
A Meta-Analysis of Co-Teaching Research

Where Are the Data?

WENDY WEICHEL MURAWSKI AND H. LEE SWANSON

ABSTRACT

Students with special needs are increasingly being served in the general education classroom. Co-teaching is one service delivery option designed to meet those needs. The purpose of this article is to synthesize data-based articles pertaining to co-teaching between general and special education personnel. Of 89 articles reviewed, only 6 provided sufficient quantitative information for an effect size to be calculated. Effect sizes for the individual studies ranged from low (0.24) to high (0.95), with an average total effect size of 0.40. Dependent measures were varied and included grades, achievement scores, and social and attitudinal outcomes. Results indicate that further research is needed to substantiate that co-teaching is an effective service delivery option for students with disabilities.

AS PROFESSIONALS WORKING WITHIN THE field of special education search for increasingly effective ways to meet the needs of students with mild and moderate disabilities, service delivery options within the general education classroom have become more and more necessary. Although the Education for All Handicapped Children Act of 1975 mandated that students be served in the least restrictive environment (LRE), the extent to which LRE connoted the general education classroom remained a relatively controversial issue (Bauer & Shea, 1999). The Individuals with Disabilities Education Act (IDEA) of 1990, and, more specifically, the amendments to IDEA in 1997, emphasize the need to serve students with disabilities in the general education setting whenever possible. This new emphasis was included based on the principle that students are best served in settings

most like those of their nondisabled peers (Vaughn, Bos, & Schumm, 2000), a notion that has also come to be known as *inclusion*. For inclusion to be possible, students must be provided with services and supports within the general education environment. One such service delivery option for students is through the use of co-teaching between general and special education teachers.

Bauwens, Hourcade, and Friend (1989) described a pragmatic merger between general and special educators in which direct educational programming to all students would be provided by having a special educator within a general education setting. They coined the term *cooperative teaching* to represent this relationship. They also presented ways to implement cooperative teaching at that time and included complementary instruction, team teaching, and supportive learning activities. Based on the philosophy and politics of inclusion—in essence the desire to meet the needs of all students in the LRE—co-teaching makes intuitive sense.

Cook and Friend (1995) shortened the term *cooperative teaching* to *co-teaching* and further clarified the characteristics inherent in a true co-teaching relationship. They defined co-teaching as “two or more professionals delivering substantive instruction to a diverse or blended group of students in a single physical space” (p. 2). Since Bauwens, Hourcade, and Friend’s seminal article in 1989, the current educational literature has been replete with anecdotal experiences as well as suggestions for implementation and guidelines for setting up co-teaching situations (Cook & Friend, 1995; Reinhiller, 1996; Sevakis & Harris, 1992). Although there are a variety of co-teaching options (e.g., one teaching, one assisting; station teaching; parallel teaching; alternative teaching; team

teaching; see Note 1; Cook & Friend, 1995), the impact of such procedures on student outcomes is unclear. Reinhiller (1996) provided a descriptive summary of studies done in co-teaching, but to the authors' knowledge, there has not been a quantitative analysis of research in this area.

The purpose of this article is to provide a synthesis of the quantitative data on the effectiveness of co-teaching, using a meta-analytic procedure. Meta-analysis is a statistical reviewing technique that provides a quantitative summary of findings across an entire body of research. The results of individualized studies are converted to a standardized metric or effect size. The scores are then aggregated across the sample of studies to yield an overall estimate of effect size. Particular attention is given to the magnitude of the effect size estimate. According to Cohen (1988), 0.80 is considered a large effect estimate, 0.50 a moderate estimate, and 0.20 a small estimate. Thus, we hope to quantify the co-teaching literature in terms of the magnitude of treatment outcomes, as well as address two specific questions:

1. Does the magnitude of co-teaching outcomes vary as a function of grade, gender, length of study, or severity or type of disability?
2. Do studies that produce the largest effect size vary from other studies as a function of the type of dependent measure of focus (e.g., grades, social outcomes, achievement)?

METHOD

For this meta-analysis, a comprehensive literature search was conducted, using three methods. First, ERIC, PsychLit, and EdInfo databases were searched for all pertinent articles related to co-teaching between general and special educators. Descriptors included co-teaching, collaborative teaching, cooperative teaching, team teaching, mainstreaming, inclusion, pull-in, teaming, and supportive learning. Because co-teaching is a delivery model relatively recent to the field of special education (Bauwens et al., 1989), the computerized search was limited to the 10 years from 1989 to 1999. Although students were certainly included in various degrees prior to 1989, the past decade, in which the terms *cooperative teaching* and *co-teaching* have been used to specify a specific service delivery model, has been most appropriate for study inclusion. All articles that identified some form of co-teaching in the title or abstract, as well as those that indicated activities or actions that might be considered co-teaching between special and general educators, were analyzed.

Second, a hand search was done on all articles cited in review articles (e.g., Bauwens et al., 1989; Cook & Friend, 1995; Reinhiller, 1996; Vaughn, Elbaum, Schumm, & Hughes, 1998). This is sometimes referred to as "footnote chasing." All journal articles or ERIC documents that focused

on co-teaching were analyzed. No restrictions were placed on dates for inclusion or type of publication.

Finally, the past 10 years of *Exceptional Children*, *Teacher Education and Special Education*, and *Remedial and Special Education* were hand searched.

Using these three methods to research the literature, 89 articles were identified that mentioned a form of service delivery that included both a general and a special education teacher in the same classroom. Because the purpose of this study was to conduct a quantitative meta-analysis of co-teaching, position papers, program descriptions, and other articles on co-teaching that lacked data were eliminated from the analysis. This left 37 articles.

The 37 remaining articles were analyzed based on the following criteria for possible inclusion in the meta-analysis:

1. *The study included sufficient quantitative data that would enable the researcher to calculate effect sizes for the intervention.* For this to be possible, either the mean(s) or standard deviation(s) of the dependent measures, or tests of significance between the experimental and control groups (e.g., *F* tests, *t* tests, or *p* values) needed to be present in the article.

2. *The study included four characteristics that identify the intervention as a form of co-teaching* (see Note 2). The four characteristics were (a) general education teachers and special service providers (e.g., special education teachers, speech-language specialists, school psychologists, Title I teachers) were working together; (b) the intervention was occurring in the same physical space (e.g., the general education classroom); (c) an element of co-planning was included (so that the special educator was not relegated to the role of assistant); and (d) the intervention involved delivering instruction to a heterogeneous group of students, with and without disabilities (see Note 3).

3. *The co-teaching treatment condition lasted for more than a 2-week period, not including pretesting and post-testing.* Because of the nature of co-teaching and the necessity for educators to co-plan, collaborate, and co-instruct via open communication, we assumed that studies with less than 2 weeks' intervention were insufficient to assess treatment effects. All studies in this synthesis except one (Rosman, 1994) were conducted over the course of 1 academic year. Rosman's study took place over the course of 3 weeks, the length of time it took to teach one chapter of math in the designated general education math class.

Of the original 37 studies, 15 were eliminated due to a lack of quantitative data. Ten articles were eliminated because the quantitative information given was not sufficient to calculate an effect size (e.g., only percentages of survey results were reported). Finally, 6 studies were removed from the final analysis because they did not describe the requisite characteristics to denote the use of co-teaching (see Note 4). Therefore, based on these three criteria, only 6 of the 37 studies analyzed were retained for potential inclusion in the meta-analysis.

Codes

Each of the six studies included in the analysis was coded for the following information: study characteristics, sample characteristics, outcome measures on the dependent variables, and effect sizes.

Study Characteristics. Each study was coded by number of authors, date of publication, country/state of study, funding source reported, and length of study. The assessment by researchers of the study's treatment integrity—defined by Gresham (1989) as whether the intervention was implemented as intended—was also noted.

Sample Characteristics. Within each study, sample characteristics were coded. These included age, grade, gender, ethnicity, socioeconomic status (SES), disabilities included in the study, and the setting of the intervention. Total sample size, as well as a breakdown of students in special and general education when reported, was recorded, as was the number of treatment agents, as determined by the general education teachers and special service providers (e.g., special education teachers) involved in the study.

Outcome Measures. All dependent measures in which effect size could be calculated were coded. Dependent measures varied. Some studies focused on the outcome in terms of grades; others focused on attitudes, minimum competency results, achievement scores, and/or social outcomes.

Effect Sizes. Effect sizes were calculated on each dependent measure using Cohen's *d* for two studies in which means and standard deviations were present. For these studies, the difference in outcome for the experimental group (co-teaching) and control group were calculated. This difference was calculated by subtracting the mean of the control group's posttest score from the mean of the experimental group's posttest score. The difference obtained was then divided by the pooled standard deviation of the two groups. The equation for obtaining effect sizes using Cohen's *d* is as follows:

$$d = \frac{M_E - M_C}{SD_{\text{pooled}}}$$

In this equation:

M_E = mean posttest score on the dependent measure for the experimental group

M_C = mean posttest score on the dependent measure for the control group

SD_{pooled} = pooled standard deviation for the experimental and control groups

The formula for the pooled standard deviation is:

$$SD_{\text{pooled}}^2 = \frac{(n_e - 1) SD_E^2 + (n_c - 1) SD_C^2}{n_e + n_c - 2}$$

In this equation:

n_e = number of participants in the experimental group

n_c = number of participants in the control group

SD_E = standard deviation of the posttest score on the dependent measure for the experimental group

SD_C = standard deviation of the posttest score on the dependent measure for the control group

Not all effect sizes in this study were calculated using Cohen's *d* formula. Two studies (Lundeen & Lundeen, 1993; Self, Benning, Marston, & Magnusson, 1991) did not include a control group and utilized a pretest/posttest research design on the treatment group only. In these situations, the effect sizes were calculated using *t* test scores provided. Thus, *d* was calculated as follows:

$$d = t \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

In this equation:

d = effect size

t = values obtained through a *t* test

n_1 = number of participants in the control or nontreatment condition

n_2 = number of participants in experimental or treatment condition with co-teaching.

The Rosman (1994) study provided only *F* values. The formula for finding an effect size using an *F* value is:

$$d = \sqrt{F} \sqrt{\frac{n_e + n_c}{n_e n_c}}$$

In this equation:

d = effect size

F = *F* value obtained by comparing two variance estimates

n_e = number of participants in the experimental group

n_c = number of participants in the control group.

For one article in which only p values were available (Walsh & Snyder, 1993), those p values were used to estimate the lower limit of the effect sizes. For p values reaching a significance level of 0.05, r was calculated by using the formula $r = \frac{1.64}{\sqrt{n}}$. Effect size (d) is then obtained by using $d = \frac{2r}{\sqrt{1-r^2}}$.

For p values reaching a significance level of 0.01, r was calculated by using the formula $r = \frac{2.33}{\sqrt{n}}$. Effect size (d) is also

obtained by using $d = \frac{2r}{\sqrt{1-r^2}}$.

Effect sizes provide a standardized score for which the magnitude of a difference can be determined. Effect sizes, often denoted as ES, or Δ , are the z scores between the mean of the experimental group and the mean of the control group, expressed in standard deviation units (Glass & Hopkins, 1996).

RESULTS

Study Characteristics

Table 1 provides an overview of the characteristics of each study included in the analysis. Three of the articles are from journals and three are ERIC documents. Publication dates range from 1991 to 1998, and the number of authors range from one to five. Two of the studies (Klingner, Vaughn, Hughes, Schumm, & Elbaum, 1998; Vaughn et al., 1998) were conducted in the same state (Florida) and share com-

mon authors. One study cited a funding source (Klingner et al., 1998). As previously reported, all studies but one took place over the course of 1 academic year. None of the six studies reported measures of treatment integrity (i.e., directly analyzing whether the intervention was implemented as intended).

Only one study (Rosman, 1994) randomly assigned participants to instructional conditions. The other studies either neglected to report any information on sampling procedures or indicated that intact groups were used. All studies were conducted in public schools.

Sample Characteristics

The sample characteristics of each study are displayed in Table 2, listed by the primary author. Although all studies reported sample sizes, only three studies provided information on both general and special education students (Klingner et al., 1998; Lundeen & Lundeen, 1993; Vaughn et al., 1998). Reports of the teachers involved in the studies also varied. Four of the six studies reported the number of general education teachers and special service providers that participated in the study. Of those studies that did not report information, Walsh and Snyder (1993) stated that six schools participated in their study but did not give specific details regarding the number of teachers involved or the types of students with whom they primarily worked (i.e., general education versus special education). Vaughn and colleagues (1998) provided information on the seven classrooms involved from two elementary schools. Three classes implemented co-teaching with a full-time special education teacher and general education teacher in each room, and four of the classes implemented consultation/collaboration teaching. In their

TABLE 1. Study Characteristics

Primary author	Journal	No. of authors	Date of pub.	State of study	Funding source	Length of study	Dependent measures
Self	<i>Exceptional Children</i>	4	1991	Minnesota	None cited	1 acad. year	Attitudes, referrals, achievement (reading)
Lundeen	ERIC document	2	1993	West Virginia	None cited	1 acad. year	Grades
Walsh	ERIC document	2	1993	Maryland	None cited	1 acad. year	Grades, referrals, absences, min. competency
Rosman	ERIC document	1	1994	South Dakota	None cited	3 weeks	Grades, attitudes, achievement
Vaughn	<i>Journal of Learning Disabilities</i>	4	1998	Florida	None cited	1 acad. year	Friendships, social skills, self-concept, peer acceptance
Klingner	<i>Learning Disabilities Research & Practice</i>	5	1998	Florida	U.S. Dept. of Education	1 acad. year	Academic achievement (reading & math)

depiction of consultation/collaboration teaching, each class included a general education teacher, a part-time teaching assistant for 4 hours per day, and a special education teacher for 1 to 2 hours per day. Unfortunately, we were unable to determine whether the same special education teacher worked with more than one general education teacher.

The articles included in this analysis covered kindergarten through 3rd grade (Self et al., 1991), 3rd through 6th grades (Klingner et al., 1998; Vaughn et al., 1998), and high school (9th through 12th grades; Lundeen & Lundeen, 1993; Rosman, 1994; Walsh & Snyder, 1993). No studies included students from grades typically considered middle school or junior high (i.e., 7th and 8th grades). All six studies cited the general education classroom as the setting of the co-teaching intervention.

Gender breakdown was reported in only three of the six studies (Klingner et al., 1998; Rosman, 1994; Vaughn et al., 1998). These studies indicated that 49.4% of students included in the sample were girls.

Four of the six studies reported information on the types of disabilities of students involved. Students with learning disabilities (LD) and students with low achievement (students who were in the lowest 25th percentile of the class) were most frequently reported as being served. Two studies reported the effects of the intervention on students who were considered average or high achieving. No studies in this analysis included data on the inclusion of students with moderate or severe disabilities. Thus, there were not enough specific data available to further analyze the effects of the intervention by

type or severity of disability. There was also insufficient information to report a breakdown of studies by age, ethnicity, or SES.

As criteria for inclusion, all studies indicated that general and special education personnel were co-planning and co-instructing a heterogeneous group of students within the same physical space. However, studies were also coded for other recommended, but not requisite, characteristics of effective co-teaching (Friend & Cook, 2000), including parity, voluntariness, professional status, shared resources/accountability/responsibility, and the use of a variety of approaches for cooperative teaching. Three of the six studies reported that the general and special education teachers volunteered for their involvement and worked with parity in the classroom, with both instructors contributing to the planning and implementation of instruction. All six studies indicated that the persons involved in the interaction were professional educators. Four studies stated that teachers shared resources, accountability, and/or responsibility for the education of students with special needs. Two of the six studies described the ways in which they implemented co-teaching; one implemented a team-teaching model (Lundeen & Lundeen, 1993), and the other utilized the supplementary and alternative models of co-teaching (Self et al., 1991).

Dependent measures varied among the studies. For instance, dependent measures included grades, achievement, social outcomes, attitudes, absences, and referrals (see Table 3). Whereas some studies focused on one domain (e.g., social outcomes; Vaughn et al., 1998), others (e.g., Self et al.,

TABLE 2. Sample Characteristics

Primary author	Sample size (n)	No. of teachers	Grades	Gender	Disabilities included	Basic results
Self	170 total	14 GE 7 SSP	K-3	No info	LA (> 25%)	Gains in reading for co-taught students; positive teacher reports
Lundeen	134 SE 249 ND	8 GE 5 SE	9-12	No info	Majority LD; also ED, MMR, DHH, ESL	Grades for team-taught program increased 1st sem.; same overall by 2nd sem.
Walsh	343 exper. 363 control	6 schools	9	No info	No info	No difference for attendance, discipline, or grades; improved scores for minimum competency tests for co-taught classes
Rosman	59 total	2 GE 1 SE	9-12	34F, 25M	No info	Students in co-taught condition had higher math achievement scores
Vaughn	59 SE 126 ND	7 classes	3-6	88F, 97M	LD; also LA, average, & HA	More peer acceptance/friendship, increases in consult./collab. condition, than in co-teaching
Klingner	25 SE 89 ND	4 GE 2 SE	3-6	55F, 59M	LD; also LA, average, & HA	Gains for LD in reading but not math; lowest readers did not improve

Note. SE = special education; ND = nondisabled; GE = general education; SE = Special education; SSP = special services providers, including special education teachers, speech-language specialists, Title I teachers, etc.; LD = learning disabled; LA = low achieving; HA = high achieving; ED = emotionally disturbed; MMR = mild-moderately retarded; DHH = deaf/hard of hearing; ESL = English as a second language.

TABLE 3. Characteristics of Co-Teaching Intervention Studies

Domain	No. of studies (n)	No. of effect sizes	Mean effect size	Standard deviation
Dependent measures				
Grades	2	3	0.32	0.04
Math achievement	3	4	0.45	0.02
Reading/LA achievement	3	4	1.59	1.94
Social outcomes	1	8	0.08	0.32
Attitudinal outcomes	1	1	0.00	—
Absences	1	1	0.37	—
Referrals	1	1	0.43	—
Grade of participants				
K-3	1	2	0.95	0.12
3-6	2	11	0.19	0.28
9-12	3	9	0.30	0.01
Article type				
Refereed journal	3	13	0.51	0.24
ERIC document	3	8	0.30	0.10

Note. LA = language arts; K = kindergarten.

1991; Walsh & Snyder, 1993) reported the results on a variety of student outcomes.

Effect Sizes

Mean effect sizes for each dependent measure are reported in Table 3. The mean effect size reported for reading and language arts achievement is the highest (ES = 1.59) and incorporates four effect sizes from three studies. Although both absences and referrals report effect sizes (0.37 and 0.43, respectively), it is important to note that they both derive from the same study. Math achievement (0.45) is considered a moderate effect size and includes four measures from three studies. A small-to-moderate effect size (0.32) is reported for grades, which derives from three measures over two studies. Measures related to social and attitudinal outcomes report low effect sizes. Social outcomes, which included measures on peer acceptance, friendship quality, self-concept, and social skills, yield an effect size of 0.08. Although eight measures were included in the analysis of social outcomes, those measures were obtained for one study (Vaughn et al., 1998) and represented only scores obtained from students with disabilities. Attitude toward math was not found to be statistically significant by Rosman (1994), with a mean effect size of 0.00.

For co-teaching to be considered a viable service delivery option for students with disabilities within general education, statistical information disaggregating the effects of students served in co-teaching situations from those served in control situations (usually on a consulting or pull-out basis) is critical. Unfortunately, not all of the studies included in this analysis offered such data. Lundeen and Lundeen (1993) pro-

vided pre- and posttest data on co-teaching but did not provide an *F* value differentiating the effects on students with disabilities from those without identified disabilities. Walsh and Snyder (1993) and Self et al. (1991) also neglected to differentiate results of students in general education from those of students in special education (or with identified disabilities). Although the Rosman (1994) study provided pre- and posttest data on students with disabilities, no comparison data on general education students were provided. Two studies (Klingner et al., 1998; Vaughn et al., 1998) report statistics for both pre- and posttest situations, as well as comparison data for students with and without disabilities.

In Klingner et al.'s (1998) article, scores for students with LD were compared with those for students with low to average ability (LAA) and those for students with high ability (HA) for both fall (pretest) and spring (posttest) semesters on the academic outcomes of reading and math in a co-teaching situation. In the area of reading (as measured by the Basic Academic Skills Samples-Reading [BASS]; Espin, Deno, Maruyama, & Cohen, 1989), Klingner et al. (1998) reported the effect sizes of students with LD, LAA, and HA (0.78, 0.58, and 1.08, respectively), all of which were statistically significant. For math computation and math application, effect sizes for students with LD were considerably lower (0.47 and 0.26) than those of students with LAA (0.73 and 0.76) and with HA (0.82 and 0.79).

To determine the magnitude of effect size by grade, studies were combined and effect sizes were averaged based on students' grade levels. Only one study provided information on students in kindergarten through third grade (Self et al., 1991). This study yielded a large effect size (0.95). Klingner et al. (1998) and Vaughn et al. (1998) reported

information on students in Grades 3 through 6. Eleven measures were provided for these two studies, and a total effect size of 0.19 indicates a low effect for the elementary students. Finally, three studies (Lundeen & Lundeen, 1993; Rosman, 1994; Walsh & Snyder, 1993) were analyzed for their effects on students at the high school level (Grades 9–12). The total effect size on the nine measures was low to moderate (0.30). No studies included in this analysis reported information for students in Grades 7 and 8.

Table 4 reflects aggregated effect sizes as a function of each study. The number of dependent measures per study, as well as the means and standard deviations calculated for each, are reported. Only one study obtained high (> 0.80) effect sizes. Self et al. (1991) received a high overall effect size for their study (0.95). Two studies obtained average effect sizes; the Klingner et al. (1998) study yielded a mean effect size of 0.50, and the Walsh and Snyder (1993) study yielded an effect size of 0.41. The three remaining studies showed low effect sizes overall: Lundeen and Lundeen (1993) with 0.25, Rosman (1994) with 0.24, and Vaughn and colleagues (1998) with 0.08. The total effect size for all studies averaged in this analysis resulted in an average effect size of 0.40.

DISCUSSION

This article provides a quantitative synthesis of the intervention research on co-teaching between general and special education professionals. The overall mean effect size was 0.40, suggesting that co-teaching is a moderately effective procedure for influencing student outcomes. However, this

finding should be interpreted cautiously because only three studies included effect sizes related to students with reported disabilities. The results obtained spanned a gamut of dependent measures and a range of student grade levels. More importantly, although numerous authors currently espouse co-teaching as an effective alternative to service delivery for students with disabilities within the general education setting, very few provide experimental data. Of the 89 articles reviewed on co-teaching, only 6 met the criteria set for selection in this meta-analysis. We will now address the two research questions that directed this synthesis.

Does the magnitude of co-teaching outcomes vary as a function of grade, gender, length of study, or severity or type of disability? Insufficient data were provided in these studies to further analyze the effects of co-teaching on gender, length of study, or disability type. Although all authors reported that students with disabilities were included in the interventions, few provided clear definitions of the disabilities and none reported working with students with moderate to severe disabilities. Many studies did not group the results by ability or disability type, so it is unclear whether the intervention effects reflect a trend as a function of a particular disability type (e.g., learning disability versus mild cognitive delay).

Vaughn et al. (1998) and Klingner et al. (1998) provided examples of how co-teaching data can be effectively distributed by ability levels. Although they did not further differentiate the effects on varying disability types (e.g., mild cognitive delay, emotional disturbance), their studies demonstrated clear comparisons of students with school-identified disabilities versus those without such classifications. Also, because the literature is replete with examples of how schools tend to classify students with all forms of low achievement as having LD (MacMillan, Gresham, & Bocian, 1998), this type of comparison may be sufficient to demonstrate how co-teaching will affect students placed in general education settings with the label LD, despite potential discrepancies in how research might identify these students.

Although differentiation by ability was not generally possible, the trends found by grade level provide for interesting discussion. Large effect sizes (ranging from 0.87 to 3.67) were obtained for the one study focused on students in kindergarten through third grade and one of the studies conducted with high school students. The magnitude of these effect sizes suggests that co-teaching may be an appropriate service delivery option for students in these grades. Interestingly, none of the articles synthesized here focused on students at the middle or junior high school level. This is notable given that much of the literature on teaming (another method of providing collaborative instruction) and clustering of teachers or subjects (e.g., Gable, Hendrickson, & Rogan, 1996; Gable & Manning, 1997; Howell, 1991; White & White, 1992) focuses on these grade levels.

Do studies that produce the largest effect size vary from other studies as a function of the type of dependent measure of focus (e.g., grades, social outcomes, achievement)? Unfor-

TABLE 4. Effect Size Average Within Each Study

Study	K	M	SD
Self, Benning, Marston, & Magnusson (1991)	2	0.95	0.12
Lundeen & Lundeen (1993)	2	0.25	0.21
Walsh & Snyder (1993)	4	0.41	0.00
Rosman (1994)	2	0.24	0.08
Vaughn, Elbaum, Schumm, & Hughes (1998)	8	0.08	0.33
Klingner, Vaughn, Hughes, Schumm, & Elbaum (1998)	3	0.50	0.26
Total (N = 6)	21	0.40	0.17

Note. K = number of dependent measures; M = mean of effect sizes; SD = standard deviation of effect sizes.

tunately, this question cannot be answered because of the variability in what was measured and reported in the studies. For example, achievement testing included mathematics achievement scores (Rosman, 1994), minimum competency results in math and language arts (Walsh & Snyder, 1993), curriculum-based measurements (Self et al., 1991), and standardized achievement measures in both reading and mathematics (Klingner et al., 1998). Despite this variability in measures, however, the limited data suggest that co-teaching can have a positive impact on student achievement. Achievement in reading and language arts resulted in the largest effect size overall (1.59). Achievement in mathematics and reduction of referrals both received moderate effect sizes of (0.45 and 0.43, respectively). These results indicate that there is a potential for positive results in the area of achievement using co-teaching as a service delivery option for students with special needs in a general education setting.

Positive social outcomes for students are frequently cited as a potential benefit to including students with disabilities in a general education setting (Hunt, Alwell, Farron-Davis, & Goetz, 1996; Jones & Carlier, 1995; Pugach & Wesson, 1995). However, only one study in our synthesis focused on social outcomes (Vaughn et al., 1998). Further, this study did not yield higher effect sizes for students with LD in the co-teaching experimental setting than in the control condition. Although attitudes toward co-teaching are frequently referenced in program descriptions and discussions of co-teaching (e.g., Bergren, 1997; Johnson, Test, & Algozzine, 1995; Pugach & Wesson, 1995), only one study here included a measure of attitudinal outcomes, and they were directed specifically toward math (Rosman, 1994).

Because of the meager number of studies included in our analysis, caution must be exercised in interpreting the results. Specifically, three of the articles included in this analysis were ERIC documents not published in refereed journals. Although Krathwohl (1998) recommended that a meta-analysis not confine itself to published materials because nonsignificant results are most likely to emerge in unpublished studies, refereed journals provide some insurance that a study has been conducted correctly and that it reports quality research. Studies that are not reviewed by experienced researchers and authors retain a questionable status regarding reliability and/or validity.

Of greater concern to us, however, is that none of the studies reported explicit measures of treatment integrity. Without a measure of treatment integrity, it is difficult to determine whether the studies genuinely adhered to their reported interventions as described. If, in the course of the academic year, treatment agents determine that it is easiest to have the special services provider work with the students with special needs in the back of the room as the general educator continues to work with the rest of the class, the study has been invalidated because co-teaching is no longer truly occurring. Measures of treatment integrity would ensure against such a change. Thus, without these measures, it is

impossible to determine whether the studies in question maintained the intervention as intended.

Future Research

Although these studies provide limited results that must be generalized with caution, one thing is certain: For co-teaching to be considered a valid service delivery option for students with disabilities in the general education or least restrictive placement, more experimental research must be conducted. Specifically, additional studies with experimental and control groups need to be coordinated in an effort to determine how co-teaching differs from other service delivery options, such as consultation and pull-out services or no assistance whatsoever. Future studies should analyze outcomes as a function of gender, age, grade, and subject matter. Further, if the primary purpose of co-teaching is to increase the success of students with special needs within the general education environment, the effects on those students need to be analyzed by disability type and severity level in order to optimize the way co-teaching is used and with whom. More research needs to be conducted with students with moderate to severe disabilities because inclusion involves these students as well (Downing & Bailey, 1990; Goessling, 1998; Jones & Carlier, 1995). Finally, research reporting qualitative information on co-teaching should be synthesized and reported (Weichel, 1999).

Following a review and an analysis of the studies related to co-teaching, Weiss and Brigham (2000) expounded on the six basic problems that appear to be thematic within the research on co-teaching:

1. The authors leave out vital information on the measures used in their studies.
2. The authors interview teachers where co-teaching is already considered "successful," thus incurring potential bias.
3. The major variable in the success or failure of a co-teaching program appears to be the teachers' personalities.
4. The general and special education teachers do not have a clear and/or similar definition of co-teaching or collaboration.
5. Due to the design of the research, behavioral and grade changes (as well as other outcomes) are frequently stated qualitatively (e.g., "improved," "better"), rather than quantitatively or definitively.
6. Few studies describe the actions of the special education teacher during the process of co-teaching.

Clearly, for a fruitful investigation on co-teaching to continue, researchers must attempt to address each of these concerns.

In summary, experimental research supporting the use of co-teaching as an appropriate and effective intervention is sparse. Only six studies were identified that provided sufficient quantitative information. Effect sizes for the individual studies varied considerably (0.08 to 0.95), suggesting at best that the effectiveness of co-teaching is moderately successful. The sheer dearth of experimental research in the area of co-teaching between general and special educators emphasizes the need for future research. The evidence clearly points the direction for educational researchers, but what does this imply for educators themselves?

Practical Implications

The fact that the research on co-teaching is lacking does not mean that co-teaching in schools should be eliminated altogether. In fact, for researchers to collect the needed data, teachers who are employing co-teaching as a service delivery option at all grade levels should open their classrooms for study. It is imperative that data be gathered in classes where the merger between general and special education faculty members has been successful, as well as in those where it has not worked well. This type of information will lead us to a better understanding of how the collaborative teaching between general and special education teachers can best meet the needs of students with disabilities who are served in the general education setting. In addition, the collection of data on student outcomes will also allow researchers to provide teachers with a more cogent understanding of when, where, and with whom co-teaching is best implemented. Thus, the more educational researchers work with teachers in schools, utilizing co-teaching at various grade levels, the more the educational community will learn of the impact this type of working relationship between general and special educators has on students and what our next steps should be. ■

WENDY A. WEICHEL MURAWSKI, EdS, is an assistant professor in special education at California State University, Northridge. Her interests are the academic outcomes of secondary students with learning disabilities in co-taught settings and the potential impact of staff development and block scheduling on students with mild-moderate disabilities. **H. LEE SWANSON**, PhD, is a professor of educational psychology and special education at University of California, Riverside. He was formerly the editor of *Learning Disability Quarterly*. His research interests are memory, intelligence, and intervention models as they apply to students with learning disabilities. Address: Wendy Weichel Murawski, c/o Department of Special Education, California State University, Northridge, 18111 Nordhoff Street, Northridge, CA 91330-8265; e-mail: wweichel@yahoo.com

NOTES

1. Many terms have been used to identify co-teaching, including *team teaching*, *collaborative teaching*, and *cooperative teaching*. In some instances, team teaching is considered a subset of co-teaching (Bauwens & Hourcade, 1991; Cook & Friend, 1995; Friend & Cook, 1992), and in

other cases it denotes co-teaching itself (Jones & Carlier, 1995). This confusion of terms can hinder the research process, making it difficult to determine the effectiveness of a co-teaching intervention (Reinhiller, 1996).

2. Although the criteria for inclusion in this analysis mandated the presence of four particular characteristics of co-teaching, other characteristics have been cited as being important predictors of effective co-teaching as well. These include parity; voluntariness; professional status; shared resources, accountability, and responsibility; and the use of a variety of models for co-teaching instruction (Friend & Cook, 2000). Studies were coded by these suggested characteristics of effective co-teaching as well.
3. Because the terminology used for co-teaching often varies considerably (e.g., *cooperative teaching*, *teaming*, *collaborating*, *mainstreaming*), the use of these characteristics allowed the researcher to identify appropriate articles with or without the authors' specific use of the term *co-teaching*. These characteristics were based on reviews by Bauwens et al. (1989), Cook and Friend (1991), and Friend and Cook (1992).
4. Some authors used terms such as *consultation*, *collaboration*, or *pull-in instruction* but did not define what was meant by these terms; others described their interpretation of co-teaching as working with homogeneous groups within the general education setting, or as planning together but delivering instruction separately. Although these may be viable service delivery options for students with special needs, they do not meet the criteria for co-teaching as defined for our purposes, and they were thus eliminated from the study.

REFERENCES

Note: Asterisks denote the articles included in the meta-analysis.

- Bauer, A. M., & Shea, T. M. (1999). *Inclusion 101: How to teach all learners*. Baltimore: Brookes.
- Bauwens, J., & Hourcade, J. J. (1991). Making co-teaching a mainstreaming strategy. *Preventing School Failure*, 35(4), 19-24.
- Bauwens, J., Hourcade, J. J., & Friend, M. (1989). Cooperative teaching: A model for general and special education integration. *Remedial and Special Education*, 10(2), 17-22.
- Bergren, B. A. (1997). *Teacher attitudes toward included special education students and co-teaching*. (ERIC Document Reproduction Service No. ED 408 754)
- Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cook, L., & Friend, M. (1991). Principles for the practice of collaboration in schools. *Preventing School Failure*, 35(4), 6-9.
- Cook, L., & Friend, M. (1995). Co-teaching: Guidelines for creating effective practices. *Focus on Exceptional Children*, 28(3), 1-16.
- Downing, J., & Bailey, B. R. (1990). Sharing the responsibility: Using a transdisciplinary team approach to enhance the learning of students with severe disabilities. *Journal of Educational and Psychological Consultation*, 1, 259-278.
- Education for All Handicapped Children Act of 1975, 20 U.S.C. § 1400 *et seq.*
- Espin, C., Deno, S., Maruyama, G., & Cohen, C. (1989, April). *The Basic Academic Skills Samples (BASS): An instrument for the screening and identification of children at-risk for failure in regular education classrooms*. Paper presented at the annual American Educational Research Association Conference, San Francisco.
- Friend, M., & Cook, L. (1992). The new mainstreaming: How it really works. *Instructor*, 101(7), 30-36.
- Friend, M., & Cook, L. (2000). *Interactions: Collaboration skills for school professionals* (3rd ed.). White Plains, NY: Longman.
- Gable, R. A., Hendrickson, J. M., & Rogan, J. P. (1996). TEAMS supporting students at risk in the regular classroom: Teacher educational assistance for middle school students. *Clearing House*, 69, 235-239.
- Gable, R. A., & Manning, M. L. (1997). In the midst of reform: The changing structure and practice of middle school education. *Clearing House*, 71, 58-62.

- Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). Needham Heights, MA: Simon & Schuster.
- Goessling, D. P. (1998). Inclusion and the challenge of assimilation for teachers of students with severe disabilities. *Journal of the Association for Persons with Severe Handicaps*, 23, 238–251.
- Gresham, F. M. (1989). Assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review*, 18, 37–50.
- Howell, P. (1991). Taking AIM to assist middle school students with special needs. *Preventing School Failure*, 35, 43–47.
- Hunt, P., Alwell, M., Farron-Davis, F., & Goetz, L. (1996). Creating socially supportive environments for fully included students who experience multiple disabilities. *Journal of the Association for Persons with Severe Handicaps*, 21(2), 53–71.
- Individuals with Disabilities Education Act of 1990, 20 U.S.C. § 1400 *et seq.*
- Individuals with Disabilities Education Act Amendments of 1997, 20 U.S.C. § 1415.
- Johnson, R. R., Test, D. W., & Algozzine, B. (1995). Special education in general education classrooms: Cooperative teaching using supportive learning activities. *Special Services in the Schools*, 10(1), 25–43.
- Jones, M. M., & Carlier, L. L. (1995). Creating inclusionary opportunities for learners with multiple disabilities: A team-teaching approach. *Teaching Exceptional Children*, 27, 23–27.
- *Klingner, J. K., Vaughn, S., Hughes, S. T., Schumm, J. S., & Elbaum, B. (1998). Outcomes for students with and without learning disabilities in inclusive classrooms. *Learning Disabilities Research & Practice*, 13, 153–161.
- Krathwohl, D. R. (1998). *Methods of educational and social science research: An integrated approach* (2nd ed.). Menlo Park, CA: Longman.
- *Lundeen, C., & Lundeen, D. J. (1993, November). *Effectiveness of mainstreaming with collaborative teaching*. Paper presented at the annual convention of the American Speech–Language–Hearing Association, Anaheim, CA. (ERIC Document Reproduction Service No. ED 368 127.
- MacMillan, D. L., Gresham, F. M., & Bocian, K. M. (1998). Discrepancy between definitions of learning disabilities and school practices: An empirical investigation. *Journal of Learning Disabilities*, 31, 314–326.
- Pugach, M. C., & Wesson, C. L. (1995). Teachers' and students' views of team teaching of general education and learning-disabled students in two fifth-grade classes. *The Elementary School Journal*, 95, 279–295.
- Reinhiller, N. (1996). Co-teaching: New variations on a not-so-new practice. *Teacher Education and Special Education*, 19, 34–48.
- *Rosman, N. J. S. (1994). *Effects of varying the special educator's role within an algebra class on math attitude and achievement*. Masters' thesis, University of South Dakota, Vermillion. (ERIC Document Reproduction Service No. ED 381 993).
- *Self, H., Benning, A., Marston, D., & Magnusson, D. (1991). Cooperative teaching project: A model for students at risk. *Exceptional Children*, 58, 26–33.
- Sevakis, P., & Harris, G. (1992). Co-teaching: Using the expertise of regular and special education for language arts instruction and remediation. *Reading and Writing Quarterly: Overcoming Learning Disabilities*, 8, 57–70.
- Vaughn, S., Bos, C. S., & Schumm, J. S. (2000). *Teaching exceptional, diverse, and at-risk students in the general education classroom* (2nd ed.). Boston: Allyn & Bacon.
- *Vaughn, S., Elbaum, B. E., Schumm, J. S., & Hughes, M. T. (1998). Social outcomes for students with and without learning disabilities in inclusive classrooms. *Journal of Learning Disabilities*, 31, 428–436.
- *Walsh, J. M., & Snyder, D. (1993, April). *Cooperative teaching: An effective model for all students*. Paper presented at the annual convention of the Council for Exceptional Children, San Antonio, TX. (ERIC Document Reproduction Service No. ED 361 930)
- Weichel, W. A. (1999). *A qualitative look at the research on co-teaching*. Manuscript in preparation, California State University, Northridge.
- Weiss, M. P., & Brigham, F. J. (2000). Co-teaching and the model of shared responsibility: What does the research support? In T. E. Scruggs & M. A. Mastropieri (Eds.), *Educational interventions: Advances in learning and behavioral disabilities* (Vol. 14, pp. 217–245). Stamford, CT: JAI.
- White, A. E., & White, L. L. (1992). A collaborative model for children with mild disabilities in middle schools. *Focus on Exceptional Children*, 24, 1–10.

Received May 31, 2000
 Initial Acceptance August 14, 2000
 Final Acceptance October 11, 2000