2008

Determining the relative efficacy of reciprocal and non-reciprocal peer tutoring for students identified as at-risk for academic failure

Keri F. Menesses
Louisiana State University and Agricultural and Mechanical College, kfishe3@lsu.edu

DETERMINING THE RELATIVE EFFICACY OF RECIPROCAL AND NON-RECIPROCAL PEER TUTORING FOR STUDENTS IDENTIFIED AS AT-RISK FOR ACADEMIC FAILURE

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Arts

in

The Department of Psychology

by
Keri F. Menesses
B.S., Louisiana State University, 2005
December 2008

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The current study directly compared the academic and social gains of reciprocal peer tutoring, non-reciprocal peer tutoring, and traditional classroom instruction with elementary students. Participants included 59 students who performed below-average on class-wide screenings using curriculum-based measurement math probes. Students involved in peer tutoring were trained to tutor basic math facts using a constant time delay procedure. Both types of peer tutoring produced significantly larger academic gains than traditional classroom instruction; the two types of tutoring produced comparable academic and social results, although reciprocal tutoring resulted in marginally larger academic gains. Reciprocal peer tutoring is recommended based on the fact it remediates twice the number of students simultaneously as compared to non-reciprocal tutoring.

INTRODUCTION

Effective and correctly implemented pre-referral interventions can decrease the number of special education referrals and prevent inappropriate placement in special education (Kovaleski, Gickling, Morrow, & Swank, 1999; McDougal, Clonan, & Martens, 2000). Direct instruction (DI) has been largely documented as an effective intervention for children with academic and/or behavioral difficulties (Kroesbergen & Van Luit, 2003). A meta-analysis conducted by Adams & Engelmann (1996) containing 34 studies showed that gains were nearly a standard deviation higher for students receiving direct instruction than for those in comparison non-DI approaches. A mean effect size of d =.87 confirms that the overall effect of DI is quite large across a variety of students (regular and special education ranging from elementary to secondary level) and subject matter (reading, math, science, spelling social skills; Adams & Engelmann, 1996). Furthermore, students who have failed to respond to conventional teacher-led instruction are able to make rapid and continual progress when taught one-on-one by a qualified teacher (Delquadri, 1978). However, taking into account high student-to-teacher ratios and lack of resources, teachers may not have the time or the materials to carry out such instruction. In fact, the lack of teacher implementation of prescribed interventions has been noted as a significant problem (Noell, Witt, Gilbertson, Ranier, & Freeland, 1997; Witt, Noell, LaFluer, & Mortenson, 1997). As a result, interventions that employ alternative agents of change need to be considered.

One possible agent of change is another student within the classroom. Peer tutoring is the process by which a student helps one or more students learn a skill or concept (Thomas, 1993). A substantial advantage of peer tutoring is the decreased amount of teacher responsibility in conducting the invention, as the teacher fills a supervisory role over the students who are the

primary conductors of the intervention. As a result, peer tutoring is an efficient method by which to provide individualized instruction for many students simultaneously, thereby making it extremely cost-effective. In fact, it has been shown to be more cost-effective than computer aided instruction, class reduction, and increasing the length of the school day (Levin, Glass, & Meister, 1987).

Previous research has documented the benefits of peer tutoring, including acquisition of academic skills (Cohen, Kulik, & Kulik, 1982), development of appropriate social skills (Mathur & Rutherford, 1991), enhancement of peer relations (Greenwood, Carta, & Hall, 1988), improved classroom behavior (Fuchs, Fuchs, Phillips, Hamlett, & Karns, 1995), increased school attendance (Miller, Kohler, Ezell, Hoel, & Strain, 1993), and positive socioemotional outcomes, such as a sense of belonging and internal attributions for success (Nazzal, 2002). These gains are evident in both tutors and tutees. Furthermore, there are additional benefits for tutees in the peer tutoring relationship. These include more time academically engaged and on-task (Ginsburg-Block & Fantuzzo, 1997), increased opportunities to respond (Delquadri, Greenwood, Whorton, Carta, & Hall, 1986), immediate feedback (Topping, 2005), continuous progress monitoring (Greenwood et al., 1988), and ability to progress at an individualized pace (Dupaul, Ervin, Hook, & McGoey, 1993). There are also multiple benefits for tutors beyond increased academic performance, which include positive attitude toward subject (Cohen et al., 1982), improved self-esteem (Byrd, 1990; Cardenas, Harris, del Refugio, & Supik, 1991), improved locus of control (Lazerson, Foster, Brown, & Hummel, 1988), and improved attitude toward school (Cardenas et al., 1991).

Noting the advantages of being a tutor, Fantuzzo and colleagues developed Reciprocal Peer Tutoring (RPT), a procedure in which students alternate between roles of tutor and tutee so

that both children have access to all the advantages of peer tutoring (Pigott, Fantuzzo, Heggie, & Clement, 1984; Wolfe, Fantuzzo, & Wolter, 1984). The tutoring pair works together to prompt, monitor and evaluate each other in order to earn group rewards while learning a specific academic task (Fantuzzo, Polite, & Grayson, 1990). Greenwood and colleagues utilize the same reciprocal tutoring method in Classwide Peer Tutoring (CWPT) in which all students within a classroom participate in peer tutoring simultaneously. This method was originally designed to prevent future academic failure in low socioeconomic status and culturally diverse schools (Delquadri, Greenwood, Stretton, & Hall, 1983), but has also been successful with students with mild disabilities (Maheady, Harper, & Mallette, 2001). In CWPT, the class is divided into two teams which are then partitioned into tutor-tutee pairs, and the two teams compete to earn the most points each week. The tutee earns points by correctly responding on the academic task while the tutor can earn points by tutoring correctly. RPT and CWPT require both members of the dyad to contribute to their success; therefore both are responsible for attaining the goal, and the accountability of achieving the goal does not rest with just one student. Both reciprocal tutoring methods have been effective across various academic subjects, including mathematics, reading, spelling, vocabulary, social studies, and science (Greenwood, Carta, & Maheady, 1991), and have also resulted in increased socialization among peers (Wolfe, Fantuzzo, & Wolfe, 1986) and decreased behavior problems at school (Greenwood et al., 1991).

Extensive research has documented the effectiveness of peer tutoring. Greenwood et al. (1984) compared traditional teacher-mediated instruction to Classwide Peer Tutoring and measured the effectiveness of each using weekly tests of the subject matter taught or tutored that week. A within-subjects design was used in which the type of instruction (teacher- or peer-mediated) was assigned to various weeks during the experimental trial. Controlling for order

effects and the amount of time in each type of instruction, all students scored significantly higher on the weekly tests when the peer tutoring procedure was in effect as compared to weeks when teacher-developed procedures were being utilized. Furthermore, low-achieving students performed as well as average students when CWPT was use, but scored lower than average students when teacher-based instruction was in place. Therefore, peer-mediated instruction resulted in more academic gains than teacher-mediated instruction, especially for students who were below-average.

Greenwood, Delquadri, and Hall (1989) conducted a longitudinal study with 94 teachers and 416 students using a between-subjects design in which classrooms received either conventional classroom-based instruction or classwide peer tutoring. Classrooms that participated in peer tutoring showed less inappropriate classroom behavior and performed better on standardized tests of arithmetic, reading, spelling, and language as compared to the control group. Furthermore, at-risk students who participated in peer tutoring performed better than non-risk control students, thereby closing the gap in achievement levels that had previously existed.

In 2005, Robinson, Schofield, and Steers-Wentzell conducted a literature review of peer tutoring programs designed to teach mathematical facts, which included 28 articles published after 1988. Effect sizes, using Cohen's (1988) *d* procedure, were calculated for academic, attitudinal, and socioemotional outcomes. Peer tutoring had positive effects on a variety of math skills, including arithmetic computation, conceptual understanding, and problem solving skills. Moreover, this finding held true for all levels of socioeconomic status. Effect sizes of academic achievement ranged from $d = .48$ to $d = 1.17$ for tutees, with the binominal effect size displays (BESD; Rosenthal, Rosnow, & Rubin, 2000) ranging from 62% to 76% improvement rate for the

treatment group. Effect sizes ranged from d = .48 to d = 1.37 for tutors, with the success rate for the treatment group ranging from 62% to 79% using the BESD. Results showed improved classroom behavior (d = .71, BESD = 67%) and increased school attendance (d = .90, BESD = 71%) for both tutors and tutees. Peer tutoring also produced increased social skills for all students involved (d = .89, BESD = 70%).

The Robinson et al. review also examined how specific features of tutoring programs and characteristics of the students influenced their outcomes (Robinson et al., 2005). For example, results from previous reviews indicate that programs with a longer duration do not necessarily produce better academic outcomes. In fact, Cohen et al (1982) concluded that shorter programs (less than 4 weeks) had the largest effect sizes (d = .95, BESD = 72% for tutees and d = .56, BESD = 64% for tutors), while medium length programs (5-18 weeks) had medium effect sizes (d = .42, BESD = 61% for tutees and d = .38, BESD = 60% for tutors), and longer programs (19-36 weeks) resulted in the smallest effect sizes (d = .16, BESD = 54% for tutees and d = .10, BESD = 53% for tutors). Robinson et al (2005) found similar results, with shorter tutoring programs resulting in larger effect sizes (d = 1.01, BESD = 73%) than longer programs (d = .38, BESD = 60%), and proposed the "newness" of brief programs as a possible cause of producing better results.

Robinson et al. (2005) also investigated the differential effects of tutor training v. no tutor training, reciprocal v. non-reciprocal tutoring, reward-based v. non-reward-based tutoring, and same-gender v. mixed-gender dyads. The authors underscored the importance of tutor training, as training has been shown to result in proper tutoring behavior in students as young as preschoolers (Balenzano, Agte, McLaughlin, & Howard, 1993; Tabacek, McLaughlin, & Howard, 1994). Furthermore, tutees showed more academic improvement when their tutors had

been trained as opposed to having non-trained tutors (Polirstok & Greer, 1986). When comparing outcomes of reciprocal v. non-reciprocal programs, they concluded that reciprocal peer tutoring resulted in larger effect sizes, ranging from medium (d = .5) to large (d = .96) and a BESD ranging from 62% to 72%, while non-reciprocal tutoring showed small (d = .2) to medium (d = .48) effect sizes and a BESD ranging from 55% to 62%. Rewards that are contingent on performance have also been shown to be influential on results, with the combination of rewards and tutoring producing better academic performance than either condition alone (Fantuzzo, King, & Heller, 1992). When investigating the effects of various dyad compositions, same-gender dyads produced better academic results than mixed-gender pairs (Topping & Whiteley, 1993; Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003).

Another aspect of peer tutoring that can take on many different forms is the instructional focus of the program. Most procedures focus on teaching basic skills and facts that are essential and prerequisite to complex, higher order concepts. Fundamental math facts, sight words, spelling, and vocabulary are commonly target materials in peer tutoring procedures. Several studies have demonstrated the utility of constant time delay (CTD; Touchette, 1971; Stevens & Schuster, 1988) for teaching such basic academic skills (Gast, Ault, Wolery, Doyle, & Belanger, 1988; Hughes, Fredrick, & Keel, 2002; Koscinski & Gast, 1993). CTD is a systematic procedure in which basic skills are presented, typically on flashcards, and corrective prompts are given after a pre-determined lapse in time, such as 3 or 5 seconds. This is a form of direct instruction that includes response prompting, increased academic engaged time, frequent opportunities to respond, immediate feedback, and teaching to mastery, which have all been shown to increase student achievement (Beirne-Smith, 1991; Brophy & Good, 1986; Keel, Dangel, & Owens, 1999). In fact, CTD has been effective in the acquisition of sight words (Keel & Gast, 1992),

math facts (Cybriwsky & Schuster, 1990; Mattingly & Bott, 1990), written spelling words (Stevens, Blackhurst, & Slaton, 1991), and vocabulary definitions (Schuster, Stevens, & Doak, 1990). However, this direct approach to instruction is quite time-demanding, especially for one teacher with many students. As a result, CTD has recently been integrated into more efficient peer tutoring procedures.

Telecsan, Slaton, and Stevens (1999) examined the effectiveness of combining reciprocal peer tutoring and CTD to teach written spelling skills to 6 students identified as learning disabled. Results indicated that peer tutors accurately and reliably implemented the time delay procedure, which was highly effective in teaching written spelling words, as the six students learned to spell all the target words presented by their peers. Similarly, Hughes and Fredrick (2006) trained six grade students to teach each other vocabulary definitions using CTD. This procedure was also implemented at high levels of integrity, both by students with and without learning disabilities. Furthermore, all students mastered the vocabulary words, maintained them over time, and generalized the words across context (Hughes and Fredrick, 2006).

Summary and Research Questions

The efficacy of peer tutoring in enhancing academic achievement is evident based on previous research showing average effect sizes of $d = .86$ for tutees $d = .80$ for tutors, and average success rates of 70% and 69% for tutees and tutors, respectively. Likewise, peer tutoring has been credited with improving problem behaviors for both tutors and tutees, with effect sizes ranging from $d = .69$ to $d = 1.25$ and the BESD ranging from 66% to 77%. Although peer tutoring is widely accepted as an effective intervention, some aspects of peer tutoring remain uninvestigated experimentally. For instance, although the Robinson et al. (2005) review reported larger effect sizes for reciprocal programs (ranging from $d = .5$ to $d = .96$) than non-reciprocal

programs (ranging from $d = .2$ to $d = .48$), the two interventions have not been directly compared within the same experiment previous to this study. Thus, the major purpose of the current study was to compare reciprocal and non-reciprocal peer tutoring to determine which program results in more significant gains, and in so doing, ensure the most effective peer tutoring intervention is implemented for and by students in need. The comparison consisted of employing identical CTD procedures previously shown to be successful when implemented by peer tutors, which differed only in whether or not students changed roles between tutor and tutee. It was hypothesized that reciprocal tutoring would produce marginally larger gains in academic scores as compared to those produced in non-reciprocal tutoring.

Another purpose of this study was to determine the effectiveness of peer tutoring as a remedial intervention for students who were identified as at-risk for academic failure. It was hypothesized that both tutoring programs would increase the students' academic scores to an average range, thereby proving successful in remediating at-risk students. Furthermore, both peer tutoring programs were compared to a treatment-as-usual control group in an attempt to add to the literature on the effectiveness of peer tutoring as compared to traditional teacher-led instruction. It was predicted that both tutoring procedures would result in greater academic gains than conventional classroom instruction. These results are expected based on previous research findings that have established the effectiveness of each peer tutoring procedure in isolation.

Although previous research has documented the academic gains of tutors as well (Britz, Dixon, & McLaughlin, 1989; Cohen et al., 1982), the tutors in the current study were not expected to show significant academic improvement. Because tutors were only exposed to single-number answers, they were not subject to any instruction or practice whatsoever, nor did

they provide instruction to tutees when incorrect answers were given. As a result, it was hypothesized that tutors would not make significant gains in academic scores.

Another commonly cited result of peer tutoring is improved social interactions in both tutees and tutors (Greenwood, Carta, & Hall, 1988; Mathur & Rutherford, 1991), with an average effect size of d = .89 and improvement rate of 71% for the treatment group using the BESD. In order to measure the quality of socializations between classmates in the current study, teachers completed Direct Behavior Ratings (DBR; Chafouleas, Christ, Riley-Tillman, Briesh, & Chanese, 2007) which targeted specific peer interaction behaviors of each participant. It was hypothesized that reciprocal peer tutoring would have significantly more social gains based on its mutual nature as compared to the sense of hierarchy present in non-reciprocal tutoring, which may result in less social confidence for tutees. It was also predicted that both tutoring interventions would produce larger social gains than the control classroom-instruction group, which is based on prior research.

METHOD

Participants, Screening, and Setting

Seven general education classrooms (two 2nd grade, three 3rd grade, and two 4th grade classrooms) in an East Baton Rouge Public School were administered curriculum-based measurement (CBM) math probes with addition, subtraction, multiplication, or division problems. The types of problems distributed to each classroom were determined during a teacher interview in which each teacher reported the two types of math problems that should have been mastered most recently by their students. Therefore, this procedure screened for students who did not acquire or maintain facts that were no longer being directly taught by their teacher.

Digits correct per two minutes were scored for each student according to the Aimsweb scoring instructions for CBM math probes (Math-CBM, available from www.aimsweb.com). A student in 2nd or 3rd grade scoring 0 through 19 digits correct are considered at a "frustrational" level as determined by national benchmarks (Deno & Mirkin, 1977), while a fourth grade student who produces 0 through 39 digits correct also falls in the frustrational range. Sixty-seven students who fell within the frustrational range were screened for eligibility in this study. Written parental consent and student assent were obtained before the screening procedure began.

In order to meet criteria for inclusion, students were required to name numbers 0-18 for addition and subtraction tutoring and 0-81 for multiplication and division tutoring. Also, participants had to read completed mathematical equations that were to be included in their tutoring program (e.g. 7 x 5 = 35 if participating in multiplication tutoring). Students were then given two more probes with the type of problem on which they had scored in the frustrational range. The median of the three probes served as a baseline score for each participant; if that score fell above the frustrational range, the student was no longer eligible for participation. Finally,

students could not be receiving any remedial instructional services in mathematics outside of the classroom. All students who completed the screening procedure were allowed to choose a reward for their cooperation from a plastic box containing candy, pencils, stickers, and small toys. Sixty-two students met criteria and were included in the study. Those who did not meet criteria were either receiving other services outside the classroom or their median baseline score fell above the frustrational range.

A power analysis was computed to determine an adequate sample size for the statistical analyses used in the current study. Sample size was determined using the G*power version 3.0.5 (Faul, 2006). For an alpha level of .05 and a standardized effect size of d = .25, a total of 42 students was required to reach a power level of .80. Therefore, each condition needed to contain 14 students; however because the tutors in the one-way peer tutoring (PT) condition did not receive any tutoring, this condition had to contain 28 students total in order to have 14 tutees receiving the intervention. Therefore a total of 56 students were required.

All sessions took place in a quiet hall of the school containing tables and chairs with a supervisor (the experimenter and/or a research assistant) present. Times for sessions were determined by each teacher based on the daily classroom schedule.

Materials

During the intervention, students in the tutoring conditions were presented with 3”x 5” index cards which had math problems printed on one side and the answers to the problems on the other side. A timer was used to keep the time of each session consistent. Each tutoring dyad had their own folder that contained a green square/red square game board, a progress monitoring chart, a treatment integrity checklist, and a point tracker chart, all of which will be described

below. A reward box containing candy, stickers, pencils and small toys was used to reinforce students' tutoring behaviors and academic gains.

Dependent Measures

CBM probes were used to measure each participant's fluency level in mathematics by scoring the amount of digits correctly produced within two minutes. CBM was chosen as the academic dependent measure due to extensive research establishing its psychometric properties as well as its ability to monitor progress and its quick administration and scoring procedures (Hosp, Hosp, & Howell, 2007). The computer-generated probes (Math Computation Probes, available from www.interventioncentral.org) were given immediately before and after the intervention to determine the acquisition of math facts, as well as three weeks later to evaluate maintenance of the facts.

A Daily Behavior Report (DBR) was completed for each student by their classroom teacher to assess the quality of peer social interactions. A DBR was chosen as the social dependent measure because completing one is less time intensive than direct observation, and DBR scores can be used for monitoring progress throughout an intervention unlike broad behavior rating scales which are typically not change-sensitive. Furthermore, a DBR has the capability of targeting very specific behaviors, also unlike broad behavior rating scales. The DBR created for this study consisted of one item which read “Interacts cooperatively with peers,” which was further defined as “Student enters work/play situations and participates appropriately with peers, without anyone bullying or teasing.” The rating scale for this item ranged from 1-9 with a rating of 1 meaning “Never,” 5 meaning “Sometimes,” and 9 meaning “Always.” Based on recent research, a DBR is considered to approximate a reliability coefficient of .90 after 10 ratings (Chafouleas et al., 2007). Therefore, classroom teachers

completed a rating two times a day, in the morning and afternoon, for five days to reach a total of 10 ratings for each participant. These 10 scores were averaged for each student for a total pre, post, and follow-up social interaction score.

Experimental Design

Two 4 x 2 (Group x Time) mixed-model designs were used to determine the effectiveness of each intervention in increasing math and social skills. Each design consisted of a between-subjects factor (group) with four levels (reciprocal tutoring, tutees in non-reciprocal tutoring, tutors in non-reciprocal tutoring, and control group), and a within-subjects factor (time) with two levels (post-score, and follow-up score). The dependent variables were academic scores on math CBMs and social scores on teacher-completed DBRs.

Sample Selection

Seven teachers were approached at the beginning of the spring semester and all agreed to have their classrooms screened for participants. Seven classrooms were screened based on the estimate that 1/3 of each class would fall in the frustrational range on the math CBM probe; 1/3 of 25 students equals approximately 8 students per class for a total of 56 students, the amount required in the power analysis. Sixty-seven students performed in the frustrational range on the first CBM probe, passing the first gate of screening; 62 students were determined eligible to participate after passing the second gate of screening previously described. Throughout the course of the study, no students dropped out and 3 students moved to a different school. As a result, a total of 59 participants (27 females, 32 males) were included in all data analyses.

Procedure

The three types of instruction (peer tutoring, reciprocal peer tutoring, and control group/ traditional classroom instruction) were randomly assigned to classrooms in which teachers were

willing to allow their students to participate. The Peer Tutoring (PT) condition consisted of one-way peer tutoring using a CTD procedure in which one student was always the tutor and the other student was always the tutee. In Reciprocal Peer Tutoring (RPT), students switched roles between tutee and tutor within the same session, also utilizing the same CTD procedure. Finally, students in the control group received conventional classroom instruction throughout the entire intervention and had no knowledge of peer tutoring. The three different instructional conditions were randomly assigned to classrooms so that all eligible students within a classroom received the same treatment in order to control for preference effects. If students within the same classroom had received different interventions, they may have discovered a more favorable condition by talking with the other participants. Preferring a different condition can lead to a lack of cooperation in the assigned condition or even attrition, which was avoided by assigning one condition to all eligible students in a classroom. This random assignment resulted in the control condition being assigned to one 2nd and one 3rd grade classroom, peer tutoring to one 2nd, one 3rd, and one 4th grade classroom, and reciprocal peer tutoring to one 3rd and one 4th grade classroom. Each student in the peer tutoring conditions (PT and RPT) was randomly assigned to another eligible student within the classroom in order to form a dyad. If at any time during the intervention a pair of students was not able to participate correctly, they were to be re-assigned to form a new dyad. Such re-assignment was not warranted during the study however, as all dyads participated correctly.

Tutor Training. Tutors were trained using a "tell, show, do" model, modified from the Telecsan et al. (1999) study. Following an explanation of CTD and peer tutoring, the process of presenting cards using CTD was demonstrated by the experimenter. Tutors then participated in role-play with the experimenter, who acted as a tutee. Training trials were identical to the three

minute tutoring sessions that took place during the intervention. The experimenter answered some problems correctly to ensure the tutor provided praise and answered other problems incorrectly to ensure the tutor provided corrective feedback when necessary. Tutors' performance was measured using the treatment integrity checklist shown in Table 1. In order to ensure tutors were consistently presenting each card for 3 seconds, they were instructed to count by *silently* saying "1-Mississippi, 2-Mississisppi, 3-Mississipi," a procedure most of the students were very familiar with in order to time seconds. Tutors were trained until they reached 100% accuracy (performing all 6 behaviors throughout the trial) on 3 consecutive training trials.

Table 1

Treatment Integrity Checklist

Tutor: _____ Tutee: _____ Date: _____

Session Component	Step
Tutoring	1. Waited for timer to begin
	2. Showed each card for 3 seconds
	3. Corrected any wrong answers
	4. Said "Good job!" or "That's right!" after each card
	5. Shuffled cards
	6. Continued until timer sounded
Progress Monitoring	1. Showed each card for 3 seconds
	2. Did not talk
	3. Sorted cards unto green and red squares
	4. Counted and recorded correct cards with tutee

Tutors were also trained to monitor the progress of their tutees in a progress monitoring session that took place immediately after each practice session. During the progress monitoring session, tutors did not provide feedback but rather, measured independent performance of the tutees. Tutors were again trained until they reached 100% accuracy (performing all 3 behaviors throughout the trial) on 3 consecutive training trials as assessed by the treatment integrity checklist (Table 1).

Tutoring Procedure. Students assigned to the tutoring conditions (PT and RPT) participated in an average of three tutoring sessions per week until they reached the session requirement of 15 sessions total. If one student in a dyad was absent, his or her partner also did not participate that day. The total amount of weeks spent tutoring differed depending on each student's attendance; however, all participants completed the intervention in an average of 7 school weeks.

Each CTD tutoring session was three minutes in duration and included ten different math problems printed on index cards. The tutoring program was explained to the students as a game (Delquadri et al., 1983), with the tutor described as the "coach" and the tutee as the "player," as adopted from CWPT. In RPT, students alternated filling the role of the coach first. For example, if student A coached first on Monday, student B coached first on Wednesday, and so forth throughout the week.

Two to three pairs (four to six students) were removed from their classroom at the time designated by the teacher and brought to a quiet hallway with a table and chairs. Each pair had a pocket folder stored in their classroom which contained all the materials needed for tutoring. Each session began by the coach filling out the "name" and "date" blanks on the treatment integrity checklist (See Table 1) and subsequently reading over the required tutoring steps. Once

all pairs were ready to begin, the supervisor (the experimenter or research assistant) started the timer and said “Begin.” Each coach presented ten cards to each player using 3-second CTD. If the player provided the correct response within three seconds, the coach praised the player and presented the next card. If the player answered incorrectly or did not provide an answer within three seconds, the coach said the correct answer which the player had to repeat before being presented with the next card. Once all 10 cards were presented, the coach shuffled them and continued the process until the timer sounded.

Progress Monitoring Procedure. Immediately following the tutoring procedure, coaches presented the ten cards once again using 3-second CTD but without providing any verbal feedback. Coaches used their green square/red square game board to sort correct and incorrect answers, respectively. Once all ten cards were presented and sorted into the two piles, the player counted the correct amount of responses on the green square while the coach supervised to make sure the player was counting correctly. The player then marked the amount of cards he/she answered correctly on the progress monitoring chart kept in the pair's folder. Once the player provided ten correct responses on two consecutive progress monitoring sessions, a new set of ten different cards were introduced in the next session.

Interdependent Contingencies. Based on Reciprocal Peer Tutoring and Class-wide Peer Tutoring, both members of all dyads were responsible for working toward their goal and earning points for their “team.” Because the accountability of achieving the goal does not rest with just one student, group contingencies have been shown to increase student cooperation and performance (Fantuzzo et al., 1992). During each session, every pair of students had the opportunity to earn four points, which equated the possible amount of reinforcement in each condition. In the PT condition, coaches were able to earn two points by tutoring correctly and

players could earn two points by beating their previous score on their progress monitoring chart. In RPT, the student who coached first could earn one point for tutoring correctly and the first player could earn one point for beating his/her previous score. When they switched roles, they each had the opportunity to earn another point, which sums to four possible points. Once a pair accrued 5 team points, which were tracked on their "Team Points" chart in their pocket folder, both members were allowed to choose a prize from the reward box.

Procedural Integrity. Coaches monitored their own treatment integrity using the coach checklist each day. Immediately following both the tutoring and the progress monitoring procedures, coaches were prompted to check the steps they completed correctly. The experimenter and research assistants closely monitored the integrity of all tutors and provided corrective prompts whenever necessary. Performance feedback was given to tutors after each session, as feedback contingent on poor implementation has been shown to immediately increase appropriate peer tutoring behaviors (Dufrene, Noell, Gilbertson, & Duhon, 2005). Booster training sessions were provided when tutors had to be prompted more than once on any step during the procedure, which only occurred in two sessions during the study. Data from these sessions were not included in any analyses. Tutors were re-trained on the incorrectly implemented steps and required to correctly complete three consecutive tutoring sessions with the experimenter before tutoring a peer again. It was crucial to maintain high procedural integrity in all tutoring conditions in order to legitimately compare the outcomes they produced.

RESULTS

CBM Data

A 4 x 2 mixed-model ANCOVA was conducted with pre-intervention score as the covariate. The two independent variables were Group (between-subjects factor) and Time (within-subjects factor). The dependent variable was the number of digits correct within two minutes on a CBM math probe. The main effect for Group was significant, F (3, 54) = 11.41, p < .001, partial η^2 = .39. A Bonferroni post-hoc test revealed reciprocal tutoring produced higher scores (M = 34.88) than the control group (M = 19.72) and the tutees produced higher scores (M = 32.90) than the control group (M = 19.72), collapsing across time. The main effect of Time was insignificant, F (1, 54) = .04, p > .05, partial η^2 = .001. The interaction between Group and Time was also not significant, F (3, 54) = .75, p < .05, partial η^2 = .31.

The adjusted means of the post-intervention scores were ordered as expected across the four instructional groups (Table 2). The reciprocal group had the largest adjusted mean (M = 34.41), the tutees had a slightly smaller adjusted mean (M = 34.18), the tutors had an even smaller adjusted mean (M = 26.26), and the control group had the smallest adjusted mean (M = 19.76). The Bryant-Paulson post-hoc procedure was conducted to evaluate differences among these adjusted means. There were significant differences in the adjusted post-score means between the reciprocal group and the control group and between the tutee group and the control group.

The adjusted means of the follow-up scores were also ordered as expected across the four instructional groups (Table 2). The reciprocal group had the largest adjusted mean (M = 35.35), the tutees had a smaller adjusted mean (M = 31.61), the tutors had an even smaller adjusted mean (M = 25.87), and the control group had the smallest adjusted mean (M = 19.69). The Bryant-

Paulson post-hoc procedure indicated significant differences in the adjusted follow-up means between the reciprocal group and the control group, the tutee group and the control group, and the reciprocal group and the tutor group.

Table 2

Means, Standard Deviations, Adjusted Means, and Ranges for CBM scores

	Group			
	Reciprocal (N = 15)	Tutees (N = 14)	Tutors (N =14)	Control (N = 16)
Time				
PRE				
M	19.20	17.21	17.00	15.81
(*SD)*	(9.34)	(8.79)	(8.15)	(4.07)
Adjusted *M*	19.20	17.21	17.00	15.81
Range	24	23	28	11
POST				
M	37.33	34.07	25.71	17.50
(*SD)*	(16.33)	(19.05)	(12.94)	(6.96)
Adjusted *M*	34.04	34.18	26.16	19.76
Range	54	62	37	23
FOLLOW-UP				
M	38.27	31.50	25.43	17.44
(*SD)*	(20.01)	(17.73)	(11.49)	(7.19)
Adjusted *M*	35.35	31.61	25.87	19.69
Range	58	59	42	24

DBR Data

A 4 x 2 mixed-model ANCOVA was conducted with the pre-score as the covariate. The main effect of Time was not significant, F (1, 54) = .08, $p > .05$, partial $\eta^2 = .002$, when controlling the effects of the pre-score. There was a significant main effect of Group on participants' social score across post and follow-up, F = (3, 54) = 3.60, p = .02, partial $\eta^2 = .17$.

A Bonferroni post-hoc test revealed there were no significant differences between groups collapsing over time. However, there was a significant interaction between Group and Time, F (3, 54) = 2.80, p = .04, and tests of simple effects revealed Group significantly influenced social scores at post measurement, F (3, 55) = 4.22. p < .01, and at follow-up measurement, F (3, 55) = 8.39, p < .001.

The post-score adjusted means were not ordered as expected (Table 3). The reciprocal group had the largest adjusted mean (M = 7.35), the tutees had a smaller adjusted mean (M = 6.93), the control group had an even smaller adjusted mean (M = 6.51), and the tutors had the smallest adjusted mean (M = 5.46). Pairwise comparisons were conducted using the Bryant-Paulson post-hoc test to evaluate the differences among the adjusted post-score means. There were significant differences between the reciprocal group and the control group and the tutee group and the control group at post-score. The post-score means among the three groups involved in tutoring were not significantly different.

The adjusted means of the follow-up scores were also not ordered as expected (Table 3). The tutees had the largest adjusted mean (M = 7.35), the reciprocal group had a smaller adjusted mean (M = 6.90), the tutors had an even smaller adjusted mean (M = 6.06), and the control group had the smallest adjusted mean (M = 5.79). A post-hoc analysis using the Bryant-Paulson procedure indicated significant differences in the adjusted follow-up score means between the reciprocal group and the control group, the tutee group and the control group, and the tutor group and the control group. The three groups involved in tutoring were not significantly different at the follow-up measurement.

Table 3

Means, Standard Deviations, Adjusted Means, and Ranges for DBR scores

	Group			
	Reciprocal (N = 15)	Tutees (N = 14)	Tutors (N =14)	Control (N = 16)
Time				
PRE				
M	6.91	7.87	8.54	5.54
(*SD)*	(2.03)	(1.29)	(.54)	(1.87)
Adjusted *M*	6.91	7.87	8.54	5.54
Range	5.9	3.7	1.5	7
POST				
M	7.15	7.54	6.63	5.16
(*SD)*	(1.92)	(1.50)	(2.00)	(2.38)
Adjusted *M*	7.35	6.93	5.46	6.51
Range	6	4.8	5.7	8
FOLLOW-UP				
M	6.70	7.93	7.17	4.50
(*SD)*	(2.49)	(1.19)	(1.30)	(2.47)
Adjusted *M*	6.90	7.35	6.06	5.79
Range	6.7	3.8	3.7	7.4

DISCUSSION

Academic Findings

As hypothesized, all students receiving the intervention (reciprocal and tutee groups) produced a significant increase in the number of digits correct on a math CBM probe, the academic dependent variable, at both post and follow-up measurements (Figure 1). Tables 4 and 5 provide effect sizes for each tutoring group compared to the control group at post and follow-up measurement, respectively. These large effects contribute to the empirical data establishing

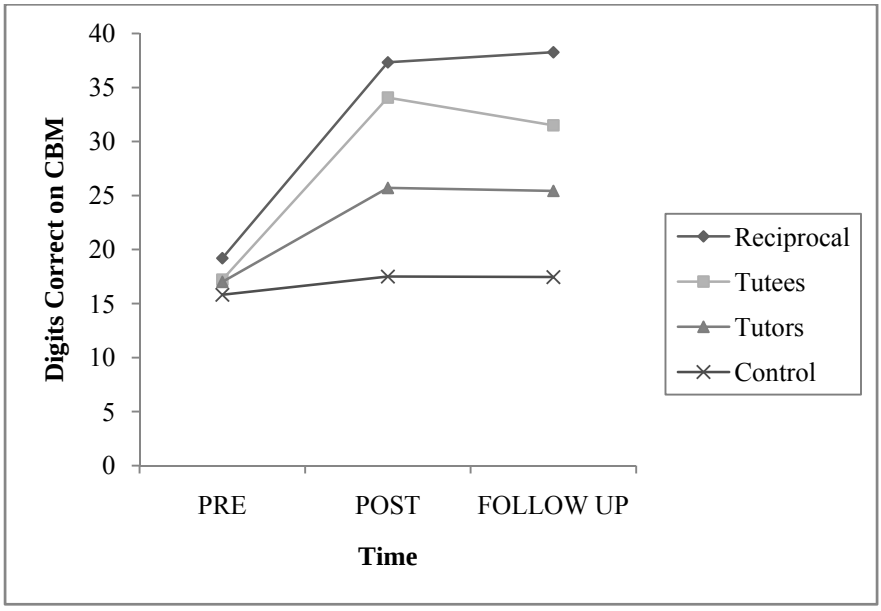

Figure 1. Mean digits correct on Curriculum-Based Measurement math probe.

the effectiveness of peer tutoring as an academic intervention. It is important to note that students in the tutoring role produced significantly more digits correct as compared to their pre-intervention score. Although this increase in academic achievement for tutors was not expected in the current study, it has been found by previous researchers (Britz et al., 1989; Cohen et al., 1982). Because tutors were not exposed to any form of math instruction or practice in this study, the fact that their math scores increased lends to the theory that fulfilling a tutoring role may

Table 4

Overall Effect Size Display for Post-Intervention CBM Scores

	Effect Size expressed as Cohen's *d*	Effect Size expressed as *r*	Success rate as percentage of subjects showing improvement r_{besd}	
Group			Treatment Group	Control Group
Reciprocal	*d* = 1.58	*r* = .62	81%	19%
Tutee	*d* = 1.16	*r* = .50	75%	25%
Tutor	*d* = .79	*r* = .37	69%	31%

Table 5

Overall Effect Size Display for Follow-up CBM Scores

	Effect Size expressed as Cohen's *d*	Effect Size expressed as *r*	Success rate as percentage of subjects showing improvement r_{besd}	
Group			Treatment Group	Control Group
Reciprocal	*d* = 1.39	*r* = .57	79%	21%
Tutee	*d* = 1.05	*r* = .46	73%	27%
Tutor	*d* = .83	*r* = .38	69%	31%

increase students' self-confidence. Perhaps being a math "coach" for another student caused the tutors to feel more experienced in the subject matter and to be more effortful when completing math assignments. This possibility needs to be further investigated in future research.

While the reciprocal, tutee, and tutor groups all showed significant increases in their academic scores, the control group did not. These results show the effectiveness of peer tutoring above and beyond that of traditional classroom instruction. It is important to note that although the control students were identified as at-risk for failure at the beginning of the study, they only participated in conventional teacher-led instruction and had only improved slightly across 4

school months (mean increase of 1.69 digits and 1.63 digits at post and follow-up measurements, respectively). This lack of academic gains produced by the control group further enhances the efficacy of peer tutoring by removing the possibility of history effects and showing that only minimal, non-significant gains were achieved over time as a result of traditional school instruction. Hence, a considerable amount of the academic increases produced by all other participants can be attributed to their involvement in peer tutoring rather than to the usual math instruction they received in the classroom.

When evaluating the maintenance of facts over time, results indicate that there were no significant differences between post and follow-up scores in any of the groups. Thus, students did not improve significantly within the three weeks after the intervention ended, nor did they show a significant decrease in maintenance of the facts. It is important to note, however, that the tutee group had a visible, although insignificant, drop in number of digits correct at the follow-up measurement.

When examining the difference between groups, which only existed at post and follow-up scores, the group means were ordered as predicted; the reciprocal group had the highest means at post and follow-up measurements, followed closely by tutees, then tutors, and lastly, the control group. The two groups of students being tutored did not significantly differ in their academic gains though reciprocal students had slightly higher scores than tutees, as expected. This information suggests the two different forms of peer tutoring are comparable in effectiveness, as they did not produce significantly different results in the students being tutored.

Although all students involved in peer tutoring outperformed the control group, pairwise comparisons indicated that only the reciprocal and the tutee groups produced significantly more gains than the control group. In other words, while tutors had a significant increase from pre to

post-scores, their gains were not significantly different than those of the control group. These findings somewhat support the hypothesis that tutors would not show substantial increases in math scores. Also of importance, the reciprocal group showed significantly more gains than the tutor group. The fact that tutors did not gain significantly more than the control group and gained significantly less than the reciprocal students provides reason to utilize reciprocal peer tutoring. The results imply reciprocal tutoring successfully remediates more at-risk students concurrently as opposed to non-reciprocal tutoring, which only produced significant gains in tutees.

When considering the capability of peer tutoring in remediating at-risk students, the results indicate an efficacious pre-referral intervention as expected. When individual scores were compared to national benchmarks, 12 of the 15 reciprocal students and 10 of the 14 tutee students fell within the "instructional" range rather than the "frustrational" range at both post and follow-up scores. In contrast, only 6 of the 16 tutors and 5 of the 16 control students reached the instructional range at post and follow-up data collections. Peer tutoring can and should be implemented as an academic intervention with at-risk students, as it is evident that students being tutored by their peers make enough gains to be considered in the average achievement range. Also, it is important to note the relatively low frequency and duration of this intervention, as tutoring sessions were 3 minutes in length and took place for an average of 3 times a week until the 15 session limit was met. In the case of students who did not respond to this intervention, perhaps a more intense peer tutoring intervention (more frequent and/or longer duration) would be more likely to result in attaining an average level of achievement.

Social Findings

In terms of social benefits associated with peer tutoring, the findings in this study were less clear than in previous research. To begin with, pre-scores differed vastly between groups.

(Figure 2). Analyses at post and follow-up scores, however, showed that all three tutoring groups no longer differed, and they were all significantly higher than the control group, which is reflected in the effects sizes displayed in Tables 6 and 7.

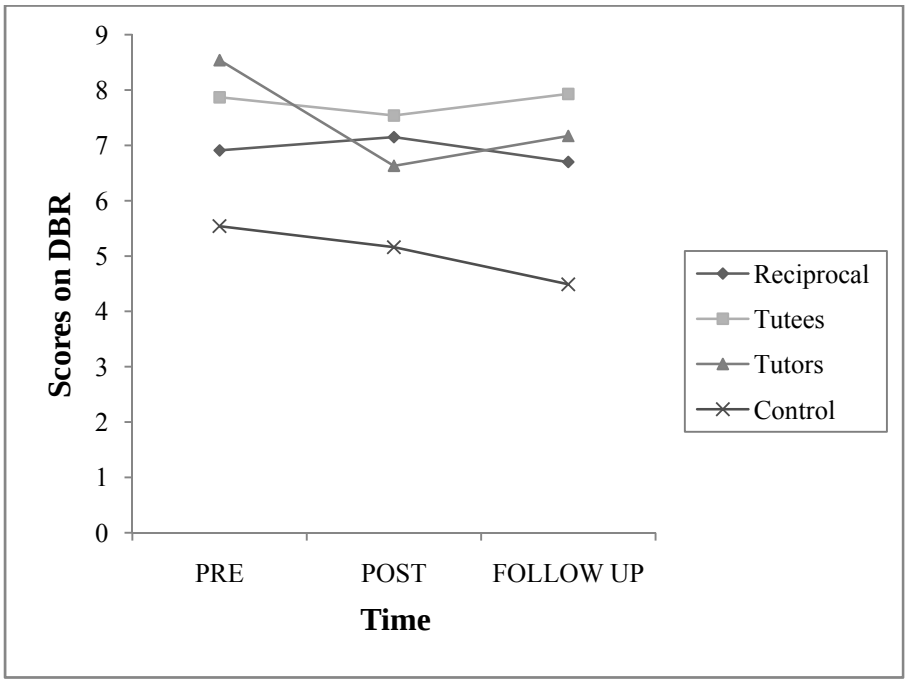

Figure 2. Mean social scores on Direct Behavior Report (1 = very poor, 9 = excellent).

Table 6

Overall Effect Size Display for Post-Intervention DBR Scores

	Effect Size expressed as Cohen's *d*	Effect Size expressed as *r*	Success rate as percentage of subjects showing improvement r_{besd}	
Group			Treatment Group	Control Group
Reciprocal	*d* = .92	*r* = .42	71%	29%
Tutee	*d* = 1.20	*r* = .51	76%	24%
Tutor	*d* = .67	*r* = .32	66%	34%

Table 7

Overall Effect Size Display for Follow-up DBR Scores

	Effect Size expressed as Cohen's *d*	Effect Size expressed as *r*	Success rate as percentage of subjects showing improvement r_{besd}	
Group			Treatment Group	Control Group
Reciprocal	*d* = .89	*r* = .41	71%	29%
Tutee	*d* = 1.77	*r* = .66	83%	17%
Tutor	*d* = 1.35	*r* = .56	78%	22%

The hypothesis that students participating in reciprocal tutoring would have the greatest social gains was not supported. However, due to the fact the groups varied to such a dramatic extent before the interventions were implemented, we cannot make any confident conclusions based on the comparisons between groups.

When analyzing the results within each group, the social effects over time did not confirm the author's initial hypothesis. None of the groups' social ratings increased significantly as a result of the peer tutoring interventions. When considering these social portions of the results however, it is vital to note the subjective nature of the dependent measure, the Direct Behavior Report (DBR). Although it is designed to assess behavior at one point in time, there is a substantial possibility that past behavior influences teachers' ratings on a DBR. For instance, a teacher can stigmatize certain students, which can result in consistently lower ratings for students that have had past behavior problems but may not have been misbehaving during the day or time of the rating. Furthermore, teachers may have considered students' social behavior over the course of the entire day and rate it as an average rather than only selecting one interactive activity in the morning and afternoon, as directed. Likewise, teachers may have considered other types of behavior when rating students rather than purely their social interactions with peers. For

these reasons and perhaps many others, the DBR ratings were more variable than expected and clear conclusions could not be drawn from the results.

Anecdotally, all students involved in peer tutoring were excited to participate each day and were visibly disappointed when their teacher did not allow them to participate or when their partner was absent. They also asked each day after the intervention ended when they could continue peer tutoring. As a result, after all data was gathered for this study, each participant was given a set of flashcards used in the study to peer tutor on his/her own time, which many reported doing. It was apparent the students enjoyed participating in peer tutoring, while the teachers and school administrators were very pleased to report gains in math class and on standardized tests. The only complaints came from the non-reciprocal students, who asked quite frequently when they would switch roles as tutor and tutee without knowing students from other classrooms were doing so. They were allowed to switch roles once all data were collected.

Limitations and Directions for Future Research

There are several limitations in this study that warrant discussion. First, although classrooms were randomly assigned to groups, this did not result in random assignment of students to groups. As previously described, all students within a classroom participated in the same group to control for preference effects and attrition. This may have been beneficial, as many tutees asked when they were going to be a tutor and may have been very disappointed to know there was a condition in which both students were allowed to tutor. Although this may have prevented some dropping out or lack of cooperation in the assigned condition, it came at a cost in that assignment was only semi-random due to classrooms being pre-existing groups. It is important to note, however, that multiple classrooms were randomly assigned to each group to try to account for any individual classroom differences that may have affected the overall group.

Follow-up studies can address this issue by randomly assigning students to conditions and informing them that they can participate in another condition once the experimental intervention is completed, e.g. “You can be a tutor after 10 more sessions.”

Another limitation is the difference between the tutoring material (flashcards) and the academic dependent measure (CBM probes). This difference assumes that the students could generalize the individual facts on the tutoring flashcards to completing multiple problems on a probe worksheet. There is no apparent reason students would have difficulty answering the same math problems in a different format; however, some students could not answer the same problems on the probe that they had mastered with the flashcards. Furthermore, after mastering multiple sets of flashcards, some students still performed below-average on the post and follow-up CBM probes. Hence, the generalization of peer-tutored facts requires further investigation.

A third limitation is the variability in social scores among groups throughout the intervention. This inconsistency may have reflected true variability among students, in which case several of the control participants exhibited poor social skills that may have contributed to their lack of academic growth (DiPerna & Elliott, 2002). The variability could also be attributed to differences in the way teachers rated their students using the DBR, with some teachers being more lenient raters than others. Aside from student and teacher variability, setting events may also affect ratings on DBRs. For example, the students participated in class-wide field trips and a school-wide field day the week DBR post-scores were gathered. A few teachers noted the unusually poor behavior of some of their students during that fairly unstructured week, which may have resulted in unrepresentative ratings. Likewise, the follow-up score ratings were gathered in the last few weeks before summer break, in which many teachers complained of students frequently misbehaving. And although the DBR may not be the best progress

monitoring tool for measuring social behavior, it is the best to date for use in large group designs, especially compared to direct observation which is time-intensive and not representative of behavior unless numerous observations are completed (Hintze & Matthew, 2004). Regardless of the tool used to measure changes in behavior, there will undoubtedly be extraneous variables that are affecting behavior but are difficult to control. Nevertheless, there is room for improving the instruments used to measure change in targeted behaviors throughout an intervention.

The last limitation concerns elements of the procedure that were arbitrarily chosen. Selecting a session length of 3 minutes, the number of flashcards in a set as 10, and the criteria for achieving mastery as two consecutive scores of 10 had no scientific basis. The possibility of manipulating these parameters and measuring any effects produced by variations is open to future research.

In summary, the current study demonstrated that reciprocal and non-reciprocal peer tutoring were comparable in their efficacy as academic interventions. However, the larger amount of students reciprocal tutoring is able to remediate simultaneously speaks to its efficiency as compared to one-way tutoring. This distinction is also important in the fact that below-average achievers were able to accurately and effectively tutor each other to the extent that they both achieved significant academic gains. In other words, there are situations in which at-risk students do not need to be paired with average or above-average students to produce increases in performance. As a result, twice the number of children can receive the instruction they require concurrently. Moreover, the outcomes of this study demonstrate there are feasible peer-mediated interventions that can prevent future academic failure and inappropriate referrals to special education. Although the social effects of these two tutoring programs require further examination before solid conclusions can be made, the results of this study further validate the

academic utility of peer tutoring, especially when it is reciprocal in nature. The overall findings advocate for more wide-spread use of students, even those identified as at-risk, as primary delivery agents of interventions.

REFERENCES

Adams, G. L. & Engelmann, S. (1996). *Research on direct instruction: 25 years beyond DISTAR.* Seattle, WA: Educational Achievement Systems.

Balenzano, S., Agte, L. J., McLaughlin, T. F., & Howard, V. F. (1993). Training tutoring skills with preschool children with disabilities in a classroom setting. *Child and Family Behavior Therapy, 15,* 1–36.

Beirne-Smith,M. (1991). Peer tutoring in arithmetic for children with learning disabilities. *Exceptional Children*, *57,* 330–337.

Britz, M. W., Dixon, J., & McLaughlin, T. F. (1989). The effects of peer tutoring on mathematics performance: A recent review. *B. C. J. Spec. Educ.* 13: 17–33.

Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375) New York: MacMillan.

Byrd, D. E. (1990). Peer tutoring with the learning disabled: A critical review. *Journal of Educational Research, 84,* 115-118.

Cardenas, J. A., Harris, R., del Refugio Robledo, M., & Supik, J. D. (1991, April). *Valued youth program dropout prevention strategies for at-risk students*. Paper presented at the annual meeting of the American Education Research Association, Chicago, IL.

Chafouleas, S. M., Christ, T. J., Riley-Tillman, T. C., Briesh, A. M., & Chanese, J. A. (2007). Generalizability and Dependability of Direct Behavior Ratings to Assess Social Behavior of Preschoolers. *School Psychology Review, 36,* 63-79.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cohen, P. A., Kulik, J. A., & Kulik C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings**.** *American Educational Research Journal, 19,* 237-248.

Cybriwsky, C. A., & Schuster, J. W. (1990). Using constant time delay procedures to teach multiplication facts. *Remedial and Special Education, 11,* 54-59.

Delquadri, J. C. (1978). An analysis of the generalization effects of four tutoring procedures on the oral reading responses of eight learning disability children (Master's thesis, University of Kansas, 1978). OCLC's Experimental Thesis Catalog.

Delquadri, J. C., Greenwood, C.R., Stretton, K., & Hall V. R. (1983). The peer tutoring spelling game: A classroom procedure for increasing opportunity to respond and spelling performance. *Education and Treatment of Children, 6,* 225-239.

Delquadri., J., Greenwood C. R., Whorton, D., Carta J. J., & Hall, R. V. (1986). Classwide peer tutoring. *The Exceptional Child, 52,* 535-542.

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

DiPerna, J. C. & Elliott, S. N. (2002). Promoting academic enablers to improve student achievement: An introduction to the mini-series. *School Psychology Review, 31*, 293-297.

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.

Dupaul, G, J., Ervin, R. A., Hook, C. L., & McGoey, K. E. (1993). Peer tutoring for children with attention deficit hyperactivity disorder: Effects on classroom behavior and academic performance. *Journal of Applied Behavior Analysis, 31,* 579-592.

Fantuzzo, J. W., King, J. A., & Heller, L. R. (1992). Effects of reciprocal peer tutoring on mathematics and school adjustment: A component analysis. *Journal of Educational Psychology, 84,* 331–339.

Fantuzzo, J.W., Polite, K, & Grayson, N. (1990). An evaluation of reciprocal peer tutoring across elementary school settings. *Journal of School Psychology, 28,* 309-323.

Faul, F. (2006). G*power version 3.0.5 [Computer software and manual]. Retrieved from http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/

Fuchs, L. S., Fuchs, D., Phillips, N. B., Hamlett, C. L., & Karns, K. (1995). Acquisition and transfer effects of classwide peer-assisted learning strategies in mathematics for students with varying learning histories. *School Psychology Review, 24,* 604-620.

Gast, D. L., Ault, M. J.,Wolery, M., Doyle, P. M., & Belanger, S. (1988). Comparison of constant time delay and the system of least prompts in teaching sight word reading to students with moderate retardation. *Education and Training in Mental Retardation*, *23,* 117–128.

Ginsburg-Block, M. D., & Fantuzzo, J.W. (1998). An evaluation of the relative effectiveness of NCTM standards-based interventions for low-achieving urban elementary students. *Journal of Educational Psychology, 90,* 560–569.

Greenwood, C. R., Carta, J. J., Kamps, D., & Hall, R. V. (1988). The use of classwide peer tutoring strategies in classroom management and instruction. *School Psychology Review, 17,* 258-275.

Greenwood, C. R., Carta, J. J., & Maheady L. (1991). Peer tutoring programs in the regular education classroom. In G. Stoner, M. R. Shinn, & H. M. Walker (Eds.), *Interventions for Achievement and Behavior Problems* (pp. 179-200). Bethesda, MD: National Association of School Psychologists.

Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1989). Longitudinal effects of classwide peer tutoring. *Journal of Educational Psychology, 81,* 371–383.

Greenwood, C. R., Dinwiddie, G., Terry, B., Wade, L., Stanley, S. O., Thibadeau, S., & Delquadri, J. C. (1984). Teacher- versus peer-mediated instruction: An ecobehavioral analysis of achievement outcomes. *Journal of Applied Behavior Analysis, 17,* 521–538.

Hintze, J. M., & Matthew, W. J. (2004). The generalizability of systematic direct observations across time and setting: A preliminary investigation of the psychometrics of behavioral observation. *School Psychology Review, 33,* 258-270.

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

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.

Hughes, T. A., Fredrick, L. D., & Keel, M. C. (2002). Learning to effectively implement constant time delay procedures to teach spelling. *Learning Disabilities Quarterly*, *25,* 209–222.

Keel, M. C., Dangel, H. L., & Owens, S. H. (1999). Selecting instructional interventions for students with mild disabilities in inclusive classrooms. *Focus on Exceptional Children*, *31*, 1–16.

Keel, M. C., & Gast, D. L. (1992). Small group instruction for students with learning disabilities: Observational and incidental learning. *Exceptional Children, 58*, 357–368.

Koscinski, S. T., & Gast, D. L. (1993). Use of constant time delay in teaching multiplication facts to students with learning disabilities. *Journal of Learning Disabilities*, *26,* 533–567.

Kovaleski, J. F., Gickling, E. E., Morrow, H., & Swank, P. R. (1999). High versus low implementation of instructional support teams: A case for maintaining program fidelity. *Remedial and Special Education, 20*, 170-183.

Kroesbergen, E. H. & Van Luit, J. E. H. (2003). Mathematics interventions for children with special educational needs: A meta-analysis. *Remedial and Special Education, 24,* 97-114.

Lazerson, D. B., Foster, H. L., Brown, S. I., & Hummel, J. W. (1988). The effectiveness of cross-age tutoring with truant, junior high school students with learning disabilities. *Journal of Learning Disabilities, 2,* 253-255.

Levin, H. M., Glass, G. V., & Meister, G, R. (1987). Cost-effectiveness of computer-assisted instruction. *Evaluation Review, 11,* 50-72.

Maheady, L., Harper, G. F., & Mallette, B. (2001). Peer-mediated instruction and interventions and students with mild disabilities. *Remedial and Special Education, 22,* 4-14.

Math – Curriculum Based Measurement (M-CBM). Retrieved August 23, 2007, from www.aimsweb.com/uploaded/files/MCBM6pp.ppt.pdf

Math Computation Probes – Basic Math Operations. Retrieved August 23, 2007, from http://www.interventioncentral.org/htmdocs/interventions/cbmwarehouse.php

Mathur, S. R. & Rutherford, R. B. (1991) Peer-mediated interventions promoting social skills of children and youth with behavioral disorders. *Education and Treatment of Children, 14,* 227-242.

Mattingly, J. C., & Bott, D. A. (1990). Teaching multiplication facts to students with learning problems. *Exceptional Children, 56,* 438-449.

McDougal, J. L., Clonan, S. M., & Martens, B K. (2000). Using organizational change procedures to promote the acceptability of prereferral intervention services: The school-based intervention team project. *School Psychology Quarterly, 15,*149-171.

Miller, L., Kohler, F., Ezell, H., Hoel, K., & Strain, P. (1993). Winning with peer tutoring. *Preventing School Failure*, *37,* 14-18.

Nazzal, A. (2002). Peer tutoring and at-risk students: An exploratory study. *Action in Teacher Education, 24,* 68–80.

Noell, G. H., Duhon, G. J., Gatti, S.L., & Connell, J.E. (2002). Consultation, follow-up and implementation of behavior management interventions in general education. *School Psychology Review, 31,* 217-234.

Pigott, E. H., Fantuzzo J. W., Heggie, D. L., & Clement P. W. (1984). A student administered group-oriented contingency intervention: Its efficacy in a regular classroom. *Child and Family Behavior Therapy, 6,* 41-55.

Polirstok, S. R., & Greer, R. D. (1986). A replication of collateral effects and a component analysis of a successful tutoring package for inner-city adolescents. *Education and Treatment of Children, 9,*101–121.

Robinson, D. R., Schofield, J. W., & Steers-Wentzell, K. L. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. *Educational Psychology Review, Vol. 17,*327-362.

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,* 240–257.

Rosenthal, R., Rosnow, R., & Rubin, D. R. (2000). Contrasts and effect sizes in behavioral research. New York: Cambridge University Press.

Schuster, J. W., Stevens, K. B., & Doak, P. K. (1990). Using constant time delay to teach word definitions. *Journal of Special Education, 24,* 306-318.

Stevens, K. B., Blackhurst, A. E., & Slaton, D. B. (1991). Teaching memorized spellings with a microcomputer: Time delay and computer assisted instruction. *Journal of Applied Behavior Analysis, 24,* 153-160.

Stevens, K. B., & Schuster, J. W. (1988). Time delay: Systematic instruction for academic tasks. *Remedial and Special Education*, *9,* 11–16.

Tabacek, D. A., McLaughlin, T. F., & Howard, V. F. (1994). Teaching preschool children with disabilities tutoring skills: Effects on pre-academic behaviors. *Child and Family Behavior Therapy, 16,* 43–63.

Telecsan, B. L., Slaton, D. B., & Stevens, K. B. (1999). Peer tutoring: Teaching students with learning disabilities to deliver time delay instruction. *Journal of Behavioral Education*, *9,* 133–154.

Thomas, R. L. (1993). Cross-age and peer tutoring. *ERIC Clearinghouse on Reading, English, and Communication, Digest 78.*

Topping, K. J. (2005). Trends in peer learning. *Educational Psychology, 25,* 631-645.

Topping, K. J., & Whiteley, M. (1993). Sex differences in the effectiveness of peer tutoring. *School Psychology International, 14,* 57–67.

Touchette, P. E. (1971). Transfer of stimulus control: Measuring the moment of transfer. *Journal of the Experimental Analysis of Behavior*, *15,* 347–354.

Witt, J. C., Noell, G. H., LaFleur, L. H., & Mortenson, B. P. (1997). Teacher use of interventions in general education settings: Measurement and analysis of the independent variable. *Journal of Applied Behavior Analysis, 30,* 693-696.

Wolfe, J. A., Fantuzzo, J., W., & Wolfe, P. K. (1986). The effects of reciprocal peer management and group contingencies on the arithmetic proficiency of underachieving students. *Behavior Therapy, 17,* 253-265.

Wolfe, J. A., Fantuzzo, J. W., & Wolter, C. (1984). Student-administered group-oriented contingencies: A method of combining group-oriented contingencies and self-directed behavior to increase academic productivity. *Child and Family Behavior Therapy, 6,* 45-60.

VITA

Keri F. Menesses is a candidate for the Master of Arts degree in the school psychology program at Louisiana State University. She graduated *summa cum laude* with a Bachelor of Science degree in psychology from Louisiana State University in 2005. Upon receiving her Master of Arts in December 2008, she will pursue a doctorate degree under the supervision of Dr. Frank M. Gresham.