THE EFFECTIVENESS OF PEER TUTORING PROGRAMS
IN ELEMENTARY SCHOOLS

by

Melinda Gee

A Plan-B paper submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Psychology

UTAH STATE UNIVERSITY
Logan, Utah

2004
ABSTRACT

The Effectiveness of Peer Tutoring Programs in Elementary Schools

by

Melinda Gee, Master of Science

Utah State University, 2004

Major Professor: Dr. Gretchen A. Gimpel
Department: Psychology

The present review examined the effectiveness of three peer tutoring programs: cross-age peer tutoring, Classwide Peer Tutoring (CWPT), and Peer-Assisted Learning Strategies (PALS), for elementary students in the academic areas of math and reading. The research reviewed indicates students who participated in cross-age peer tutoring and CWPT had improved test scores on basic math facts as well as increased math scores on standardized assessments. Students also showed improvement in reading fluency, reading comprehension, spelling, and reading level after participating in cross-age peer tutoring and CWPT. High-, average-, and low-achieving students as well as students in special education benefitted from these peer tutoring programs. In addition, both tutors and tutees received benefits from the tutoring programs, although fewer studies examined outcomes for tutors. Although cross-age peer tutoring, and CWPT resulted in mostly positive outcomes, this was not the case for the PALS program. Students participating in this program demonstrated little change from pre- to post-assessment. Since the PALS studies
were better designed with tighter experimental control, it is less likely results were due to another variable. The cross-age and CWPT studies lacked the tight experimental control of the PALS studies.
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>REVIEW OF REVIEWS</td>
<td>5</td>
</tr>
<tr>
<td>TYPES OF TUTORING</td>
<td>13</td>
</tr>
<tr>
<td>Reciprocal Peer Tutoring</td>
<td>13</td>
</tr>
<tr>
<td>Reverse-Role Tutoring</td>
<td>14</td>
</tr>
<tr>
<td>Classwide Student Tutoring Teams</td>
<td>15</td>
</tr>
<tr>
<td>Collaborative Strategic Reading</td>
<td>15</td>
</tr>
<tr>
<td>Numbered Heads Together</td>
<td>16</td>
</tr>
<tr>
<td>MAIN TUTORING PROGRAMS</td>
<td>18</td>
</tr>
<tr>
<td>Cross-Age Peer Tutoring</td>
<td>18</td>
</tr>
<tr>
<td>Classwide Peer Tutoring</td>
<td>26</td>
</tr>
<tr>
<td>Peer-Assisted Learning Strategies</td>
<td>39</td>
</tr>
<tr>
<td>CRITIQUE OF CURRENT STUDIES</td>
<td>49</td>
</tr>
<tr>
<td>BEST PRACTICES IN PEER TUTORING</td>
<td>53</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>61</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Metropolitan Achievement Test Normal Curve Equivalent (NCE) Means Summary</td>
</tr>
<tr>
<td>2</td>
<td>Comprehensive Test of Basic Skills (CTBS-U) Summary</td>
</tr>
<tr>
<td>3</td>
<td>Summary of Results for the Comprehensive Reading Assessment Battery-Revised (CRAB-R) and the Woodcock Reading Mastery Tests-Revised (WRMT-R)</td>
</tr>
<tr>
<td>4</td>
<td>Summary of Current Studies for Cross-Age, CWPT, and PALS Tutoring Programs</td>
</tr>
</tbody>
</table>
INTRODUCTION

Peer tutoring has been frequently used over the years in attempts to increase student academic performance. Classrooms are becoming more diverse with children with a wide range of learning abilities in each classroom. Teachers must discover ways to ensure all their students receive the best education possible, but many are finding this difficult. Peer tutoring is one way to help spread the teacher's effectiveness, by teaching students to be teachers.

Peer tutoring involves students teaching each other. The student who teaches is usually called the tutor, while the student who is taught becomes the tutee. The students are paired by the teacher. One way for the students to be paired is by, first, ranking the students' ability in the subject area from the highest ability to the lowest ability. After students are ranked, they are then put in pairs with one higher skilled and one lower skilled student. Another way students may be paired is by random assignment or with older peers. The students in each dyad are trained in the peer tutoring process. Training may involve one individual student being trained or the entire class (as in group peer tutoring programs). Training involves instructing the tutor on the explanations to be given to the tutee.

In the past, peer tutoring has “represented an economical means of providing individualized, intensive instruction to academically needy pupils” (Fuchs, Fuchs, Mathes, & Simmons, 1997, p. 179). Peer tutoring has also been used as an intervention to increase students’ ability to actively respond to prompts in math, reading, and spelling (Ezell,
Kohler, & Strain, 1994), as well as to increase peer interactions with students with social

Recent reviews (Maheady, Harper, & Mallette, 1991, 2001; Utley & Mortweet, 1997) have examined peer tutoring as a component of peer mediated instruction and interventions, but not as an isolated intervention. In addition to peer tutoring, peer mediated interventions include peer modeling (usually used for social skills, a child will demonstrate an appropriate skill while the other child imitates it), peer initiation training (a teacher trains peers how to elicit appropriate social behaviors from a target child), and peer monitoring (e.g. token systems) (Utley & Mortweet, 1997). A meta-analysis conducted by Cohen, Kulik, and Kulik (1982) twenty years ago found positive results (i.e. increased scores for tutees) on achievement tests when peer tutoring was used. It was also found that tutored students' positive attitudes toward the subject matter increased as well. The tutors also received positive outcomes from tutoring. Tutors gained a better understanding of the information they were covering in the program which was measure by examinations on the subject matter. An update of the research of that review is needed, since no comprehensive review of the academic effects of peer tutoring has been completed since 1982.

The purpose of this paper is to examine and summarize the literature on the effectiveness of peer tutoring interventions for elementary-aged, special and general education students. Based on this evaluation, recommendations for changes in current programs and ideas for the future will be made.
This paper will begin with a brief overview of what peer tutoring is and how it is currently being used in the classrooms. After the overview, a review of recent reviews about peer mediated interventions will be discussed. Then, other types of peer tutoring not thoroughly researched will be covered. Next, the main tutoring programs will be examined, cross-age peer tutoring, Classwide Peer Tutoring (CWPT) and Peer-Assisted Learning Strategies (PALS). A review of the effectiveness of these programs in mathematics and reading or writing curriculum will then be presented. Next, a critique of these programs which will include difficulties and limitations of peer tutoring programs will be presented. Finally, recommendations for peer tutoring in the future will be given based on the findings from this review.

Descriptors and Inclusion/Exclusion Criteria

To locate studies for inclusion in this review, databases PsychInfo and ERIC were searched. Various terms for peer tutoring were matched with different terms for academic improvement. The specific descriptors for peer tutoring were: peer tutoring, cross-age peer tutoring, peer mediated interventions, classwide peer tutoring, peer assisted learning strategies, and peer helpers. The descriptors for academic improvement were: school performance, academic improvement, and academic performance. Other terms were used in combination with the above descriptors, such as: elementary students, special education, general education and regular education. Other references were located by conducting a bibliography search of the sources previously found.

Inclusion Criteria

The following list were used as the criteria for the selection of articles:
1) Articles must include peer tutoring programs that include training sessions for the teachers and the students to instruct them on how to conduct the peer tutoring in the classroom. It was important that each program have similar tutoring procedures so comparisons to other studies can be made.

2) Peer tutoring must focus on math and reading or writing. This criterion was to narrow the search for studies and the areas where most of the research is focused.

3) Academic improvement must be the focus of the peer tutoring program.

4) Only peer tutoring programs aimed at elementary students (grades K-6) were included because of the limited research on secondary students.

5) Only articles with empirical data were included.

6) Studies must include peer tutoring alone or results for peer tutoring alone. Studies that used peer tutoring in combination with other interventions were excluded.

7) Only studies published after the last review, (i.e. 1982 and later), were included.

8) Only studies using the programs of cross-age peer tutoring, Classwide Peer Tutoring (CWPT), and Peer-Assisted Learning Strategies (PALS) were included. These programs are the ones most frequently evaluated by researchers in this area.
REVIEW OF REVIEWS

Recent reviews (Maheady, Harper, & Mallette, 1991, 2001; Utley & Mortweet, 1997) have examined peer tutoring as a component of peer mediated instruction and interventions but not as an isolated intervention. The next section will contain a discussion on the academic outcomes of the recent reviews on peer mediated interventions, as well as other outcomes such as social and behavioral changes.

Maheady, Harper, and Mallette (1991, 2001) conducted two of the reviews on peer mediated interventions. Both of the reviews examined the effects of peer mediated interventions on special education students. All of the studies reviewed used students with disabilities as the peer tutors. Both reviews reported the same basic conclusions, but in the 2001 review, the authors included additional studies that were not in their 1991 review. The outcomes from the additional studies were the same as the outcomes discussed in the first review. The 2001 review also include two new programs, Classwide Student Tutoring Teams (CSTT) and Numbered Heads Together (NHT), which had not been developed at the time of the 1991 review.

The 1991 review separated the articles into two different groups, one looking at cooperative learning and the other at peer tutoring. Brief descriptions were given about various cooperative learning and peer tutoring programs that have been used to effectively help students having academic problems. After the authors review the outcomes of the different programs, they discuss advantages and disadvantages of cooperative learning and peer tutoring compared to teacher-led instruction.
Cooperative learning involves allowing groups of three to six students to work together on different projects. It was noted that although cooperative learning has been thoroughly researched the findings for mildly handicapped students were mixed. Half of the studies favored cooperative learning and the other half produced “nonsignificant differences” (p. 83) based on the examination of six previous reviews. Maheady et al. (1991) concluded that the key to the effectiveness lies in how structured the groups are (i.e. having objectives for the group along with individual accountability of each group member).

Peer tutoring was also discussed as a part of peer mediated interventions. One peer tutoring program mentioned was cross-age peer tutoring, where the tutor was mildly handicapped and the tutees were disabled as well. The researchers reviewed 25 studies on cross-age peer tutoring that had been examined in two previous reviews by Osguthorpe and Scruggs (1986) and Scruggs, Mastropieri, and Richter (1985). One conclusion from these studies was that the tutors with disabilities were able to learn teaching skills. In addition, the tutors and the tutees increased their positive social interactions with others following their tutoring experiences. The tutees and the tutors both made academic gains as well. Although, it was not a variable examined in the studies, some researchers reported decreases in absenteeism and office referrals with participation in cross-age peer tutoring (Maheady et al., 1991).

Only one study was reviewed that examined reverse-role tutoring, in which the student with a disability switched roles (i.e. from tutee to tutor) with a nondisabled peer. The results of the study indicated that the student with a disability was able to learn
teaching skills. The study did not report whether or not the student improved academically on the subject. The self-concepts of the students with disabilities did not show statistically significant improvements as the researchers from the study hypothesized.

Classwide Peer Tutoring (CWPT) was shown to be an effective program as well. Based on the eight studies examined it was concluded that CWPT increased students’ performance in spelling, reading, and math. However, specific numerical data was not given.

One difficulty that was encountered in many of the studies was that of implementation. Time restraints on teachers (e.g. the length of time it took teachers to teach the students the tutoring procedures) along with an increase in the noise level in the classroom were some of the difficulties reported. However, the researchers did not examine these variables, so it was not clear whether the variables influenced the outcomes in any of the studies included in the review.

One conclusion from the initial (Maheady et al., 1991) review was that peer mediated interventions and peer tutoring can be powerful teaching strategies which should be used more often in classrooms with students who have both academic and behavioral challenges. In the studies reviewed, both the tutee and the tutor showed improvements, although not all of the studies reported outcomes for the tutors. Improvements were noted not only in academic areas, such as reading, math, and spelling, but also in interpersonal areas and social development, such as appropriate peer interactions and reduction in the number of disciplinary referrals. Maheady et al. (1991) also concluded that, in their opinions, the effectiveness of peer mediated interventions were due to the lowering of
student-teacher ratios, increasing the number of opportunities to respond, the immediacy of positive, corrective feedback, and the opportunity for one-to-one instruction.

In Maheady, Harper, and Mallette's (2001) second review, they conclude that the literature since 1991 shows the same outcomes for peer mediated interventions as in their first review with no additional information. The authors report a number of new studies as well as replication studies were included in the 2001 review but a specific number of studies is not given. Two peer mediated programs that had been developed since the last review were briefly mentioned: Classwide Student Tutoring Teams (CSTT) and Numbered Heads Together (NHT).

In CSTT, students are placed in groups of three to four people. The groups are given stacks of cards with answers that correspond with questions on a study guide. The groups practice going through the questions and answer as a group. Points are given for correct answers and for correcting incorrect answers. In NHT, students are placed in groups of four members. After the teacher gives some instruction, the teams are told to put their heads together to find a solution.

Only two studies examined the effectiveness of CSTT. One of the studies stated that students made "substantial" gains on mathematics performance. However, substantial was never defined specifically. The second study reported that students showed an average of 30% improvement on solving math word problems. NHT was shown to be effective when compared with whole group questioning (i.e. when the teacher questions the entire class at one time). In one of the two studies, one group of third graders showed a 15% increase in correct answers on daily social studies quizzes. Students were on task
twice as much during NHT than whole group. During a sixth grade science class, NHT class averages increased by one letter grade (i.e. from a C to a B) following the implementation of NHT.

Maheady, Harper, and Mallette (2001) also discussed future directions for research evaluating peer mediated interventions. One direction is in connecting the intervention to specific instructional outcomes. Many times it is not explained when to use peer mediated interventions. The researchers stated it would be helpful to explain when these interventions work best. For example, many peer mediated interventions seem to work better for basic math facts, such as multiplication facts, rather than problems requiring multiple steps to complete. Research on methods of implementing programs in the classrooms is also needed, as they indicate “a clear gap remains between our empirical knowledge and applied practice” (p. 11, 2001). The final research direction would be to match the curriculum with peer mediated interventions. Maheady et al. (2001) indicate that education seems to encourage the development of new curricular materials (e.g. textbooks and videos) instead of improving the current curriculum and how it is presented. For example, a student who has limited reading skills may be given different books on different grade levels to try and help improve his reading fluency. Implementing peer mediated interventions, such as peer tutoring, may be a better alternative to improve the student’s fluency.

Utley and Mortweet (1997) also reviewed peer mediated interventions by first, examining and explaining the components of peer mediated interventions, and, second, describing various programs that use the components of peer mediated interventions. A
discussion of the implications for research and practice conclude the review. Throughout the review, the number of studies examined is not specified.

Utley and Mortweet (1997) list the following as components of peer mediated interventions: peer modeling, peer initiation training, peer monitoring, peer networking, peer tutoring, and group-oriented contingencies. Peer modeling involves having a student demonstrate a skill or concept to a less skilled student. Peer modeling is usually used for social skills training, but can be for academic skills as well. Utley and Mortweet (1997) noted that the research shows that peer modeling shows slight improvements in social behavior, but significant changes will not occur unless it is used in combination with other procedures, such as peer tutoring.

Peer initiation training has been used most frequently with children with disabilities. Teachers instruct the peers on how to keep the target child performing the appropriate social interaction. For example, teachers may teach peers how to ask for help or describe social interactions while playing with the target child. Some behaviors the peers will need to master are maintaining eye contact and initiating conversation while interacting with the target child. These are common behaviors that are necessary for appropriate social interactions. The studies on peer initiation training show an increase in social interactions for the target child, but the research indicated that it is unlikely skills will generalize without additional unspecified programmed interventions (Utley & Mortweet, 1997).

Another component, peer monitoring, entails having a peer monitor appropriate behavior of another student. The peer monitor gives points or tokens to the target student
when an appropriate behavior is displayed. A series of three studies indicated that peer monitoring decreased inappropriate behaviors during transition times in a kindergarten classroom and decreased the amount of time students were off task in a fifth grade classroom (Utley & Mortweet, 1997).

The next component, peer networking, is used most frequently with children with disabilities, particularly autism. The goal of networking is to create a support system of current friends along with socially competent peers. For example, a group of students (network) may be preselected to befriend a student with autism. The network of peers discuss methods and times of the day to interact with the target student. The network peers are taught how to reinforce positive social interactions. Only two studies examined the effectiveness of peer networking. Utley and Mortweet (1997) stated that one finding was that friendships were enhanced between children with autism and their peers. The other finding indicated that children with autism learned social and communication skills.

Group contingencies, the final component, are those programs that involve the entire classroom. According to Utley and Mortweet (1997), group contingencies are effective due to peers being a powerful influence on behavior in the natural classroom setting. One study reviewed examined a group contingency with a class of preschoolers with disabilities. Three of the participants had autism and six had other disabilities not specified. The results showed that social interactions increased between the students with autism and their classmates. Results from three other studies indicated that group contingencies reduced the number of disruptive behaviors and increased school work production and efficiency (Utley & Mortweet, 1997).
The Utley and Mortweet review (1997) indicated that peer mediated interventions are effective for students with and without disabilities, although students with disabilities were the focus of this review. Peer modeling, peer initiation training, peer monitoring, and peer networking resulted in improvements in appropriate social interactions. Research showed that peer tutoring and group-oriented contingencies lead to both academic and social improvements. Students seemed to benefit academically and socially when the components of peer mediated interventions were used in both the regular and special education classrooms. Some advantages of peer mediated interventions were an increase in the amount of time engaged as well as an increase in the number of opportunities to respond. Some disadvantages were the time required to train the peers and the difficulty monitoring the quality of the instruction the target child received.
TYPES OF TUTORING

Peer tutoring, one type of peer mediated intervention addressed in the above reviews, has been used as an intervention for many years. For instance, Aristotle had many students and implemented peer tutoring to help him teach (Enright & Axelrod, 1995). With peer tutoring being used for such a long period of time, it seems logical that many different kinds of programs would be developed throughout the years. Even though the focus of this paper will be on cross-age peer tutoring, Classwide Peer Tutoring (CWPT), and Peer-Assisted Learning Strategies (PALS), there are many other types of peer tutoring programs currently being used. These are each briefly reviewed here.

Reciprocal Peer Tutoring

Reciprocal peer tutoring can be done with one dyad or multiple dyads. It involves having each member of the dyad take turns being the “student” or tutee and the “teacher” or tutor. The teacher assigns one of the students to be the tutor first. The tutor then teaches the tutee about a certain concept. Once the concept is mastered, then the tutee becomes the tutor for the next concept. Ginsburg-Block and Fantuzzo (1997) demonstrated that reciprocal peer tutoring lead to some positive outcomes such as a reduction of inappropriate behaviors and an increase in students’ positive self-perception. It was interesting to note that the use of reciprocal peer tutoring was not linked to an increase in basic math calculation scores. Another study showed similar results in basic math calculation scores (no significant increase in the number of problems correct per minute), although there was an improvement in the average attendance of the children participating in reciprocal peer tutoring (82.3%) compared to the control group (64.2%).
who had regular classroom instruction (Fantuzzo, Polite, & Grayson, 1990). One study, using fifth graders, did show mathematic improvements when compared to a control group. Ninety-two percent of those students using reciprocal peer tutoring improved their scores from the pretest by 10% or better compared to 38% of the students who improved from the control group (Fantuzzo, King, & Heller, 1992).

Preschool children may also benefit from reciprocal peer tutoring. Brady (1997) reported that preschool children increased in basic academic skills (such as letter and number recognition, identification of shapes and colors, and writing some letters) following reciprocal peer tutoring. The subjects also increased the time spent interacting with one another when peer tutoring was not being implemented. The researcher indicated more research with preschoolers needs to be conducted since the study did not have a control group.

**Reverse-Role Tutoring**

Reverse-role tutoring is similar to cross-age peer tutoring except the tutor is an older child with mild disabilities and the younger tutee may have a disability. One study reported the tutors involved in such a program interacted more than before the intervention with regular classroom peers and showed gains on standardized academic tests (Eiserman, Shisler, & Osguthorpe, 1987). In another study Top and Osguthorpe (1987) reported that tutors and their first grade tutees experienced improvements in reading achievement as measured by the Woodcock-Johnson Psycho-Educational Battery. The researchers also noted improvements in the tutors’ self-esteem as measured by the Piers-Harris Children’s Self-Concept Scale and the Student’s Perception of Ability Scale.
Classwide Student Tutoring Teams

The Classwide Student Tutoring Teams (CSTT) program was designed to be used at the secondary school level (i.e., seventh grade through twelfth grade) in specific content areas. The goal of CSTT is to help students master skills or concepts that have already been introduced. In CSTT, students are placed in teams of three consisting of a high, average, and low performing student. Each week the teams are given a study guide and cards with numbers that correspond to the study guide. Each student takes a turn being the teacher. The teacher draws a card and reads the question from the study guide. The other two members write the answer down and the teacher checks their answers. The teacher then gives points for each correct response and gives the correct answer for each incorrect response. Another student gets to be the teacher next and repeats the process. In a study done with CSTT, researchers found that CSTT was more effective than teacher-led instruction (Maheady, Sacca, & Harper, 1988). Another study found CSTT was more effective when used after the teacher has previously presented the instructional material (Harper, Mallette, Maheady, & Brennan, 1993).

Collaborative Strategic Reading (CSR)

Collaborative Strategic Reading (CSR) is a classwide peer tutoring reading program. The purpose of CSR is to teach four reading comprehension strategies. First, the preview strategy, which requires retrieving previous knowledge and making predictions is taught. Next, the click and clunk strategy, which works to expand the student’s vocabulary by monitoring that student’s reading is covered. Get the gist strategy which helps students locate the main idea during reading, is taught next. The final strategy, wrap-
up, requires the student to identify the main ideas of the reading passage (Vaughn, Klingner, & Bryant, 2001). CSR has been used with groups of students as well as pairs. In CSR, the four strategies are first taught to the entire class by the classroom teacher. The teacher then selects several students to model each strategy. Once the students have become proficient with the strategies during the teacher-led activities, the students are divided into groups or pairs. Once in their groups, the students are assigned roles (e.g. leader, reporter, clunk expert). CSR is relatively new program, but early findings indicate students participating in this program made greater improvements in understanding their texts compared to a control group (Vaughn, Klingner, & Bryant, 2001). Another study examined the effects of CSR compared to reading with a partner (control group). No differences in either groups’ vocabulary scores were found, but students in the CSR group did improve their reading fluency compared to students in the control group. No gains were made in the students’ reading comprehension in either group (Vaughn, Chard, Bryant, Coleman, Tyler, Linan-Thompson, & Couzakanani, 2000).

**Numbered Heads Together (NHT)**

Numbered Heads Together (NHT, Kagan, 1992) was designed as an alternate question asking strategy. Students are placed in groups of four, assigned a number from 1 to 4, and sit together during the lessons. Students are told to “put their heads together” to come up with the best answer to the teacher’s question. Students are randomly selected to give a response and/or state whether they agree with the previous group’s response. One study, using NHT, reported students scored higher on daily quizzes after participating in this program (Maheady, Harper, & Mallette, 2001). Another study examined the effects of
NHT compared to two other methods of instruction in a science class, response cards (i.e., the teacher asks a question and the students write their answer on a card and hold it up for the teacher to see) and whole group question and answer (i.e., the teacher asks a question and calls on a student who volunteers to answer). The researchers found that more students would answer questions when using NHT compared to the other methods. Student performance increased more using NHT (82.1% questions correct on science quizzes) and response cards (81.38% questions correct on science quizzes) than the whole group question and answer method (73.24% questions correct on science quizzes, Maheady, Michielli-Pendl, Mallette, & Harper, 2002).
MAIN TUTORING PROGRAMS

As can be seen, there are many different types of peer tutoring programs. Now the tutoring programs of focus, in this review cross-age peer tutoring, Classwide Peer Tutoring (CWPT), and Peer-Assisted Learning Strategies (PALS) will be discussed.

Cross-age Peer Tutoring

Cross-age peer tutoring involves pairs of students in which the tutor is approximately two or more years older than the tutee. It is assumed that the tutors receive benefits from tutoring as well as effectively teach skills the teacher would normally have to teach one-on-one with a student (Utley & Mortweet, 1997). The tutors may be lower achieving students (Giesecke, Cartledge, & Gardner, 1993), older students from nearby high schools or even older students with disabilities (Barbeta, Miller, Peters, Heron, & Cochran, 1991). To help with the selection of tutors, the tutors may be given a test to assess their proficiency in the subject area. One advantage to using cross-age peer tutors rather than same-age tutors is the older peers are usually more competent in the concepts being taught (Beirne-Smith, 1991; Vacc & Cannon, 1991). Seven studies were reviewed to examine the effectiveness of cross-age peer tutoring.

In a study focusing on reading outcomes, 25 high school students tutored 25 children in grades 1-3 who were reading below their grade level. Tutoring took place over a 6-week period (Barbeta et al., 1991). The tutors were trained in peer tutoring procedures in two 45-minute sessions. Of the 25 pairs, only 6 were selected as target students for the study. The target students were selected by their teacher because the students had more severe difficulties with word recognition. The number of vocabulary
words mastered (i.e., reading the word correctly for two or three consecutive days) as well as the percentage of words maintained were examined over six weeks. A multiple probe baseline design across word sets was used to evaluate outcomes. The word sets were established from lists of words constructed by the classroom teachers. The word sets were a group of 10 words from the teachers' lists that the tutees incorrectly responded to before the tutoring was started. Two variations to standard multiple probe baseline designs were made. First, it was not possible to give a pretest of all the word sets before the study began because it was difficult to know how many word sets would be mastered throughout the study. Also, some of the words would be learned in the classroom when the peer tutoring was not happening. So, word sets were created during each week of the study and given intermittently as probes throughout the study. Due to the speed that some of tutees mastered the word sets, the second variation was to give only one pretest measure before tutoring rather than three consecutive pretest measures. The average number of words mastered each day increased 78% (2.1 pretest mean to 9.4 tutoring mean). The tutees were also given maintenance probes beginning one week after tutoring began until four months after tutoring ended. The tutees maintained an average of 97% of the vocabulary words even four months after the initial acquisition of the words. It was also noted that tutees were able to read the words in context (Barbetta et al., 1991).

A separate study with third and low achieving fourth graders showed reading improvements for the tutors as well as the tutees (Giesecke, Cartledge, & Gardner, 1993). There were four tutoring pairs from an inner-city elementary school located in a low socioeconomic area. The tutees were selected based on their poor reading performance
and low peer social status according to their teachers’ judgments. The tutees’ reading ability was measured by the number of vocabulary words (picked from a third and fourth grade reading series) each student could identify. A multiple baseline design across word sets was used for the tutees along with a pretest/post-test procedure to measure the gains of the tutors on sight words and self concept. At the pretest, the tutees had mastered only a few of the sight words (tutee 1 = 10 words, tutee 2 = 9, tutee 3 = 18, and tutee 4 = 5). Each tutee increased the number of words mastered from the pretest by at least 75%. The tutors’ sight word recognition increased by at least 41%. The researchers indicated that the tutors’ learning rate may have been lower than the tutees because the tutees’ responses set the pace for learning. The tutors also knew more of the words at the pretest (tutor 1 = 28 words, tutor 2 = 15, tutor 3 = 15, and tutor 4 = 37). The tutors also reported an increase in their self-concept as rated on the Piers-Harris Children’s Self-Concept Scale from the pre-assessment to the post-assessment (the tutees’ self-concept was not assessed, Giesecke et al., 1993).

Taylor and Hanson (1999) examined the effects of cross-age tutoring in addition to another reading intervention, Early Intervention in Reading, throughout a school year. Seven and eight year olds were paired with older (9-10 year olds) tutors. For the tutees, 12 students were selected for the intervention plus tutoring group, 7 for the intervention only, and 12 for the control group. The tutors met the following criteria: identified by teachers as being behind in reading, not having scheduling conflicts, and able to read a third grade basal reader with at least 85% accuracy. The tutors were trained in a class that met for 45 minutes each day over 14 weeks. Mondays and Tuesdays were used to train the
tutors by modeling and practicing tutoring procedures. On Wednesdays and Thursdays the cross-age peer tutoring was implemented. Fridays were for debriefings for the tutors. Each tutor kept a journal of his/her tutoring experience for the week. Problems that occurred were discussed along with solutions for those problems. After the 14 weeks of training and tutoring, the tutors only tutored once a week and no longer received training. By the end of the school year, 75% (9 out of 12) of the tutees in the reading group plus peer tutoring group could read a book at grade level with 90% accuracy. Less than 30% (2 out of 7) of the tutees in the reading program only were able to read the same book with 90% accuracy. None of the tutees in the control group could read the same book with 90% accuracy. All of the groups were given the Metropolitan Achievement Test during the fall of second grade and the fall of third grade. The scores of the tutoring group improved from the fall of second grade (12th percentile) to the fall of third grade (19th percentile). The scores for the students in the reading program only condition decreased from the 12th percentile to the 11th percentile. The control group’s scores also decreased from the 9th percentile to the 8th percentile. In addition, all of the tutors, who were reading below grade level at the beginning of the study, were able to read at grade level by the end of the year (Taylor & Hanson, 1999).

Another study examined the reading fluency and comprehension of third graders. Six children were randomly selected to participate, three were placed in the tutoring condition and the other three were in the control condition. All of the children were approximately 18 months below their reading level. Three fifth graders were randomly selected to be the tutors with three others selected to be in the control condition. Tutors
were trained through modeling and practice in two 15-minute sessions. A three-phase multiple baseline was used. The first phase was used to obtain a baseline on classroom reading for both the tutees and the tutors. The baseline phase ran for three weeks for one dyad, five weeks for another, and six weeks for the last pair. In phase 2, the tutee was instructed to read aloud for ten minutes and ask the tutor for help if needed. This phase was used to ascertain whether any tutoring behaviors would occur spontaneously. Phase 2 lasted for nine sessions for all of the pairs. Phase 3 began with the training for the tutors and the tutoring of the tutees three times a week. This phase lasted for six, eight, and ten weeks respectively for each pair. By the end of phase 3, each tutee answered at least 70% of the comprehension questions correctly compared to less than 50% at the baseline. The researchers indicated that the students in the control condition did not increase their percentage correct on comprehension questions, but no specific data were given.

Comprehension questions were taken from Cloze exercises on classroom reading material. Cloze exercises are when a student is given a passage with words deleted, then he/she fills in the missing words. On the Neale Analysis of Reading Ability, the tutees gained an average of 11 months on accuracy and 23 months on comprehension from the pretest to the post-test compared to the control group gains of 6.3 months on accuracy and 4.6 months on comprehension. The tutors showed significant improvements as well with gains of 19 months on accuracy and 25 months on comprehension compared to the control group’s gains of 3.6 months on accuracy and 6.0 months on comprehension. The tutors also gained an average of 8.0 months on the post-test in reading accuracy and
comprehension on the Metropolitan Reading Achievement Test compared to 6.0 months for the tutors in the control group (Limbrick, McNaughton, & Glynn, 1985).

Cross-age tutoring has also been shown to be effective for students struggling with basic math skills. One study (Beirne-Smith, 1991) examined the effects of tutoring on the acquisition of basic addition facts (e.g., $4 + 6$, $2 + 5$, $3 + 1$). The subjects were 40 students, 20 were tutors and 20 were tutees, from 4 different schools (2 urban and 2 rural). Half of the students were randomly selected for the tutoring condition and the other half were selected for the control condition. The study was conducted over 6 weeks. The tutoring occurred over 4 weeks with one week before and after for the administration of pretests and post-tests. The tutees' ages ranged from six to ten years old and they were functioning at least one grade level below average in basic math skills. The average number correct on a pretest of 60 addition facts for the tutees was 23.8 (40%), while the average number correct for the control group was 22.9 (38%). On the post-test, the tutees' average score was 40.8 (68%) correct compared to the control group's score of 25.9 (43%) correct. The tutors showed a little improvement from the pretest (56.8 correct or 94.7%) to the post-test (59 correct or 98.3%) as well (Beirne-Smith, 1991).

Another study used sixth graders to tutor four special education students who were six, eight, nine, and twelve years old. The length of the intervention was 6 weeks with a follow-up 2 years later. Each tutee was tutored in rote counting; matching number words and numerals; counting objects, identifying number words; and identifying month, day, and days of the week. Baselines were not given for each tutee, but the number of correct responses increased over each session for all of the tutees. On rote counting and
counting objects, three of the subjects improved their correct responses from three correct at the first session to seven correct by the last session. The fourth subject improved from eight to ten correct responses. On matching number words and numerals three of the subjects again improved from zero correct responses to three correct responses. Subject 4 improved from three correct responses to five correct responses on matching number words and numerals. When identifying number words, subjects 1, 2, and 3 improved from zero to two correct responses. Subject 4 improved from two to four correct responses. On identifying the month, day, and days of the week, subjects 1 and 2 improved from zero to three correct responses. Subject 3 improved from zero to five correct responses and subject 4 improved from one to six correct responses. A follow-up of 5 sessions was conducted 2 years later and unfortunately, in most of the areas there was little or no change from the average number of correct responses of the last 5 tutoring sessions (Vacc & Cannon, 1991).

Tutors have also benefitted academically from tutoring younger students. Two middle schools were used in a recent study that examined tutor outcomes. Both schools had a reading class which helped struggling readers improve their reading skills. One class of 22 students was assigned to tutor first and second graders at a nearby elementary school. Another class of 23 students was used as the control group. The researcher did not indicate whether the students in each class were at similar levels in their reading abilities prior to the tutoring. This study extended over an entire school year. During the first 9 weeks of class, the teacher selected the same book for all of the tutors to use. By the tenth week, the tutors were encouraged to select their own book to use. The teacher developed
and modeled scripts for the tutors to use. Each Monday, the scripts were modeled and the
tutors practiced with each other. On Tuesdays and Thursdays, the tutoring took place. On
Wednesdays and Fridays, the tutors wrote about their experiences in their journals.
Observations were done on the tutors throughout the school year. Classroom observations
showed that the tutors were reading difficult texts with increasing speed. It was observed
that books the students had difficulties with at the beginning of the year were read
smoothly and accurately later on in the year. However, no numerical data were presented
to support this outcome. Tutors outperformed the control group on a standardized reading
assessment (Gates-MacGinitie Reading Tests), which was given at the beginning and end
of the school year, as well as the Stanford Achievement Test. The author did not report
specific data or whether the results were statistically significant. No assessment was done
on the tutees in this study (Fisher, 2001).

Summary. The seven studies examined suggest that cross-age peer tutoring is effective in
improving academic skills. Tutors and tutees both benefit from cross-age peer tutoring. In
reading, tutees improved their reading fluency and comprehension (Limbrick et al., 1985)
as well as their overall reading level (Taylor & Hanson, 1999). Tutees also increased the
number of vocabulary words mastered as well as maintained 97% of those words after
four months (Barbetta et al., 1991). In math, tutees improved their test scores on basic
math facts (Beirne-Smith, 1991) and increased the number of correct responses when
counting (Vacc & Cannon, 1991). The research indicated that the tutors improved their
fluency, comprehension (Limbrick et al., 1985), and reading level as well (Taylor &
Hanson, 1999). Tutors were also observed reading more difficult texts (Fisher, 2001) and
improved their sight word recognition (Giesecke et al., 1993). Most of the studies examined improvements in basic academic skills, such as learning sight words and basic math facts. Only two studies examined tutoring for math skills, so more research is needed examining the effects of peer tutoring on math.

Classwide Peer Tutoring (CWPT)

Classwide Peer Tutoring (CWPT) was developed at the Juniper Gardens Children Project through the work of Delquadri, Greenwood, Whorton, Carta, and Hall (1986). In CWPT, students match up with their partner each day for a tutoring session. The tutor follows along while the tutee reads for ten minutes. During this time period, the tutor gives points and corrects errors made by the tutee. In reading, for example, word substitutions, omissions, and hesitations are counted as errors. The tutee receives two points for correctly responding to the tutor’s prompt (e.g., correctly spelling or reading a word). One point can still be earned if a correction is made when the tutor points out the error. For example, if a word is spelled incorrectly, the tutor spells it correctly, then the tutee must write the word correctly three times to receive one point. Throughout the tutoring session, the teacher walks around and gives bonus points to the tutors who are using appropriate tutoring behaviors (learned through training) and the tutees for quick responses and cooperating with the tutor. CWPT is not a program used during the entire school day, but for about 30 minute blocks of time. The time allows for each student to be the tutee and the tutor for 10 minutes each. The remaining 10 minutes is used to total and post the points for each team. It is also important to change teams, not the pairs, each week, so each student will eventually be on the winning team.
It is important for all students to be trained before tutoring begins. The teacher first explains how the activity will work. The teacher then acts as the tutor and selects one student to be the tutee. The teacher shows how to give the points as well as how to correct errors made by the tutee. Next, two students are selected to be the tutor and the tutee. The teacher gives feedback throughout the demonstration. A few other students are selected to demonstrate before the entire class attempts it. The teacher circulates the room giving corrective feedback. After the whole class has attempted CWPT, students are usually capable of following the correct procedures. Younger students (i.e. first and second graders) may need a few extra sessions (Delquadri et al., 1986).

CWPT is based on the following concepts: increasing the number of chances a student has to respond to teacher and/or classroom prompts, using important concepts the teacher selects to determine a student's progress, and implementing behavioral principles to facilitate responding (Delquadri et al., 1986). In a regular classroom of 25 students, the opportunity for students to answer questions, read aloud, or ask questions is limited. CWPT increases the amount of interactions between a teacher and student. Each student (tutee) has his/her own teacher (tutor) which allows the student numerous opportunities to engage in academic interactions. In one study, students' opportunities to respond during a 60-minute reading class increased from 28% to 78% when using CWPT (Delquadri et al., 1986).

The second concept, functionality of key academic skill areas, is used when teachers select skills needed to show a student's progress. For example, to assess a student's reading rate, a teacher may select a passage from the school district's reading
materials for the student to read. CWPT would be used to help students read words together in sentences correctly rather than helping students learn phonetic rules and decoding skills. The student’s progress is measured by how many words he or she can read in a set amount of time. CWPT is not used with a specific curriculum, but can be adapted to the curriculum already used in various school districts.

CWPT is also based on various behavior analysis principles, such as reinforcement. The CWPT program entails having tutors reward tutees with points for correct responses. Each dyad is placed on a team and the dyad’s points are totaled for team points. Each week the team points are totaled with the winning team receiving a reward. Having each pair be part of a team, adds a powerful reinforcer—the influence of peers, or in other words, peer mediated contingencies. The other part of the behavioral principles consists of posting points as well as the teacher verbally acknowledging to the group the points received, due to individual and team efforts.

Nine studies were reviewed to examine the effectiveness of CWPT. Many of these studies (n = 8) were conducted by Greenwood and colleagues from the Juniper Gardens Children’s Project.

Greenwood and colleagues from the Juniper Gardens Children’s Project conducted a longitudinal study of CWPT that spanned grades 1 to 6 (Greenwood, Terry, Utley, Montagna, & Walker, 1993; Greenwood, 1991; Greenwood, Delquadri, & Hall, 1989). The original pool of schools was selected from 35 different schools in a Kansas City school district. Ten of the schools were eliminated after the district relocated entire grades to achieve racial balance between schools. From the pool, the final sample was randomly
selected to include six Chapter 1 schools and three non-Chapter 1 schools. Chapter 1 schools consist of a high population of students from low-income families who qualify for federal resources. The resources are used for materials and teachers’ salaries to provide remedial educational programs for disadvantaged and at-risk students in these schools. Non-Chapter 1 schools do not receive these resources. The schools were placed into three different groups. Students from the three non-Chapter 1 schools were the comparison group and received regular teacher-led instruction. The other six schools were randomly assigned either to the control group or the treatment (CWPT) group. A total of 416 students participated in the study. All students were given the Metropolitan Achievement Test (MAT) in the first grade (pretest) and in the fourth grade (post-test). The MAT has three subtests that were used: reading, math, and language. Normal curve equivalent mean scores were used. The treatment group showed a decrease on the reading subtest from the pretest to the post-test (See Table 1), but increases on math and

<table>
<thead>
<tr>
<th>Subtests</th>
<th>Treatment</th>
<th>Control</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest Mean (SD)</td>
<td>Post-test Adjusted</td>
<td>Pretest Mean (SD)</td>
</tr>
<tr>
<td>Reading</td>
<td>50.1 (12.9)</td>
<td>44.0</td>
<td>44.0 (14.7)</td>
</tr>
<tr>
<td>Math</td>
<td>42.2 (12.5)</td>
<td>49.5</td>
<td>42.1 (15.1)</td>
</tr>
<tr>
<td>Language</td>
<td>39.2 (17.2)</td>
<td>53.9</td>
<td>39.4 (16.8)</td>
</tr>
</tbody>
</table>

Pretest = Grade 1
Post-test = Grade 4

Note. Table is adapted from Greenwood et al., 1989, p. 380.
language from the pretest to the post-test. The control group also decreased on the reading subtest from the pretest to the post-test. On the math subtest, the control group showed little change from pretest to post-test, and in language, the control group improved from the pretest to the post-test. The comparison group also showed a decrease on the reading subtest from the pretest to the post-test. The comparison group stayed relatively the same on the math subtest from pretest to post-test and improved on the language subtest from the pretest to the post-test. At post-test, significant group differences were noted using Analyses of Covariance (with pretest performance as a covariate). Significant differences in all academic areas were noted between the control and treatment groups (with the treatment group demonstrating better skills) as well as the control and comparison groups (with the comparison group performing better). There were not significant differences between the treatment group and the comparison group at the post-test. The mean effect sizes between the treatment and control groups using adjusted post-test scores were 0.57 in reading, 0.37 in math, and 0.60 in language. For the comparison and control groups, the mean effect sizes were 0.93 in reading, 0.59 in math, and 0.90 in language. (Greenwood et al., 1989).

The next study (Greenwood, 1991) examined the trends that were appearing from the 1989 data, in regards to academic engagement. All the groups improved in the percent of academic engagement time. The treatment group increased from 32.2% at the first assessment in the first grade to 49.6% at the last assessment in the fourth grade. In the control group engaged time increased from 35.9% to 53.0%, and in the comparison group, it increased from 36.6% to 54.9%. The final study (Greenwood et al., 1993) was
conducted when the subjects were in the sixth grade. The Comprehensive Test of Basic Skills (CTBS-U) rather than the MAT was used as an assessment. The only explanation for the change of assessments was that the district used the CTBS-U as part of the yearly evaluation process. The CTBS-U's subtests included the same subtests as the MAT as well as additional subtests in social studies and science. Normal curve equivalent scores were again used. On all three subtests (reading, math, and language) the comparison group's average score was higher than the treatment and the control groups' average scores. The treatment group's scores were higher than the control group's scores on all three subtests, as well (See Table 2). The differences between all of the groups were statistically significant. The group differences on the science and social studies subtests were also statistically significant. On the science and social studies subtests, the comparison group again scored higher than the treatment and the control group. In

<table>
<thead>
<tr>
<th>Subtests</th>
<th>Treatment Mean (SD)</th>
<th>Control Mean (SD)</th>
<th>Comparison Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>46.17 (15.78)</td>
<td>40.77 (14.99)</td>
<td>53.70 (19.47)</td>
</tr>
<tr>
<td>Math</td>
<td>54.61 (17.84)</td>
<td>46.08 (17.35)</td>
<td>58.10 (23.99)</td>
</tr>
<tr>
<td>Language</td>
<td>52.02 (17.61)</td>
<td>47.02 (17.34)</td>
<td>57.83 (19.33)</td>
</tr>
<tr>
<td>Social Studies</td>
<td>48.09 (17.95)</td>
<td>41.94 (17.39)</td>
<td>51.52 (21.38)</td>
</tr>
<tr>
<td>Science</td>
<td>45.23 (15.33)</td>
<td>38.55 (15.25)</td>
<td>52.03 (21.08)</td>
</tr>
</tbody>
</table>

*Note: Table is adapted from Greenwood et al., 1993, p. 506.*
addition, the treatment group outperformed the control group. The mean effect sizes between the treatment and control groups using adjusted post-test scores were 0.39 in reading, 0.57 in math, and 0.35 in language. For the comparison and control groups, the mean effect sizes were 0.94 in reading, 0.80 in math, and 0.75 in language. Effect sizes were not reported for the science and social studies subtests. In addition to the academic gains seen in the CWPT group, fewer students in the CWPT group (25.7%) were placed in more restrictive special education classes compared to those in the control group (55.6%) at the final follow-up. The treatment group also had more special education students (34%) remaining in the regular classroom compared to the control group (7%) and the comparison group (32%; Greenwood et al., 1993).

A number of studies have also examined the effects of CWPT on spelling (Greenwood, Terry, Arreaga-Mayer, & Finney, 1992; Madrid, Terry, Greenwood, Whaley, & Webber, 1998; Mortweet et al., 1999; Sideridis, Utley, Greenwood, Delquadri, Dawson, Palmer, & Reddy, 1997; Greenwood, Arreaga-Mayer, Utley, Gavin, & Terry, 2001). One study (Greenwood et al., 1992) examined pretest and post-test spelling scores of 88 students from the second, fourth, and fifth grades. CWPT was implemented over 19 weeks. Each week the students were placed into different groups based on their pretest and post-test scores. The success group included students who scored less than 40% on the pretest and scored more than 80% on the post-test. The challenged/undermastery group contained students who scored less than 40% on the pretest and less than 80% on the post-test. The underchallenged/undermastery group scored more than 40% correct on the pretest and less than 80% on the post-test. The students could qualify for different
groups or stay in the same group each week based on their pretest and post-test scores. From the sample of 88 students, the success group, improved by an average of 66.2% on spelling accuracy on the post-test. The challenged/undermastery group showed an average gain of 38.2% from the pretest to the post-test, the underchallenged/mastery group improved by 28.2% from the pretest to the post-test, and the underchallenged/undermastery group improved by 0.9% from the pretest to the post-test. Overall the students showed an average gain from the pretest to the post-test of 39.7%. A control group was not used in the study, so it is unknown whether the amount of growth from the pretest to the post-test would have been different if CWPT had not been used.

In another study examining the effects of CWPT on spelling scores (Madrid et al., 1998), the participants were 16 first graders from an inner city Title I school. The students selected were identified as being at risk for school failure based on a history of poor spelling test scores and below-average scores on the Metropolitan Achievement Test (MAT). The participants were randomly assigned to one of three conditions, which were randomly alternated each week. The three conditions were CWPT, passive peer tutoring which entailed having the tutee listen and observe while the tutor spelled the word aloud and wrote each spelling word, and teacher-mediated instruction which required the tutors and tutees to listen and observe the teacher leading a spelling discussion followed by an assignment in a workbook. The study was conducted over 12 weeks with each condition being implemented for one week before using a different condition. During the CWPT condition, the percentage gained on words correct from the pretest to the post-test ranged from 12% to 53% with an average gain of 34%. When using passive peer tutoring,
students' gains from the pretest to the post-test ranged from 29% to 47% with an overall mean of 40%. The teacher-mediated instruction led to gains of 20% to 38% from the pretest to the post-test. The overall mean gain for the teacher-mediated instruction condition was 28%. All the conditions improved the students' spelling scores, but the passive peer tutoring condition resulted in the greatest improvements.

Another study examined the effect of CWPT on the spelling as well as the sight word acquisition of English language learners (ELL; Greenwood et al., 2001). Five elementary teachers participated in the study for 21 weeks. A total of 117 ELL students from the first to the fifth grades participated. There was one class from each grade. The first grade students used CWPT on sight word acquisition, while in the other grades CWPT was implemented during spelling instruction. No control or comparison group was used in this study. Pretests and post-tests were given weekly on spelling words. Post-tests were given on Fridays and consisted of the words learned during CWPT as well as the pretest for the next set of words. After CWPT, the first grade students improved from 17.1% words correctly identified on the pretest to 71.7% correct on the post-test of sight words. The second graders improved their spelling scores from 15.5% correct on the pretest to 81.4% correct on the post-test; third graders improved from 21.6% correct on the pretest to 87.0% correct on the post-test; fourth graders improved from 15.5% correct on the pretest to 73.3% correct on the post-test; and the fifth graders improved from 24.6% correct on the pretest to 77.3% correct on the post-test. The overall gain across all of the classes from the pretest to the post-test was 59.8% (Greenwood et al., 2001).
Along with improvements in spelling, two studies showed improvements in academic engagement following CWPT (Mortweet et al., 1999; Sideridis et al., 1997). Academic engagement was calculated, through observation, as a compilation of student behaviors such as writing, reading aloud and silently, task participation, and academic talk. In Mortweet et al. (1999) two classrooms of 27 students were used in the study. Eight students were selected as the target students, four typical peers and four with mild mental retardation (MMR). The typical students were selected for the study by the teachers based on previous test scores in spelling. The typical peers consisted of high and low achievers in spelling. Data were not given to differentiate between the high and low achievers. The students with MMR were chosen based on their Individualized Education Programs' (IEPs) goals which specified an inclusive classroom placement. A withdrawal treatment design was used to compare the effectiveness of CWPT and teacher-led instruction on spelling test performance over eleven weeks. All of the students demonstrated greater gains from the pretest to the post-test when CWPT was used instead of teacher-led instruction. The students with MMR had average pretest-post-test gains (i.e. post-test minus pretest scores) of 58%, 35%, 27%, and 42% during teacher-led instruction. The average gain scores were 60%, 52%, 73%, and 72%, respectively, during CWPT. The typical peers also had higher gain scores during CWPT (69%, 68%, 57%, and 48%, respectively) than during teacher-led instruction (27%, 66%, 48%, and 43%). Students using CWPT were academically engaged 56% of the time, whereas, those not using CWPT were only academically engaged 44% of the time (Mortweet et al., 1999).
In a study by Sideridis et al. (1997), six students from a sixth grade general education classroom were selected to participate. One student each was identified as low-, average-, and high-achieving in spelling, based on their teacher's judgment. The other three students had disabilities, two were diagnosed as learning disabled (LD) and one with educable mental retardation (EMR) along with attention deficit disorder (ADD). An ABAB design was used with the baseline condition being conventional teaching procedures and CWPT used as the treatment. Mean spelling gains were used, which were found by taking the percentage correct on the pretest and subtracting it from the percentage correct on the post-test and then averaging the differences for each condition. The mean spelling gains between the pretest and the post-test for the first baseline (A1) was 66% and at the last treatment (B2) was 83% for the low-achieving student, 38% at the first baseline (A1) and 73% at the last treatment (B2) for the average-achieving student, and 29% at the first baseline (A1) and 37% at the last treatment (B2) for the high achieving student. The researchers indicated the reason the high-achieving student improved less was because he had a smaller range to improve. The students with disabilities also improved following CWPT. The two students with LD improved from 26% and 28% at the first baseline (A1) to 68% and 77% at the last treatment (B2), while the third student with disabilities improved from 37% at the first baseline (A1) to 73% at the last treatment (B2). Three of the students increased their academic engagement, which is a compilation of student behaviors such as writing, reading aloud and silently, task participation, and academic talk, during CWPT. The low-achieving student improved the level of academic responding from 13% during baseline to 63.25% during CWPT, one
student with LD improved from 8% during baseline to 65% during CWPT, and the
student with EMR/ADD improved from 24% during baseline to 70% during CWPT. The
other three students had similar percentages of engaged behavior during baseline and
CWPT (Sideridis et al., 1997).

Another study examined the reading fluency and reading comprehension of three
students with autism as well as 14 of their classroom peers (Kamps, Barbetta, Leonard, &
Delquadri, 1994). The study was conducted over 20 weeks. A multiple baseline design
across subjects was used to measure the students’ performances. Each day after the
tutoring session was completed, each student would read that session’s reading passage to
the experimenter for a 2-minute timed assessment. The experimenter would record errors,
but would not give corrective feedback. The experimenter would then ask five
comprehensive questions (who, what where, when, and why) developed by the
experimenter. The researcher also examined, through classroom observations, the
frequency and duration of social interactions between peers using the Social Interaction
Code developed by Niemeyer and McEvoy. Every student, except one, improved when
CWPT was being implemented. The reading rates increased by an average of twenty
words read correctly per minute during CWPT. The mean percentage correct on
comprehension questions also improved from 70.8% at baseline to 94% during CWPT.
The student who did not improve was the most fluent reader in the classroom. All students
increased their social interaction time (i.e., the total amount of time the student interacted)
with an average gain of 58.5 seconds from the first baseline to the last treatment. The
length of the social interactions (i.e., how long each single interaction lasted) had an
average increase gain of 57.7 seconds from the first baseline to the second treatment (Kamps et al., 1994).

Summary. CWPT is a classwide intervention with positive effects on students' academic skills. The nine studies examined indicate that CWPT is effective in improving students' reading (including spelling). Five of the studies demonstrated that students improved their spelling scores from pretests to post-tests. Students also improved their reading fluency and comprehension more when using CWPT than during teacher-led instruction (Kamps et al., 1994). Two studies indicated that CWPT increases the percent of time a student is engaged academically (Mortweet et al., 1999; Sideridis et al., 1997). The effects of CWPT seem to be long term as well, according to one study, but additional studies need to be done to further solidify the long term outcomes (Greenwood et al., 1993). Lower achieving students seemed to benefit the most from CWPT whereas higher achieving students benefitted the least. CWPT also appears to benefit students with limited English proficiency in spelling and sight word acquisition (Greenwood et al., 2001). Overall, no negative effects, only positive, were discovered for CWPT, which indicates it is an effective intervention to help improve academic skills. Although there were many positive outcomes, some of the studies did not use a comparison or control group so it is difficult to know if the improvements were due to CWPT or to another variable. It should also be noted that Charles Greenwood was one of the researchers on eight of the nine studies examined. Given this, it would be important for independent evaluations of this program to be conducted.
Peer-Assisted Learning Strategies (PALS)

The Peer-Assisted Learning Strategies (PALS) program was developed in the 1990s as a classwide intervention in the regular education classroom to help students with learning disabilities (Fuchs, Fuchs, Phillips, Hamlett, & Karns, 1995). Before beginning the PALS program, the teacher conducts training sessions for the class. There are 6 to 10 scripted lessons to teach the principles of PALS, which entail teacher presentations and demonstrations. The students practice implementing the principles while receiving teacher feedback. Each training session lasts approximately 45 minutes.

After the training sessions are complete, each student is paired with another student. The pairs each have a higher and a lower performing student, which the teacher has determined from each student’s competence in reading. In each PALS activity, the higher performing student reads first to serve as a model. The reading level of the material is based on the reading level of the lower performing student. The pairs are also assigned to one of two teams similar to CWPT. As discussed in the following paragraphs, points are given throughout the tutoring sessions by either the tutor or the teacher. The team with the most points is rewarded each week. The teams and dyads are changed every four weeks (Fuchs et al., 2001).

The PALS reading program involves three different activities: Partner Reading, Paragraph Shrinking, and Prediction Relay. Partner Reading, the first activity, involves each student reading aloud for five minutes with the higher performing student reading first. After the first student reads, the lower performing student rereads the same section. While the second student (tutee) reads, the tutor immediately corrects mistakes made by
the tutee. After the correction, the tutee rereads the sentence from the beginning of the sentence with the correction. The lower performing student then retells the section just read. Pairs receive one point for each sentence read correctly. Ten points are given for retelling the text read (Fuchs et al., 2001).

The purpose of Paragraph Shrinking, the second activity, is to bring about comprehension through summarizing and identifying the main idea. The dyad continues to read out loud, but stops at the end of each paragraph to identify the main idea. To help the reader identify the main idea, the tutor asks the tutee to, first, identify the subject of the paragraph, and, second, identify the most important thing about the subject. This summary of each paragraph should not be more than ten words. If the summary is not correct, the tutee is told to skim the paragraph and try again. The tutor decides if the summary is correct, no answer key is given to the tutors. Three points can be earned for the summary: one point for the identification of the subject, one point for the most important thing about the subject, and one point for using ten or fewer words. Corrections are still made while the tutee is reading, but points are not given for correctly reading sentences during Paragraph Shrinking (Fuchs et al., 2001).

Prediction Relay, which requires the pairs to make and confirm or disconfirm predictions, has four steps. First, the reader predicts what information will come in the next section. Next, the tutor monitors (identifies and corrects reading mistakes) while the tutee reads the next section. After reading the next section, the reader then states whether or not the prediction was correct. The final step requires the reader to summarize the main idea of the section, again in ten or fewer words. Students can earn one point for correctly
completing each of the four steps (Fuchs et al., 2001). Five studies were reviewed to examine the effectiveness of the PALS program.

PALS has been investigated as a method to help improve reading skills. In a study by Mathes, Howard, Allen, and Fuchs (1998), PALS was added to an already existing reading program. The researchers examined the effects of PALS on low-, average-, and high-achieving first grade students. The subjects came from the classrooms of 20 different teachers in six different schools. Ten of the classrooms were selected to use PALS and 10 served as the control classrooms. Five participants were selected from each classroom based on their performance on a Curriculum-Based Measurement (CBM) probe on phonological segmentation and an oral reading fluency probe at the first grade reading level. Three of the five selected were low-achieving students, one was average-achieving, and one was high-achieving. The intervention was implemented three times per week for 35-minute sessions over a 16-week time period. Students in each of the groups increased their scores on the Comprehensive Reading Assessment Battery-Revised (CRAB-R) and the Woodcock Reading Mastery Tests-Revised (WRMT-R) from the pretest to the post-test (See Table 3). On the CRAB-R, the low- and average-achieving students in the PALS group showed greater average raw score gains on words correct (low-achieving ES = 0.30, average-achieving ES 0.27) and questions correct (low-achieving ES = 0.20, average-achieving ES 0.18) than the control group. However, the high-achieving students in the control group had greater gains on words correct and questions correct than the high-achieving students in the PALS group on the CRAB-R (ES = -0.25 words correct, -0.08 questions correct). On the WRMT-R, the low-, average-, and high-achieving
Table 3.
Summary of Results for the Comprehensive Reading Assessment Battery-Revised (CRAB-R) and the Woodcock Reading Mastery Tests-Revised (WRMT-R)

**CRAB-R**

<table>
<thead>
<tr>
<th>Treatment (CWPT) Students</th>
<th>CRAB-R Subtests</th>
<th>Average Raw Score Gains (std dev)</th>
<th>Control Students</th>
<th>Average Raw Score Gains (std dev)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-Achieving</td>
<td>Words Correct</td>
<td>62.71 (74.25)</td>
<td>Low-Achieving</td>
<td>45.12 (39.20)</td>
</tr>
<tr>
<td></td>
<td>Questions Correct</td>
<td>.79 (1.87)</td>
<td></td>
<td>.49 (1.11)</td>
</tr>
<tr>
<td>Average-Achieving</td>
<td>Words Correct</td>
<td>123.75 (83.75)</td>
<td>Average-Achieving</td>
<td>105.7 (41.40)</td>
</tr>
<tr>
<td></td>
<td>Questions Correct</td>
<td>2.55 (2.07)</td>
<td></td>
<td>2.25 (1.11)</td>
</tr>
<tr>
<td>High-Achieving</td>
<td>Words Correct</td>
<td>122.4 (58.17)</td>
<td>High-Achieving</td>
<td>137.95 (66.40)</td>
</tr>
<tr>
<td></td>
<td>Questions Correct</td>
<td>2.05 (2.19)</td>
<td></td>
<td>2.20 (1.45)</td>
</tr>
</tbody>
</table>

**WRMT-R**

<table>
<thead>
<tr>
<th>Treatment (CWPT) Students</th>
<th>WRMT-R Subtests</th>
<th>Standard Score Points Gained (std dev)</th>
<th>Control Students</th>
<th>Standard Score Points Gained (std dev)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-Achieving</td>
<td>Word Identification</td>
<td>11.29 (12.45)</td>
<td>Low-Achieving</td>
<td>2.06 (13.89)</td>
</tr>
<tr>
<td></td>
<td>Word Attack</td>
<td>14.61 (13.25)</td>
<td></td>
<td>5.32 (10.48)</td>
</tr>
<tr>
<td></td>
<td>Passage Comprehension</td>
<td>2.08 (12.83)</td>
<td></td>
<td>- 1.7 (15.06)</td>
</tr>
<tr>
<td>Average-Achieving</td>
<td>Word Identification</td>
<td>13.8 (16.38)</td>
<td>Average-Achieving</td>
<td>12.3 (16.73)</td>
</tr>
<tr>
<td></td>
<td>Word Attack</td>
<td>13.2 (9.70)</td>
<td></td>
<td>.10 (20.59)</td>
</tr>
<tr>
<td></td>
<td>Passage Comprehension</td>
<td>8.4 (4.30)</td>
<td></td>
<td>1.9 (6.67)</td>
</tr>
<tr>
<td>High-Achieving</td>
<td>Word Identification</td>
<td>9.83 (8.81)</td>
<td>High-Achieving</td>
<td>12.20 (8.34)</td>
</tr>
<tr>
<td></td>
<td>Word Attack</td>
<td>6.83 (12.47)</td>
<td></td>
<td>5.3 (10.88)</td>
</tr>
<tr>
<td></td>
<td>Passage Comprehension</td>
<td>2.22 (4.34)</td>
<td></td>
<td>- 1.10 (8.31)</td>
</tr>
</tbody>
</table>

**Note.** Table is adapted from Mathes et al., 1998, p. 76.
students in the PALS group had greater gains on the word identification, word attack, and passage comprehension subtests than the control group. On the word identification subtest the high-achieving students in the control group showed a bigger increase than the high-achieving students in the PALS group. The difference between the low-achieving students in the PALS group and the low-achieving students in the control group were statistically significant on the word attack (ES = 0.78) and word identification (ES = 0.70) subtests. There was a statistically significant difference between the average-achieving students in the PALS group and in the control group on the passage comprehension subtest (ES = 1.12). There were no statistically significant differences between the high-achieving PALS and control groups on any assessments (ES = -0.25, 0.13, 0.5 on WRMT-R). There were no statistically significant results for scores on the CRAB-R. The outcomes of this study suggests that PALS may not be as effective for high-achieving students as other traditional teaching methods (Mathes et al., 1998).

Another study also examined the effects of PALS on three learning levels (learning disabled, low performing, and average achievement) for students in reading (Fuchs, Fuchs, Mathes, & Simmons, 1997). This study examined students in grades second to sixth. Twenty-two schools were selected for the study. The schools were placed into groups based on the percentage of students on free or reduced lunch and the average reading scores on a state standardized test. High-level schools had a high average reading score and a low percentage of students on free or reduced lunch; low-level school were just the opposite with low reading scores and a high percentage on free or reduced lunch; middle-level schools fell in between the two extremes. After stratifying the schools into these
categories, they were randomly assigned to a PALS or No-PALS condition. Three
students were selected from each class: one who was classified as learning disabled (LD)
according to special education regulations, another student, called low performing (LP),
fell into the lowest quartile in reading in the class, and one considered average-achieving
(AA). Sixty students were assigned to each condition, PALS and no-PALS, twenty of
each student type (i.e. LD, LP, and AA). The PALS intervention was implemented for
fifteen weeks. All students using PALS had improved scores on the CRAB
(Comprehensive Reading Assessment Battery). On words correct, the students with
learning disabilities in PALS improved by 51.08 (sd = 29.6) words correct compared to
28.68 (sd = 28.91) words correct for those students with disabilities not using PALS.
Low-performing students in the PALS condition improved by 56.25 (sd = 36.82) words
correct compared to 40.35 (sd = 28.38) words correct for their no-PALS counterparts.
The average-achieving students in PALS showed a similar increase of 59.50 (sd = 47.32)
words correct compared to 37.38 (sd = 45.25) words correct for those average-achieving
students not using PALS. On questions correct, the students with learning disabilities in
PALS improved by 1.90 (sd = 1.24) questions correct compared to .43 (sd = 1.60)
questions correct for those students with disabilities not using PALS. Low-performing
students in the PALS condition also improved by 1.90 (sd = 1.88) questions correct
compared to 1.08 (sd = 1.35) questions correct for their no-PALS counterparts. The
average-achieving students in PALS showed an increase of 1.20 (sd = 1.77) questions
correct compared to 1.00 (sd = 1.42) questions correct for those average-achieving
students not using PALS. Although these changes were present, there were no statistically
significant differences in overall changes across student type between groups for words correct or questions correct. Effect sizes were small for words correct (ES = 0.22) but moderate for questions correct (ES = 0.55; Fuchs et al., 1997).

Although most of the focus of PALS has been on reading, Fuchs, Fuchs, Karns, and Phillips (1995) have developed Math PALS to help students improve their math skills. Each student is paired with another student based on curriculum-based measurements (CBM). Twice a month, the teacher evaluates the dyads and changes the pairs based on CBM scores. The tutoring assignments are changed to allow every student the chance to be the tutor. During the tutoring, the tutor has a question sheet which guides the tutee through each problem. The tutor demonstrates how to complete each problem by answering the questions on the question sheet. The tutors correct each digit that the tutee writes. For example, if the problem is adding two digit numbers, such as 52 + 31, then the tutor would give feedback after the tutee adds 2 + 1. If an error is made, then the tutor provides an explanation for the error. The question sheet is gradually faded out as the tutor becomes more familiar with the procedure. Throughout the tutoring, the teacher moves around the room awarding points for good tutoring interactions as well as assisting dyads who need help. After the tutee completes 12 problems on one concept, a 3-minute drill sheet containing the type of problem just practiced as well as easier problems from the curriculum is given. The drill sheet is corrected with one point awarded for each correct problem. The pair with the most points receives applause for their win (Fuchs et al., 1995).
Fuchs, Fuchs, and Karns (2001) examined the effect of PALS on math skills in twenty kindergarten classrooms. Half of the classrooms were assigned PALS and the other half were the control group. Within the classrooms, 168 students were selected to participate, 84 for each condition. The PALS intervention was in place for 15 weeks. The students examined were placed into three different groups: low-, average-, or high-achieving. A few students (eight in PALS and seven in no-PALS), who had been identified for special education, were placed in a separate group. Both average- and low-achieving students as well as the students with disabilities in the PALS condition showed improvement from the pretest scores to post-test scores on the Stanford Early School Achievement Test (SESAT). Average-achieving students in the PALS condition improved by 6.80 (sd = 4.52) compared to 4.76 (sd = 4.44) by the control group. The low-achieving students in the PALS condition improved by 8.62 (sd = 5.14) compared to 6.50 (sd = 4.31) by the control group. The students with disabilities in the PALS condition improved by 10.50 (sd = 2.33) compared to 6.86 (sd = 2.79) for the control group. However, none of the differences in change scores between the groups were statistically significant. Those high-achieving students in the no-PALS condition showed a slightly greater, but not statistically significant, improvement (gain of 3.53) on the SESAT compared to those in the PALS condition (gain of 3.00). The researchers suggest one reason PALS may not seem as effective for high-achieving students is due to the fact that high-achieving students have already mastered much of the curriculum (Fuchs, Fuchs, & Karns, 2001).

Another study showed that higher achieving first graders using PALS showed the most improvement on the Stanford Achievement Test compared to a control group in
math. Twenty classrooms participated over 16 weeks with 10 placed in the PALS condition and 10 in the No-PALS condition. All students who were present for the pretest and the post-test were sampled, which was 327 for the PALS condition and 184 for the control group. There were also 18 students with disabilities, 9 in each condition. PALS was implemented three times each week for 30 minutes over sixteen weeks. The high-achieving PALS students showed raw score gains on the SAT of 20.06 (sd = 7.08) whereas the control group gained 17.87 (sd = 7.04). The average-achieving PALS students improved their raw scores by 19.77 (sd = 7.25) compared to 17.14 (sd = 8.48) for the control group. The low-achieving PALS students showed raw score gains of 16.87 (sd = 9.31) compared to the control group gains of 13.49 (sd = 10.62). Even though there were increases in scores, none of the changes were statistically significant between groups (Fuchs, Fuchs, Yazdian, & Powell, 2002).

Another study reported that PALS helped students improve their skills in basic math facts as well as in mathematical applications. Forty classrooms from nine different schools were used for the study with half in the PALS condition and half in the No-PALS condition. Teachers implemented the treatments for all students in their classrooms, but only three students were selected from each class to examine treatment effects over a 25 week time period. Students using PALS improved on basic math facts by 7.7 (sd = 3.61) problems correct on a CBM compared to an improvement of 5.05 (sd = 4.36) problems correct by the control group. On math applications the improvement was smaller with the PALS group improving by 5.63 (sd = 3.18) problems and the control group improving 4.4
(sd = 3.16) problems correct. None of the outcomes were statistically significant (Fuchs et al., 1995).

**Summary.** PALS is a newer classwide intervention being used to help increase students' academic achievement. Even though the studies showed an increase in different academic domains (i.e. reading, math, and language), four of the five studies did not report statistically significant differences between the PALS group and the control group outcomes. Only one study (Mathes et al., 1998) had statistically significant results. In this study, on the WRMT-R subtests of word attack and word identification, the difference between low-achieving students in the PALS group and the control group were significant. Two of the studies (Fuchs et al., 2001; Mathes et al., 1998) reported that the high-achieving students in the control groups showed greater improvement (although not statistically significant) than the high-achieving students in the PALS treatment groups. One possible explanation for high-achieving students improving less than the control group is the pace of learning is dependent on their partner. PALS is a relatively new intervention and from the current research it appears to not be as effective as CWPT and cross-age peer tutoring. More research is needed to more fully evaluate the effectiveness of PALS. It should be noted that although many of the studies did not have statistically significant results, those students using PALS did not show a decline in their academic performance. The studies on the PALS program had tight experimental control due to random assignment of individual students rather than random assignment of classrooms, so the results were less likely due to chance.
CRITIQUE OF CURRENT STUDIES

Based on the examination of the research, peer tutoring seems to have many positive effects on students' academic performance. Even with these positive outcomes, some limitations were found in the current studies.

Sampling is a limitation found in most of the studies. With the classwide programs, PALS and CWPT, even when classrooms are randomly placed in different conditions, it is difficult to know whether the most effective teachers are placed in one condition and less effective teachers placed in another condition. In four of the studies, schools were randomly assigned to different conditions rather than the classes in the schools. This makes it difficult to know whether the differences were due to the particular school or to the treatment. In one study (Fuchs et al., 1997), different schools were used because the researchers wanted to avoid having teachers in the control group implement the PALS procedures learned from other teachers in the school. Although this would prevent treatment condition contamination, there still is the possibility there were differences between the control and treatment schools. Another critique of the CWPT programs is the unit of analysis used. The CWPT studies used entire classrooms, rather than individual students, as the unit of analysis. Another limitation in some of the studies was the small sample size (Madrid et al., 1998). Two of the studies (Giesecke et al., 1993, Vacc & Cannon, 1991) only used four dyads and three studies (Barbetta & Miller, 1991, Limbrick et al., 1985, Sideris et al., 1997) used six pairs. One of the studies with four dyads did not randomly assign the students to a condition or use baseline design. Although the students improved over the sessions, it is difficult to know whether the results were due to the
treatment or other factors (Vacc & Cannon, 1991). In two of the studies (Giesecke et al., 1993, Barbetta & Miller, 1991), tutees were selected by their teachers to participate, not randomly selected. Another study (Sidieris et al., 1997) selected the comparison students based on the teacher's judgment, not on the results of an objective measure. The results of those studies that did not randomly assign may not generalize to other students or other settings, especially if the sample size is small. The remaining study (Limbrick et al., 1985) did randomly assign students to the treatment and control groups.

Another limitation, which was also considered a benefit, was the amount of technical assistance received in the PALS and CWPT conditions. The assistance was a benefit because of the support the teachers needed to accurately implement the procedures. When beginning the intervention, it is important to make sure the students are implementing the procedures correctly (i.e. awarding points, corrective feedback, praise, Mathes et al., 1998). The limitation comes when that assistance is not available for all of the teachers. It is difficult to know whether the same results would occur if a teacher implemented PALS or CWPT without any assistance or if teachers using regular teaching methods would do better with more support. Some of the teachers in the treatment condition may have had more difficulties implementing the treatment and so they may have received more assistance than another teacher in the treatment condition with less implementation difficulties (Mortweet et al., 1999).

Another limitation discovered in one study (Mathes et al., 1998) was that the low-achieving students in the PALS and the No-PALS group were not equivalent on the pretest on reading fluency. No information was given on if this limitation was controlled
for. Another difficulty occurred on one of the studies involving improving spelling performance; the words were either too difficult or too easy during the pretest. It is unknown whether the outcomes were due to the treatment or to the set of words given at the pretest. If one student is given words that are too easy, performance improvement would most likely be minimal. Whereas, a student given the more difficult words at the pretest would improve more (Greenwood et al., 1992).

A limitation noted in one cross-age peer tutoring study (Barbetta & Miller, 1991) was the instability of pretest measures, for example, some tutees read words correctly from the unknown word sets that had been missed on the pretest. It was unknown whether or not the subjects were learning the unknown words in another setting besides in the tutoring sessions. This makes it difficult to know how effective the treatment is. Also, teachers reported that the tutoring helped in other academic areas, but data were not collected to substantiate the reports. Without data, it is difficult to know whether this is true or not.

Five studies (Giesecke et al., 1993, Greenwood et al., 1992, Sideridis et al., 1997, Kamps et al., 1994, & Greenwood et al., 2001) did not use a comparison group which would have strengthened the study’s outcomes. The researchers also recommended extending the length of the intervention (it was only a 6-week program) to further validate the positive outcomes on academic achievement. All of the studies, except one (Greenwood et al., 1993), lacked long-term outcome data which would give more information on the effectiveness of peer tutoring programs.
One limitation in a CWPT study (Greenwood et al., 1993) was the researchers used a different assessment at the follow-up than was used in the original study. Also, in this study, it was reported that strength of the treatment was varied across the length of the study.

Another limitation involves the monitoring of the points as well as tutors giving verbal encouragement. It is not possible to watch all of the pairs in the classwide programs. The researchers in two studies (Madrid et al., 1998; Greenwood et al., 1992) were not sure if points were given too conservatively or too liberally to the tutees. One explanation for this limitation, according to Greenwood et al. (1992), was the tutor training was insufficient. Another training problem occurred when some tutors were not prepared for behavior problems displayed by the tutees. The subjects may have shown greater academic improvements if the tutors did not have to spend time dealing with inappropriate behaviors. It is not clear how much time during tutoring sessions was spent managing inappropriate behaviors and how much time was spent on tutoring behaviors (Vacc & Cannon, 1991).

Another possible limitation is that the developers of PALS and CWPT program wrote many of the articles being examined. It is not clear whether this influenced any of the results, but it is important to have independent evaluations completed of the programs.
BEST PRACTICES IN PEER TUTORING

Based on the literature reviewed, some overall conclusions for best practices for implementing a successful peer tutoring program will now be discussed. The programs of discussion will be cross-age peer tutoring, which involves an older peer tutoring a younger peer, Classwide Peer Tutoring (CWPT), in which each student in a class has the opportunity to be both the tutor and the tutee, and Peer-Assisted Learning Strategies (PALS), which also allows each student to be the tutor and the tutee.

Peer tutoring can be used to improve skills in many academic areas at the elementary level for the tutee. In cross-age peer tutoring, students improved their reading skills, including fluency, comprehension, accuracy, and the number of vocabulary words read. Students also improved their basic addition facts along with identification of numbers, days of the week, and months. In CWPT, students seemed to improve their spelling, reading fluency, and reading comprehension skills. Some of the CWPT studies did not include control or comparison groups, so it is not clear whether CWPT was the reason or another variable. Students' performance on standardized assessments increased in reading, math, language, social studies, and science. Students also increased the amount of time they were engaged in academic tasks. Students using the PALS program also showed little differences in reading fluency, reading comprehension, word attack, and word identification as well as math facts and applications.

Tutors also benefit from the peer tutoring. In cross-age peer tutoring, the tutors improved their reading skills along with the tutees. The tutors also increased their self-concepts, which was not the focus of this paper. Five of the studies examined the positive
effects on the tutors during cross-age peer tutoring. Although only one of the studies examined the tutors' self-concept, so more research would be needed in tutors' self-concept to make definite conclusions. In CWPT and PALS each student has the opportunity to be both the tutee and the tutor, so the academic benefits would be the same as previously mentioned.

Each peer tutoring program can be used for students of different abilities. Cross-age peer tutoring improved the skills of both special education and regular education students. CWPT students helped low-, average-, and high-achieving students as well as special education students and students with limited English skills. PALS was also used with low-, average-, and high-achieving students along with special education students.

Since peer tutoring programs can be used with students of various abilities, it is important to consider how to pair the dyads. In cross-age peer tutoring, the tutors will always be older than the tutees. In CWPT and PALS the students are all the same age. In the PALS program, the dyads consist of a higher performing student and a lower performing student (Fuchs et al., 2001), but CWPT does not follow the same rule. Teachers, implementing CWPT, must decide whether to randomly assign students to pairs or to assign based on ability in the different academic areas. None of the studies indicated which way was more effective. Teachers also need to be aware of possible personality conflicts between students. Some problems may arise with different pairs throughout the program, so teachers may need to change some of the pairs. The PALS program also suggests changing the dyads regardless of problems every four weeks (Fuchs et al., 2001).
All of the peer tutoring programs trained the students on the procedures used in each program. It is important for all of the tutors to be trained before the tutoring begins. Tutors need to be taught the importance of positive verbal feedback. It may help to give students a list of common positive statements. Students also need to be told how frequently to give positive verbal feedback. It is important for the tutors to give sincere positive feedback. Tutors also need to be taught on how to respond when an incorrect answer is given, or in other words corrective feedback. When an incorrect answer is given the tutor should quickly give and explain the correct answer without being critical. After giving the correct answer the tutor then allows the tutee repeat the correct answer (Smith, 1995). One study indicated (Vacc & Cannon, 1991) that tutors were not taught how to manage misbehavior. None of the programs indicated how dealing with inappropriate behaviors should be taught.

After explaining about positive verbal feedback and corrective feedback, the teacher then models the behavior for the students. The teacher may give correct and incorrect examples to see if the students understand the two techniques previously explained. Throughout the tutoring it is important to allow the students to ask questions. Role playing between the teacher and a student is next. The teacher must play both roles to see if the students completely understand. If there are only a few tutors (possible in cross-age tutoring), the teacher should role play with each student. The students will then role play with other students while the teacher circulates giving corrective feedback and positive verbal feedback (Smith, 1995).
Although there were many positive outcomes for the peer tutoring programs, some of the outcomes showed very small gains or did not report statistically significant results. Cross-age peer tutoring programs need more research on improving math skills. The CWPT program had some outcomes that were positive, but some of the studies did not use a control group which makes it difficult to know what variable contributed to the academic gains. Also, five of the CWPT studies did not randomly assign participants to the different conditions. The outcomes for the PALS program generally lacked the most statistically significant results, but the studies were more controlled than the CWPT studies. Small improvements were made in many academic areas, but small improvements were also made in the comparison groups. Students in the PALS program did not show a decline in the academic performance, but many of the gains were very small or lack statistical significance.

Given the results of the studies (See Table 4), it is important to examine whether a peer tutoring program would be the best intervention to use when working with children with academic problems. It seems peer tutoring is effective for students of all elementary ages (i.e. grades K-6). It also appears peer tutoring is more effective when teaching basic skills, such as word recognition, spelling, and reading fluency. Basic math skills also seemed to improve due to peer tutoring, but limited research supported this. Due to the tutors being older, cross-age peer tutoring would appear to be more effective with higher order skills (e.g., problem solving), but no research supports this. The studies also did not indicate what specific mechanisms make peer tutoring programs effective (e.g. one-on-one instruction, increased opportunities to respond, etc.). The PALS program appeared not to
have the positive outcomes that were expected. CWPT and cross-age tutoring had more positive outcomes and seem to be the peer tutoring interventions that should be used. Still, more research is needed in peer tutoring, particularly since some of the studies reviewed did not report statistically significant gains or clear outcomes due to the lack of control or comparison groups.

Table 4.
Summary of Current Studies for Cross-Age, CWPT, and PALS Tutoring Programs

<table>
<thead>
<tr>
<th>Study</th>
<th>Program</th>
<th>Participants</th>
<th>Design</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbetta et al., 1991</td>
<td>cross-age</td>
<td>-6 pairs&lt;br&gt;-Tutees (1-3 Grade) chosen based on word recognition scores&lt;br&gt;-Tutors (High School)</td>
<td>-Multiple Probe Design&lt;br&gt;-Participants not randomly selected (teacher selected target students)&lt;br&gt;-No control group</td>
<td>-Number of words mastered increased&lt;br&gt;-Read words in context&lt;br&gt;-Maintained 97% of words after 4 months&lt;br&gt;-Tutors not examined</td>
</tr>
<tr>
<td>Gieseecke et al., 1993</td>
<td>cross-age</td>
<td>-4 pairs&lt;br&gt;-3rd graders and low-achieving 4th graders</td>
<td>-Multiple Baseline with pretest/posttest&lt;br&gt;-Participants not randomly selected (teachers selected on poor reading performance)&lt;br&gt;-No control group</td>
<td>-Tutees increased words mastered by 75% from pretest&lt;br&gt;-Tutors increased sight word recognition by 41%&lt;br&gt;-Tutors increased self-concept</td>
</tr>
<tr>
<td>Taylor &amp; Hanson, 1999</td>
<td>cross-age</td>
<td>-31 pairs&lt;br&gt;-Tutees 7-8 yrs old&lt;br&gt;-Tutors 9-10 yrs old</td>
<td>-Control Group&lt;br&gt;-Participants not randomly selected</td>
<td>-75% of tutees could read book at grade level w/90% accuracy&lt;br&gt;-Improved from 12th to 19th percentile on MAT&lt;br&gt;-Tutors read on grade level</td>
</tr>
<tr>
<td>Study</td>
<td>Age</td>
<td>Participants</td>
<td>Tutors</td>
<td>Improvements</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>Limbrick et al., 1985</td>
<td>cross-age</td>
<td>-6 pairs 3rd graders 18 months below reading level -Tutors 5th graders</td>
<td>-3-phase multiple baseline -Participants randomly selected from lowest reading group</td>
<td>-Tutees answered 70% of comprehension questions compared to 50% at baseline -Tutees and tutors improved on standardized assessments</td>
</tr>
<tr>
<td>Beirne-Smith, 1991</td>
<td>cross-age</td>
<td>-40 pairs 6-10 yrs old -One grade level below average in math</td>
<td>-Pretest/Posttest -Control group -Participants randomly assigned to treatment and control group</td>
<td>-Tutees and tutors improved more than control group from pretest to posttest</td>
</tr>
<tr>
<td>Vac &amp; Cannon</td>
<td>cross-age</td>
<td>-4 pairs 6-12 yrs old special ed students -Tutors 6th graders</td>
<td>-Participants not randomly selected -No baseline</td>
<td>-Improved number of correct responses in rote counting, matching numbers with words, counting objects, identifying words -Tutor not examined</td>
</tr>
<tr>
<td>Fisher, 2001</td>
<td>cross-age</td>
<td>-45 tutors (only tutors examined)</td>
<td>-Control group -Participants not randomly selected</td>
<td>-Outperformed control group on standardized reading assessment</td>
</tr>
<tr>
<td>Greenwood et al., 1989; Greenwood, 1991; Greenwood et al., 1993</td>
<td>CWPT</td>
<td>-9 schools, 416 students -Treatment group included Chapter 1 schools</td>
<td>-Pretest/posttest -Control and comparison groups -Participants randomly assigned by school -Followed students from 1st to 6th grade</td>
<td>-Statistically significant improvements on standardized assessments in reading, math, and language -Academic engagement increased</td>
</tr>
<tr>
<td>Greenwood et al., 1992</td>
<td>CWPT</td>
<td>-88 students from 2nd, 4th, and 5th grades -Grouped by spelling pretest scores</td>
<td>-Pretest/posttest -Participants not randomly selected -No comparison group</td>
<td>-Increased spelling scores 39.7% from pretest to posttest</td>
</tr>
</tbody>
</table>
| Madrid et al., 1998 | CWPT | 16 at-risk students  
- Low spelling test scores and low scores on MAT | 3 conditions  
(CWPT, passive peer tutoring, teacher-mediated instruction  
- Random assignment to each group) | Both passive peer tutoring and CWPT increased spelling scores, with passive tutoring showing greatest increase |
|-------------------|------|--------------------------------------------------|--------------------------------------------------|--------------------------------------------------|
| Greenwood et al., 2001 | CWPT | 5 classes from 1st to 5th grade  
- 117 LEP students | Pretest/posttest  
- Participants not randomly selected  
- No control or comparison group | All grades improved from pretest to posttest |
| Mortweet et al., 1999 | CWPT | 8 target students  
- 4 with mild mental retardation (MMR)  
- Selection based on IEP  
- 4 typical peers selected by teachers based on spelling test scores | Withdrawal treatment design  
- Participants not randomly selected  
- Comparison group was typical peers | All students improved more on spelling accuracy during CWPT than teacher-led instruction  
- Academic engagement time increased during CWPT as well |
| Sideridis et al., 1997 | CWPT | 6 students  
- 3 with disabilities  
- 1 low-, average-, and high-achieving students selected by teacher’s judgment | ABAB design  
- Participants not randomly selected  
- No control group | Increased spelling scores for all students as well as academic engagement time |
| Kamps et al., 1994 | CWPT | 17 students  
- 3 w/autism  
- 14 classroom peers | Multiple baseline  
- Participants not randomly selected  
- No control group | All students, except one, improved during CWPT  
- Social interaction time also increased |
<table>
<thead>
<tr>
<th>Study</th>
<th>PALS</th>
<th>Pretest/posttest</th>
<th>Improved on standardized assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathes et al., 1998</td>
<td>100 students from 20 different 1st</td>
<td>Pretest/posttest</td>
<td>low- and average-achieving students had statistically significant results on word attack, word identification, &amp; passage comprehension</td>
</tr>
<tr>
<td></td>
<td>classrooms</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 students from each class including</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>one low-, average-, and high-achieving student</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-Selection based on CBM given</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fuchs et al., 1998</td>
<td>120 students from 22 classrooms</td>
<td>Pretest/posttest</td>
<td>All groups improved on standardized assessment, but not statistically significant differences</td>
</tr>
<tr>
<td></td>
<td>-40 students w/ learning disabilities</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-40 low-performing students</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-40 average-achieving students</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fuchs et al., 2001</td>
<td>168 kindergarten students</td>
<td>Pretest/posttest</td>
<td>All groups except high-achieving group improved more on SESAT, but not statistically significant differences</td>
</tr>
<tr>
<td></td>
<td>-Placed in low-, average-, or high-achieving groups with a few special education students</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fuchs et al., 2002</td>
<td>20 classrooms (10 for each condition)</td>
<td>Pretest/posttest</td>
<td>All groups improved more on posttest than control group on SAT, but not statistically significant differences</td>
</tr>
<tr>
<td></td>
<td>-Placed in low-, average-, and high-achieving groups based on SAT scores</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
REFERENCES


Smith, K. (1995). Special project to provide technical assistance, inservice training and site development for positive behavioral support strategies for students with disabilities. Minneapolis: University of Minnesota, Institute on Community Integration.


