers are responsible. As a result, it is important to evaluate resource efficient methods for delivering evidence-based intervention procedures in RtI systems. Along those lines, researchers are investigating alternative methods for implementing LPP and RR. For example, recent investigations have evaluated the effectiveness of LPP and RR when implemented in a small group format (Begeny et al. 2009; Begeny and Silber 2006; Klubnik and Ardoin 2010). Results from those initial investigations are promising and suggest that small group administered LPP and RR may be effective for increasing elementary students' immediate reading fluency, as well as performance on retention and generalization passages. Moreover, LPP and RR in small group formats are more resource efficient than individualized protocols, and as a result may be more feasible for RtI systems.

In addition to evaluating LPP and RR in small group formats, a limited number of studies have evaluated peer mediated LPP and RR. Yurick et al. (2006) found that peer mediated RR was effective for improving elementary and middle school students' reading fluency and comprehension for trained passages. Additionally, the authors provided some evidence for improvements for untrained passages. Salend and Nowak (1988) demonstrated the effectiveness of peer implemented LPP on reading accuracy for elementary students with learning disabilities in reading. Dufrene et al. (2006) conducted a study in which elementary and middle school students implemented a reading intervention package that included LPP, RR, and contingent reward. Results indicated that implementation of the peer tutoring package resulted in substantial increases in reading fluency for trained passages. Unfortunately, the authors did not include data for untrained passages. Finally, the Yurick et al. and Dufrene et al. studies demonstrated that trained peer tutors implemented the interventions with moderate to high levels of integrity, which suggests that peer mediated reading fluency interventions may be feasible for use in RtI systems.

Peer tutoring is a cost-effective student-mediated instructional procedure in which student dyads or small learning groups work together on instructional tasks. Peer tutoring positively impacts student outcomes because it produces increases in the time students spend actively engaged with academic tasks (Greenwood et al. 1993). In addition to increased academic engagement, peer tutoring also increases students' opportunities to respond and immediate feedback regarding performance. As a result, peer tutoring offers the potential of individualized instruction with modest demands on educators. Furthermore, peer tutoring is an enormously resource efficient procedure because students are in abundant supply in our schools.

Numerous studies have demonstrated the beneficial effects of peer tutoring in reading on students' reading accuracy, fluency, and vocabulary (Atherley 1989; Dufrene et al. 2006; Ezell et al. 1997; Mathes and Fuchs 1994; Mathes et al. 1999; Salend and Nowak 1988). Peer tutoring procedures for reading have included a variety of intervention components. For example, peer tutoring studies have included RR (Yurick et al. 2006), LPP (Wright and Cleary 2006), contingent reward, and performance feedback (Dufrene et al.). Peer tutoring for reading is effective for students with and without disabilities (Mathes and Fuchs 1994; Sutherland and Snyder 2007; Wright and Cleary) across a wide range of ages. Moreover, peer tutoring for reading has been demonstrated effective for ethnically and linguistically diverse students (Calhoon et al. 2006).

As stated previously, peer tutoring is a relatively resource efficient procedure as students are abundant in schools. Moreover, it appears that students may be relied on to accurately and consistently implement academic interventions. Dufrene et al. (2005) evaluated the extent to which students implemented a peer tutoring procedure for mathematics that included curriculum-based measurement (CBM). Results indicated that 32 of 37 students implemented the tutoring procedure with acceptable integrity. Additionally, for the five tutors with unacceptable integrity, performance feedback substantially improved their implementation. In another study, Dufrene et al. (2006) evaluated the extent to which trained peer tutors implemented a reading intervention package that included LPP, RR, contingent

reward, CBM, and performance feedback. Additionally, effects on tutees' oral reading fluency were evaluated. Results indicated that tutors implemented the package with moderate to high integrity and tutees benefited in terms of increased oral reading fluency. In addition to the work done by Dufrene and colleagues (2005, 2006), others have found similarly promising results with regard to students' accurate and consistent implementation of academic interventions (e.g., Hughes and Fredrick 2006; Yurick et al. 2006). Given recent research demonstrating quality intervention implementation by trained peer tutors, peer tutoring appears to hold promise as an alternative intervention delivery model for use in RtI systems.

This study was designed to inform the RtI literature with regard to the degree to which peer tutors could (a) implement a fluency-based intervention with integrity and (b) reliably conduct assessments on trained passages and untrained progress monitoring probes using CBM procedures. Furthermore, the study was designed to replicate previous research demonstrating the positive effects of LPP and RR on students' oral reading rate. Based on previous literature, it was believed that students would implement the reading fluency intervention package with at least a moderate degree of integrity (i.e., 60% or greater) and they would be able to accurately monitor other students' reading progress using CBM. Again, previous research (Dufrene et al. 2005, 2006) indicates that the trained students may implement peer tutoring procedures with integrity while monitoring their peers' progress using CBM. Additionally, it was hypothesized that students would respond positively to a peer mediated reading intervention designed to increase reading fluency. Specifically, it was believed that student's reading fluency would increase for instructional passages following implementation of the peer mediated intervention package.

Method

Participants and Setting

Seven students from the rural southeastern United States participated in this study. Three students (Chelsea, Jason, and Tessie) served as tutors and four were tutees (Jan, Donna, Darian, and Tanya). All participants were in the 6th grade and were between 11 and 12 years of age. None of the participants indicated any prior experience with peer tutoring, and none received special education services. Tutors Chelsea and Jason were identified as gifted under their state education agency's criteria for gifted. The four students receiving peer tutoring were identified for Tier II intervention services within their school's RtI system based on universal screening using the Dynamic Indicators of Basic and Early Literacy Skills (DIBELS; Good and Kaminski 2002). Specifically, 6th grade students scoring below the 15th percentile for the DIBELS Oral Reading Fluency subtest were selected for Tier II intervention. Chelsea provided tutoring to Jan and Donna. Jason provided tutoring to Tanya. Tessie provided tutoring to Darian. All sessions were conducted in a quiet location outside of the classroom depending on availability. Specifically, some sessions were conducted in the school's library and other sessions were conducted in an empty classroom.

Materials

Instructional and progress monitoring reading passages were selected from the DIBELS (Good and Kaminski 2002) and AIMSweb (Edformation 2003) sets of reading passages, respectively. Sixth-grade passages were used for all procedures in the study. Reliability and validity of DIBELS oral reading fluency passages are high, with test–retest reliability coefficients ranging from .92 to .97 and alternate form reliability coefficients ranging from .89 to .94. Criterion-referenced validity ranged from .52 to .91 (Good and Kaminski). Test–retest reliability for AIMSweb probes was .92, alternate form reliability was .89, and inter-rater agreement was .99, using students from grades 1st through 6th as the participant pool. Criterion-referenced validity of the AIMSweb probes was .91 for Ginn and Scott-Foresman basal readers as the criterion measure, using a participant pool of 1st through 6th graders (Shinn and Shinn 2002). DIBELS and AIMSweb probes were randomly selected prior to each instructional or progress monitoring session.

In addition to instructional probes, tutors and tutees were provided with all materials necessary for conducting a peer tutoring session. Additional tutoring materials included a digital audio recorder, digital timer, pencils, and data recording forms. Finally, tutors were provided with tickets (from their school's Positive Behavior Interventions and Supports [PBIS] Tier I system) that were used as rewards for tutees' participation in peer tutoring. Tutees could exchange tickets at the end of each week for access to a variety of backup rewards (e.g., entry into weekly lottery, school supplies, access to preferred activities) as part of the school's PBIS system.

Tutor Training

Students were taught to act as tutors prior to implementation of the peer tutoring procedure using modeling, numerous opportunities for practice, feedback, and cumulative reviews of the skills presented. Researchers provided training and used training protocols to ensure consistency. A detailed protocol for tutor training for this peer tutoring package is provided by Dufrene et al. (2006).

Tutor training was conducted in two sessions on consecutive days. On the first day, tutors were taught to conduct all peer tutoring components in isolation. First, tutors were taught to use the digital timer. Next, they were taught to use the digital audio recorder. Afterward, tutors were trained to read at a comfortable rate (i.e., approximately 120 words per min). Subsequently, tutors were trained to score instructional probes for words correct per min (WCPM) and errors per min (EPM) using standard CBM procedures (Shinn 1989). Tutors learned to listen to other students read while simultaneously recording errors. Then they were taught to count the number of correct and incorrect words read and to record WCPM and EPM at the top of the reading probe. After learning to accurately score reading probes, tutors were taught to record scores on the data recording form. Tutor training for each tutor continued until the tutor was able to complete the following tutoring steps independently and accurately: (a) correctly use the digital recorder and timer, (b) appropriately model passages, (c) score a reading probe with 80% or greater

agreement with an adult scorer, (d) accurately record the score, and (e) accurately record the score on the data recording form. Experimenters used a checklist (available from the first author) to record tutors' accuracy for each component of the peer tutoring procedure.

On the second day of training, tutors implemented the peer tutoring procedure with a researcher acting as a tutee. Tutors were allowed to implement peer tutoring after completing one session with 100% integrity (see description of tutor integrity checklist later). All tutors completed tutor training in 2 days.

Peer Tutoring Procedure

The peer tutoring procedure included LPP, RR, performance feedback to tutee, and reward for participation. Reading passages and data recording forms were placed in the tutors' classrooms each morning by an experimenter at the beginning of school. Reading passages and the data recording form were placed in a three ring binder with an illustrated cover made by the tutor and tutee prior to the beginning of data collection. Additionally, a plastic container that included the audio recorder, digital timer, pencils, and PBIS tickets was left in each tutor's classroom so that they could gather materials immediately prior to each peer tutoring session.

Tutoring sessions for each dyad were scheduled by the tutor and tutee's teacher during a non-instructional time (e.g., study hall) that was deemed feasible by the teacher. Once the tutor and tutee arrived at their designated tutoring location, the session began. First, the tutor pressed "record" on the audio recorder. Next, the tutor conducted the LPP component of the tutoring package. After this, RR was implemented in which the tutee read the same passage aloud until completion on two consecutive trials while the tutor corrected any errors that were made (i.e., told the tutee the correct word). After two complete readings, the tutee read the same passage one more time for 1 min while the tutor scored the probe for WCPM and EPM. During the 1 min probe, if a tutee paused on a word for 3 s, the tutor provided the word. However, the tutor did not correct errors. At the conclusion of 1 min, the tutor recorded WCPM and EPM at the top of the probe and on the data recording form. Additionally, the tutor shared the score with the tutee. Next, the tutor gave the tutee a PBIS ticket for participating in the session. During training, tutors were instructed to provide tutees with a ticket if they completed all aspects of the peer tutoring session. Moreover, tutors earned tickets provided by an experimenter for conducting the session. Finally, tutors were instructed to press "stop" on the audio recorder at the conclusion of the session. Audio recordings from sessions were analyzed at the completion of the study and those data indicate that average session length for all dyads was approximately 9.5 min.

Progress monitoring was conducted approximately once per week by tutors using an untrained progress monitoring probe. As stated previously, progress monitoring probes were randomly selected from the 6th grade AIMSweb reading progress monitoring set. Progress monitoring probes were administered by tutors using standard CBM procedures in a quiet location outside of the classroom. During progress monitoring sessions, tutors did not correct errors; however, if a tutee paused on a word for 3 s then the tutor provided the word to the tutee. Progress

monitoring sessions occurred immediately after peer tutoring sessions on those days designated as progress monitoring days. The presence of a progress monitoring probe inside a peer tutoring binder signaled to the tutor that progress monitoring was to occur on that day.

Dependent Measures and Data Collection

Dependent measures included tutor integrity scores along with WCPM and EPM for instructional and progress monitoring probes. Tutor integrity was defined as the percentage of correctly completed peer tutoring steps. Tutor integrity was scored based on review of permanent products and audio recordings. A 10-item checklist (available from the first author) was used to record the number of correctly completed tutoring steps. Percentage of steps completed was calculated by dividing the number of tutoring steps completed accurately by the total number of steps. Peer tutoring steps were based on instructional relevance and amenability to permanent product measurement. The tutoring steps were (a) tutor audio recorded the entire tutoring session, (b) tutor read passage aloud to the tutee, (c) tutee practiced reading the passage aloud three times, (d) tutee read the entire passage aloud for the first two trials, (e) tutor corrected at least 80% of tutee errors during the two practice trials, (f) tutor administered 1-min timed probe, (g) tutor marked errors on probe, (h) tutor wrote correct WCPM on probe, (i) tutor recorded tutee's WCPM and EPM on data recording form, and (j) tutor provided PBIS ticket to tutee if tutee completed all session requirements. Administration of progress monitoring probes was not included as an item in the procedural integrity checklist for the peer tutoring package. However, progress monitoring probes were included in the probes selected for evaluation of interscorer agreement.

Tutor integrity data were collected for 73.68, 100, 44.44, and 100% of the tutoring sessions for Jan, Donna, Darian, and Tanya, respectively (see Table 1). It was not possible to code integrity for all sessions for Jan and Darian because several audio files from their digital recorders were inaudible and corrupted and could not be transferred to a computer. As a result, tutor integrity data were available for 79.53% of the total tutoring sessions across all dyads. Chelsea failed to record one of Donna's sessions in its entirety (i.e., session 6). During that session, Chelsea began recording the session after peer tutoring had already begun. For that session, any peer tutoring component that could not be scored by permanent product measurement was scored as not occurring as there was no evidence to document implementation of that step.

Table 1 Tutor integrity data by Dyad

Dyad (tutor/tutee)	Percentage of sessions integrity data collected (%)	Mean integrity (%)	Range (%)
Chelsea/Jan	73.68	82.86	40–90
Chelsea/Donna	100	84	80–90
Tessie/Darian	44.44	87.5	80–90
Jason/Tanya	100	85	70–100

WCPM and EPM were scored by tutors immediately after each instructional and progress monitoring session. For the instructional passages, the score that was recorded by the tutor for each tutee's session was the score from the third reading trial (i.e., 1-min timed probe from instructional passage). A word was scored as read correctly if it was correctly pronounced within 3 s. Errors included omissions, substitutions, misread words, and pauses that lasted more than 3 s. WCPM and EPM reported by the tutor (i.e., score recorded on the probe and data recording form) were used for data entry.

Experimental Design

A multiple baseline design across participants was used to evaluate the effectiveness of the peer tutoring procedure on tutees' oral reading fluency. A multiple baseline design was selected due to the small number of tutees and because of the potentially irreversible nature of students' learning following implementation of the peer tutoring package. Visual analysis of level, trend, and variability was used to inspect data. Baseline for all students began on the same school day. During baseline, experimenters administered reading fluency probes to students using standard CBM procedures. Probes were randomly selected from the 6^{th}-grade DIBELS reading set. Baseline data were collected until stability or a downward trend was obtained. During baseline, tutees were not provided with corrective feedback regarding errors nor were WCPM or EPM reported to the student. Following baseline, the peer tutoring package was implemented on the next school day in a staggered fashion across tutees according to conventions of the multiple baseline design. Teachers were instructed to allow dyads to conduct peer tutoring sessions on Monday, Wednesday, and Friday. However, practice for state-wide testing and other activities sometimes resulted in teachers not allowing dyads to conduct peer tutoring. Jillian received three tutoring sessions per week, Donna received an average of 2.1 tutoring sessions per week, Darian received an average of 1.7 tutoring sessions per week, and Tanya received an average of 2 tutoring sessions per week.

Interscorer Agreement

An experimenter independently scored 30.4% of the instructional and progress monitoring probes from peer tutoring sessions for interscorer agreement. Agreement between tutors and the experimenter for instructional and progress monitoring probes was calculated on a word-by-word basis by dividing the number of agreements by the number of agreements and disagreements and multiplying by 100. Interscorer agreement for WCPM was 99% (range = 98–100%). Another experimenter collected interscorer agreement data for 36.75% of the integrity checklists that were completed. Interscorer agreement for integrity checklists was calculated by dividing the number of agreements by the number of agreements and disagreements and multiplying by 100. Interscorer agreement for integrity checklists was 100% for all agreement checks.

Results

Tutor integrity data, by dyad, are presented in Table 1. It was hypothesized that tutors would implement the intervention package with moderate to high integrity. Tutor integrity data indicate that tutors generally implemented the peer tutoring package with high integrity overall ($M = 84.84\%$), with the lowest mean integrity score across tutors being 82.86%.

Table 2 shows the mean integrity scores for each of the individual peer tutoring steps. Seven of the 10 steps were implemented with 90% or greater integrity. However, implementation for two steps in particular was substantially below implementation for all other steps. Specifically, tutors performed most poorly for requiring tutees to read the passage three total times and correcting 80% or more of tutor errors. Mean integrity scores for those steps were 32.5 and 57.32%, respectively. Evaluation of audio recordings indicated that on many occasions tutors required their tutee to read the entire instructional passage twice while timing the first min of the second reading to obtain a 1-min score for that session. With regard to correcting errors, tutees' errors occurred at relatively low rates throughout the study, so when tutors missed an opportunity to correct an error it often resulted in them failing to correct 80% or more of total tutee errors. All tutors administered, timed, and scored the CBM probes and provided the PBIS ticket for participation with 100% integrity across all sessions.

Tutees' performance for WCPM and EPM is illustrated in Fig. 1. It was predicted that tutees would increase their oral reading fluency on instructional passages following intervention and would show an increasing trend on untrained progress monitoring probes throughout the intervention phase. Visual inspection of the WCPM and EPM data suggests that all four tutees demonstrated improvements in

Table 2 Mean integrity scores for each peer tutoring step by Dyad

Tutoring Dyad	Chelsea/ Jan (%)	Chelsea/ Donna (%)	Tessie/ Darian (%)	Jason/ Tanya (%)	Mean for step across dyads (%)
Tutoring step					
Step 1—Recorded entire session	100	90	100	100	97.5
Step 2—Read passage to tutee	92.85	100	100	70	90.7
Step 3—Tutee read passage three total times	0	10	100	20	32.5
Step 4—Tutee read entire passage aloud two times	78.57	70	50	90	72.13
Step 5—Corrected 80% > errors	64.29	70	25	70	57.32
Step 6 Administered 1 min timed probe	100	100	100	100	100
Step 7—Marked errors on probe	100	100	100	100	100
Step 8—Wrote WCPM on probe	100	100	100	100	100
Step 9—Recorded scores on data recording form	100	100	100	100	100
Step 10—Provided ticket	100	100	100	100	100

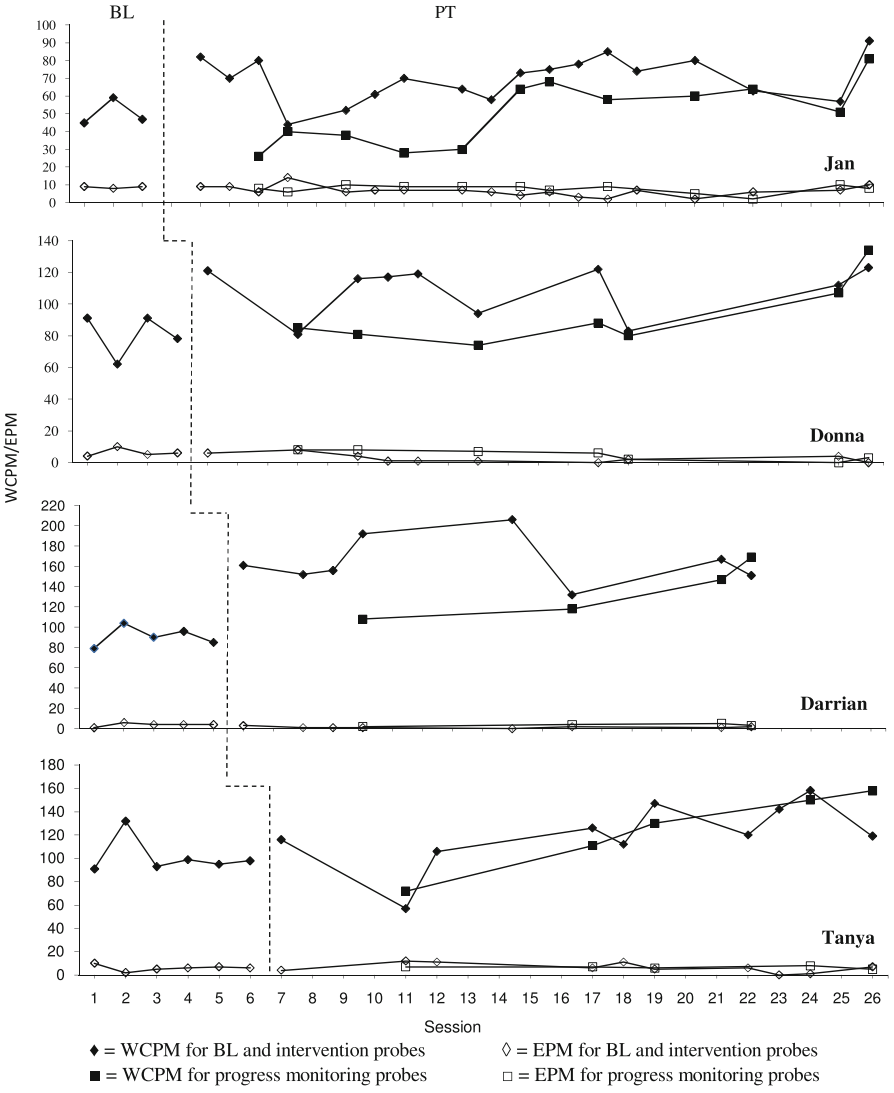

Fig. 1 Words correct per minute and errors per minute by tutees. *WCPM* words correct per minute, *EPM* errors per minute, *BL* baseline, *PT* peer tutoring

oral reading rate following intervention implementation on the trained passages, and three students showed an increasing trend on the progress monitoring probes.

Jan

During baseline Jan's WCPM was somewhat stable with a mean of 50.33 WCPM and 8.67 EPM. Following intervention implementation, WCPM for instructional probes immediately increased to 87 WCPM. After the initial substantial increase in

WCPM for instructional probes, performance was variable with some sessions resulting in WCPM that were below levels observed during baseline. However, mean WCPM on instructional passages across all intervention sessions was 69.53 WCPM, which was greater than the level observed during baseline. Moreover, EPM decreased from a mean of 8.67 during baseline to 6.42 during intervention. Jan's WCPM for progress monitoring probes was initially low (i.e., 26 WCPM); however, over the course of the intervention phase Jan's performance for progress monitoring probes increased such that by the end of the intervention phase scores on progress monitoring probes yielded a mean of 50.67 WCPM. Mean EPM for progress monitoring passages was 7.67.

Donna

During baseline, Donna's WCPM were relatively stable with a mean of 80.5 WCPM and 6.25 EPM. Following intervention implementation, WCPM for instructional probes immediately increased to 121 WCPM. Over the course of the intervention phase, Donna's WCPM for instructional probes was variable but never dropped below baseline levels. Further, WCPM for instructional probes ended on an upward trend. Donna's mean scores on instructional passages were 107.67 WCPM and 2.89 EPM. Donna's WCPM for progress monitoring passages initially demonstrated a slight downward trend. Then, performance began to trend upward, and mean WCPM for progress monitoring passages was 91 with 4.86 EPM.

Darian

During baseline, Darian's reading fluency was stable with mean scores of 90.8 WCPM and 3.8 EPM. When intervention was implemented, her reading performance on instructional probes immediately and substantially increased in level to 161 WCPM, and performance for instructional probes remained above those observed during baseline for the duration of intervention (mean = 164.63). Moreover, EPM decreased to a mean of 1.38 for instructional probes. Darian's WCPM for progress monitoring trended upward for the four progress monitoring sessions conducted, and mean WCPM was 135.5. EPM for progress monitoring probes remained stable and low over the course of implementation with a slight decrease to a mean of 3.5.

Tanya

During baseline, Tanya's reading performance was relatively stable, and mean WCPM was 101.33 while mean EPM was 6. Following intervention implementation, reading fluency for instructional passages immediately increased to 116 WCPM. Although variable over the course of the study, Tanya's reading fluency for instructional passages included a mean of 120.3 WCPM and a mean of 6.3 EPM. WCPM for progress monitoring passages trended upward throughout the course of progress monitoring and resulted in a mean of 124.2 WCPM with 6.6 EPM.

Discussion

RtI calls for an increased focus on a continuum of academic supports for students who are experiencing academic difficulties. Specifically, students who fail to make adequate progress are immediately provided with high efficiency supplemental supports or intense individualized supports depending on the severity of academic difficulties. As a result, schools must become more creative with regard to identifying resources for providing a continuum of academic supports. Results from this study are encouraging in light of the increased need for academic intervention resources. First, this study provides preliminary support for the use of students as primary Tier II intervention agents in an RtI system. Previously, peer tutoring studies indicated that students may accurately and consistently implement a variety of peer tutoring protocols (Dufrene et al. 2005, 2006; Hughes and Fredrick 2006; Yurick et al. 2006). This study provides a unique contribution to the RtI and peer tutoring literatures in that trained peer tutors accurately implemented an intervention package that served as tutees' Tier II intervention in an RtI system. In the current study, peer tutors implemented 8 of 10 steps in a reading fluency intervention package that included LPP, RR, performance feedback, and reward for participation with a high level of integrity (i.e., greater than 80% for all tutors). This is encouraging for schools as they seek novel resources for intervention implementation in an RtI framework.

Tutor integrity findings from this study are consistent with those of Dufrene et al. (2006). Specifically, tutors in Dufrene et al. were trained in a similar manner and implemented a nearly identical peer tutoring protocol with high overall integrity. However, tutors in that study failed to consistently correct tutee errors. In the current study, a frequent integrity breakdown was again failure to consistently correct student errors. Therefore, future research should evaluate training and feedback procedures that improve tutors ability to accurately detect and correct tutee errors. Additionally, these findings may highlight the need to provide tutees with instructional level reading passages rather than grade level passages. Use of instructional level passages may have resulted in fewer tutee errors, thereby decreasing the frequency with which tutors would have been required to detect and correct tutees' errors.

In addition to demonstrating tutors' ability to accurately and consistently implement the majority of substantive intervention steps, tutors in this study accurately assessed other students' reading using CBM. As a result, this study extends previous research by Bentz et al. (1990) and Dufrene et al. (2006) by again demonstrating that trained students can monitor their peers' progress in reading using standard CBM procedures. In this study, middle school students were trained to monitor their same grade peers' reading fluency using CBM procedures, and they did so while achieving greater than 90% agreement with an adult scorer. Previous research (Bentz et al., Dufrene et al.) included progress monitoring by cross-age peers with older peers monitoring the reading of peers in lower grades. Results such as these are promising given the progress monitoring requirement of RtI systems. These results are especially important for applied practice in middle school settings that do not have the benefit of teacher assistants who might bear some of the

responsibility for routine progress monitoring duties. Should future studies continue to show that trained students can reliably monitor other students' reading performance with procedures similar to standard CBM procedures, progress monitoring demands on educators may be substantially reduced.

Second, this study extends previous research demonstrating the effectiveness of a peer mediated intervention for increasing students' oral reading rate. Following intervention, all tutees demonstrated increased reading fluency on instructional passages. Additionally, tutees' performance generally trended upward on untrained progress monitoring probes during the intervention phase, albeit less so for some tutees. Unfortunately, we cannot be sure that these effects were primarily a function of intervention given the lack of baseline data on progress monitoring passages, as well as the fact that tutees were also receiving classroom instruction. Dufrene et al. (2006) demonstrated gains for instructional passages, but did not include data for untrained progress monitoring probes.

It is also important to note that the Dufrene et al. (2006) study was conducted during a summer academic camp at a university-based school psychology clinic. In effect, that study was conducted under conditions that might be considered analogue. The current study was conducted in a traditional public middle school with the myriad stimulus conditions found in schools (e.g., departmentalized schedules, multiple teacher and student demands, routine activities). Additionally, the reading fluency intervention package served as an SP Tier II intervention for the tutees in this study. Consequently, results from this study are quite encouraging in that evidence is now provided demonstrating the feasibility and effectiveness of this peer tutoring package in a naturalistic setting with middle school students in the context of RtI.

This study includes limitations that warrant further discussion and future research. First, visual inspection of Jan's data indicates some of the lowest scores for progress monitoring probes and the most variability. This may have occurred because of poor implementation of certain peer tutoring components and limited motivation to participate in peer tutoring. With regard to implementation, Jan's tutor, Chelsea, never required Jan to read the passage three times and also regularly failed to correct Jan's reading errors. Jan's performance on progress monitoring probes may have suffered as a result of these integrity failures. With regard to motivation, data collectors anecdotally reported that Jan appeared less motivated to participate in peer tutoring than the other tutees. While the peer tutoring procedure used in this study included tutees earning tickets for participation, those tickets were not provided contingent upon meeting a WCPM performance criterion. The decision not to include a performance criterion was done to keep demands on tutors modest. Future research may evaluate the effectiveness of this peer tutoring package with a modified reward component that includes reward contingent upon meeting some pre-specified performance criterion. Additionally, this study did not include a preference assessment to identify preferred rewards for tutees. As a result, the rewards that were available to Jan may not have been sufficiently strong to motivate her to participate. Future research may improve the motivation component of this peer tutoring package by conducting preference assessments with tutees prior to intervention implementation.

Second, while all tutees demonstrated improved performance for WCPM, EPM did not decrease substantially across tutees. Specifically, EPM decreased only marginally for Jan and Darian, while EPM increased slightly for Tanya. Failure to substantially impact EPM was likely the result of an intervention package that did not include an intense error correction procedure and tutors failing to consistently correct tutees' reading errors. Mean implementation for error correction across tutors was only 57.32%. The intervention package in this study might have been more effective at reducing EPM if there had been a specific error correction procedure (e.g., word error correction, phrase error correction) embedded in the package or if performance feedback had been used to improve tutors' ability to more consistently correct tutees' reading errors. Future research may address this by adding a more intense error correction procedure to intervention or ensuring that tutors consistently correct tutees' errors.

Third, with regard to Donna's performance for WCPM, scores trended downward across the first four progress monitoring sessions, and mean WCPM for all progress monitoring probes was 91. However, it is important to note that performance generally trended upward beginning with the fifth progress monitoring session. In fact, WCPM for the final two progress monitoring sessions were 107 and 123. Donna's data highlight the need to routinely progress monitor student's performance while allowing intervention implementation to occur over a reasonable time period prior to making hasty decisions regarding intervention termination or modification.

Fourth, for some tutees, only a small number of instructional and progress monitoring sessions were conducted, and no data on the progress monitoring probes were collected at baseline. This is problematic given measurement error of progress monitoring probes due to variability in passage difficulty within probe sets. As measurement error increases, the accuracy of our decisions regarding students' RtI decreases (Ardoin and Christ 2009). Moreover, we must be cautious in evaluating progress monitoring data included in this study as measurement error may have contributed to some of the increases in student performance. This also has important applied implications as the accuracy of decision making in an RtI process is obviously important given that students' are provided or denied access to special education, in part, based on progress monitoring data. Related to this limitation, the peer tutoring intervention was introduced to subsequent tutor/tutee dyads following only a small number of staggered data points in baseline. Although this weakened the demonstration of experimental control, immediate gains were obtained on the intervention passages for three of the four students. Fifth, this study was conducted in a school that received support for RtI and PBIS from a university-based school psychology training program. As a result, the extent to which these findings generalize to schools receiving fewer supports is not known. Future research may include a variety of school settings with varying levels of support.

Finally, while tutor intervention integrity was high for all tutors, opportunities for tutors to provide tutoring were not always available. Evaluation of intervention data indicates that some dyads experienced fewer sessions than other dyads (range = 1.7 sessions per week to 3 sessions per week). Decreased tutoring opportunities resulted from tutors not being allowed to conduct tutoring due to conflicting activities and

tasks such as practice for state-wide tests and other school-wide activities. Consequently, while we can say that tutors implemented intervention with high integrity when given the opportunity, we must acknowledge that peer tutoring implementation was dependent upon adults allowing tutors to conduct sessions. Previous peer tutoring research has encountered similar difficulties (Noell et al. 2000), and adults facilitated peer tutors sessions with more consistency when various follow-up procedures were implemented (e.g., performance feedback, administrator support). Therefore, this study could have benefited from monitoring of adults role in the peer tutoring process and appropriate follow-up (e.g., performance feedback).

As mentioned previously, the expansion of RtI systems necessitates a feasible continuum of academic supports for struggling students. The current study provides data demonstrating students' ability to implement a multi-component reading intervention that included progress monitoring. Further, accurate implementation of the intervention package resulted in improved student reading performance. Taken together, this study has important implications for research and applied practice related to peer tutoring as an SP approach to Tier II intervention in the context of RtI.

References

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*, 268–283.

Atherley, C. A. (1989). Shared reading: An experiment in peer tutoring in the primary classroom. *Educational Studies, 15*, 145–153.

Begeny, J. C., Krouse, H. G., Ross, S. G., & Mitchell, R. C. (2009). Increasing elementary-aged students' reading fluency with small-group interventions: A comparison of repeated reading, listening passage preview, and listening only strategies. *Journal of Behavioral Education, 18*, 211–228.

Begeny, J. C., & Silber, J. M. (2006). An examination of group-based treatment packages for increasing elementary-aged students' reading fluency. *Psychology in the Schools, 43*, 183–195.

Bentz, J., Shinn, M. R., & Gleason, M. M. (1990). Training general education pupils to monitor reading using curriculum-based measurement procedures. *School Psychology Review, 19*, 23–32.

Calhoon, M. B., Otaiba, S. A., Greenberg, D., King, A., & Avalos, A. (2006). Improving reading skills in predominantly Hispanic Title I first-grade classrooms: The promise of Peer Assisted Learning Strategies. *Learning disabilities research & Practice, 21*, 261–272.

Daly, E. J., Martens, B. K., Hamler, K. R., Dool, E. J., & Eckert, T. L. (1999). A brief experimental analysis for identifying instructional components needed to improve oral reading fluency. *Journal of Applied Behavior Analysis, 32*, 83–94.

Daly, E. J., Murdoch, A., Lillenstein, L., Webber, L., & Lentz, F. E. (2002). An examination of methods for testing treatments: Conducting brief experimental analyses of the effects of instructional components on oral reading fluency. *Education and Treatment of Children, 25*, 288–316.

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49*, 36–45.

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 Practice for Schools, 7*, 118–137.

Dufrene, B. A., Noell, G. H., Gilbertson, D. N., & Duhon, G. J. (2005). Monitoring implementation of reciprocal peer tutoring: Identifying and intervening with students who do not maintain accurate implementation. *School Psychology Review, 34*, 74–86.

Dufrene, B. A., & Warzak, W. J. (2007). Brief experimental analysis of Spanish reading fluency: An exploratory evaluation. *Journal of Behavioral Education, 16*, 143–154.

Edformation. (2003). *AIMSweb early literacy measures.* Retrieved March 2, 2007 from http://aimsweb.com/.

Ezell, H. K., Hunsicker, S. A., & Quinque, M. M. (1997). Comparison of two strategies for teaching reading comprehension. *Education & Treatment of Children, 20*, 365–383.

Freeland, J. T., Skinner, C. H., Jackson, B., McDaniel, E. C., & Smith, S. (2000). Measuring and increasing silent reading comprehension rates: Empirically validating a repeated readings intervention. *Psychology in the Schools, 37*, 415–430.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness to intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice, 18*, 157–171.

Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Greenwood, C. R., Terry, B., Utley, C. A., Montagna, D., & Walker, D. (1993). Achievement, placement, and services: Middle school benefits of classwide peer tutoring used at the elementary school. *School Psychology Review, 22*, 497–516.

Hintze, J. M., Callahan, J. E., James, E., & Matthews, W. J. (2002). Oral reading fluency and prediction of reading comprehension in African American and Caucasian elementary school children. *School Psychology Review, 31*, 540–553.

Hughes, T. A., & Fredrick, L. D. (2006). Teaching vocabulary with students with learning disabilities using classwide peer tutoring and constant time delay. *Journal of Behavioral Education, 15*, 1–23.

Individuals with Disabilities Education Improvement Act. (2004). Public Law 108-446 (20 U.S.C. 1400 *et seq.*).

Klubnik, C., & Ardoin, S. P. (2010). Examining immediate and maintenance effects of a reading intervention package on generalization. *Journal of Behavioral Education, 19*, 7–29.

Marston, D. (2005). Tiers of intervention in responsiveness to intervention: Prevention outcomes and learning disabilities identification patterns. *Journal of Learning Disabilities, 38*, 539–544.

Mathes, P. G., & Fuchs, L. S. (1994). The efficacy of peer tutoring in reading for students with mild disabilities: A best-evidence synthesis. *School Psychology Review, 23*, 59–80.

Mathes, P. G., Grek, M. L., Howard, J. K., Babyak, A. E., & Allen, S. E. (1999). Peer assisted learning strategies for first graders: A tool for preventing early reading failure. *Learning Disabilities Research & Practice, 14*, 50–60.

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 (NIH Publication No. 00-4769).* Washington, DC: U.S. Government Printing Office.

No Child Left Behind Act. (2001). Public Law 107-15.

Noell, G. H., Witt, J. C., LaFleur, L. H., Mortenson, B. P., Ranier, D. D., & LeVelle, J. (2000). A comparison of two follow-up strategies to increase teacher intervention implementation in general education following consultation. *Journal of Applied Behavior Analysis, 33*, 271–284.

Noell, G. H., Witt, J. C., Slider, N. J., Connell, J. E., Gatti, S. L., Williams, K. L., et al. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review, 34*, 87–106.

O'Connor, R. E., Fulmer, D., Harty, K. R., & Bell, K. M. (2005a). Layers of reading intervention in kindergarten through third grade: Changes in teaching and student outcomes. *Journal of Learning Disabilities, 38*, 440–455.

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

Pinnell, G. S., Pikulski, J. J., Wixson, K. K., Campbell, J. R., Gough, P. B., & Beatty, A. S. (1995). *Listening to children read aloud: Data from NAEP's Integrated Reading Performance Record (IRPR) at grade 4. National assessment of educational progress*. Princeton: Educational Testing Service.

Salend, S. J., & Nowak, M. R. (1988). Effects of peer-previewing on LD students' oral reading skills. *Learning Disabilities Quarterly, 11*, 47–53.

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

Shinn, M. R., & Shinn, M. M. (2002). *AIMSweb training workbook: Administration and scoring of reading curriculum-based measurement (R-CBM) for use in general outcome measurement*. Eden Prairie, MN: Edformation, Inc.

Sindelar, P. T., Lane, H. B., Pullen, P. C., & Hudson, R. F. (2002). Remedial interventions for students with reading decoding problems. In M. R. Shinn, H. M. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Preventive and remedial approaches* (pp. 703–730). Bethesda: National Association of School Psychologists.

Skinner, C. H., Cooper, L., & Cole, C. L. (1997). The effects of oral presentation previewing rates on reading performance. *Journal of Applied Behavior Analysis, 30*, 331–333.

Sutherland, K. S., & Snyder, A. (2007). Effects of reciprocal peer tutoring and self-graphing on reading fluency and classroom behavior middle school students with emotional behavioral disorders. *Journal of Emotional and Behavioral Disorders, 15*, 103–118.

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading. *Remedial and Special Education, 25*, 252–261.

Valleley, R. J., & Shriver, M. D. (2003). An examination of the effects of repeated readings with secondary students. *Journal of Behavioral Education, 12*, 55–76.

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*, 99–107.

Ysseldyke, J. (2005). Assessment and decision making for students with learning disabilities: What if this is as good as it gets? *Learning Disability Quarterly, 28*, 125–128.

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 and Treatment of Children, 29*, 469–506.