DOCUMENT RESUME

ED 344 352 EC 301 078

AUTHOR Mathes, Patricia G.; Fuchs, Lynn S.
TITLE The Efficacy of Peer Tutoring in Reading for Students with Disabilities: A Best-Evidence Synthesis.
INSTITUTION Vanderbilt Univ., Nashville, TN. Peabody Coll.
SPONS AGENCY Special Education Programs (ED/OSERS), Washington, DC.
PUB DATE 91
CONTRACT H023B0026
NOTE 97p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Disabilities; Elementary Secondary Education; *Instructional Effectiveness; Meta Analysis; *Peer Teaching; Program Effectiveness; *Reading Instruction; Remedial Reading; *Research Methodology; *Tutoring

ABSTRACT

This review examined the literature on peer tutoring in reading with students with disabilities, using the methodology of best evidence synthesis. This approach combines positive aspects of meta analysis with traditional integrative review procedures. It requires all studies reviewed to meet stringent a priori standards. Analysis of the 11 studies in this review indicates that peer tutoring in reading with students with disabilities can be effective. Peer tutoring was found to have an overall effect size of instruction students with disabilities typically receive. However, it was not found to be more effective than teacher-led instruction when the teacher implemented another research intervention. Average effect sizes across dimensions were consistent; however, individual treatments evidenced variable effect sizes. Treatments in which students with disabilities were paired with normally achieving peers and allowed to serve in the role of tutor some of the time consistently produced strong effect sizes and significant findings. An appendix lists the methodological reasons that 19 additional studies were excluded from analysis. Another appendix details the coding process and compares the 11 selected studies in detail. (94 references) (DB)

The Efficacy of Peer Tutoring in Reading for Students with Disabilities: A Best-Evidence Synthesis

Patricia G. Mathes

Lynn S. Fuchs

George Peabody College of Vanderbilt University

Inquires should be addressed to Patricia G. Mathes, Department of Special Education, Box 328, George Peabody College of Vanderbilt University, Nashville, TN 37203. This research was supported, in part, by the U.S. Department of Education, Office of Special Education Programs (Grant #H023B0026). This information does not necessarily reflect the position or policy of the funding agency and no official endorsement by them should be inferred.

Running head: PEER TUTORING IN READING

Abstract

Peer tutoring is generally accepted as an effective instructional alternative (Gerber & Kauffman, 1981; Jenkins & Jenkins, 1985; Topping, 1988). However, the literature of its efficacy specifically in reading with students with disabilities has not been systematically integrated. This review comprehensively examines the literature on peer tutoring in reading with students with disabilities using the methodology of best-evidence synthesis (Slavin, 1986) which combines positive aspects of meta-analysis (Glass, 1976) and traditional integrative review procedures. Best-evidence synthesis reduces the number of studies included in a review to only those which meet stringent, a priori criteria. Eleven studies met the criteria set forth for this review. The results of these studies indicate that peer tutoring in reading with students with disabilities can be effective. Peer tutoring was found to have an overall effect size of .36 and was found to be more effective than the reading instruction students with disabilities typically experience. However, it was not found to be more effective than teacher-led instruction when the teacher implemented another research intervention. Many formulations of peer tutoring were described in the literature. Average effect sizes across various dimensions were very consistent; however, individual treatments evidenced variable effect sizes. Treatments in which students with disabilities were paired with normally achieving peers and

which allowed the students with disabilities to serve in the role of tutor, at least some of the time, consistently produced strong effect sizes and significant findings.

The Efficacy of Peer Tutoring in Reading for Students with Disabilities: A Best Evidence Synthesis

Peer tutoring is accepted widely as an effective tool for enhancing the academic achievement of all types of students (Cohen, Kulik, & Kulik, 1982; Delquadri, Greenwood, Whorton, Carta, & Hall, 1986; Gerber & Kauffman, 1981; Jenkins & Jenkins, 1985; Kalfus, 1984; Topping, 1988). Peer tutoring refers to an alternative teaching arrangement in which students mediate instruction for other students (Maheady, Harper, & Sacca, 1988). It occurs whenever a teacher arranges for students to be instructed by other students and represents an efficient and feasible use of available classroom resources.

Peer tutoring is not a new idea. Its history has been traced back as early as the first century A.D. to Quintilian in his Institutio Oratoria in which he described an early cross-age tutoring program (e.g., Eiserman, Shisler, & Osguthorpe, 1987). Peer-tutoring strategies were resurrected in this century within the context of the anti-poverty and compensatory education movement of the 1960's (e.g., Elliott, 1991). Since that time, peer tutoring has been acclaimed as an intervention designed to correct underachievement and improve life outcomes of children at-risk for school failure, including students with disabilities (e.g., Gerber & Kaufman, 1981). Empirical evidence to support this acclaim has been slow to accumulate; however, at the present time, a sizable body of empirical literature indicates that peer

tutoring may be beneficial to underachievers and students with disabilities.

The purpose of this review is to synthesize the literature in order to determine the efficacy of peer-tutoring strategies on the reading achievement of students identified as disabled. Although other academic and social benefits of peer tutoring are recognized, they are not included in this review so that the effects of peer tutoring on reading can be separated from other domains. This review sheds light on two main questions:

1. What are the effects of peer-tutoring strategies on the reading achievement of students with disabilities?
2. Under what conditions are peer-tutoring procedures effective for students with disabilities?

Toward this end, (a) the need for peer tutoring in reading for students with disabilities is explored, (b) previous reviews on peer tutoring are examined, and (c) experimental studies investigating the effects of peer tutoring on the reading achievement of students with disabilities are reviewed systematically and integrated using the methodology of best-evidence synthesis (Slavin, 1986).

The Need for Peer Tutoring in Reading

Opportunity to Respond

It is estimated that one in three children experience significant problems in learning to read (National Commission of Excellence in Education, 1983). Of these students, a large

portion qualify for special education services (Will, 1986). A major reason why many of these students do not develop adequate reading skill is that they are not afforded adequate opportunity to practice reading (Allington & McGill-Franzen, 1989; Haynes & Jenkins, 1986; Leinhardt, Zigmond, Cooley, 1981; Nagy & Anderson, 1984; O'Sullivan, Ysseldyke, Christenson, & Thurlow, 1990; Simmons, Fuchs, Fuchs, Mathes, & Pate, 1990).

A major premise of special education is that how teachers structure the learning environment makes a difference in how students spend their time, and how students spend their time affects their level of reading proficiency (Leinhardt et al., 1981, p. 357). Research on effective instruction repeatedly illustrates that students' opportunities to respond academically is a critical factor related to achievement (Brophy & Good, 1986; Greenwood, Delquadri, & Hall, 1984; Rosenshine & Stevens, 1986). The opportunity to respond is defined as "the interaction between: (a) teacher formulated instructional antecedent stimuli (the materials presented, prompts, questions asked, signals to respond, etc.), and (b) their success in establishing the academic responding desired or implied by the material" (Greenwood et al., 1984, p. 64).

The importance of opportunities to respond has been explored empirically by researchers at Juniper Garden's Children's Project of the University of Kansas. Their research suggests that the opportunities students have to respond to academic tasks is a

causal and direct factor in their academic achievement: Greater opportunities to respond result in greater achievement, while few opportunities to respond result in little academic attainment (e.g., Greenwood et al., 1984). Thus, it is imperative that teachers implement strategies which structure the learning environment so that students will respond actively to academics. Peer-mediated instruction is one arrangement that structures the environment to increase students' opportunities to respond (Greenwood et al., 1984, Greenwood, Carta, & Kamps, 1990). Applications of some peer-tutoring strategies in reading have resulted in students receiving double or triple the amount of reading practice (Greenwood, Delquadri, & Carta, 1988; Greenwood, Delquadri, & Hall, 1989). For example, students' opportunities to respond in a reading period were observed to increase from 28% to 78% when peer tutoring was implemented (Elliot, Hughes, & Delquadri, 1984).

Current State of Reading Instruction

Descriptive studies indicate that the regular and special education reading instruction, as it is currently structured, does not provide an environment in which students with disabilities are afforded the necessary opportunities to practice reading to facilitate reading growth (Haynes & Jenkins, 1986; Gelzheiser & Meyers, 1991; O'Sullivan et al., 1990; Simmons et al., 1990). Observations of special education reading teachers indicate they provide less group instruction and more individual

seatwork than their regular education counterparts (Allington & McGill-Franzen, 1989; Haynes & Jenkins, 1986; O'Sullivan et al., 1990). Studies indicate that special education students spend a large percentage of their time waiting, off-task, and working independently on indirect reading activities such as worksheets (Allington & McGill-Franzen, 1989; Gelzheiser & Meyers, 1991; Haynes & Jenkins, 1986; Leinhardt et al., 1981). Haynes and Jenkins (1986) found that children with disabilities sent to resource rooms for reading instruction spent 52% of their time doing worksheets and only 25% of their time actually reading.

Although the picture looks bleak for special education reading classes, time usage in regular reading classes is no better (Gelzheiser & Meyers, 1991). Low performing and mainstreamed students spend approximately two-thirds of their reading periods independent of the teacher and engaged in nonreading or indirect reading activities (Allington & McGill-Franzen, 1989; Haynes & Jenkins, 1986). Additionally, when students are being instructed directly by the teacher in reading, they spend about 70% of their time passively watching and listening to the teacher, with little or no opportunity to respond; they spend only a small fraction of time actually reading (O'Sullivan et al., 1990; Simmons et al., 1990). In one observational study, low-performing fourth graders were given less than 10 seconds of actual reading practice in a 2-week period (Delquadri et al., 1986).

Data also suggest that students most at risk for school failure may receive less reading instruction and practice than their higher-performing peers (Allington, 1984; Hall, Delquadri, Greenwood, & Thurston, 1982). Allington (1984) observed that as early as the first week of first grade, students at risk for qualifying for special education or remedial services received less reading practice and instruction. This translated into the at-risk students having the opportunity to read only 16 words of print as compared to higher achieving students being afforded the opportunity to read 1,933 words while being instructed by the same teachers. Similarly, it has been observed that at-risk first-graders averaged no more than 20 seconds of direct reading practice during a reading instructional period (Hall, Delquadri, & Harris, 1977) and that teachers spent disproportionately more time with high performers, leaving little or no time for reading instruction for low performers (Hall et al., 1982). This trend continues as the years increase, resulting in an ever-increasing gap between the reading proficiency of different ability groups (Nagy & Anderson, 1984).

The result of these differences in reading experiences has been labeled by Stanovich (1986) as the "Matthew Effect" after a verse in the bible which discusses how the rich get richer and the poor get poorer (Matthew 25:29). He comments that, "The very children who are reading well and who have good vocabularies will read more, learn more word meanings, and hence will read even

better. Children with inadequate vocabularies, who read slowly and without enjoyment will read less, and as a result have slower development of vocabulary knowledge, which will inhibit further growth in reading ability" (p. 381).

Given that the children who need the most seem to receive the least, it is not surprising that estimates indicate that 20-30% of the school-age population fail to achieve when provided traditional instruction (Will, 1986). Juel (1988) conducted research indicating that the probability of remaining a poor reader at the end of fourth grade, given a child was a poor reader at the end of first grade, was .88.

In sum, it appears that students with disabilities are not afforded necessary opportunities to read and that they actually receive less instruction than their higher achieving peers. Regardless of setting, students with disabilities appear to spend a good portion of their reading instruction waiting for the teacher, engaged in indirect reading activities and noninstructional activities, and passively watching and listening to the teacher. Opportunities to respond to the teacher's instruction are few, and active engagement in the act of reading is low.

Given that opportunities to respond are essential for academic growth, interventions are necessary to increase students' opportunities to respond during reading instruction. Peer tutoring offers one instructional arrangement which has been

shown to increase students' academic responding significantly (Greenwood et al., 1990; Greenwood et al., 1984; Greenwood et al., 1989). Thus, peer tutoring may represent a promising alternative to rectifying the current state of reading instruction.

Previous Reviews of Peer Tutoring

Many reviews of peer-tutoring interventions have appeared in the literature over the past 20 years. These reviews have evaluated (a) studies with general education populations exclusively (Cohen et al., 1981; Devin-Sheehan, Feldman, & Allen, 1976; Feldman, Devin-Sheehan, & Allen, 1976), (b) studies with disabled populations exclusively (Cook, Scruggs, Mastropieri, & Casto, 1985-86; Eiserman et al., 1987; Osguthorpe & Scruggs, 1986; Scruggs, Mastropieri, & Richter, 1985; Scruggs & Richter, 1985), and (c) studies on peer tutoring regardless of learner type (Gerber & Kaufman, 1981; Kalfus, 1984; Topping, 1988). Collectively, reviewers have agreed consistently that peer tutoring can be an effective instructional alternative. In the following section, the conclusions of past reviews are examined briefly.

Reviews Focusing on General Education Students

In general, reviews agree that peer tutoring is an effective technique for promoting academic gains in normally achieving and low-performing students. Additionally, there is agreement that both tutors and tutees benefit academically, but that the effects

on self-esteem and social behavior are less dramatic (Cohen et al., 1982; Devin-Sheehan et al., 1976; Ellson, 1976; Feldman et al., 1976; Gerber & Kaufman, 1981; Kalfus, 1984; Topping, 1988). Ellson (1976) reported that academic gains were reported only for well structured and cognitively oriented programs. However, Kalfus (1984) concluded that unstructured programs were more effective than independent seatwork, but agreed that structured programs probably promote greater academic gain. Gerber and Kauffman (1981) noted, "In general, the results indicate that peer tutoring may be at least as effective as teacher-led instruction under certain conditions, and that peer tutoring as a supplement to teaching may be better than teaching alone" (p. 160). They point out that the use of peer tutoring represents a different allocation of existing resources, which may or may not result in better academic outcomes. They argue that peer tutoring should be used as a supplement to teacher-directed instruction and that peer-tutoring schemes need to be well planned and incorporated carefully into the ongoing instructional process. Kalfus (1984) explored the role that tutors have played and concluded that tutors can serve as successful mediators of academic instruction, reinforcing agents, and as facilitators of retention.

All of the previously discussed reviews have used traditional integrative review techniques. In contrast, Cohen et al. (1982) used the systematic methodology of meta-analysis

(Glass, McGaw, & Smith, 1981), which may be less subject to bias (e.g., Glass, 1976; Slavin, 1984). An important feature of the Cohen et al. (1982) review is that effect sizes were reported rather level of significance. The average effect size for tutee performance on academic measures was .40, while mean effect size for tutor performance on academic measures was .33. The average effect size for self-esteem measures was only .18. Thus, the effect sizes reported in this review provide support for the conclusions of previous reviews.

Unfortunately, the conclusions of the discussed reviews have not dealt specifically with disabled populations. Therefore, no generalizations about students with disabilities can be inferred, since achievement gains generally are harder to obtain for this type of learner. However, if reviews focusing on disabled populations yield similar results, greater confidence can be held that peer tutoring may be effective for all learner types.

Reviews Focusing on Special Education Populations

Reviews of peer tutoring dealing specifically with students with disabilities are limited in number. In an exhaustive search of the literature, four reviews of peer tutoring with students with disabilities were found (Cook et al., 1985-86; Osguthorpe & Scruggs, 1986; Scruggs et al., 1985; Scruggs & Richter, 1985). It should be noted that Thomas E. Scruggs was an author on each of these reviews, thus they may all reflect one specific perspective.

Handicapping condition. Two reviews on the efficacy of peer tutoring with disabled students focused on specific handicapping conditions. Using traditional review methodology, Scruggs et al. (1985) reviewed 17 studies to determine the efficacy of peer tutoring with students with behavior disorders (BD students). Only studies with an academic focus and with BD students were included, regardless of methodological adequacy. However, methodological considerations were addressed and conclusions were stated with caution. This review yielded four major conclusions. First, BD tutees invariably make academic gains. The amount of gain depends on the level of structure, the type of content, and the ability level of the tutee. Second, BD tutors gain academically when the material taught is academically appropriate for their skill levels. The reviewers conclude that tutors are likely to exhibit gain when they are tutoring in an area in which they need fluency development. Third, BD tutors and tutees appear to benefit socially from peer tutoring, as manifested by increased positive comments and initiations during tutoring time. Fourth, BD students do not appear to gain on global measures of self-esteem or social functioning.

In another review, Scruggs and Richter (1985) evaluated 24 empirical studies on the effects of tutoring interventions on the academic performance of students identified as learning disabled (LD students). Again, studies were included regardless of methodological adequacy, but methodological flaws were addressed.

Twenty studies focused on academics, but reported social data as well. Scruggs and Richter concluded that the effects of peer tutoring on LD students are equivocal. They point out that "it is not clear whether tutoring interventions are more effective than other instructional activities" (p. 285). They go on to state, "It is hard to imagine another instructional intervention in the field of learning disabilities which meets with such unqualified enthusiasm and, yet, is so lacking in empirical evidence" (p. 286). However, in their final analysis they conclude that peer tutoring may have utility for increasing the academic achievement of students with disabilities.

Reverse-role tutoring. Two reviews have focused on the effects of having handicapped students serve as tutors for their peers. Osguthorpe and Scruggs (1986) reviewed 26 studies that measured the effects of tutoring on the academic performance and social development of both tutors and tutees. Only studies in which students with disabilities served in the role of tutor were included; however, in several instances students with handicaps also served as tutees. As with Scruggs et al. (1985), this review was limited to studies with an academic focus. A unique feature was that students with LD, BD and mental retardation (MR students) were examined separately. From the research reported, Scruggs and his colleges concluded that: (a) students with LD, BD, and MR are capable of serving as tutors to both handicapped as well as normally achieving peers, (b) careful training and

supervision are necessary for students with disabilities to serve successfully as tutors, (c) handicapped students experience academic growth by functioning in either tutor or tutee role, and (d) serving as a tutor does not seem to improve handicapped students' self-esteem as traditionally measured.

Cook et al. (1985-86) examined 19 studies in which handicapped students served as tutors using the methodology of meta-analysis (Glass et al., 1981). Studies in which the same students served as both tutor and tutee (i.e., reciprocal tutoring) were excluded, as were nonacademic studies. This review yielded many interesting effect sizes. The mean effect size for participating in peer tutoring was .53 for tutors and .58 for tutees. These effect sizes are larger than those reported earlier by Cohen et al. (1982) for normally achieving students. Moreover, when tutoring was used as a supplement to regular instruction, the mean tutor effect size was .96 and tutee effect size was .69. When tutoring substituted for part of the instructional time the effects for the tutor were less dramatic (ES = .63), however; the difference was negligible for the tutee (ES = .66). Specifically in reading, the average effect sizes for tutors and tutees were respectable, but less impressive (tutor ES = .30; tutee ES = .49). As reported elsewhere, the effects on self-concept were negligible (tutor ES = -.06; tutee = .12). The findings and conclusions of the Cook et al. (1985-86) review were consistent with the findings and conclusions of

other reviews of peer tutoring with students with disabilities. All four reviews agree that peer tutoring holds promise as a tool for increasing the academic gains of students with disabilities.

Generalizations and unresolved issues. Across all learner types, reviews of peer tutoring yield remarkably consistent results. In general, peer tutoring is seen as an effective technique for raising the academic skills of both tutors and tutees. These conclusions are true for students with and without disabilities. Additionally, reviewers agree that self-esteem effects have not materialized.

The reviews of peer tutoring leave a number of questions unanswered. First, previous reviewers have pooled studies from several academic areas together and made generalizations about the effects of peer tutoring on all academics areas. However, there is no evidence that such pooling is warranted. In fact, Cook et al. (1985-86) computed average effect sizes for different academic areas and found the effect sizes to be very different. Unfortunately, Cook et al. (1985-86) included many studies which had serious methodological problems. The present review limited the studies included to only methodologically adequate studies. Thus, the present review yields a more accurate picture of the effects of peer tutoring on reading. Second, previous reviews have treated all tutoring treatments as equal. However, it is not clear that all variations of peer tutoring are of equal effectiveness. The present review will examine individual

tutoring treatments for efficacy and identify common features which seem to be related to effectiveness. Third, peer tutoring has been espoused as a tool for facilitating mainstreaming (Gerber & Kaufman, 1981; Jenkins & Jenkins, 1985; Simmons et al., 1990). However, the impact of the setting of peer tutoring has never been explored. Thus, it is not clear if peer tutoring with disabled students in mainstream setting truly is beneficial. In this review, the impact of peer tutoring on reading achievement in mainstream classrooms and special education classrooms is investigated.

The present review represents a unique addition to the literature. It is the first review to examine comprehensively the literature on the efficacy of peer tutoring for students with disabilities in reading and the first to synthesize this literature using the methodology of best-evidence synthesis.

Method

This review uses the methodology of best-evidence synthesis described by Slavin (1986) for integrating research findings from a body of literature. This methodology incorporates features of both meta-analysis (Glass, 1976) and traditional integrative review procedures. It was developed to draw on the strengths, while avoiding the weaknesses, of both meta-analysis and traditional narrative review (Slavin, 1986).

Best-evidence synthesis has several characteristics that

differentiate it from meta-analysis and traditional narrative review. First, studies are selected carefully as representing the "best evidence" on a given topic through the consistent application of clearly stated, defensible, a priori criteria. Criteria are selected to ensure that all studies included for analysis meet standards of germaneness to the topic and methodological adequacy of the resear Although only selected studies actually are included for analysis, it is incumbent on the researcher to conduct an exhaustive search of the literature to find every study relevant to the topic under examination.

Second, best-evidence synthesis uses both effect size and statistical significance to determine where the weight of the evidence lies. Studies that meet inclusion criteria, but for which effect size cannot be computed, are included in the review with an indication of the level and direction of statistical significance (Slavin, 1986).

Third, effect sizes are used carefully in a best-evidence synthesis. Unlike meta-analysis, average effect sizes are not the primary outcome. Instead, they are presented only as adjuncts to the literature review. Additionally, averaging of effect sizes is conducted only for categories of dependent measures where it is logical to aggregate.

Fourth, the primary emphasis of best-evidence synthesis is the literature review (Slavin 1986). In this section, the reviewer summarizes the findings of each study and makes

conclusions about the topic (Slavin, 1986). Additionally, methodological and substantiative issues of individual studies are discussed.

The "Method" section of this best-evidence synthesis outlines the procedures used in conducting the review and integrating the findings. It delineates (a) how studies were located, (b) what criteria were used for selecting studies, (c) how effect sizes were computed, (d) how studies were coded and categorized, and (e) how averaging of effect sizes was handled.

Literature Search

The first step in conducting this best-evidence synthesis was to conduct an exhaustive search of the literature to locate as complete a set of studies on peer tutoring in reading with students with disabilities as possible (Glass et al., 1981; Light & Pillemer, 1984; Slavin, 1986). This search proceeded through five stages. First, multiple descriptors were generated from key topic-related terms using the Thesaurus of ERIC Descriptors, 12th Edition (Educational Resources Information Center, 1990). The terms generated were peer tutoring, peer teaching, cross-age tutoring, peer-mediated instruction, dyad reading, and paired reading.

Second, these terms were used to facilitate a computer search of two on-line data bases: (a) ERIC, a database on educational materials from the Educational Resources Information Center, consisting of the files Research in Education and Current

Index of Journals in Education; and (b) Comprehensive Dissertation Abstracts. The ERIC database was searched from April 1991 back to 1966 and the Comprehensive Dissertations Abstracts database was searched from April, 1991 to 1861. The descriptors initially were entered into the computer as isolated phrases to promote a wider search (e.g., Dusek & Joseph, 1983). However, for the descriptors "peer tutoring" and "peer teaching," this procedure yielded hundreds of citations. Thus, these two descriptors were restricted by including the requirement that the citation also relate to reading.

Third, a manual search was conducted of pertinent journals from 1980 to April 1991. Journals searched were: American Educational Research Journal, Education and Training in Mental Retardation, Education and Treatment of Children, Exceptional Children, Focus on Exceptional Children, Learning Disability Quarterly, Remedial and Special Education, Journal of Educational Psychology, Journal of Learning Disabilities, The Journal of Special Education, and Reading Research Quarterly.

Fourth, the reference section of all articles collected from the previous three stages were examined for other relevant references. Last, researchers who either had been involved with peer-tutoring research in the past or who were involved presently with peer tutoring research were telephoned and asked if they had any unpublished technical reports, unpublished manuscripts, or in press manuscripts which were not yet available which should be

represented in this review. Contacted researchers were Joseph Jenkins at the University of Washington, Thomas Scruggs at Purdue University, Paul Sindelar at the University of Florida, Deborah Kamps at the University of Kansas - Juniper Gardens Children's Project, and Joseph Delquadri also at the University of Kansas - Juniper Gardens Children's Project.

This search yielded over 130 articles related to peer tutoring in reading. Additionally, multiple reviews and meta-analyses related to peer tutoring both with regular and special education populations were located. All citations were examined and all studies which incorporated reading treatments with low-achieving readers (n = 44) were reviewed further. In keeping with the focus of this review, only studies with populations of disabled learners were maintained in the final review. Moreover, in keeping with best-evidence synthesis methodology, only studies which used group designs were maintained in the final review and coding process (n = 30).

Criteria for Study Inclusion

In keeping with best-evidence synthesis methodology, the studies on which this review is based had to meet a set of a priori criteria with respect to germaneness and methodological adequacy.

Germaneness. To be germane to the review, all studies had to evaluate peer-tutoring treatments designed to address reading deficits of students identified as LD, MR, or BD. Peer tutoring

was operationalized as an instructional arrangement in which school-age students are taught by other school-age students on a one-to-one basis. To be included, studies had to have the following characteristics.

1. The peer-tutoring treatments had to be carried out by school-age students in grades 1 - 12. Teacher involvement could include initial training and monitoring throughout the study. Studies examining tutoring by parents, paraprofessionals, or other adults were excluded.

2. Although the study did not have to focus solely on special education students, the number of participants with disabilities had to be delineated clearly and the effects of peer tutoring on students with disabilities had to be determined separately from effects on nonhandicapped students. This criterion excluded several studies (e.g., Azcoitia, 1989; Brown, 1971; Greenwood, Delquadri, & Hall, 1989; Melberg, 1981; Slavin, 1980; Strother, 1984).

3. The study had to have been conducted in schools during the regular school day rather than in laboratory settings.

4. Study duration must have been at least 6 weeks or a total of 18 sessions; thus, ensuring that the peer-tutoring procedures were employed in schools for extended periods. Only one study was excluded because it did not meet this criterion (e.g., Jenkins, Mayhall, Peschka, & Jenkins, 1974).

5. Peer-tutoring sessions must have been carried out at

least 2 times per week for at least 10 minutes a session; thus ensuring that procedures were part of a regularly occurring classroom activity.

6. The role of the students with disabilities must have been stated explicitly. Studies were included regardless of the role of the special education student; however, their role during tutoring was coded for further analysis.

7. Peer tutoring in reading had to be the principal intervention. Studies in which students were involved in multiple treatments were excluded.

Methodological adequacy. Criteria for methodological adequacy were as follows.

1. The study must have included a control group to evaluate the effects of peer tutoring on reading achievement. This criterion was part of the initial screening of studies; however, two studies had to be examined more closely to determine that no control group was present for the handicapped population (e.g., Eiserman, 1988; Maher, 1986).

2. Evidence of initial equivalence on pretest measures between experimental and control grouped must have been demonstrated. If groups were not initially equivalent, then the degree of nonequivalence must have been quantified or statistically adjusted. Examples of studies excluded based on this criterion are Epstein (1978) and Lombardo (1975).

3. In nested designs in which whole classes or schools were

assigned to treatments, there must have been at least two teachers or schools assigned to each treatment group. If the design was not nested, then subjects must have been assigned randomly to comparison groups. Thus, possible confounds of teacher or school effects were controlled. This criterion excluded Jenkins, Jewell, and Leceister (in press) and Shisler, Top, and Osguthorpe (1986).

4. Dependent measures must have been reported for reading achievement. Although studies may have included other measures such as attitudinal or social skills data, reading achievement must have been reported. This criterion was used to initially screen studies.

5. Reading dependent measures directly tied to the reading peer-tutoring process were permitted only when the control group also followed a curriculum tied to the dependent measures. Otherwise dependent measures had to be more global. This criterion excluded the Maher (1982) study, which used grades received during tutoring as the dependent measure.

It should be noted that almost all studies excluded from this review were excluded on the basis of more than one criterion. A list of excluded studies and reasons for their exclusion can be found in Appendix A.

Computation of Effect Size

Effect sizes were computed for each study to determine the size and direction of effects of peer tutoring in reading on the

reading achievement of students with disabilities. In general, effect sizes were computed using procedures explicated for meta-analysis by Glass et al. (1981), Hedges (1981, 1982), and Hedges and Olkin (1983) As described by Glass et al., (1981), effect size is defined as the difference between the mean final status scores of the experimental group and control group divided by the standard deviation of the control group. The basic effect size formula was adapted as recommended by Hedges (1981) to yield an unbiased estimate of the underlying population effect when sample sizes are small. In this review, it was possible to compute effect sizes for each study that met inclusion requirements.

The specific formula used to determine each individual effect sizes was determined based on the information available in each study. The formula used to determine effect size based on final status scores was:

$$\frac{X_e - X_c}{SD_{pooled}}$$

The pooled standard deviation was defined as:

$$\frac{\sqrt{[(N_e - 1)S^2 + (N_c - 1)S^2]}}{N_e + N_c - 2}$$

Effect sizes were determined using the recommended final status scores only when there was evidence that the experimental and control group scores were equivalent at pretest. When they were not initially equivalent other procedures were followed as recommended by Glass et al. (1981) and Hedges and Olkin (1983).

When analysis of covariance was conducted, effect size was determined using the following formula:

$$\frac{(\text{Regressed adjusted}_e - \text{Regressed adjusted}_c)}{\sqrt{[MS_w\ (df_w - 1) \div (1 - r^2xy)\ (df_w - 2)]}}$$

Because the correlation between the experimental and control group was rarely available, it was estimated to be .80 when not provided. However, in Simmons et al.(1990) and Simmons, Fuchs, Fuchs, Pate, and Mathes (1991) the actual correlations were available and thus used.

In some cases, gain scores were presented for analysis. In these cases, effect sizes was determined in the following manner:

$$X_{e\ diff} - X_{c\ diff} / SD_{gain} \sqrt{(1 - rxy)}$$

In some cases, final status scores were not statistically different at pretest, but examination of the scores indicated that the final status scores were practically different. In these cases, the effect size was determined using gain scores to better represent the magnitude of the effect.

In a few studies, t tests were reported for gain scores while the standard deviation of the gain scores was not reported. In these cases, effect size was estimated based on t-tests based on gain scores. The procedure used was:

$$t_{gain} \sqrt{[2(1 - rxy)\ (1/n_e + 1/n_c)]}$$

In some instances, the effect size had to be determined from the F score as recommended by Glass et al. (1981). In these cases the effect size was determined using the following formula:

$$\frac{2\sqrt{[F\ (1 - r^2xy)\ (df_w - 1)}}{(n_e + n_c)\ (df_w - 2)]}$$

Coding Studies

Studies were reviewed to (a) determine if they met inclusion criteria and (b) code specific features of the studies. A coding form was developed (see Appendix B). All studies were coded by the author. Reliability of the coding process was checked by having a second coder code 10 of the 30 studies included in the final review. Intercoder reliability was determined using the following formula: Percentage agreement = $\frac{\text{agreements}}{\text{agreements + disagreements}}$. In cases of disagreement, discussion occurred until 100% agreement was achieved. Reliability across categories ranged from 80% to 100% with an overall average agreement of 97% (see Table 1).

Insert Table 1 about here

Beyond basic inclusion criteria, studies were coded for: type of subject disability; number of subjects with each handicap; type of reading taught during peer tutoring (i.e., phonics, sight words, decoding, fluency, comprehension, or mixed); setting in which tutoring occurred; type of tutoring (i.e., classwide, individual, cross-age, same age, expert, or reciprocal); role of the handicapped subjects (i.e., tutor, tutee, or both); structure of the tutoring procedures (i.e., structured or nonstructured); type of reinforcement system; dependent measures and their proximity to the type of tutoring conducted; and location of tutoring (i.e., mainstream or special

education). The coding scheme and percentage of agreement (before resolution of disagreements) is presented in Table 1.

Averaging of Effect Sizes

In keeping with best-evidence synthesis methodology, averaging effect sizes was done only when logical. Because this best-evidence synthesis restricted the range of studies to include only those of peer tutoring in reading with students with disabilities, pooling across a variety of characteristics was logical. First, an average effect size for all reading measures and all learner types was calculated. Additionally, effect sizes were pooled to determine the average effect sizes of: (a) role of the student with a disability (i.e., tutor, tutee, or reciprocal); (b) age difference between tutor and tutee (i.e., cross-age or peer); (c) type of reading activity tutored; (d) classroom arrangement (i.e., classwide or individual pairs); (e) location of tutoring (i.e., mainstream classrooms or special education classrooms; (f) prerequisite tutor skill (i.e., expert or reciprocal); and (g) proximity of the dependent measures to the reading content covered during tutoring. The results of pooling are presented in Table 2.

Results

Eleven studies examining peer tutoring in reading with students with disabilities met the inclusion criteria discussed earlier. These 11 studies yielded a total of 74 effect sizes. Table 2 summarizes the aggregated data for major characteristics

of these studies.

Patterns in the Data

Table 2 indicates that the average effects of peer tutoring in reading with students with disabilities, for the most part, were respectable. The average unbiased effect size across all 11 studies was .36, with effect sizes ranging from .07 to .75. This unbiased effect size increased to .40 when calculated from studies comparing peer tutoring to teacher-led instructional control classrooms in which teachers provided reading instruction without any intervention from research staff (i.e., no-treatment control). However, when comparing peer tutoring to teacher-led treatment groups in which teachers implemented a specific reading intervention under examination by the researcher, the effect size decreased to a negligible .14. This indicates that peer tutoring (a) has a greater effect on reading achievement than reading instruction typically implemented by teachers with students identified as disabled and (b) is equally effective to other researcher-guided interventions which are implemented by the teacher. A tenable conclusion is that peer tutoring is as effective as teacher-led instruction, and can be even more effective than teacher-led instruction, depending on the quality of that teacher-led instruction and the tutoring treatment.

Insert Table 2 about here

As can be seen from Table 2, the average effect sizes for the various characteristics generally are consistent; most characteristics have an effect size close to the overall effect size of .36. Additionally, the average effect sizes for all characteristics (excluding comparisons to teacher-led instructional control, i.e., no-treatment control, groups) are significantly different from zero ($p < .01$).

Looking at the results of the pooled effect sizes for specific study characteristics, it would seem that different formulations of peer tutoring are about equally effective. For instance, cross-age tutoring appears equally effective as same-age peer tutoring. This appears to be the case for the tutors' academic knowledge level as well, with effects being approximately equal when the tutor has expert or similar knowledge. Interestingly, treatments with a decoding focus and treatments with a more holistic, multiple focus also appear equally effective.

The setting in which tutoring occurs appears to make a difference. Tutoring treatments conducted in general education appear to have stronger effects than those occurring in special education classrooms. However, it must be noted that subjects in these studies were not assigned randomly to setting. In two of four studies in which tutoring occurred in the mainstream, students with disabilities in the tutoring treatment were already mainstreamed (Simmons et al., 1990; Simmons et al., 1991).

Importantly, these two studies account for 26 of 33 effect sizes pooled for this characteristic. Having only previously mainstreamed subjects is problematic because the average effect size for tutoring in the mainstream may reflect the effect of peer tutoring on more capable learners rather the effect of setting. This is because students with disabilities, who have been previously mainstreamed, may represent a more capable group of learners than students with disabilities who have not been mainstreamed (e.g., Slavin, 1984). Thus, at the present time, it is not clear if setting has a true effect on the strength of peer tutoring in reading with students with disabilities.

A last factor which appears important is the role students with disabilities play during tutoring. From the results of the pooling, it appears that handicapped students make greater gains when they serve as tutors. However, caution is warranted in interpreting this finding. First, tutors were not selected randomly in any study reviewed. In several studies, they were selected because their knowledge on the skills to be tutored was greater than tutees (Carlton, Litton, & Zinkgraf, 1985; Lamport, 1982). Thus, they represented a more capable group of learners who could be expected to benefit more from intervention. Additionally, this finding is biased by one treatment package (i.e., Beginning Decoding) developed by researchers at Brigham Young University and Utah State University. This curriculum was used in three of five studies in which students with disabilities

served as tutors (Scruggs & Osguthorpe, 1986; Top & Osguthorpe, 1985; Top & Osguthorpe, 1987). Thus, no generalizations should be made.

While pooling based on studies characteristics yielded similar results, the average effect sizes of individual studies were not similar. Table 3 presents the major characteristics of each study and presents the average effect size for each treatment. Some treatments produced strong effect sizes, while others yielded negligible results. In the following section, the individual studies are examined and reasons for treatment differences explored.

Insert Table 3 about here

Review of Individual Studies

Studies comparing similar treatments delivered by peer tutors and teachers. Three studies compared reading interventions which were delivered by peers and by teachers. In these studies, the instructional activities and the amount of instructional time were comparable for both tutoring and teacher-led groups (McCracken, 1979; Sindelar, 1982; Russell & Ford, 1983). Thus, these three studies represent a class of studies distinct from the rest. They examined whether a similar treatment is more or less effective when delivered by peers or when delivered by teachers.

McCracken (1979) compared peer tutoring provided by LD students to one-on-one teacher tutoring using the same procedures for teaching sight words to students identified as LD and EMR in special education resource rooms. Of the 11 studies included, this study is the only one to compare peer tutoring to teacher-led, one-on-one tutoring using the same intervention. Additionally, because the peer tutors were identified as LD, it provides an interesting test for examining the feasibility of implementing peer tutoring in resource rooms, with students with disabilities serving as tutors as well as tutees.

The treatment examined by McCracken (1979) consisted of teachers or tutors teaching sight words found on the Slosson Oral Reading Test (SORT) (Slosson, 1963) and then testing word recognition on this same test. Both groups provided one-to-one instruction. Obviously, a serious flaw of this study was that the dependent measures were linked directly to the treatments. However, because both groups taught the same curriculum, effect sizes could be calculated. The reported average effect size of .08 reflects the comparison of the peer-tutored group (tutee scores) to the teacher-tutored group based on the SORT tutee scores and indicates that the two conditions were equally effective. The effect size for each comparison is presented in Table 4. Tests of statistical significance indicated no differences between the teacher-tutored and the peer-tutored tutees on word recognition (t =.61, p =.54) or comprehension (t

= .28, p =.78). Data for the handicapped tutors were not presented.

The McCracken (1979) study also presented comparisons to a no-treatment control group. However, effect sizes were not computed for this control group because the dependent measures were linked to the treatments. As would be expected, both the teacher-tutored and the peer-tutored groups performed significantly better on posttest measures than the control group which did not receive instruction on SORT sight words (t = 7.98, $p < .0001$).

The McCracken study had two major flaws which should lead to cautious interpretation of the results. First, because the dependent measures were linked to the treatments, it is not possible to ascertain whether the treatments promoted general reading ability as indexed by other measures. Secondly, the statistical analyses presented for determining statistical significance were flawed. In order to control for initial pretest differences, McCracken ran a series of t tests for every comparison, rather than using multivariate analysis of covariance; thus, increasing the likelihood of Type 1 error (Glass & Hopkins, 1984). For the purposes of the present review, tests for statistical significance were not used, and effect sizes were calculated based on pretest/posttest gains.

Of major importance to the present review, this study provides evidence that learning disabled peers can be as

effective in providing instruction to their learning disabled classmates as teachers under the same conditions of one-to-one instruction. This study represents the only evidence in the literature that handicapped peers can be as effective as teachers in providing an equivalent instructional intervention. However, it must be recognized that the treatments were highly structured and relatively simple; thus, no generalizations can be made about the ability of handicapped peers to tutor other handicapped peers under other, more complex conditions.

Insert Table 4 about here

Sindelar (1982) presents a study in which peer tutoring with LD students was compared to teacher-directed, small-group instruction utilizing similar procedures. This study differs from the McCracken study in three important ways: a) the tutors were general education students, b) the teacher delivered the treatment to a group of six students, and c) three tutoring treatments were compared to one teacher-led treatment. The teacher-led treatment was similar to one of the tutoring treatments. The study's average effect size of .07 indicates that the peer tutoring provided by students without disabilities produced similar effects as a teacher-directed, small-group intervention.

In this particular study, teachers were assigned randomly to

one of four treatment groups: (a) a hypothesis/test (H/T) peer-tutoring group, (b) a sight-word peer-tutoring group, (c) a sustained oral reading practice tutoring group, and (d) a H/T teacher-directed small group instruction group. The tutors were recruited from regular education classes to tutor their handicapped peers in special education resource rooms. The treatments were based on previous work by Samuels, Dahl, and Archwatemy (1974) which found the H/T method superior to repeated reading and word recognition drill when conducted by the teacher. However, Sindelar (1982) extended the research by implementing the treatments with peers.

It must be noted that the H/T procedure had already been validated as an effective technique when used as a teacher-directed method. Thus, it is reassuring that Sindelar found that peers could conduct this procedure as well as teachers. Additionally, findings indicated that the H/T tutorial group had significantly superior performance to the word recognition tutorial group ($t = 2.92$, $p < .005$). This finding suggests that more complex treatments may be superior to simple sight-word treatments; however, the question of treatment complexity has not been addressed empirically.

Russell and Ford (1983) presented another study comparing peer tutoring conducted by general education peers to similar procedures conducted by teachers to a small group of students. This study is distinct from the Sindelar (1982) and the McCracken

(1979) study in two important ways. First, the focus of the tutoring intervention was more complex and focused on decoding, fluency, and comprehension. Second, the tutoring treatment was patterned after the traditional Directed Reading Activity (DRA) consisting of three phases: preparation for reading, guided reading of a selection, and follow-up activities (Harris & Sipay, 1985). The teachers in the teacher-led group had previously taught reading following this format. Rather than requiring these teachers to implement a different instructional intervention, they continued conducting reading as they normally had. Thus, this study examined whether peer-tutoring was more effective than teacher-led small group instruction when both groups provided reading instruction similar to instruction typically implemented by teachers.

The focus of this study was EMR students. Both tutoring and teacher-led students were assigned to the same classrooms and were assigned randomly to treatments. Both groups conducted daily hour-long sessions for a 3-month period. The lessons included 20 minutes of word introduction and review; 15 minutes of oral reading and drill on word attack skills; 10 minutes of worksheets; and 15 minutes of grading work, recording grades, and charting daily progress. Students in the tutoring condition received all of their reading instruction from their tutors; thus, this treatment totally replaced teacher-led instruction.

The results of this study were impressive; the average

effect size of .75 was the largest obtained for any study in this review. Reading growth was measured using the Peabody Individual Achievement Test (PIAT) reading scales (Dunn & Markward, 1970). The EMR students who were tutored by their general education peers made significantly greater reading growth on both reading recognition (t =3.21, $p < .05$, ES = .52) and comprehension (t = 2.16, $p < .05$, ES = 1.00). This study provides evidence that peer tutors can be more effective than teachers providing small group instruction when the instructional activities are very similar and instructional time is held constant.

Comparisons of peer tutoring to a no-treatment control group: Treatments that focused on decoding. Four studies focused on decoding using a highly structured decoding program modified from Beginning Reading, a curriculum originally designed to be used by parents or paraprofessionals (Harrison, 1982). In each of these studies, the tutoring treatment group was compared to a control group of students receiving small group, teacher-led instruction from special education teachers, without any intervention from the researchers.

The first two studies examined the efficacy of peer tutoring when students with disabilities tutored other students with disabilities in special education resource rooms. Both were reported by Scruggs and Osguthorpe (1986). A more detailed version of both studies was located in an ERIC document (Osguthorpe, Eiserman, Shisler, Top,& Scruggs, 1984).

The first study reported by Scruggs and Osguthorpe (1986) examined the effect of cross-age tutoring; older elementary LD and BD students served as tutors for LD and BD students in lower grades. Tutors were paired with tutees who had less knowledge of the material to be tutored. Tutoring occurred in special education resource rooms; special education teachers paired the tutors and tutees, but researchers supervised the peer-tutoring sessions. The control students received instruction in the same experimental classrooms from the same teachers.

Dependent measures included a criterion-referenced test of decoding skills (Harrison, 1982) and the Woodcock-Johnson Psycho-educational Battery (WJPB) reading subtests comprised of letter-word identification, word-attack, and passage comprehension subtests (Woodcock & Johnson, 1977). The Criterion-referenced test of decoding skills was not linked directly to the treatment; however, it was more closely matched to the tutoring treatment than the WJPB. Thus, it is not surprising that statistical significance based on gain scores was found on all criterion-referenced test measures for tutees in favor of the tutorial group ($p < .05$). However, this was not replicated with the tutors ($p > .05$) On the WJPB, statistical significance was found for only the word attack subtest for both tutors and tutees ($p < .01$).

Despite finding of statistical significance, the overall effect size for this study was a negligible .07. There are two

reasons why this study produced a weak overall effect size. First, statistically significant comparisons had borderline effect sizes due to large control group standard deviations on gain scores, which were used to calculate effect size (WJPB word attack for tutee ES = .20, for tutors ES = .22; Criterion-referenced Decoding for tutees ES = .31). Second, the effect sizes calculated on nonsignificant measures were all below .15 (see Table 4).

The results of this study provide weak evidence that having older students with disabilities tutor younger students with disabilities is academically beneficial to both tutors and tutees with disabilities. However, it does appear that tutees did make reasonable gain on decoding skills. Unfortunately, transfer to other areas of reading was not evidenced. Thus, it is not clear if the treatment itself was inadequate, if cross-age tutoring was not beneficial, or if having students with disabilities serve as tutors for other students with disabilities was ineffective.

The second study reported by Scruggs and Osguthorpe (1986) was similar in design and treatment to the first study. Again the Beginning Reading curriculum was used (Harrison, 1982). However, this study examined reciprocal peer tutoring. LD and BD students were paired with other LD and BD students who had similar decoding skills. One student served as the tutor for half of each session, then the students reversed roles and the second student served as the tutor. Tutoring occurred in special

education classrooms. An important dimension was that students received training on to-be-tutored skills from the teacher or aide before tutoring sessions. Thus, students received both teacher-directed instruction and peer instruction on the content of the tutoring program. Additionally, pairs were supervised by the researchers during each session .

The average effect size obtained in this study was stronger (ES = .23) than in the cross-age tutoring study (ES = .07), even though the actual tutoring curriculum was the same (Harrison, 1982). However, the average effect size of .23 was still not reliably different from zero. As in the first study, dependent measures included a criterion-referenced test of decoding skills (Harrison, 1982) and the Woodcock-Johnson Psycho-educational Battery (WJPB) reading subtests comprised of letter-word identification, word-attack, and passage comprehension subtests (Woodcock & Johnson, 1977).

Results were similar to the cross-age tutoring study. Statistical significance was demonstrated for gain scores on all criterion-referenced decoding test measures ($p < .003$) and for the WJPB word-attack subtest ($p < .01$). Results on other WJPB subtest measures were not significant. The borderline average effect size of .23 reflects a combination of weak effect sizes on the WJPB reading subtests and respectable effect sizes on the criterion-referenced decoding test (see Table 4).

As in the cross-age study, results of this study are

equivocal. Based on this and the first Scruggs and Osguthorpe study (1986), no conclusions can be drawn about the efficacy of peer tutoring in reading with students with disabilities when the tutoring pair consists of two students with disabilities.

Two other studies (Top & Osguthorpe, 1985; Top & Osguthorpe, 1987) made use of the modified Beginning Reading curriculum (Harrison, 1982). The effects achieved in the Top and Osguthorpe studies were greater than those reported by Scruggs and Osguthorpe (1986), even though the research design, treatment duration, tutoring procedures, and frequency were similar in all four studies. However, the two Top and Osguthorpe studies were conducted with both special and general education students working together, while the two Scruggs and Osguthorpe (1986) studies involved only students with disabilities in the tutoring treatment. Thus, providing evidence that peer tutoring in reading with students with disabilities may be more effective when conducted with general education students and disabled students working together.

The Top and Osguthorpe (1987) study actually was conducted before the Top and Osguthorpe (1985) study. Earlier and more detailed versions of the Top and Osguthorpe (1987) report were found in dissertations abstracts (Top, 1984) and an ERIC document (Osguthorpe, Eiserman, Shisler, Top, & Scruggs, 1985). This study examined the effects of reverse-role tutoring in which LD and BD students tutored younger general education students in

basic decoding skills. The tutoring sessions occurred in general education classes and were supervised by instructional aides trained in the procedures. The disabled tutors were in the fourth through the sixth grades, while the general education tutees were all first graders.

A design strength of this study was that students in the tutoring condition and control condition received equivalent amounts of reading instructional time; thus, additional instructional time for the tutoring group was not an issue. Both tutors and tutees were measured for reading growth, and results indicated that both groups made significant reading growth. However, only measures gathered on the disabled tutors are reported in this review.

The average effect size for this study was .48. This effect size reflects only the WJPB (Woodcock & Johnson, 1977) comprised of letter-word identification, word-attack, and passage comprehension subtests. The criterion-referenced decoding test (Harrison, 1982) was reported only for the first grade tutees. The fact that the effect size reflects only WJPB scores is noteworthy, since previous reports using the same curriculum yielded negligible effects on the same measures. As in Scruggs and Osguthorpe (1986), the greatest gain was evidenced on the word attack subtest (F =49.75, $p < .01$, ES =.96); a logical result since the treatment focused primarily on word attack skills. WJPB passage comprehension also was significantly

greater for the tutor group (F = 8.99, p = .01, ES = .41), as was the WJPB Total Reading score (F = 17.79, p < .01, ES = .58). Thus, it appears that the decoding skills transferred to other areas of reading not directly taught during tutoring.

Top and Osguthorpe (1985) essentially replicated their earlier study. Again, LD and BD middle-school students served as tutors for students in primary, general-education classrooms. The general education students included both kindergarten and first grade students. Again, instructional time was held constant across conditions and the curriculum used during tutoring was a modified version of Beginning Reading (Harrison, 1982). However, a major difference of the 1985 study was that the BD students conducted the tutoring sessions in special education classrooms, while the LD students conducted tutoring in general education classrooms. Thus, results for LD and BD students are reported separately.

Results indicated that both groups of tutors benefited from tutoring younger general education students and that the general education students also made significant reading achievement gains. The average effect size for the study was .63. This reflects LD, BD, and LD/BD student scores on the WJPL (Woodcock & Johnson, 1977) reading subtests. All measures significantly favored the students who served as cross-age tutors for normally achieving students (p < .01) (see Table 4).

The results of this second study are striking considering

that the curriculum and basic procedures were essentially the same as those reported by Scruggs and Osguthorpe (1986). Given the disparity in results between the two Scruggs and Osguthorpe (1986) studies and the two Top and Osguthorpe (1985, 1987) studies, it appears that the Beginning Reading curriculum may be best suited for cross-age tutoring, with handicapped students tutoring general education students rather than other special education students. Additionally, the two Top studies demonstrate that students with disabilities can serve effectively, under controlled conditions, as tutors for younger general education students and that this inverse role can be beneficial for the students with disabilities.

A last study examining the effects of a peer tutoring decoding treatment on the reading achievement of students with disabilities was Carlton et al. (1985). A more detailed version of this study was located from Dissertation Abstracts (Carlton, 1981). This study is unique because EMR students served as tutors for EMR tutees and the treatment was conducted with entire classes at the same time (i.e., classwide). This study is the only study in the literature in which EMR students served as tutors.

The treatment presented in the Carlton study is similar to the treatment described by McCracken (1979), in which sight words were taught. The sight words taught in the Carlton study were listed in the Brigance Inventory of Basic Skills (Brigance, 1977)

and presented for 6 weeks. Tutee and tutor roles were determined by administering a sight word inventory and assigning the lower scoring half of each class to the tutee condition and the higher half to the tutor condition. Subjects were assigned randomly to either the tutoring condition or to a teacher-led, no-treatment control in which teachers taught reading using their typical method.

Results of this study are impressive when one considers the simplicity of the treatment. Subjects in the tutoring condition demonstrated significantly greater gains on all measures of the Gates-McGinitie (MacGinitie, 1978). This was true for tutors and tutees. Perhaps most impressive is the finding that significant comprehension gains were evidenced ($F = 3.30$, $p < .05$). The average effect size calculated for this study was a respectable .38. The effect sizes for tutee and tutor role were essentially the same (see Table 4).

It must be recognized that this study had one serious flaw which should lead to cautious interpretation of results. It appears that the tutoring group received more overall instructional time in reading. Thus, it cannot be concluded that the tutoring treatment alone was responsible for results. However, a design strength was that the dependent measures were not related to the treatment, as they were measures of general reading ability.

Comparisons of peer tutoring to teacher-led, no-treatment

control: Treatments with a multiple focus. The last three studies which met inclusion criteria focused the peer tutoring treatment on both reading fluency and comprehension (Lamport, 1982; Simmons et al., 1990, 1991). Lamport (1982) presented a cross-age tutoring study similar in design to the cross-age tutoring study reported by Scruggs and Osguthorpe (1986). Sixth-grade students identified as LD tutored second- through fourth-grade LD students. As in Scruggs and Osguthorpe (1986), students in the tutoring group and students in the control group were assigned to the same teachers. Students in each group were matched according to pretest performance on the Stanford Diagnostic Reading Test (Karlsen, Madden, & Gardner, 1977) so that there was a matched control group for both tutors and tutees.

The tutoring treatment occurred in special education resource classrooms and was similar to the treatment described by Russell and Ford (1983). However, the students with disabilities served as tutors of their disabled peers. Each session lasted 55 minutes twice a week for 8 weeks and consisted of the following sequence: (a) word study, (b) oral reading of a selection, (c) discussion of the selection, (d) skills activities, and (e) record keeping activities. Thus, the tutoring treatment followed a sequence similar to the traditional DRA lesson (Harris & Sipay, 1985). Unfortunately, it could not be determined from the manuscript what reading instruction was like in the control

condition. Thus, it could not be assumed that c ntrol students received teacher-led instruction following a similar DRA format.

The average effect size for the Lamport study was .44 on Stanford Diagnostic Reading Test (SDRT) measures (Karlsen et al., 1977). Tutors significantly outperformed controls on the phonetic analysis subtest ($F = 5.43, p = .03$, ES = 1.04) and on the auditory vocabulary subtest ($F = 5.60$, $p = .03$, ES = .69). No other SDRT subtest scores were statistically significant. Even so, respectable effect sizes were obtained for the tutees' performance on the phonetic analysis (ES = .33) and comprehension (ES = .29) subtests, indicating that this particular treatment had favorable effects for both the tutee and tutor groups (see Table 4).

The results of the Lamport (1982) are similar to those reported by Russell and Ford (1983), which made use of a similar treatment. These two studies present evidence that the traditional DRA reading lesson sequence can be utilized effectively as a peer tutoring treatment. Additionally, this study provides evidence that students with disabilities can serve effectively as tutors for younger students with disabilities, even when the tutoring procedures are relatively complex. This finding lends support to the equivocal findings of Scruggs and Osguthorpe (1986) and provides evidence that students with disabilities can serve as tutors in resource rooms, even when the tutoring procedures are complex, an issue left unanswered by

Carlton et al. (1985). It may be that the effects for the tutors with disabilities were better in this study than in the Scruggs and Osguthorpe (1986) study due to the nature of the treatment. The wider focus of the tutoring treatment of the Lamport study may have represented a better match to the needs of the tutor, while the Scruggs and Osguthorpe's (1986) focus on decoding may not have been an appropriate match.

The final two studies of this review (Simmons et al., 1990, 1991) were conducted by a team of researchers at George Peabody College at Vanderbilt University of which the author was a member. These studies both examined treatments implemented in the mainstream, by mainstream teachers as part of efforts to modify mainstream reading instruction to accommodate the needs of students with mild disabilities.

In Simmons et al. (1990), experimental teachers implemented an instructional model based on teacher effectiveness literature during their daily teacher-directed reading instruction. In half of these classrooms peer tutoring occurred 3 days per week for 8 weeks as a supplement to the teacher-effectiveness model. Tutors were selected by the teachers from among higher performing students in the same general education classroom.

The treatment consisted of two components: (a) fluency development based on repeated reading (O'Shea, Sindelar, & O'Shea, 1987; Samuels, 1987) and (b) comprehension development

utilizing paragraph restatement (Jenkins, Heliotis, Haynes, & Beck, 1987). Tutoring occurred 3 times per week for 20 minutes each session. Unfortunately, teachers had little involvement with the tutoring treatment. Tutors and tutees were trained and supervised by research staff, and the tutoring sessions were conducted away from teachers (usually in the hall, outside the teachers' classrooms) while they worked with other students. Importantly, tutoring replaced independent seatwork time; thus, tutoring did not represent an addition to the overall instructional time allotted to reading instruction. Rather it represented only a different use of already allocated time.

Results on the Comprehensive Reading Assessment Battery (CRAB) (Fuchs, Fuchs, & Hamlett, 1989) indicated that the tutoring group significantly outperformed both the teacher effectiveness treatment group and the no-treatment control group on the number of words read ($F = 4.28$, $p < .05$) and on comprehension questions answered correctly ($F = 4.67$, $p < .05$). On words read correctly, the effect size was .35 when the tutoring group was compared to the control group and .30 when the tutoring group was compared to the teacher-effectiveness treatment group. The effect size for comprehension was .44 for the control group comparison and .28 for the teacher-effectiveness group comparison. The peer-tutoring group also outperformed the control group on a story summarization measure ($F = 3.76$, $p < .05$, ES = .52) and a maze task ($F = 3.49$, $p < .05$,

ES = .77), and outperformed the teacher-effectiveness group on the reading comprehension subtest of the Stanford Achievement Test (Gardner, Rudman, Karlsen, & Mervin, 1982) ($F = 3.21$, $p < .05$, ES = .37). The overall average effect size of .35 reflects both the teacher-led and the no-treatment control group comparisons and includes several nonsignificant comparisons (see Table 4). It should be noted that the over all average effect size for peer tutoring compared to the control condition was .45. However, when compared to the teacher effectiveness treatment group, the effect size dropped to .24.

The peer-tutoring treatment developed by Simmons et al. (1990) was modified and extended in the Simmons et al. (1991) study. Repeated reading and paragraph restatement were modified for use with entire classrooms (i.e., classwide). The repeated reading and paragraph restatement treatment (i.e., Peabody model) was compared to an already established classwide peer-tutoring model developed at the University of Kansas Juniper Garden's Children's Project (i.e., Kansas model). The Kansas model consisted of sustained, oral reading practice followed by tutor generated comprehension questions.

Additionally, the role that the students with a disability played during the process was examined. In some classrooms, LD students served as tutees for half of each session and as tutors for half of each session. In other classrooms, LD students always served as tutees. Thus, there were four versions of

classwide peer tutoring compared in this study: (a) version 1 representing the Kansas model with LD students serving both as tutees and tutors, (b) version 2 representing the Kansas model with LD students always serving as the tutee, (c) version 3 representing the Peabody model with LD students serving both as tutee and tutor, and (d) version 4 representing the Peabody model with LD students serving a tutee only. Tutoring sessions were conducted 3 times per week, for approximately 35 minutes per session for 14 weeks. Each of the 4 treatments was similar in focus (i,.e., fluency and comprehension) and time requirements. However, in the static-role versions, LD students received more oral reading practice. Additionally, the tutoring treatments were compared to a teacher-led, no-treatment control.

This study is distinctive because it compared four versions of peer tutoring and a control group. Additionally, this study independently tested the Kansas model, an established tutoring method. The Kansas model has been validated with low-achieving students and students with disabilities. However, studies examining the Kansas model were not included in this review because investigations with disabled students have used only single-subject design methods and group design studies did not provide data in which the disabled students could be examined separately, thus, these studies not meet inclusion criteria.

As in Simmons et al. (1990), the Comprehensive Reading Assessment Battery (CRAB) (Fuchs et al., 1989) was used to assess

reading achievement in this study. The reported results included LD as well as average achieving students and low-achieving students. The findings indicated that all four peer-tutoring treatments significantly outperformed the control group on words read correctly ($F = 3.26$, $p < .05$). On comprehension questions answered correctly, only the reciprocal Peabody version (version 3) reliably exceeded controls. There were no reliable difference between tutoring groups and the control group on any other measures. Additionally, no tutoring treatment reliably outperformed any other treatment on any CRAB measure.

Effect sizes were calculated from reanalyzed data to include only the LD students. The effect size trends look different from the tests of statistical significance based on all subjects. As can be seen from Tables 3 and 4, role reciprocity appeared to be an important factor in the effect sizes. The two reciprocal versions (i.e., versions 1 and 3) yielded significantly greater effect sizes than the two static role versions (i.e., version 2 and 4). The average effect sizes from the two reciprocal treatments were not significantly different from each other (Kansas - reciprocal ES =. 76, Peabody - reciprocal ES = .65). Thus, superiority between these two treatments with LD students can not be inferred based on their effect sizes. However, statistical significance was evidenced between the average effect sizes of the two static role versions (Kansas - static role ES = -.07, Peabody - static role ES = .25) ($P < .05$). Thus, it can be

inferred that given the static tutee role, the Peabody version is superior with LD students.

Based on these findings, it appears that LD students do benefit from classwide applications of peer tutoring and that this benefit is greater when they serve as both tutors and tutees. It is not clear why reciprocity of role makes a difference. In actuality, students who always served as tutees were afforded more supervised reading practice. It can be speculated that reciprocal versions allowed for modeling of fluent reading and comprehension, or that this formulation may increase students' investment in the process.

Discussion

The results of this review indicate that peer tutoring in reading with students with disabilities can be effective. In general, students with disabilities who participated in peer-tutoring reading interventions made greater reading achievement gains than control students who experienced typical teacher-directed reading instruction without researcher intervention. The effect sizes generated were generally educationally relevant and significantly different from zero. However, caution is necessary when interpreting the results. While peer tutoring in reading was generally effective, certain formulations resulted in impressive gains, while others evidenced only negligible results. Thus, it must be recognized that the effectiveness of peer tutoring in reading is dependent on the actual tutoring

treatment and the needs of the learner.

Interestingly, peer tutoring in reading was not found to be superior to teacher-direct instruction when teachers implemented a researcher driven intervention (i.e., McCracken, 1979; Simmons et al., 1990; Sindelar, 1982). However, peer tutoring was found to be more effective than the traditional DRA reading lesson format (Harris & Sipay, 1985) typically followed by teachers in group instruction, even when the peer tutoring procedures followed the same basic format (Russell & Ford, 1983).

Unfortunately, none of the studies compared peer tutoring to other empirically validated methods such as cooperative learning (Stevens, Madden, Slavin, & Farnish, 1987) or Direct Instruction (Becker, 1984). Thus, as pointed out by Scruggs and Osguthorpe (1985) it remains unclear whether peer tutoring interventions with students with disabilities are equally or more effective than other validated interventions.

It is important to keep in mind that peer tutoring in reading represents only one of many interventions which probably deserve a place in the repertoire of both special education and general education teachers. It would be foolish to think that peer tutoring could or should represent students' total reading program. What is clear from the present review is that peer tutoring in reading with students with disabilities can promote significant reading growth as compared to typical reading instruction occurring in most special education and general

education classrooms. As such it represents one promising methodology for improving the current state of reading instruction.

Consistency with Previous Reviews

Like the present review, previous reviews have concluded that students with disabilities usually evidence greater academic gain as a result of peer tutoring than they do from typical instruction (Cook et al., 1985-86; Osguthorpe & Scruggs, 1986; Scruggs et al., 1935; Scruggs & Richter, 1985). However, the findings of this best-evidence synthesis are somewhat more conservative than previous reviews. The present findings probably reflect a more accurate estimate of the true effect of peer tutoring in reading with students with disabilities since only methodologically adequate studies which focused on reading were included. Other reviews have made conclusions about peer tutoring in general based on studies from a variety of academic areas and have included studies of questionable technical merit. Given that both Cook et al. (1985-86) and Cohen et al. (1982) found the effects of peer tutoring in reading to be weaker than for other academic areas, it seems a mistake to generalize aggregated findings as if they were applicable to all academics. Additionally, previous reviews have discussed peer tutoring as if it were one intervention. However, many variations and formulations exist. As with most interventions, the efficacy of peer tutoring seems to depend on the actual tutoring treatment

and the needs of the students.

Comparison to previous meta-analyses. Cook et al. (1985-86) presents the only meta-analysis in the literature to examine the effects of peer tutoring with students with disabilities. The Cook et al. (1985-86) review focused on the average effects for peer tutoring across academic areas. However, as an adjunct, average effect sizes by individual academic areas were calculated.

The overall average effect size calculated from the 11 studies included in the present review (ES =.36) represents a somewhat more reserved, yet consistent, estimate of the strength of peer tutoring in reading with students with disabilities. The Cook et al. (1985-86) meta-analysis reported an average effect size of .49 for peer tutoring on disabled tutees and .30 on disabled tutors in reading. Interestingly, the trend of these effect sizes for tutor/tutee role are opposite to the effect sizes found in the present review (ES tutee = 30; ES tutor = .42).

The Cook et al. (1985-86) meta-analysis included many studies which were excluded from the present review (i.e., Caspo, 1976; Lombardo, 1975; Melberg, 1981). These excluded studies reported very favorable tutee results. Additionally, the present best-evidence synthesis included several more recent studies not included by Cook et al. (i.e., Russell & Ford, 1983; Simmons et al., 1990, 1991). In all, the two reviews shared only the

Carlton (1981) (i.e., Carlton et al., 1985), Lamport (1982), and Top (1984) (i.e., Top & Osguthorpe, 1987) studies. Given that the pool of effect sizes for tutee and tutor effects were so different, it is clear how disparity between the findings of the two reviews occurred.

Interestingly, the findings of this best evidence synthesis and the Cook et al. (1985-86) meta-analysis yielded higher average effects sizes for peer tutoring in reading with disabled populations than the Cohen et al. (1982) meta-analysis with normally achieving populations. The average effect size for peer tutoring in reading reported by Cohen et al. (1982) was only .21. As theorized by Gerber and Kauffman (1981), it may be that peer tutoring has greater benefits for students with learning problems than for normally achieving students. One can speculate why peer tutoring appears to have stronger effects low-performing students and students with disabilities. However, given the evidence that students with disabilities are not afforded as many opportunities to practice reading as their higher performing counterparts (Allington, 1984; Hall et al., 1982) it may be that increasing opportunities to respond for these students allows them to "catch up" somewhat. Additionally, it is likely that normally achieving students have attained a level of automaticity, not yet achieved by students with disabilities; thus, as automaticity is achieved, gains are evidenced. Since the higher performing students have already achieved a high level of proficiency, achievement growth,

as traditionally indexed, is less evident. Of course this remains speculative.

Peer Tutoring Characteristics

This best-evidence synthesis yielded average effect sizes that were extremely consistent across various study dimensions (see Table 2). Given this consistency, one might conclude that different formulations of peer tutoring are equally effective. However, the consistency of average effect sizes on different tutoring dimensions did not necessarily characterize the specific treatments. For example, based on aggregated data, one might conclude that serving in the role of tutor was more powerful for students with disabilities than other roles, but that serving as the tutee or in a reciprocal role were essentially equal. However, Simmons et al. (1991) demonstrated that for two different classwide peer tutoring treatments, the effects were significantly stronger when the students with disabilities served in a reciprocal role rather than as only tutees.

Two factors did reliably result in differences in the effect size obtained: a) comparison group and b) the interaction of setting and student role. When peer tutoring in reading was compared to a teacher-directed research intervention, the average effect size was negligible; however, when compared to a no-treatment control group in which teachers presented reading in their typical manner, the peer tutoring group outperformed their counterparts by nearly half a standard deviation.

Moreover, the interaction of the setting and the role that the disabled student played during tutoring made a reliable difference. In every instance in which: a) tutoring occurred in general education settings and b) the students with disabilities were tutors at least part of the time, significant results and strong effect sizes were achieved (Simmons et al., 1991;, Top & Osguthorpe, 1985, 1987). Thus, it may be that, at least in reading, students are more likely to make significant achievement gains when they are paired with general education students and when they are allowed to participate in the role of tutor at least part of the time.

Other factors determining the effectiveness of particular peer tutoring treatments appear to be specific to the individual treatment. For instance, the Beginning Reading (Harrison 1982) treatment was utilized in four studies. In two of these studies, the effects were negligible (Scruggs & Osguthorpe 1986), while in the other two the effects were substantial (Top & Osguthorpe, 1985, 1987). The factor which seems to account for this difference was the type of students who were paired together. When disabled students were paired with other disabled students, as in Scruggs and Osguthorpe (1986), the effects of the Beginning Reading treatment were negligible. However, when disabled students were paired as tutors of younger, normally achieving students, the results with the same treatment were quite impressive. However, concluding that disabled students need to

be paired with nondisabled students for peer tutoring to be effective would be erroneous. Both Carlton et al. (1985) and Lamport (1982) achieved significant results and strong effect sizes with disabled students tutoring other disabled peers.

Based on the aggregated effect sizes, one might conclude that peer tutoring conducted in special education settings are less effective than tutoring treatments occurring in general education settings. Again, this appears to be a treatment specific phenomena. Carlton et al. (1985), Lamport (1982) and Russell and Ford (1983) all achieved impressive results in special education classrooms. Additionally, it must be remembered that the types of students who are found in special education classrooms are likely to more difficult to evidence achievement gains with, since they probably represent a lower-functioning group than students who have been mainstreamed. Thus, it would be expected for treatments to look less powerful in special education settings. However, evidence does indicate that, if the treatment is strong, impressive achievement gains can be achieved in special education settings.

Another treatment specific phenomena was evidenced in Simmons et al. (1991). In this study, widely different results on similar treatments were achieved by varying the role of the disabled learner. When the students with disabilities were allowed to be the tutor for part of the time, sound reading gains are achieved, but when the disabled students served only as the

tutee, these gains were not evidenced. However, several other treatments achieved strong effects even when the student with a disability served only as a tutee (i.e., Russell & Ford, 1983; Simmons et al., 1990). Furthermore, in the static role versions, the two treatments achieved different results, even though both treatments were conducted classwide, occurred in the mainstream, had a multiple focus and provided for similar amounts of reading practice.

Thus, while it does appear that trends exist, it seems reasonable to conclude that tutoring treatments need to be assessed individually. Specific treatments such as the reciprocal, classwide treatments presented by Simmons et al. (1991), and the treatments based on the traditional DRA (Lamport, 1982; Russell & Ford, 1983) appear to be powerful, while others appear to have less strength. Of course this is an area that needs further research. The results of this review lead to the conclusion that one cannot assume that peer tutoring, in and of itself, will effect positive growth. Treatments must be carefully planned, trained and assessed.

Unanswered Questions. Many questions about peer tutoring in reading with students with disabilities remain unanswered and future research is necessary. Only future research will be able to sort out the following issues:

1). It not known how long daily sessions of peer tutoring should be, how often they should occur in a week, as

well as how many weeks of the school year peer tutoring needs to be in effect for optimal growth.

2). Likewise, it is unknown which tutoring interventions provide the most power. However, in choosing or designing a peer tutoring intervention, it does seem likely that treatments which develop automaticity and also address comprehension are likely to promote better reading growth.

3. It is not known if peer tutoring is most effective used as a supplement to teacher-directed instruction, or if its use as an instructional replacement is most beneficial. However, evidence exists that even when tutoring totally replaces teacher-directed instruction, peer tutoring can still produces superior results (Russell & Ford, 1983).

4. Each of the studies which met inclusion criteria presented highly structured tutoring procedures in which the students received intensive training. No generalizations can be made about less structured and less trained procedures. Thus, it is not known how much structure is necessary for a peer tutoring treatment to be effective.

6. It is not known what level of complexity produces the greatest gains. For instance, Carlton et al. (1985) achieved impressive results with a very simple sight-

word treatment. However, other low-complexity treatments did not exhibit the same effects. Positive results occurred more frequently with more complex tutoring procedures (Lamport, 1982; Russell & Ford, 1983; Simmons et al., 1990,1991).

5). Perhaps more importantly, it is not known how peer tutoring in reading with students with disabilities compares to other empirically strong methodologies or if other interventions are enhanced by the addition of peer tutoring component.

Conclusion

The purpose of this review was to synthesize the literature on peer tutoring in order to determine its efficacy on the reading achievement of students with disabilities. Results indicated that the effectiveness of peer tutoring in reading with students with disabilities is not an "all or nothing proposition." Its effectiveness is dependent on the specific tutoring treatment and the needs of the students.

The results of this best-evidence synthesis lead to several interesting conclusions. First, it appears that peer tutoring in reading with students with disabilities is generally more effective than the reading instruction students with disabilities typically experience. Of course, this finding is not surprising giving the current state of reading instruction. It is likely that an important reason why peer tutoring promotes greater

reading growth is that it provides a structure in which students' opportunities to respond and become fluent on various reading skills is increased substantially. Moreover, it allows students to be carefully monitored and to receive immediate feedback and reinforcement on their reading performance.

Second, it may be that peer tutoring is more effective with disabled populations than with normally achieving populations. Of course, this is a question that needs to be examined empirically, but it does make sense that it would promote greater gain for students who are in need of greater opportunities to respond (Greenwood et al., 1984). Assuming that peer tutoring is more effective with disabled or low-achieving students, then it represents a useful methodology for narrowing the gap between normally achieving students and students with disabilities.

Third students with disabilities can effectively serve as tutors for their disabled and nondisabled peers when that tutoring procedures are highly structured and well trained.

Fourth, and relatedly, peers can provide instruction as well as teachers. This appears to be true for both disabled and nondisabled students. It is worth noting that no control group out-performed any tutoring group. Since this review included unpublished studies, the likelihood of finding a "no-effects" study was increased. Thus, it can be concluded that peer tutoring will not cause detrimental effects, even if it does not promote greater gain. This finding is important because it

indicates that teachers can use peers as proxies for adult instruction without worrying that tutoring will have detrimental effects. Essentially, peers can be used to increase the options available to teachers for meeting the individual needs and increasing the academic engaged time of their students.

Fifth, assuming that peer tutoring in reading with students identified a disabled is not more effective than carefully designed and implemented teacher-directed instruction, it does appear to be more effective than having students off-task, waiting for their teachers, or completing busy work, as is typically observed in reading classes (Haynes & Jenkins, 1986; O'Sullivan et al., 1990; Simmons et al., 1990).

Sixth, specific formulations of peer tutoring in reading seem to hold great promise for aiding with mainstreaming efforts while simultaneously increasing the reading achievement of students with disabilities. Two formulations which produced significant gains in mainstream settings were cross-age tutoring in which the disabled student served as the tutor for younger, normally achieving students and classwide peer tutoring in which mainstreamed disabled students participated in tutoring activities simultaneously with all other students in a general education classroom. The classwide application seems to hold the greatest potential because: a) it provides mainstream teachers with a structure that is highly feasible to implement since it does not require the coordination of multiple classes and

schedules, b) facilitates differentiated instruction since reading text and tutoring procedures can be individualized for specific students within the group, and c) enhances disabled students' ability to perform adequately in the mainstream by providing an environment which is highly structured, carefully monitored, highly reinforcing, and provides necessary opportunities to practice reading.

While much remains to be learned about peer tutoring in reading with students with disabilities, this best-evidence synthesis indicates that it can be a methodology of great power and utility. As such, it represents one promising intervention for the improving reading achievement of students with disabilities.

References

Allington, R.L. (1984). Content coverage and contextual reading in reading groups. Journal of Reading Behavior, 16, 85-96.

Allington, R.L., & McGill-Frazen, A. (1989). School response to reading failure: Instruction for chapter 1 and special education students in grades two, four, and eight. The Elementary School Journal, 89, 530-542.

Azcoitia, C. (1989). Structured peer tutoring in Chicago's vocational education program. Doctoral Dissertation, Northern Illinois University (ERIC Document Reproduction Service NO. ED 320 020)

Becker, W.C. (1984, May). Direct Instruction: A twenty year review. Paper presented at the 16th Banff International Conference on Behavioral Science, Alberta, Canada.

Brigance, A.H. (1977). The Brigance comprehensive inventory of basic skills (2nd Ed.). North Billerica: Curriculum Associates.

Brophy, J., & Good, T.L. (1986). Teacher behavior and student achievement. In M.C. Wittrock (Ed.), Handbook of Research on Achievement. New York: Macmillan.

Brown, J.C. (1971). Effects of token reinforcement administered by peer tutors on pupil reading achievement and tutor collateral behavior. Dissertation Abstracts International, 32, 3775. (University Microfilms No. 72-03, 020)

Carlton, M.B. (1981). The effects of an intraclass peer tutoring

program on the sight-word recognition ability of mildly mentally retarded students. Dissertation Abstracts International, 40, 4516. (University Microfilms No. 79-245, 399)

Carlton, M.B., Litton, F.W., & Zinkgraf, S.A. (1985). The effects of an intraclass peer-tutoring program on the sight-word recognition of students who are mildly mentally retarded. Mental Retardation, 23, 74-78.

Caspo, M. (1976). If you don't know it, teach it. Clearing House, 12, 237-248.

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

Cook, S.B., Scruggs, T.E., Mastropieri, M.A., & Casto, G.C. (1985-86). Handicapped students as tutors. The Journal of Special Education, 19, 155-164.

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

Devin-Sheenhan, L., Feldman, R.S., & Allen, V.L. (1976). Research on children tutoring children: A critical review. Review of Educational Research, 46, 355-385.

Dunn, L., & Markward, F. (1970). Peabody individual achievement test. Circle Pines, MN: American Guidance Service

Dusek, J.R., & Joseph, G. (1983). The bases of teacher

expectations: A meta-analysis. Journal of Educational Research, 75, 327-346.

Educational Resources Information Center. (1990). Thesaurus of ERIC descriptors (12th ed.). Reston, VA: Author.

Eiserman, W.D. (1988). Three types of peer tutoring: Effects on the attitudes of students with learning disabilities and their regular class peers. Journal of Learning Disabilities, 21, 249-252.

Eiserman, W.D., Shisler, L., & Osguthorpe, R.T. (1987). Handicapped students as tutors: A description and integration of three years of research findings. B.C. Journal of Special Education, 11, 215-231.

Elliot, M. (1991). A replication and analysis of classwide reading peer tutoring effects with high socioeconomic status fourth grade students. Unpublished Dissertation, University of Kansas.

Elliot, M., Hughes, V., & Delquadri, J. (1984, May). Experimental validation and filed testing of instructional packages that increase "opportunity to respond." In J. Delquadri (Chair), Applications of behavioral instruction models: Issues influencing adoption and use in public schools. Symposium presented at the Tenth Annual Convention of the Association of Behavior Analysis, Nashville, TN.

Ellson, D.G. (1976). Tutoring. In N. Gage (Ed.), The psychology of teaching methods (pp. 130-165). Chicago: University of Chicago

Press.

Epstein, L. (1978). The effects of intraclass peer tutoring on the vocabulary development of learning disabled children. Journal of Learning Disabilities, 11, 518-521.

Feldman, R.S., Devin-Sheenhan, L., & Allen, V.L. (1976). Children tutoring children: A critical review of research. In V.L. Allen (Ed.), Children a teachers: Theory and research on tutoring, (pp. 235-252). New York: Academic Press.

Fuchs, L.S., Fuchs, D., & Hamlett, C. (1989). Monitoring reading growth using student recalls: Effects of two teacher feedback systems. Journal of Educational Research, 21, 513-516.

Gardner, E.J., Rudman, H.C., Karlsen, B. & Merwin, J.C. (1982). Stanford Achievement Test. San Antonio, TX: Psychological Corp.

Gelzheiser, L.M., & Meyers, J. (1991). Reading instruction by classroom, remedial, and resource room teachers. The Journal of Special Education, 24, 512-526.

Gerber, M., & Kauffman, J.M. (1981). Peer tutoring in academic setting. In P. Strain (Ed.), The utilization of peers as behavior change agents (pp. 155-188). New York: Plenum.

Glass, G.V. (1976). Primary, secondary, and meta-analysis of research. Educational Researcher, 5(10), 3-8.

Glass, G.V., & Hopkins, K.D. (1984). Statistical methods in education and psychology. Englewood Cliffs, NJ:

Prentice-Hall.

Glass, G.V., McGraw, B., & Smith, M.L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage.

Greenwood, C.R., Carta, J.J., Kamps, D. (1990). Teacher versus peer-mediated instruction: A review of educational advantages and disadvantages. In M. Foot, F. Morgan, & R. Shute (Eds.), Children helping children (pp. 177-205). Sussex, England: Willey.

Greenwood, C.R., Delquadri, J.C., & Hall, R.V. (1984). Opportunity to respond and student academic performance. In W.L. Heward, T.E. Heron, D.S. Hill, & J. Trap-Porter (Eds.), Focus on behavior analysis in education (pp. 58-88). Columbus, OH: Charles E. Merrill.

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

Greenwood, C.R., Delquadri, J.C., & Carta, J.J. (1988). Classwide Peer Tutoring (CWPT). Delray Beach, FL: Educational Achievement Systems.

Hall, R.B., Delquadri, J., Greenwood, C.R., & Thurston, L. (1982). The importance of opportunity to respond to children's academic success. In E. Edger, N. Haring, J. Jenkins, & C. Pious (Eds.), Mentally handicapped children: Education and training (pp. 107-140). Baltimore, MD: University Park Press.

Hall, R.V., Delquadri, J., & Harris, J. (1977, May). Opportunity to respond: A new focus in the field of applied behavior analysis. Paper presented at the Midwest Association for Applied Behavior Analysis, Chicago, IL.

Harris, A.J., & Sipay, E.R. (1985). How to increase reading ability: A guide to developmental and remedial methods. New York: Longman.

Harrison, G.V. (1982). Beginning reading: A professional guide for the lay tutor. Blanding, UT: METRA.

Haynes, M.C., & Jenkins, J.R. (1986). Reading instruction in special education resource rooms. American Educational Research Journal, 23, 161-190.

Hedges, L.V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. Journal of Educational Statistics, 6, 107-128.

Hedges, L.V. (1982). Estimating Effect size from a series of independent experiments. Psychological Bulletin, 92, 490-499.

Hedges, L.V., & Olkin, I. (1983). Clustering estimates of effect magnitude from independent studies. Psychological Bulletin, 93, 563-573.

Jenkins, J.R., & Jenkins, L. (1985). Peer tutoring in elementary and secondary programs. Focus on Exceptional Children, 17, 1-12.

Jenkins, J.R., Heliotis, J.D., Haynes, M.C., & Beck, K. (1987).

Does passive learning account for disabled readers' comprehension deficits in ordinary reading situations? Learning Disability Quarterly, 9, 69-75.

Jenkins, J.R., Jewell, M., & Leceister, N. (in press). Development of a school building model for educating handicapped and at-risk students in general education classrooms. Journal of Learning Disabilities.

Jenkins, J.R., Mayhall, U.R., Peschka, C.M., & Jenkins, L. (1974). Comparing small group and tutorial instruction in resource rooms. Exceptional Children, 40, 245-250.

Juel, C. (1988, April). Learning to read and write: A Longitudinal study of fifty-four children from first through fourth grade. Paper presented at the American Educational Research Association annual meeting, New Orleans, LA.

Kalfus, G.R. (1984). Peer mediated intervention: A critical review. Child and Family Therapy, 6, 17-43.

Karlsen, B. Madden, R., & Gardner, E.F. (1977). Stanford Diagnostic Reading Test. New York: Harcourt Brace Jovanovich.

Lamport, K.C. (1982). The effect of inverse tutoring on reading disabled students in public school settings. Dissertation Abstracts International, 44, 729. (University Microfilms No. 83-15, 707)

Lazerson, D.B. (1980). "I must be good if I can teach!" - Peer tutoring with aggressive and withdrawn children. Journal of Learning Disabilities, 13, 43-48.

Leinhardt, G., Zigmond, N., & Cooley, W.W. (1981). Reading instruction and its effects. American Educational research Journal, 18, 343-361.

Light, R.J., & Pillemer, D.B. (1984). Summing up: The science of reviewing research. Cambridge, MA: Harvard University Press.

Lombardo, V.S. (1975). The effectiveness of retarded and nonretarded tutors for educable mentally retarded children. Dissertation Abstracts International, 37, 2113. (University Microfilm No. 76-21, 957)

Lue, M.S. (1981). The use of instructional arrangements and their relationship to on-task behaviors with learning disabled, emotionally handicapped, and regular class children. Dissertation Abstracts International, 42, 2614. (University Microfilm No. 81-27, 448)

MacGinitie, W.H. (1978). Gates-MacGinitie Reading Tests - Teacher's Manual. Boston: Houghton Mifflin.

Maheady, L., Harper, G.F., & Sacca, M.K. (1988). Peer-mediated instruction: A promising approach to meeting the diverse needs of LD adolescents. Learning Disabilities Quarterly, 11, 108-113.

Maher, C.A. (1982). Behavioral effects of using conduct problem adolescents as cross-age tutors. Psychology in the Schools, 19, 360-364.

Maher, C.A. (1986). Handicapped Adolescents as cross-age tutors: Program description and evaluation. Exceptional Children,

51, 56-63.

McCracken, S.J. (1979). The effect of a peer tutoring program utilizing data-based instruction on the word recognition and reading comprehension skills of secondary age level handicapped students. Dissertation Abstracts International, 40(08-A), 4516. (University Microfilms No. 79-24, 399)

Melberg, D.B. (1981). The effects of the handicapped and nonhandicapped tutor on the academic achievement of the economically disadvantaged adolescent tutor and the elementary age tutee. Dissertation Abstracts International, 42(02-A), 659. (University Microfilms No. 81-11, 469)

Nagy, W.E., & Anderson, R.C. (1984). How many words are there in printed school English? Reading Research Quarterly, 19, 304-330.

National Commission of Excellence in Education. (1983). A nation at risk: The imperative for educational reform, Washington, D.C.: U.S. Government Printing Office.

Osguthorpe, R., Eiserman, W., Shisler, L., Top, B., & Scruggs, T.E. (1984). Final report (1983-1984): Handicapped children as tutors. Document submitted to the office of Special Education and Rehabilitative Services, U.S. Department of Education, Washington, D.C. (ERIC Document Reproduction Service No. ED 255 018)

Osguthorpe, R.T., Eiserman, W.D., Shisler, L., Top, B.L., & Scruggs, T.E. (1985). Handicapped children as tutors: Final

report (1984-1985). Document submitted to the Office of Special Education and Rehabilitative Services, U.S. Department of Education, Washington, D.C. (ERIC Document Reproductive Service No. ED 267 545)

Osguthorpe, R.T., & Scruggs, T.E. (1986). Special education students as tutors: A review and analysis. Remedial and Special Education, 7, 15-26.

O'Shea, L.J., Sindelar, P.T, & O'Shea, D.J. (1987). The effects of repeated reading and attentional cues on reading fluency and comprehension of learning disabled readers. Learning Disabilities Research, 2, 103-109.

O'Sullivan, P.J., Ysseldyke, J.E., Christensen, S.L, & Thurlow, M.L. (1990). Mildly handicapped elementary students' opportunity to learn during reading instruction in mainstream and special education settings. Reading Research Quarterly, 25, 131-146.

Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock (Ed.), The handbook of research on teaching, (pp. 376-391). New York: MacMillan.

Samuels, S.J. (1987). Information processing abilities and reading. Journal of Learning Disabilities. 20, 18-22.

Samuels, S.J., Dahl, P., & Archwatemy, T. (1974). Effects of hypothesis/test training on reading skill. Journal of Educational Psychology, 66, 835-844.

Scruggs, T.E., Mastropieri, M.A., & Richter, L. (1985). Peer

tutoring with behaviorally disordered students: Social and academic benefits. Behavior Disorders, 10, 283-294.

Scruggs, T.E., & Osguthorpe, R.T. (1986). Tutoring interventions within special education settings: A comparison of cross-age and peer tutoring. Psychology in the Schools, 23, 148-154.

Scruggs, T.E., & Richter, L. (1985). Tutoring learning disabled students: A critical review. Learning Disability Quarterly, 8, 286-298.

Shisler, L., Top, B.L., & Osguthorpe, R.T. (1986). Behaviorally disordered students as reverse-role tutors: Increasing social group acceptance and reading skills. B.C. Journal of Special Education, 10, 101-119.

Simmons, D.C., Fuchs, D., Fuchs, L.S., Pate, J., & Mathes, P.G. (1991, April). Importance of Instructional complexity and role reciprocity to classwide peer tutoring. Paper presented at the American Educational Research Association Annual Meeting, Chicago, IL.

Simmons, D.C., Fuchs, L.S., Fuchs, D., Mathes, P.G., & Pate, J. (1990). The effects of explicit teaching and peer-mediated instruction on the reading achievement of learning disabled and low-performing students. Manuscript submitted for publication.

Sindelar, P.T. (1982). The effects of cross-age tutoring on the comprehension skills of remedial reading students. The Journal of Special Education, 16, 199-206.

Slavin, R.E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. Educational Researcher, 15, 5-11.

Slavin, R.E. (1984). Meta-Analysis in education: How has it been used? Educational Researcher, 13, 6-15, 24-27.

Slavin, R.E. (1980). Effects of student teams and peer tutoring on academic achievement and time on-task. Journal of Experimental Education, 48, 252-257.

Slosson, R.L. (1963). Slosson oral reading test. East Aurora, NY: Slosson Educational Publications.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. Reading Research Quarterly, 21, 360-406.

Stevens. R.J., Madden, N.A., Slavin, R.E., & Farnish, A.M. (1987). Cooperative integrated reading and composition: Two field experiments. Reading Research Quarterly, 22, 433-454.

Strother, M.E. (1984). Effects of automaticity training strategies on word recognition, reading speed, and comprehension. Dissertation Abstracts International. 45, 2468. (University Microfilm No. 84-24, 659)

Russell, T., & Ford, D.F. (1983). Effectiveness of peer tutors vs. resource teachers. Psychology in the Schools, 20, 436-441.

Top, B.L. (1984). Handicapped children as tutors: The effects of cross-age, reverse-role tutoring on self-esteem and reading

achievement. Dissertation Abstracts International, 46, 886. (University Microfilms No. 85-05, 580)

Top, B.L., & Osguthorpe, R.T. (1985). The effects of reverse role tutoring on reading achievement and self-concept. In Osguthorpe, R.T., Eiserman, W.D., Shisler, L., Top, B.L., & Scruggs, T.E. Handicapped children as Tutors: Final report (1984-1985). Document submitted to the Office of Special Education and Rehabilitative Services, U.S. Department of Education, Washington, D.C. (ERIC Document Reproductive Service No. ED 267 545)

Top, B.L., & Osguthorpe, R.T. (1987). Reverse-role tutoring: The effects of handicapped students tutoring regular class students. Elementary School Journal 87, 413-423.

Topping, K. (1988). The peer tutoring handbook: Promoting co-operative learning. Cambridge, MA: Brookline Press.

Will, M. (1986). Educating children with learning problems: A shared responsibility. Exceptional Children, 52, 411-415.

Willis, J.W., Morris, B., & Crowder, J. (1972). A Remedial reading technique for disabled readers that employs students as behavioral engineers. Psychology in the Schools, 9, 67-70.

Woodcock, R.W., & Johnson, M.B. (1977). Woodcock-Johnson Psycho-Edcuational Battery. MacAllen, TX: DLM Teaching Resources.

Appendix A

Appendix A

Studies Which Did Not Meet Inclusion Criteria

Study	Reason(s) Excluded
Azcoitia (1983)	a) Nested design with only 1 classroom in each cell. b) Data for students with handicaps could not be analyzed separately.
Brown (1971)	Data for students with handicaps could not be analyzed separately.
Caspo (1976)	No control group
Elliot (1990)	a) Data for students with handicaps could not be analyzed separately. b) Utilized single subject methodology as primary data source. c) dependent measures directly linked to the treatment.
Eiserman (1988)	No control group for handicapped population.
Epstein (1978)	a) No evidence of initial equivalence or adjustment for non-equivalence of groups at pretest. b) Dependent measures directly linked to treatment.
Greenwood et al., (1990)	Data for students with handicaps could not be analyzed separately.
Jenkins et al. (1974)	a) Study duration less than six weeks. b) Utilized single subject design methodology.
Jenkins et al. (in press)	a) Nested design with only 1 school in each cell. b) Schools were not randomly selected. c) Multiple interventions in reading occurred simultaneously.
Lazerson (1980)	a) Dependent measures used to assess reading growth not specified. Unable to determine if dependent measures were linked to the treatment. b) Treatment implemented less than 6 weeks.

	weeks.
Lombardo (1975)	Posttest only design. Unable to determine initial equivalence of groups.
Lue (1981)	a) Dependent measures linked to treatment. b) Dependent measures not academic.
Maher (1982)	a) Dependent measures directly linked to tutoring treatment. b) Unable to determine if groups were initially equivalent and adjustments are presented
Maher (1986)	a) Dependent measures directly linked to tutoring treatment. b) No control group.
Melberg (1981)	a) Data for students with disabilities could not be analyzed separately. b) Only 1 subject had a labeled disability.
Shisler et al. (1986)	Nested design with only one teacher in each group.
Slavin (1980)	Data for students with disabilities could not be analyzed separately.
Strother (1984)	Data for students with disabilities could not be analyzed separately.
Willis et al. (1972)	No control group.

Appendix B

Form used to code studies.

Citation: _____

Keep
Yes No

Setting: Mainstream Special Education

School Based Study? Yes No

Experimental Design:

Teacher N • Each group *(Specify by group)*

Reading: Yes No

Type:

Experimental groups › 1

N •

Handicapped N •

Type(s) of Handicap & N →

Yes No →

Treatments Compared:

Control Group Yes No

Random Assignment of controls Yes No

Type of Tutoring

Classwide Individuals

Cross-age Same age

Expert Reciprocal

Handicap Role

Tutor Tutee

Both

Structure:

Structured

Nonstructured

Random Assignment of experimental Yes No

How many types of tutoring compared? *(Specify)*

Evidence of initial equivalence: Yes No →

If no - how controlled?

Study Duration *(Weeks)*

Times per week

Time per day

Reinforcement: None Competitive Cooperative Individualistic

Dependent Measures:

Was treatment keyed to measures? Yes No

Findings: Significance: p

Effect Size

Summarize:

Table 1

Coding Scheme and Intercoder Reliability Based on 10 Studies

Characteristic	Intercoder Reliability
School-Based Study	100%
Yes	
No	
Experimental Design of Study	100%
Pre/Post Randomly Assigned	
Pre/Post Matched Groups	
Pre/Post Nested - Randomly Assigned	
Pre/Post Nested - Nonrandom	
posttest only	
no comparison group	
Number of Teachers in Each Comparison Group	100%
Presence of Control Group	100%
Yes	
No	
Random Assignment of Groups	
Control	100%
Experimental	90%
Type of Subject Disability	90%
LD	
BD	
MR	
Combination	
Number of Subjects with disabilities	100%
Type of Reading Taught During Tutoring	100%
Phonics	
Sight word	
fluency	
Comprehension	
Multiple Focus (Comprehension & Decoding)	
(This category was collapsed into decoding or multiple focus)	

Tutoring Setting	80%
mainstream	
special education	
(If both tutee and tutor were disabled location was coded special education)	
Type of Tutoring	
Classwide or Individual Pairs	100%
Cross-age or peer	90%
Expert or Reciprocal	100%
Role of the Students with Disabilities	100%
Tutor	
Tutee	
Reciprocal	
Structure of Treatment	100%
Structured	
Unstructured	
Type of Reinforcement System	100%
None Discussed	
Competitive	
Cooperative	
Individualistic	
Study Duration	100%
Time Per Week Tutoring Occurred	100%
Daily Time of Tutoring Sessions	100%
Proximity of Dependent Measures to Treatment	
Directly Linked	100%
Close match or Dissimilar	86%
Over All Agreement	97.8%

Table 2

Average Effect Sizes for Specific Study Characteristics

Characteristic	Average Unbiased Effect Size
Overall	+.36
- Compared to another treatment	.14
- Compared to teacher-led control	+.40
Setting	
- Special Education Classroom	+.27
- General Education Classroom	+.42
Role of Student with Disability	
- Tutee	+.30
- Tutor	+.42
- Reciprocal	+.34
Tutor or Reciprocal Role and Setting	
- Special Education	+.27
- General Education	+.45
Age Arrangement	
- Cross-age	+.38
- Peer	+.32
Classroom Arrangement	
- Classwide	+.35
- individual	+.36
Reading Activity	
- Decoding	+.35
- Multiple Focus	+.37
Tutor Academic Level	
- Expert	+.36
- Similar to Tutee	+.34

+ Overall Effect Size significantly different from zero.

Table 3
Characteristics of Studies

Article	Grades	Sp.Ed. n	Label	Setting	Duration (Weeks)	Design	Reading Activity	Tutoring Type	Role	Effect Size
Carlton et al. 1985	5-6	136	EMR	Sp.Ed Cl s	6	Teachers randomly assigned to treatments. No-treatment control.	Sight Words	Classwide Peer Expert Model	Tutor or tutee - Static	+.38
Lamport 1982	6	24	' D	Sp.Ed. Class	8	Subjects in matched groups based on reading ability. No-treatment control.	Multiple Focus	Individual Cross-age Expert Model	Tutor or Tutee - static	+.44
McCracken 1979	9-12	84	LD (5 • EMR)	Sp.Ed. Class	12 Approx.	Classes randomly assigned to treatment groups. Peer tutoring compared to teacher, one-on-one tutoring.	Decoding	Classwide Peer Expert model	Tutee	+.08
Russell & Ford 1983	7	32	EMR	Sp.Ed. Class	15 Approx.	Subjects randomly assigned to treatments. Same teachers in both groups.	Multiple Focus	Individual Cross-age Expert Model	Tutee	+.75
Souggs & Osguthorpe 1986 Exp #1	1-5	47	LD & BD	Sp.Ed. Class	10	Subjects randomly assigned to groups after being identified by teachers as possible tutors. No-treatment control.	Decoding	Individual Cross-age Expert Model	Tutor	.07
Scruggs & Osguthorpe 1986 #2	2-5	30	LD & BD	Sp.Ed. Class	8	Subjects in matched groups. Tutoring pairs determined by teachers before matching occurred. No-treatment control.	Decoding	Individual Peer Reciprocal	Both Tutor & Tutee	.23
Simmons et al. 1990	2-5	44	LD	General Class	8	Teachers randomly assigned to experimental groups. Teacher volunteered for control condition. No-treatment control.	Multiple Focus	Individual Peer Expert Model	Tutee	+.35 • vs ET .24 vs Control • +.45

Article	Grades	Sp.Ed. n	Label	Setting	Duration (Weeks)	Design	Reading Activity	Tutoring Type	Role	Effect Size
Simmons et al. 1991	2-5	58	LD	General Class	14	Teachers randomly assigned to 4 peer-tutoring groups. Teachers volunteered for control condition. No-treatment control.	Multiple Focus	Classwide Peer Reciprocal	Both	Total • +.38 Version 1 • +.76 Version 2 • -.07 Version 3 • +.65 Version 4 • .25
Sindelar 1982	3	53	LD	Sp.Ed. Class	6 Approx.	Teachers randomly assigned to 3 peer-tutoring groups and 1 teacher-led treatment group. Teacher group used same methodology as 1 peer-tutoring group	Varied by Condition	Individual Cross-age Expert Model	Tutee	•• .07
Top & Osguthorpe 1985	4-6	64	BD & LD	General Class	12	Teachers randomly assigned to each condition. No-treatment control.	Decoding	Individual Cross-age Expert Model	Tutor	+.63
Top & Osguthorpe 1987	4-6	78	BD & LD	LD in General Class. BD in Sp.Ed Class.	14	Teachers randomly assigned to each condition No-treatment control.	Decoding	Individual Cross-age Expert Model	Tutor	+.48

• Significantly different from zero ($p < .05$)

• Effect size based on comparison to a teacher-led treatment group.

•• Effect size based on reading rate measures only. Data not available to determine other effect sizes.

Table 4

Unbiased Effects Sizes for Individual Dependent Measures and Formulas Employed

Study	Dependent Measure	Formula	Unbiased Effect Size[a]
Carlton et al. 1985	Gates-MacGinitie		
	- Vocabulary - Tutee	Final Status	.38
	- Comprehension - Tutee	↓	.42
	- Vocabulary - Tutor	Gain Score	.33
	- Comprehension - Tutor	↓	.38
Lamport 1982	Stanford Diagnostic Reading Test		
	- Auditory Vocabulary - Tutor	F Statistic	.14
	- Comprehension - Tutor	↓	.22
	- Phonic Analysis - Tutor	ANCOVA	1.04
	- Auditory Vocabulary - Tutee	↓	.69
	- Comprehension - Tutee	F Statistic	.29
	- Phonic Analysis- Tutee	↓	.33
McCracken 1979	Slosson Oral Reading Test		
	- Word Recognition	t statistic	.11
	- Comprehension	↓	.05
Russell & Ford 1983	Peabody Individual Achievement Test		
	- Reading Recognition	Final Status	.52
	- Comprehension	↓	1.00
Scruggs & Osguthorpe 1986 #1	Harrison Criterion Decoding Test		
	- Decoding - Tutee	Gain Score	.25
	- Sight Word -Tutee	↓	.12
	- Overall Word Recognition - Tutee		.30
	- Decoding - Tutor		-.04
	- Sight Word -Tutor		-.16
	- Overall Word Recognition - Tutor		.00
	Woodcock Johnson Psycho-educational Battery		
	- Word Recognition - Tutee		.14
	- Word Attack - Tutee		.20
	- Passage Comprehension - Tutee		.15
	- Word Recognition - Tutor		-.11
	- Word Attack - Tutor		.22
	- Passage Comprehension - Tutor		.29

Study	Dependent Measure	Formula	Unbiased Effect Size
Scruggs & Osguthorpe 1986 #2	Harrison Criterion Decoding Test		
	- Decoding	Gains Score	.52
	- Sight Word	↓	.37
	- Overall Decoding		.51
	Woodcock-Johnson Psycho-educational Battery		
	- Word Recognition		-.01
	- Word Attack		.18
	- Comprehension		-.16
Simmons et al. in review	Comprehensive Reading Assessment Battery (Compared to Control)		
	- Words Read Correctly	ANCOVA	.35
	- Comprehension	↓	.44
	- Maze Correct		.77
	- Retell Matches		.52
	(Compared to Teacher Effectiveness Treatment)		
	- Words Read Correctly		.30
	- Comprehension		.28
	- Maze Correct		.25
	- Retell Matches		-.02
	Stanford Reading Achievement		
	- Compared to Control		.20
	- Compared to Teacher Effectiveness Treatment		.37
Simmons et al. 1991	Comprehension Reading Assessment Battery		
	Version 1 (Kansas Model - Reciprocal)		
	- Words Read Correctly	ANCOVA	.79
	- Comprehension	↓	.66
	- Maze Correct		.53
	- Retell Matches		1.08
	Version 2 (Kansas Model - Static Role)		
	- Words Read Correctly		-.14
	- Comprehension		-.21
	- Maze Correct		-.25
	- Retell Matches		.31
	Version 3 (Peabody Model - Reciprocal)		
	- Words Read Correctly		.41
	- Comprehension		.66
	- Maze Correct		.93
	- Retell Matches		.62
	Version 4 (Peabody Model - Static Role)		
	- Words Read Correctly		.32
	- Comprehension		.13
	- Maze Correct		.45
	- Retell Matches		.09

Study	Dependent Measure	Formula	Unbiased Effect Size
Sindelar 1982	Nonstandardized Measure of Reading Rate		
	- Word Recognition Group	ANCOVA	.13
	- Oral Reading Group	↓	-.04
	- Hypothesis/Testing Group		.13
Top & Osguthorpe (1985)	Woodcock-Johnson Psycho-educational Battery		
	BD Subjects		
	- Word Attack	F Statistic	.76
	- Passage Comprehension	↓	.87
	- Total Reading		.70
	LD Subjects		
	- Word Attack		.61
	- Passage Comprehension		.52
	- Total Reading		.98
	All Subjects		
	- Word Recognition		.32
	- Word Attack		.68
	- Passage Comprehension		.62
Top & Osguthorpe 1987	Woodcock-Johnson Psycho-educational Battery		
	- Word Recognition	F Statistic	.03
	- Word Attack	↓	.95
	- Passage Comprehension		.41
	- Total Reading		.57

[a] Effect size is positive unless otherwise denoted.