An Investigation of Co-teaching to Improve Academic Achievement of Students with Disabilities: A Meta-analysis

Mickey Losinski, Ph.D.
Kansas State University

Sara Sanders, Ed.D.
University of Alabama

Robin Parks-Ennis, Ph.D., BCBA-D
University of Alabama at Birmingham

Nicole Wiseman, Ed.D.
Jessica Nelson (3rd year Doctoral Student)
Kansas State University

Antonis Katsiyannis, Ed.D.
Clemson University

Author Note: Correspondence concerning this article should be sent to:
Sara Sanders, Graduate Teacher Assistant, Special Education, Counseling, and Student Affairs
338 Bluemont Hall, Kansas State University, Manhattan, KS 66506, sasndrs@ksu.edu, (913) 485-4502

Abstract

Co-teaching is a popular means to support students with disabilities in the general education classroom. However, despite its widespread use, there are no meta-analyses examining the effects of co-teaching on academic outcomes for secondary students that analyze both study quality (e.g., the Council for Exceptional Children's *Standards for Evidence-Based Practices in Special Education*) and effects. We addressed both of these gaps in the current review, identifying nine articles (10 experimental studies) that were analyzed for quality and effects. No studies met all of the quality indicators and study effects showed no significant impact on students' academic outcomes. Limitations and directions for research are presented. Implications for practice are also discussed, including the consideration that null effects, in this instance, do not necessarily suggest that co-teaching is ineffective.

Keywords: co-teaching, team teaching, least restrictive environment, IDEA, collaborative teaching, inclusion

An Investigation of Co-teaching to Improve Academic Achievement of Students with Disabilities: A Meta-analysis

The Individuals with Disabilities Education Improvement Act (hereafter IDEA, 2004) emphasizes educating students with disabilities with their same age peers to the maximum extent

appropriate. Additionally, the law requires that an individual education program (IEP) provides "an explanation of the extent, if any, to which the child will not participate with nondisabled children in the regular class…" (§1414 (d)(1)(A)(i)(V)). To illustrate current trends in the inclusion of students with disabilities, in 2014, 94.7% of students with disabilities ages 6 through 21 were educated in general education classrooms for at least some portion of the school day (U.S. Department of Education, 2016a). Additionally, 62.6% of students with disabilities were educated in the general education classroom 80% or more of the school day and only 5.3% were educated outside the general education classroom for the entire day. Thus, it appears the inclusion of students with disabilities in the general education classroom is happening as the law intended.

Success in integrated settings, particularly with regard to participation and progress in the general education curriculum, necessitates specific supports and modifications (McLeskey, Waldron, Spooner, & Algozzine, 2014). The recent passage of the Every Student Succeeds Act (ESSA) of 2015 mandated that all students with disabilities be held to the same academic state standards and assessments as their non-disabled peers. Historically, however, academic achievement of students with disabilities has been poor. For example, in 2013-2014, students with disabilities averaged just 32.5% proficiency in 8th grade state reading assessments, 29% proficiency in 8th grade mathematics, and have had a 64.6% graduation rate (U.S. Department of Education, 2016b).

To address the modest outcomes of students with disabilities and other historically low-performing students (e.g., students of color, those of low socio-economic status), congress and the research community, has embraced the use of evidence-based practices (EBPs; Cook & Odom, 2013). The authorization of the ESSA (2015) brought the term evidence-based to legislation, and with it, more definition than in the previous law, the NCLB, that only mandated the use of the vague term scientifically-based practice. Though the ESSA defined what an EBP is from a legal perspective in 2015, the previous decade saw organizations from the Department of Education's Institute of Education Sciences (What Works Clearinghouse, 2014) to the Council for Exceptional Children (CEC, 2014) create standards by which to determine whether a practice has a sufficient evidence-base. Utilizing evidence-based practices (EBPs) is seen as necessary to create more successful educational programs and, in turn, improve the achievement of students who have been shown to be resistant to normal practices (for a more thorough treatment of the history of EBPs, see Cook & Odom, 2013).

In 2014, the Council for Exceptional Children released their *Standards for Evidence-Based Practices in Special Education* (CEC-EBP; 2014), which provide guidelines for both determining if a given research study is of high quality (e.g., strong internal validity, replicable practices) as well as determining if a body of literature contains sufficient evidence to be considered evidence-based. The standards cover eight domains that must be described in the study in a manner that would facilitate replication. The domains include: (a) context and setting, (b) participants, (c) intervention agent, (d) description of practice, (e) implementation fidelity, (f) internal validity, (g) outcome measures/dependent variables, and (h) data analysis. For a practice, to be considered evidence-based by these standards, it must have a sufficient number of high-quality articles – single-case (5 with 20 participants), group design (2 randomized

experiments with 60 participants), or a combination – with positive outcomes (see CEC, 2014 for more exhaustive criteria).

Co-Teaching

Co-teaching is a popular means to support students with disabilities in general education classrooms. Co-teaching typically involves a special education teacher and a general education teacher delivering instruction to students with and without disabilities in the regular classroom (Cook & Friend, 1995). There are six primary co-teaching models: one teach/one observe, one teach/one assist, parallel teaching, station teaching, alternative teaching, and teaming (Friend & Cook, 2013). For co-teaching to be successful, both teachers must also engage in co-planning to facilitate both teachers' ability to take an active role in the lesson (Arguelles, Hughes, & Schumm, 2000). Likewise, teachers need to adopt a shared responsibility for classroom duties, drawing on their own areas of strength, typically the general educator's content expertise and the special educator's strategic and differentiated instruction knowledge.

The most prevalent co-teaching model is one teach, one assist (see Cook, McDuffie-Landrum, Oshita, & Cook, 2017). This model is most often observed in classrooms and is characterized by traditional instruction though, there is limited research to support the effectiveness of this model. Similarly, Scruggs, Mastropieri and McDuffie (2007) found that instructional methods with research to support their effectiveness (e.g., peer mediation, strategy instruction, self-advocacy, and study skills training) were rarely utilized in co-taught classrooms. This is discouraging as it suggests a potential disparity in student's access to specially designed instruction in co-taught classrooms.

Unfortunately, there are several potential barriers to co-teaching (Scruggs et al., 2007). First, many teachers are not explicitly trained to co-teach. Another challenge is that it can be difficult to find time to plan together, an integral component of a successful co-teaching model, as co-planning requires a common meeting time. Another barrier to successful co-teaching is finding strategies and models that support learners with diverse needs (Scruggs et al., 2007). These challenges may be of greater significance at the secondary level (i.e., middle and high school; 6-12) where collaboration, in general, can be rife with challenges. Barriers to co-teaching in secondary schools are tied directly to the complexity of the upper grade levels, such as highly specialized curricula, increased accountability measures, and increased teacher autonomy (Dieker & Murawski, 2003). In addition, by the time students reach secondary grades, many possess splinter skills, which can be attributed to removal from instruction or a failure to reach mastery, resulting in a failure to generalize content from one grade level to the next. Despite its widespread use, researchers have cautioned against assuming that co-teaching is an effective or evidence-based practice for addressing the needs of students with disabilities (e.g., Murawski & Swanson, 2001; Weiss & Brigham, 2000).

Previous Reviews

Six previous reviews exist on co-teaching, ranging from the purely narrative (Weiss, 2004), to systematic syntheses of qualitative (Scruggs et al., 2007) and quantitative (Murawski & Swanson, 2001) studies. However, none provide a clear picture of the evidence-base for co-teaching. To begin, the Weiss (2004) article lacks a systematic framework with which to draw conclusions. Although strictly narrative reviews are useful in describing a body of literature, and

Weiss's work certainly adds to the understanding of co-teaching, they are prone to bias as they lack a systematic methodology (Cook, Mulrow, & Haynes, 1997). Similarly, Welch, Brownell, and Sheridan (1999) conducted a literature review of co-teaching research. However, their review focused only on one co-teaching model, team teaching, and was not directly tied to students with disabilities. Further, most of the identified studies were based on anecdotal reports for implementing models, not causal research, and they did not include replicable search criteria. A meta-synthesis of the qualitative literature, conducted by Scruggs, Mastropieri, and McDuffie (2007), discussed mixed results for co-teaching, particularly with respect to some of the methods that under-utilize the special education teacher (e.g., one-teach, one observe). Additionally, their review was primarily directed towards qualitative research, thus there was no effort made to determine the effectiveness of co-teaching compared to other methods.

Murawski and Swanson (2001) conducted the only meta-analysis of co-teaching. This meta-analysis found a moderate effect for co-teaching compared to controls. However, their meta-analysis included only six articles, and those articles varied considerably on the outcome measures and subject areas included, limiting the generalizability of the findings. Additionally, the meta-analysis did not evaluate study quality, nor include all publicly available studies (e.g., grey literature – dissertations, other non-peer-reviewed manuscripts). Most recently, Cook, Landrum, Oshita, and Cook (2017) provided an update on the Murawski and Swanson meta-analysis. However, their review was not specifically a synthesis, as it did not provide search criteria and simply updated studies from the Murawski and Swanson meta-analysis. However, they did apply the Council for Exceptional Children's *Standards for Evidence-Based Practices in Special Education* (2014), an important step in determining the quality of a body of research.

In sum, there are no current high-quality meta-analyses of co-teaching as a means of including students with disabilities in the general education setting. Therefore, the purpose of the current meta-analysis is to add to the existing literature by: (a) conducting a systematic review with the goal of complete transparency and replicability of methods; (b) assessing all studies against the CEC (2014) *Standards for Evidence-Based Practices in Special Education*; and (c) including all publicly available studies (as recommended by the CEC, 2014 guidelines). The current meta-analysis is informed by the following research questions:

Research Question (1): What is the extent of evidence provided by the co-teaching literature-base as measured by the CEC (2014) standards for evidence-based practices?

Research Question (2): What is the relative effect of co-teaching versus control conditions on student academic outcomes?

Method

To answer the research questions, a database search of the extant literature, including all publicly available studies on co-teaching with students with disabilities, was conducted. The database search was conducted on July 18, 2017 and included all previous dates. Four databases, Academic Search Premiere, Education Full Text, ERIC, and PsychINFO were searched utilizing the boolean phrase (["coteaching" OR "co-teaching"] AND ["middle school" OR "high school]" OR "secondary school" OR "junior high"]). Hand searches were also conducted in the following

journals for the years 2007 through 2017: (a) *Exceptional Children,* (b) *Remedial and Special Education,* (c) *Teacher Education and Special Education,* and (d) *The Journal of Special Education.* These journals were identified from the database search as those most often including studies on the inclusion of students in special education in general education classroom settings. Finally, an ancestral search of the previous reviews on co-teaching was conducted by screening the reference list of included articles for any additional studies that met the inclusion criteria of this analysis.

Inclusion Criteria

To be included in the current meta-analysis, studies had to: (a) present findings from a quantitative study; (b) have co-teaching as an independent variable in isolation (e.g., not co-teaching plus PALS), (c) include a dependent variable that measured academic achievement, (d) involve an intervention that happened in secondary (grades 6 – 12) school, and (e) include students with a disability.

Coding Procedures

Studies that met the inclusion criteria were coded individually by four researchers. Each of the coders then met to compare coding, and in the event disagreements occurred, the researchers discussed the issue and arrived at a consensus, resulting in 100% agreement. The following variables were coded. First, participant characteristics were coded including the number of participants, grade, gender, race, and disability. Next, the type of article was coded as peer-reviewed journal, book chapter, dissertation or 'other.' The type of co-teaching used was coded as one of the six co-teaching approaches (Friend & Cook, 2013): (a) one teach, one observe; (b) one teach, one assist; (c) parallel teaching; (d) station teaching; (e) alternative teaching; and (f) team teaching. In the event it was not explicit, the coders met to determine the type of co-teaching, or listed it as 'did not specify'. Next, the type of design was coded as randomized control trial, quasi experiment, single-case design, or causal comparative. The means, standard deviations, and study *n* were coded for treatment and control groups to allow for statistical analyses. In the event this information was not provided, statistical analyses were not conducted for a given study.

Finally, each study was rated using the CEC-EBP (CEC, 2014), which includes eight domains most of which include sub-domains. The domains cover: (1) context and setting, (2) participants, (3) intervention agent, (4) description of practice, (5) implementation fidelity, (6) internal validity, (7) outcome measures, and (8) data analysis. For a study to be considered methodologically sound under the CEC-EBP all domains and sub-domains must be met. It is typical for research syntheses to exclude those studies that do not meet criteria from further analyses. However, we have included each to account for publication bias, and because previous research (Losinski, Sanders, Katsiyannis, & Wiseman, 2017; Losinski, Wiseman, White, & Balluch, 2016) has shown that meeting all of the CEC-EBP domains is challenging, especially in the literature that predates the standards. For a study to meet each standard, the information in question had to be explicitly stated. For example, to meet the first indicator in domain 6, the study had to clearly explain how the researcher was in control of the introduction of the independent variable. Causal comparative studies were not assessed for bias or quality as no such standard requirements exist with which to assess them.

Data Analysis

The current meta-analysis analyzed study outcomes utilizing the standard mean difference (*d*), computed as the mean of the control group subtracted from the mean of the treatment group and divided by the pooled standard deviation: $d = \frac{\bar{X}^t - \bar{X}^c}{Pooled\ SD}$. According to Cohen (1988), $d \leq 0.20$ indicates an insignificant effect and $d \geq 0.80$ a large effect. Study effect sizes were entered into Comprehensive Meta-Analysis (CMA; version 2.2.064) where omnibus effect sizes were calculated by weighting studies by the inverse of the variance and a using random-effects model. Random effects, "assumes that the true effect size itself is regarded as a random variable taking on different values in different studies," (Cooper, Hedges, & Valentine, 2009, p. 296). This method is especially useful when calculating varied effects across studies.

Publication bias. The likelihood of a treatment effect conveyed through a meta-analysis approximating the truth depends entirely on the validity of the studies included in the analysis (Maag & Losinski, 2015). Consequently, it is imperative to determine if the included studies are representative of the entire scope of studies thought to exist regarding co-teaching and its effect on the academic achievement of students with disabilities. Unfortunately, the research in the social sciences has come under scrutiny for being particularly prone to publication bias, in the form of a general lack of published studies with null results (Cook & Odom, 2013; Maag & Losinski, 2015). Thus, to account for publication bias in the current meta-analysis, the authors have included all publicly available studies in this review, not just those published in peer-reviewed outlets. Next, we applied two statistical tests to determine the extent to which the current findings may be subject to bias. It should be noted, however, that there is no agreed-upon method for accounting for publication bias, therefore we have chosen to present two methods: Duval and Tweedie's trim and fill method (T&F; Duval & Tweedie, 2000), and Egger's regression of the intercept test (ERI; Egger, Davey Smith, Schneider, & Minder, 1997). T&F makes use of a funnel plot of results to identify a hypothetical effect size. The funnel plot should be symmetrical if no bias exists. In the event the funnel is not symmetrical; effects are added to obtain symmetry in the funnel plot and the effect size is recalculated. ERI is a procedure that attempts to predict the effect size divided by its standard error. If bias is not present, the outcome would be zero, if bias is present integers exceeding zero would be noted.

Results

The initial search yielded 578 results. After screening the titles for duplicates, 463 articles remained. Four researchers screened the titles and abstracts, reaching a consensus on 56 articles. Those articles were then coded by two dyads of researchers, working together. Once coding was finished, the dyads met to compare results and discuss discrepancies until they agreed on 100% of the articles, leaving a total of 7 articles (see Figure 1). The research team then conducted a hand search of targeted journals and coded the articles in the same manner as before. One additional article was identified through an ancestral search of the reference lists of included articles and one other was identified in the hand search, resulting in a total of 9 articles included in the meta-analysis. In the event the researchers identified a study using a causal-comparative design during screening, the article was flagged and set aside for separate analysis.

Of the nine included articles, ten studies (Goldie, 2015 conducted two studies) were reported with a total of 3,532 participants. Descriptions of the participants, type of co-teaching, and

disabilities are found in Table 1. Six of the studies provided information for gender, with males being only slightly more prevalent (n = 1,314) than females (n = 1,231). Four of the nine studies specified race with White being the most commonly identified (*n* = 2,022), followed by Hispanic (*n* =204), Black (*n* = 171), Asian (*n* = 68), Multiple races (*n* = 165), and Indian (*n* = 8). Four of the studies specified the type of co-teaching used with the most common being One-Teach, One Assist (*n* = 2), followed by Team Teaching (*n* = 1), and Co-Teaching with Content Enhancement Routines (*n* = 1).

Quality of Included Studies

None of the included studies met all of CEC's (2014) evidence based standards. The average quality percentage for the 9 included studies was 66%, with a range of 29% (Walsh, 1993) to 83% (McDuffie, 2009; Whisted, 2011). The most common omission among the studies were a lack of reporting of implementation fidelity (indicator 5.1, 5.2, & 5.3). In fact, only one of the included studies reported the use of an implementation checklist throughout the study to insure fidelity (Whisted, 2011). Table 2 provides results of study quality as assessed by the CEC-EBP (2014) indicators.

Synthesis of Study Outcomes

Table 3 displays results of the included studies utilizing the standard mean difference *d* to determine effectiveness. Three studies within two manuscripts had a moderate positive effect (Maultsby, 2009; Zgonc, 2007). Two studies included in the analysis showed moderate negative effects (Whisted, 201; Goldie, 2015). Ultimately, results indicate that the overall effects of co-teaching were quite small (*d* = -0.012; 95% CI = -0.205 to 0.182).

With respect to the causal comparative studies, we chose to analyze their effects here to be as transparent as possible when discussing the relative effects of co-teaching, as these may be cited by others to attest to its effectiveness; however, no further coding was conducted. Results of this analysis is displayed in Table 4. We identified 11 articles that reported 13 outcomes across mathematics, reading, language arts, science and social studies. The omnibus effect size was *d* = 0.260 (var. = 0.017; 95% CI = 0.01 to 0.51) and ranged from a low of *d* = -0.70 (var. = 0.12; Mason, 2013) to a high of 0.74 (var. = 0.08; Owoh, 2013).

Publication Bias. To address publication bias, the following two analyses were conducted, Duval and Tweedie's trim and fill method (T&F; Duval & Tweedie, 2000), and Egger's regression of the intercept test (ERI; Egger, Davey Smith, Schneider, & Minder, 1997). Results of the T&F method showed the adjusted value to be *d* = -0.25 (95% CI = -0.46 to -0.04). With respect to ERI, the indicator resulted in an intercept of 1.68 (*se* = 1.18), which suggests that bias may be present. In all, both measures of publication bias suggest that some bias may be present.

Discussion

The current meta-analysis investigated the effects and evidence base of co-teaching interventions used to increase the academic achievement of students with disabilities. Overall, insignificant effects were observed, suggesting that co-teaching may not be an effective intervention when compared to controls for increasing academic achievement. Furthermore, no content area seemed to benefit more or less from the co-teaching strategy. In addition to few studies being

identified in our search, none of the studies met the quality standards set forth by CEC (2014) to determine an evidence-based practice, thus co-teaching to improve academic engagement in students with disabilities is far from attaining the standards established by CEC (insufficient evidence). Discussion of these findings with respect to the proposed research questions and implications of the current meta-analysis, in light of the results and what they mean to the inclusive movement, are presented. Finally, limitations and implications for future research are provided.

Co-Teaching versus Control Conditions

Our findings suggest co-teaching had little to no effect on increasing academic achievement across numerous content areas. Unlike the previous review conducted by Murawski and Swanson (2001), our analysis does not indicate co-teaching is an effective academic intervention. The calculated overall effect size (d = -0.01) falls well below the threshold of a moderate effect (d > 0.20; Cohen, 1988). Co-teaching had no effect on individual subject areas (math, reading, science, and social studies) with each overall effect size indicating little to no effect. For example, the largest effect size when examining individual subject areas was d = 0.179 for science, which still falls well below the standard set by Cohen (1988). Therefore, despite qualitative reviews that indicate teachers view co-teaching in a favorable light (Welch et al., 1999; Scruggs et al., 2007), there is little evidence to support co-teaching as means to improve academic achievement. Further, a larger number of retrospective causal comparative studies (n = 11) exist than experimental designs.

The importance of causal-comparative studies. Retrospective causal comparative, or ex post facto designs, (Busk, 2005) are not without their uses as they allow for the analysis of existing data through cohort designs. However, the ex post facto designs found in our search did not suggest that co-teaching is significantly better than controls (d = 0.2), and was only slightly better than the d = 0.01 found for the experiments. Neither the experiments nor the causal comparative designs pass beyond Cohen's (1988) standard of d >0.20 as an insignificant effect. Further, we were not able to judge the quality of the studies against any indicators because none exist, however issues of quality and convention were present in the majority of them, and it is unlikely that if guidelines did exist that these studies would meet them.

Co-Teaching and the CEC (2014) Standards for Evidence-Based Practices

The CEC-EBP (2014) were applied to all studies in the analysis (excluding the causal comparative studies) to establish the evidence of each intervention. None of the studies included met all the CEC-EBP's (2014) for quality reporting. Frequent issues that arose in study quality included failure to describe implementation fidelity (89%) and a lack of adequate information concerning the disability status of participants (67%). Failure to describe implementation quality is particularly concerning, as this makes studies vulnerable to threats to validity. Additionally, if a study is difficult to implement with 100% fidelity or measuring the implementation fidelity with accuracy is problematic, it is difficult to draw accurate conclusions to determine the effectiveness of the intervention (i.e., issues of fidelity may be influencing the outcomes). Paucity of information concerning co-teaching implementation also leads to questions concerning the ability to bridge the research–to-practice gap.

It is also important to note that only one study included in this paper was conducted after the CEC-EBP standards were released, allowing for the possibility that some of the studies did not detail specific components that were actually included in the procedures, preventing it from being considered as contributing to the evidence base. Previous reviews have reported similar findings when applying CEC-EBP's standards to studies (e.g., Cook, Cook, & Cook, 2016; Houchins, Oakes, & Johnson, 2016; Losinski, Cuenca-Carlino, Zablocki, & Teagarden 2014; Losinski et al., 2016), highlighting the importance for future studies to be guided by these standards to create a high-quality research base. Furthermore, it is critical that all studies examining the effectiveness of co-teaching be published, even if they show negative results, so we can accurately determine the appropriateness of co-teaching as an academic intervention.

It is important to emphasize that the lack of explicit information about the co-teaching strategies used within the treatment group limits the conclusions we can draw. Because co-teaching is a practice inclusive of multiple strategies (e.g., small group instruction, collaborative learning activities) it is difficult to fully understand what types of interventions were taking place in the classroom. Additionally, it is unclear the type of instruction or curriculum used in the control settings. Furthermore, the majority of studies did not specify the types of students with disabilities included. With the term students with disabilities covering a broad spectrum of student characteristics and abilities, it is possible co-teaching may be more effective for certain disability categories. However, due to the paucity of information provided about specific disability categories, this remains unknown. Without detailed information on the types of instruction and participant characteristics in both the control and treatment groups, the conclusions that can be drawn are limited.

Co-Teaching and the Inclusive Movement

Advocacy efforts by organizations and researchers to more meaningfully include students with disabilities in the general education classroom (e.g., Stainback & Stainback, 1985) along with seeming confusion of key aspects of the No Child Left Behind Act of 2001 (now the Every Student Succeeds Act of 2015) and the Individuals with Disabilities Education Improvement Act of 2004 have seen a dramatic increase in levels of inclusion of students with disabilities in the general education classroom (Mastropieri & Scruggs, 2014). However, as noted by subsequent court cases on the topic, courts commonly rule in favor of a continuum of placements over full inclusion (Yell, 1995; Yell & Katsiyannis, 2004; Zirkel, 1996). Further, the courts direct that these decisions should be prepared on an individual basis, and determined by the benefits of the respective setting. What's more, research has illustrated the complexities of inclusive practices in general, and a divide among the views of the classroom teacher and school administrator (Cook, Semmelb, & Gerber, 1999).

Co-teaching and differentiated instruction, swept in under the inclusive reform movement, are both viewed as a means for providing for the full inclusion of all students through a re-working of the general education pedagogical model (Mastropieri & Scruggs, 2014). Co-teaching has enjoyed extensive attention in the educational literature through anecdotal experiences and suggestions for implementation (Cook et al., 2017; Mastropieri, Scruggs, & McDuffie, 2007). However, empirical research on the efficacy of co-teaching on the achievement of students with disabilities is sparse at best (Murawski & Swanson, 2001). Our findings show a similar result,

with co-teaching showing no benefit compared to control conditions in the few studies that have researched it.

Our findings would seem to support inclusive practices in the general education classroom, which advocates for students to be placed in the general education classroom for a variety of reasons (e.g., socialization, content access; Fuchs & Fuchs, 1998; Fuchs et al., 2015). Specifically, there were no differences in the outcomes of students receiving their education in a co-teaching model versus the control. Thus, if academic outcomes are not affected, then co-teaching appears to be a useful way to provide for a holistic educational experience for students with disabilities, as it allows teachers to provide specially designed instruction for students alongside their peers who are typically developing. This is consistent with the least restrictive environment mandate of the IDEA (2004) which requires that students with disabilities be educated with non-disabled peers to the maximum extent possible (34 CFR § 300.114). Further, by receiving instruction in the general education setting, students are more likely to be exposed to unmodified grade level standards, which is essential for special education students seeking a traditional diploma.

What is not clear in the current meta-analysis is what methods were utilized in the control conditions, and whether resource rooms utilizing state-of-the-art special education instruction would outperform the co-teaching model. For example, Fuchs et al. (2015) found that state-of-the-art special education instruction in a small group setting outperformed the inclusive environment. Once again, however, it is not clear what the control condition was. Thus, the answer to the question of inclusion versus resource room cannot be definitively answered by current research. What is needed is a strong set of studies that clearly describe the control conditions and put well developed and state-of-the-art co-teaching versus state-of-the-art instruction in the resource classroom. Only through such an examination by an unbiased research team will this question be solved.

Limitations and Future Research

The current meta-analysis is not without its limitations. First, there are a variety of ways to discuss co-teaching (e.g., teaming); however, our analysis only searched for the words co-teaching. It is plausible that having entered in additional words may have provided a larger pool from which to choose from, though the ancestral and hand searches conducted would likely have caught any studies that may have been missed. Additionally, our search found studies not included in some of the other recent reviews (e.g., Cook et al., 2017). A second limitation was with the hand search, where we did not search other journals that may have had other studies published in them (e.g., the *Journal of Learning Disabilities*). However, our choosing of the top journals in the field and also those that had shown up in earlier works seems sufficient in finding studies through the hand search. Third, the addition of grey literature, all publicly available studies including dissertations, may have affected the results, although upon examination the dissertations were not meaningfully different from peer-reviewed studies in either quality or effects. Finally, this literature is fraught with an inability in the experiments to have a control condition that represents a true control for co-teaching. For example, what is being learned in the resource room maybe vastly different, and probably is, than what is being learned in the general education classroom. Therefore, it's hard to distinguish if co-teaching were being

investigated, or simply being in a room with different access to the general education curriculum were. In other words, are we comparing a Fuji apple to a pink lady apple, or are we comparing a Fuji apple to a watermelon?

Future researchers should engage in high-quality investigations of the co-teaching model so that we can have a better idea of the relative effects of co-teaching. Additionally, researchers should undertake investigations to determine which models of co-teaching perform better than others with regard to the achievement of students. Third, it would be important for future researchers to work with a control that was allowing the same access to the curriculum as the treatment group, perhaps investigating having students in the control group also be in the general education curriculum, but with different supports (e.g., paraprofessional support; differentiated instruction). Finally, future research should undertake a large-scale investigation of the inclusion of grey literature (non-peer-reviewed) and if it truly is inferior to its peer-reviewed counterpart.

Conclusion

Co-teaching makes intrinsic sense: two teachers in the classroom, one with content knowledge, one with pedagogy knowledge, should result in marked improvement of students. However, this has not been shown to be the case in the literature. The lack of quality of the studies, along with the null effects, would suggest that practitioners should not use co-teaching until other higher-quality investigations have been conducted. However, looking at the issue from another direction, a null effect in this case possibly suggests something different: students being served in a co-teaching environment are performing as well as those in resource rooms. This would imply that there is no legitimate reason to separate children from their peers and the general education curriculum if co-teaching can serve them with the same results.

References

Note: References with a * are those included in the literature review.

*Andrews-Tobo, R. A. (2009). *Coteaching in the urban middle school classrooms: Impact for students with disabilities in reading, math, and English/language Arts classrooms* (Doctoral dissertation, Capella University). Retrieved from http://search.proquest.com.er.lib.k-state.edu/docview/305163048?accountid=11789

Arguelles, M., Hughes, M., & Schumm, J. (2000). Co-teaching: A different approach to inclusion. *Principal*, *79*(4), 48, 50-51.

Bao, X., & Eaton, D. W. (2016). Fault activation by hydraulic fracturing in western Canada. *Science*, *354*(6318), 1406-1409. doi: 10.1126/science.aag2583

Busk, P. L. (2005). Causal-comparative study. In B. Everitt & D. Howell (Eds.), *Encyclopedia statistics in behavioral science.* Chichester, UK: John Wiley & Sons. DOI: 10.1002/0470013192.bsa171

*Chilcoat, E. R. (2011). *Co-Teaching: A mixed methods study of the effectiveness of a secondary co-taught classroom versus a traditional model* (Doctoral dissertation, Missouri Baptist University). Retrieved from https://search.proquest.com/docview/1016102165

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, NY: Lawrence Erlbaum.

Cook, B. G., & Odom, S. (2013). Evidence-based practices and implementation science in special education. *Exceptional Children, 79*, 135-144. doi: https://doi.org/10.1177/001440291307900201

Cook, D. J., Mulrow, C. D., & Haynes, R. B. (1997). Systematic reviews: synthesis of best evidence for clinical decisions. *Annals of Internal Medicine*, *126*(5), 376-380. doi: 10.7326/0003-4819-126-5-199703010-00006

Cook, S. C., Cook, B. G., & Cook, L. (2016). Classifying the evidence base of class-wide peer tutoring for students with high-incidence disabilities. *Exceptionality*, *25*, 9-25. doi: 10.1080/09362835.2016.1196448

Cook, S. C., Landrum, K. A., Oshita, L., & Cook, B. G. (2017). Co-teaching for students with disabilities: A critical and updated analysis of the empirical literature. In J. M. Kauffman, & D. P. Hallahan, *Handbook of special education.* New York, NY: Routledge. doi: 10.4324/9780203837306.ch12

Council for Exceptional Children. (2014). *Council for Exceptional Children standards for evidence-based practices in special education.* Retrieved from http://www.cec.sped.org/,/media/Files/Standards/Evidence%20based%20Practices%20and%20Practice/CECs%20EBP%20Standards.pdf

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The Handbook of Research Synthesis and Meta-Analysis* (2nd ed.). New York, NY: Russell Sage Foundation.

*Emery, L. (2009). *Academic achievement of middle school students with specific learning disabilities in inclusive cotaught classes* (Doctoral dissertation, Walden University). Retrieved from https://search.proquest.com/docview/305077011?pq-origsite=gscholar

*Ervin, V. L. (2010). *A Comparison of Coteaching Only, Pull-Out Only, and Combined Service Methods for Students with Disabilities* (Doctoral dissertation, Capella University). Retrieved from https://search.proquest.com/docview/305264976

Every Student Succeeds Act (ESSA) of 2015, 20 U.S.C. § 8101 *et seq.*

*Fontana, K. C. (2005). The effects of co-teaching on the achievement of eighth grade students with learning disabilities. *Journal of At-Risk Issues*, *11*(2), 17-23.

Friend, M., & Cook, L. (2013). *Interactions: Collaboration skills for school professionals* (4th ed.). New York, NY: Longman.

Fuchs, D., & Fuchs, L. S. (1998). Competing visions for educating students with disabilities inclusion versus full inclusion. *Childhood Education*, *74*(5), 309-316.

Fuchs, L. S., Fuchs, D., Compton, D. L., Wehby, J., Schumacher, R. F., Gersten, R., & Jordan, N. C. (2015). Inclusion versus specialized intervention for very-low-performing students: What does access mean in an era of academic challenge? *Exceptional Children, 81*(2), 134-157.

*Garrett-Rainey, S. (2014). *The achievement impact of the inclusion model on the standardized test scores of general education students* (Doctoral dissertation, Capella University).

Retrieved from https://search.proquest.com/docview/1530478294/previewPDF/1CFAE0C9E6854843PQ/1?accountid=11789

*Goldie, R. C. (2015). *The impact of co-teaching on academically struggling students: A study on how co-teaching impacts the reading and math achievement of special education, RTI Tiers two and three, and general education students* (Doctoral dissertation, University of St. Francis). Retrieved from https://search.proquest.com/docview/1734877067/69651439570247D1PQ/1?accountid=11789

Houchins, D. E., Oakes, W. P., & Johnson, Z. G. (2016). Bullying and students with disabilities: A systematic literature review of intervention studies. *Remedial and Special Education, 37*, 259-273. doi: 10.1177/0741932516648678

Individuals with Disabilities Education Improvement Act of 2004, 20 U.S.C. §1400 et seq.

*LaFever, K. M. (2012). *The Effect of Co-Teaching on Student Achievement in Ninth Grade Physical Science Classrooms*. (Doctoral dissertation, University of Missouri-Saint Louis). Retrieved from https://search.proquest.com/docview/1492668122/B6766C8EBC5B43D4PQ/1?accountid=11789

Lane, K. L., Common, E. A., Royer, D. J., & Muller, K. (2014). *Group comparison and single-case research design quality indicator matrix using Council for Exceptional Children 2014 standards*. Unpublished tool.

Losinski, M., Cuenca-Carlino, Y., Zablocki, M., & Teagarden, J. (2014). Examining the efficacy of self-regulated strategy development for students with emotional/behavioral disorders: A meta-analysis. *Behavioral Disorders, 40*, 52-67. doi: 10.17988/0198-7429-40.1.52

Losinski, M., Sanders, S., Katsiyannis, A., & Wiseman, N. (2017) A meta-analysis of interventions to improve the compliance of students with disabilities. Education and Treatment of Children, 40(4), 435-464. doi: 0.1353/etc.2017.0020

Losinski, M., Wiseman, N., White, S., & Balluch, F. (2016). A meta-analysis of video-modeling based interventions for reduction of challenging behaviors for students with EBD. *The Journal of Special Education, 49*, 243-252. doi:10.1177/0022466915602493

Maag, J. W., & Losinski, M. (2015). Thorny issues and prickly solutions: Publication bias in meta-analytic reviews in the social sciences. *Advances in Social Sciences Research Journal, 2*, 242-253. doi: 10.14738/assrj.23.1044

*Mason, P. L. (2013). *Comparing types of student placement and the effect on achievement for students with disabilities*. (Doctoral dissertation, Liberty University). Retrieved from https://search.proquest.com/docview/1469385591/40EB9545D0564584PQ/1?accountid=11789

Mastropieri, M.A., & Scruggs, T.E. (2014). *The inclusive classroom: Strategies for effective differentiated instruction* (5th ed.). New York, NY: Pearson.

*Maultsby-Springer, B. M. (2009). *A descriptive analysis of the impact of co-teaching on the reading/language arts and math achievement of selected middle school students in a Middle Tennessee school district*. (Doctoral dissertation, Tennessee State University).

Retrieved from https://search.proquest.com/docview/857128776/D77A52934B604F3EPQ/1?accountid=11789

*McDuffie, K. A., Mastropieri, M. A., & Scruggs, T. E. (2009). Differential effects of peer tutoring in co-taught and non-co-taught classes: Results for content learning and student-teacher interactions. *Exceptional Children*, *75*, 493-510. doi: https://doi.org/10.1177/001440290907500406

McLeskey, J., Waldron, N.L., Spooner, F., & Algozzine, B. (2014). What are effective inclusive schools and why are they important. In J. McLeskey, N. L. Waldron, F. Spooner, & B. Algozzine (Eds.), *Handbook of effective inclusive schools: Research and practice* (pp. 3-16). New York, NY: Routledge.

*Mote, S. Y. (2010). *Does setting affect achievement of students with disabilities: Comparing co-teaching to resource*. (Doctoral dissertation, Liberty University). Retrieved from https://search.proquest.com/docview/822647263/1074753DD96149DBPQ/1?accountid=11789

Murawski, W. W., & Swanson, H. L. (2001). A meta-analysis of co-teaching research: Where are the data? *Remedial and Special Education, 22*, 258-267. doi: https://doi.org/10.1177/074193250102200501

*Muscelli, G. B. (2012). *The effect of co-teaching on ninth grade world history students' academic achievement and classroom behavior* (Doctoral dissertation, Nova Southeastern University). Retrieved from https://search.proquest.com/docview/1221851929/F0E7D0B0E5D74759PQ/1?accountid=11789

*Owoh, K. (2013). *The effect of inclusion models on special education student performance in charter schools in Washington, DC.* (Doctoral dissertation, University of Phoenix) Retrieved from https://search.proquest.com/docview/1720064192/CF519193B6F945D8PQ/1?accountid=11789

*Packard, A. L., Hazelkorn, M., Harris, K. P., & McLeod, R. (2011). Academic achievement of secondary students with learning disabilities in co-taught and resource rooms. *Journal of Research in Education*, *21*, 100-117.

*Rosman, N. J. S. (1994). Effects of varying the special educator's role within an algebra class on math attitude and achievement. (Doctoral dissertation, University of South Dakota) Retrieved from https://search.proquest.com/docview/62701265/2EA6DFA6484E34PQ/1?accountid=11789

Scruggs, T. E., Mastropieri, M. A., & McDuffie, K. A. (2007). Co-teaching in inclusive classrooms: A metasynthesis of qualitative research. *Exceptional Children, 73*, 392-416. doi: https://doi.org/10.1177/001440290707300401

U.S. Department of Education. (2016a). 38th annual report to Congress on the implementation of the Individuals with Disabilities Education Act, 2015. Washington, DC: Author. Retrieved from https://www2.ed.gov/about/reports/annual/osep/2016/parts-b-c/38th-arc-for-idea.pdf

U.S. Department of Education. (2016b). ED data express: Data about elementary and secondary schools in the United States. Retrieved from http://eddataexpress.ed.gov/index.cfm

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.), *Advances in learning and behavioral disabilities: Vol. 14. Educational interventions* (pp. 217-245). Oxford, UK: Elsevier.

*Walsh, J. M., & Snyder, D. (1993). *Cooperative teaching: An effective model for all students.* Paper presented at the Annual Convention of the Council for Exceptional Children, San Antonio, TX.

Weiss, M. P. (2004). Co-teaching as science in the schoolhouse: More questions than answers. *Journal of Learning Disabilities, 37*, 218-233. doi: https://doi.org/10.1177/00222194040370030601

Welch, M., Brownell, K., & Sheridan, S. M. (1999). What's the score and game plan on teaming in schools? A review of the literature on team teaching and school-based problem-solving teams. *Remedial and Special Education, 20*, 36-49.

What Works Clearinghouse. (2014). *Procedures and standards handbook (Version 3.0).* Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf

*Whisted, M. L. (2011). *Does the Use of Co-teaching Models in Algebra Result in an Increase in Student Achievement Among Students with Disabilities and Their Non-disabled Peers?* (Doctoral dissertation, College of Notre Dame of Maryland). Retrieved from https://search.proquest.com/docview/870954775?pq-origsite=gscholar

*Williams, R. L. (2012). *Service delivery models: An effectiveness study in select schools* (Doctoral dissertation, Capella University). Retrieved from https://search.proquest.com/docview/1373445592/1C93B17C913D4FBDPQ/1?accountid=11789

Yell, M. L., & Katsiyannis, A. (2004). Placing students with disabilities in inclusive settings: Legal guidelines and preferred practices. *Preventing School Failure*, *49*(1), 28-35.

*Zgonc, K. C. (2007). *The impact of co-teaching on student learning outcomes in secondary social studies classrooms implementing content enhancement routines* (Doctoral dissertation, University of Central Florida). https://search.proquest.com/docview/304743873/BD12102E8883425CPQ/1?accountid=11789

About the Authors

Dr. Mickey Losinski, PhD, is an Associate Professor of special education in the College of Education at Kansas State University. His professional interests include legal issues in special education, parent issues in special education, and the identification of evidence-based practices in special education.

Dr. Sara Sanders EdD, is currently completing a post-doctoral research fellowship at the University of Alabama. Her research focuses on academic strategies for students with emotional

and behavioral disorders (EBD), specifically self-regulated strategy development (SRSD) in reading, as well as CBT interventions for students with internalizing symptoms.

Robin Parks Ennis, PhD, BCBA-D, is an Assistant Professor of special education in the School of Education at the University of Alabama at Birmingham. Her research interests include comprehensive integrated three-tiered models of prevention and low-intensity/instructional academic and behavioral strategies to support students with and at-risk for emotional and behavioral disorders.

Dr. Nicole Wiseman, EdD, is currently a behavior support teacher in Kansas. Her research interests include academic interventions and mindfulness strategies for students with and at-risk for students with EBD.

Jessica Nelson is a third year doctoral student at Kansas State University. Her research interests include academic interventions for students with intellectual disabilities and the training of pre-service teachers with students with special needs.

Dr. Antonis Katsiyannis EdD, is an *Alumni Distinguished Professor* of special education in the Eugene T. Moore School of Education at Clemson University. His research interests include legal and policy issues associated with special education, delinquency, and issues involving students with emotional or behavioral disorders

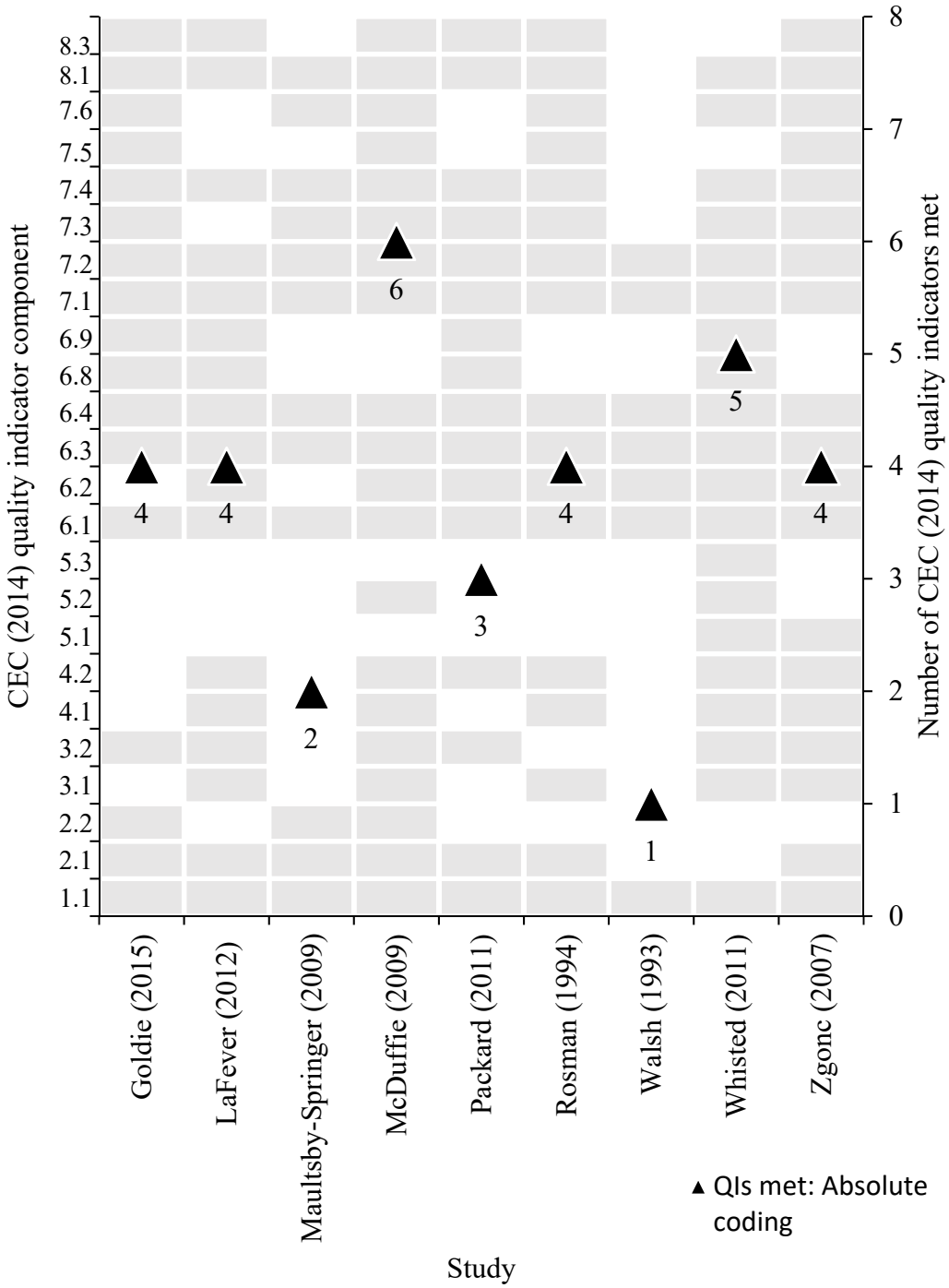
Study	N	Grade (*n*)	Gender (*n*)	Race (*n*)	Type of Co-Teaching	Disability	Results
Goldie (2015) (Math)	642	11	333(M), 306 (F)	W(533), B(26), H(45), 6 (A), 27 (MR)	DNS	DNS	Students in non-co-taught classes performed better than students in co-taught classes.
Goldie (2015) (ELA)	1305	11	689(M), 623 (F)	W(1150), B(29), H(53), 15 (A), 59 (MR), 6 (I)	DNS	DNS	Students in non-co-taught classes performed better than students in co-taught classes.
LaFever (2012)	174	9	DNS	DNS	One-Teach, One-Assist & Pull-Out	DNS	Effective if one special education teacher and one subject certified teacher are present
Maultsby (2009)	102	5-8	DNS	DNS	DNS	DNS	Increase of achievement in math with students with disabilities; decrease in reading in students without disabilities
McDuffie (2009)	194	7	106 (M), 97 (F)	137 (W), 19 (B), 22(H), 18 (A), MR (7)	Team Teach, One-Teach, One Assist, Alternative, & Station	LD, ED, OHI, Autism, SL	Students in co-teach classroom performed better than students in traditional setting
Packard (2011)	14	9	9 (M), 5 (F)	W(10), B(4)	DNS	LD	Students with LD achieved better in resource room than co-taught room
Rosman (1994)	59	9	25(M), 34(F)	DNS	Team Teaching	DNS	Students in co-teaching scored significantly higher; no significance in scores of students with IEPs
Walsh (1993)	706	9	DNS	DNS	DNS	DNS	All students benefited from the co-taught classroom
Whisted (2011)	18	9	DNS	DNS	DNS	DNS	No overall significant difference between co-taught and traditional classrooms

Zgonc (2007)	318	9-10	152(M), 166(F)	W(192), B(93), H(84), A(8), I(2), MR(9)	Co-Teaching with Content Enhancement Routines	DNS	No significant difference between co-taught and non co-taught classes

Table 1. Study Characteristics

Note. DNS = Did not specify; OHI = Other Health Impairment; SLD = Specific Learning Disability; ID – Intellectual Disability; ASD = Autism Spectrum Disorder; HI = Hearing Impaired; W = White; B = Black; A = Asian; H = Hispanic; I = Indian; MR = Multi-Racial

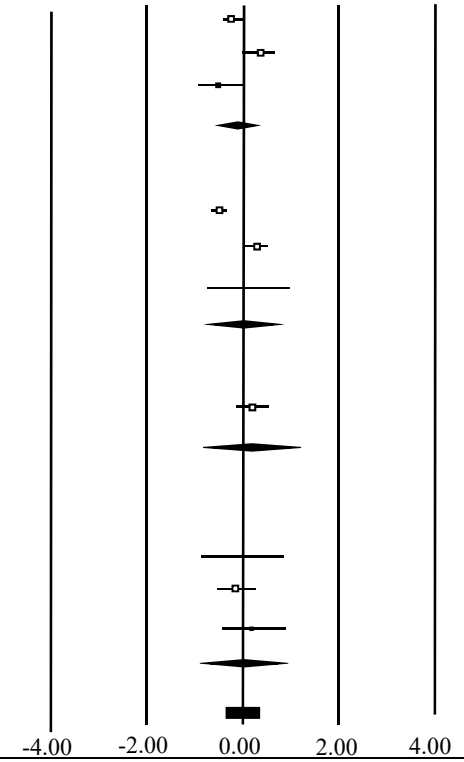
Table 2: Quality Indicators



Note: Left *y* axis displays the components of CEC (2014) quality indicators (QI). Shaded cells indicate the component was met; white cells denote the component was not met. The right *x* axis shows the number of absolute QI. Triangles represent the absolute coding of each study.
Lane, K. L., Common, E. A., Royer, D. J., & Muller, K. (2014)

Table 3: Effect Sizes

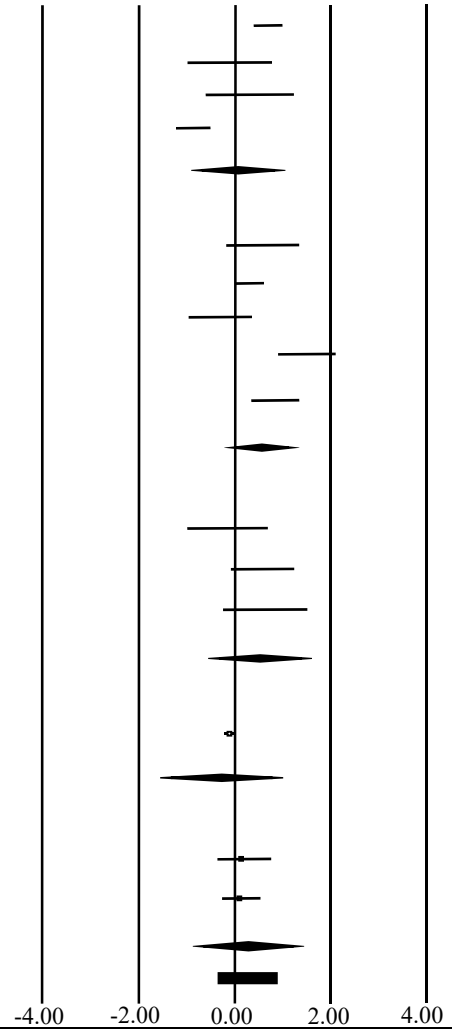
Study	Std diff in means	SE	var	Lower Limit	Upper Limit	Z	p	Std. diff in means and 95% CI
Math								
Goldie (2015)	-0.194	0.090	0.008	-0.370	-0.018	-2.156	0.031	
Maultsby (2009)	0.275	0.145	0.21	-0.009	0.559	1.897	0.058	
Whisted (2011)	-0.471	0.248	0.062	-0.957	0.015	-1.899	0.058	
Total	-0.104	0.210	0.044	-0.515	0.308	-0.494	0.621	
Reading								
Goldie (2015)	-0.336	0.076	0.006	-0.486	-0.186	-4.398	0.000	
Maultsby (2009)	0.303	0.145	0.021	0.019	0.587	2.088	0.037	
Packard (2011)	0.167	0.472	0.223	-0.758	1.092	0.354	0.723	
Total	-0.006	0.224	0.050	-0.445	0.433	-0.027	0.979	
Science								
McDuffie (2009)	0.179	0.144	0.021	-0.103	0.461	1.243	0.214	
Total	0.179	0.355	0.126	-0.517	0.875	0.504	0.614	
Social Studies								
Zgonc (2007)								
Group A	-0.072	0.155	0.024	-0.376	0.232	-0.465	0.642	
Group B	0.277	0.327	0.107	-0.364	0.918	0.847	0.397	
Group C	-0.006	0.463	0.214	-0.913	0.901	-0.013	0.990	
Total	0.047	0.253	0.064	-0.450	0.544	0.185	0.853	
Overall	-0.012	0.099	0.010	-0.205	0.182	-0.117	0.907	



Note: SE=Standard Error; var = variance; Z = Z-score; p = p-score; CI = confidence interval

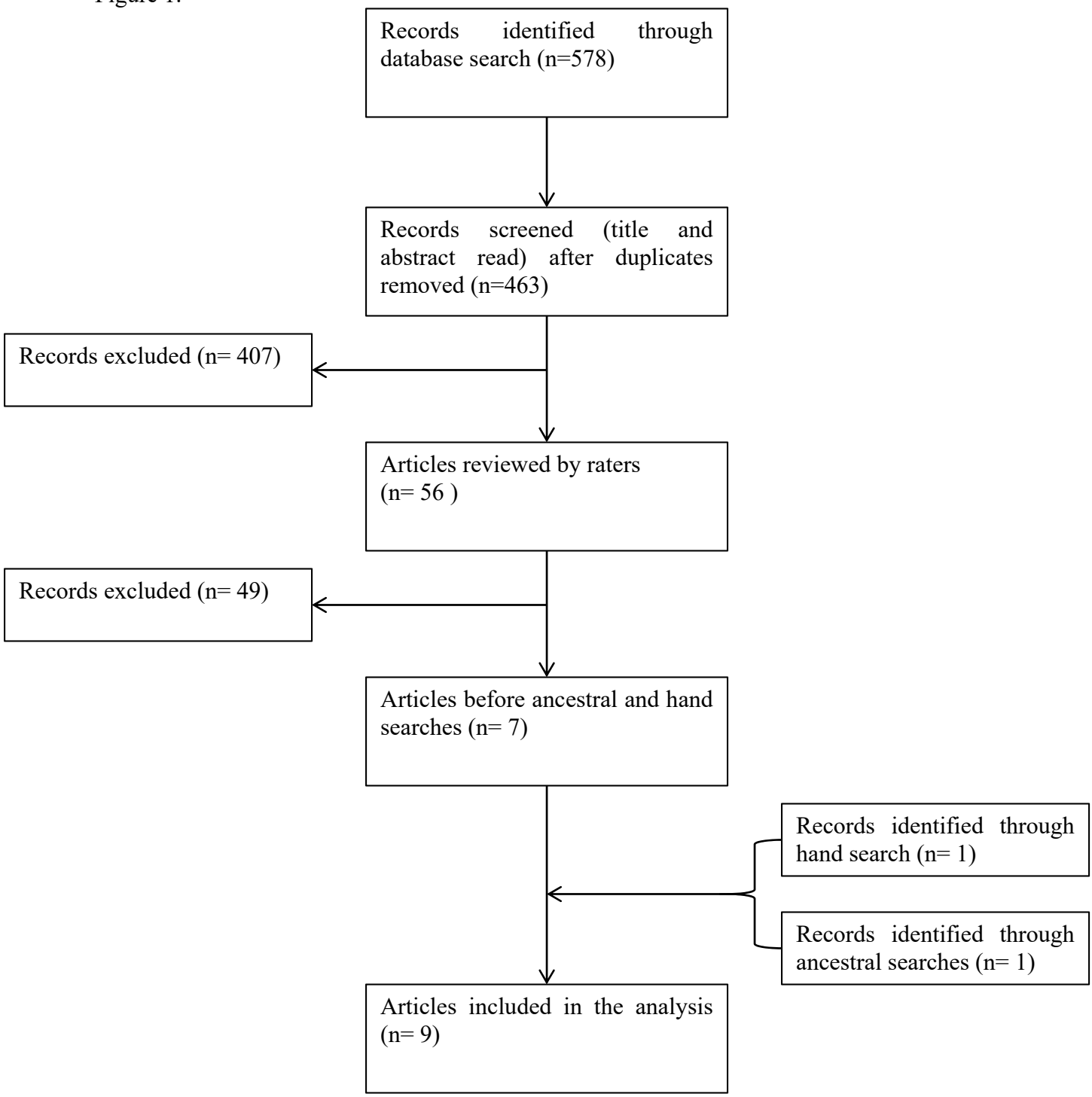
Table 4: Effect Sizes of Causal Comparative Studies

Study	Std diff in means	SE	var	Lower Limit	Upper Limit	Z	p	Std. diff in means and 95% CI
Math								
Emery (2009)	0.607	0.168	0.028	0.278	0.936	3.616	0.000	
Fontana (2005) 7th	-0.083	0.348	0.121	-0.766	0.600	-0.239	0.811	
Fontana (2005) 8th	0.277	0.350	0.122	-0.409	0.963	0.793	0.428	
Mason (2013)	-0.695	0.178	0.032	-1.044	-0.346	-3.907	0.000	
Total	0.019	0.292	0.085	-0.554	0.592	0.066	0.948	
Reading								
Andrews-Tobo (2009)	0.546	0.335	0.112	-0.110	1.202	1.630	0.103	
Emery (2009)	0.397	0.160	0.026	0.082	0.711	2.475	0.013	
Ervin (2010)	-0.309	0.301	0.091	-0.900	0.281	-1.027	0.304	
Mote (2010)	1.466	0.331	0.110	0.817	2.116	4.425	0.000	
Owoh (2013)	0.742	0.287	0.082	0.180	1.304	2.587	0.010	
Total	0.553	0.266	0.071	0.032	1.074	2.082	0.037	
English								
Fontana (2005) 7th	-0.130	0.349	0.122	-0.814	0.553	-0.374	0.708	
Fontana (2005) 8th	0.636	0.357	0.127	-0.063	1.336	1.782	0.075	
Williams (2012)	0.706	0.432	0.187	-0.141	1.554	1.634	0.102	
Total	0.385	0.344	0.118	-0.289	1.060	1.119	0.263	
Science								
Garrett-Rainey (2014)	-0.098	0.061	0.004	-0.218	0.022	-1.599	0.110	
Total	-0.098	0.465	0.216	-1.009	0.813	-0.211	0.833	
Social Studies								
Chilcoat (2011)	0.177	0.269	0.072	-0.350	0.704	0.659	0.510	
Muscelli (2011)	0.167	0.181	0.033	-0.188	0.522	0.923	0.356	
Total	0.172	0.363	0.132	-0.539	0.883	0.473	0.636	
Overall	0.255	0.184	0.034	-0.107	0.616	1.379	0.168	



Note: SE=Standard Error; var = variance; Z = Z-score; p = p-score; CI = confidence interval

Figure 1.


Records identified through database search (n=578)
Records screened (title and abstract read) after duplicates removed (n=463)
Records excluded (n= 407)
Articles reviewed by raters (n= 56)
Records excluded (n= 49)
Articles before ancestral and hand searches (n= 7)
Records identified through hand search (n= 1)
Records identified through ancestral searches (n= 1)
Articles included in the analysis (n= 9)