

## Direct and Collateral Effects of Peer Tutoring on Social and Behavioral Outcomes: A Meta-Analysis of Single-Case Research

Lisa Bowman-Perrott, Mack D. Burke, Nan Zhang, and Samar Zaini  
*Texas A&M University*

*Abstract.* This meta-analysis examined the direct (primary) and collateral (secondary) effects of peer tutoring on social and behavioral outcomes for 128 participants in prekindergarten through grade 12 across 20 studies using SCR designs. The overall TauU weighted effect size across studies was 0.62 (95% CI [0.58, 0.66]), indicating that a small to moderate effect on behavioral and social outcomes can be attributed to peer tutoring. Moderator analyses indicated that cross-age tutoring, peer tutoring interventions that did not use reward contingencies, and interventions that measured direct effects yielded higher effect sizes. The direct effect of peer tutoring on behavioral and social outcomes was moderately large ( $ES = 0.75$ ), whereas the collateral effect was relatively small ( $ES = 0.43$ ). Furthermore, peer tutoring had a greater effect on improving social skills and social interactions ( $ES = 0.69$ ) and reducing disruptive and off-task behaviors ( $ES = 0.60$ ) than academic engagement ( $ES = 0.38$ ).

There is growing recognition of the linkages between behavioral problems and academic achievement, as well as the need for interventions that address both (Barriga et al., 2002; Burke, Hagan-Burke, & Sugai, 2003; Miles & Stipek, 2006). Correlational research has examined the relation between various types of problem behaviors and academic achievements (e.g., Malecki & Elliot, 2002) and have focused on two primary areas. The first is the influence of academic underachievement as a predictor of or contributor to emotional and behavioral problems in which academic deficits, often in basic skill areas such as reading, seem to exacerbate existing behavioral difficulties and affect social and behavioral outcomes

(Darney, Reinke, Herman, Stormont, & Ialongo, 2013; McIntosh, Horner, Chard, Dickey, & Braun, 2008). The second area is the reverse; that is, behavioral deficits can adversely affect academic outcomes (Becker & Luthar, 2002; Darney et al., 2013; Lassen, Steele, & Sailor, 2006). Internalizing and externalizing behavioral deficits that may negatively affect academic outcomes include conduct disorder (Nelson, Benner, Lane, & Smith, 2004), poor self-regulation (Cleary & Zimmerman, 2004), difficulties in social and problem-solving skills (Benner, Beaudoin, Kinder, & Mooney, 2005), and problems with attention and hyperactivity (DuPaul, Ervin, Hook, & McGoey, 1998; Fantuzzo, King, & Heller, 1992).

---

Correspondence concerning this article should be addressed to Lisa Bowman-Perrott, Department of Educational Psychology, Texas A&M University, 4225 TAMU, College Station, TX 77843-4225; e-mail: lbperrott@tamu.edu

Copyright 2014 by the National Association of School Psychologists, ISSN 0279-6015, eISSN 2372-966x

Although the direction of the relationship between behavior problems and academic achievement remains unclear, the presence of academic or behavioral deficits is troublesome, particularly for children who have difficulties in both areas. Thus, learning and behavior problems put students at substantial risk over either alone (Reinke, Heman, Petras, & Ialongo, 2008). In addition, social deficits have been linked to behavior problems and lower academic achievement (e.g., poor school adjustment, grade retention, suspension and expulsion from school; Blackorby et al., 2007; Gooding, 2011; Malecki & Elliott, 2002). Thus, it is critical to implement evidence-based practices that promote behavioral and social adjustment, prevent academic failure, and remediate academic deficits. Furthermore, given the lack of clarity surrounding the relationship between behavior problems and academic achievement, interventions that address each of these areas are more desirable than those that address them separately (Algozzine, Wang, & Violette, 2011; McEvoy & Welker, 2000).

### PEER TUTORING EVIDENCE BASE

Peer tutoring, defined as “a class of practices and strategies that employ peers as one-on-one teachers to provide individualized instruction, practice, repetition, and clarification of concepts” (Utley & Mortweet, 1997, p. 9), is supported by a substantial research base for promoting improved academic outcomes across content areas (Cloward, 1967; Cohen, Kulik, & Kulik, 1982; Fuchs, Fuchs, Phillips, Hamlett, & Karns, 1995; Mastropieri, Spencer, Scruggs, & Talbott, 2001; Stainback & Stainback, 1972). Peer tutoring has led to improvements in students’ reading outcomes (e.g., fluency and comprehension; Fuchs & Fuchs, 2005), math scores on standardized assessments (Fuchs, Fuchs, & Karns, 2001) and curriculum-based measurements (Hawkins, Musti-Rao, Hughes, Berry, & McGuire, 2009), science gains on curriculum-based tests (Bowman-Perrott, Greenwood, & Tapia, 2007) and standardized tests (Topping, Peter, Stephen, & Whale, 2004), spelling accuracy (Delquadri, Greenwood, Stret-

ton, & Hall, 1983), and the acquisition of social studies content knowledge (Lo & Cartledge, 2004). Participants who benefitted from peer tutoring included students without disabilities, as well as students with disabilities, including emotional/behavioral disorders (EBD; Bowman-Perrott et al., 2007), learning disabilities (LD; Fuchs et al., 1995), attention deficit hyperactivity disorder (ADHD; DuPaul et al., 1998), autism (Kamps, Dugan, Potucek, & Collins, 1999), mental retardation (MR)/intellectual disability (ID; Maheady, Harper, & Sacca, 1988), and speech/language disorders (S/L; Sutherland & Snyder, 2007).

The smaller, yet promising body of research suggests that peer tutoring has positive direct and collateral effects on behavioral and social outcomes (Garcia-Vazquez & Ehly, 1992; Greenwood, Delquadri, & Hall, 1989; Sutherland & Snyder, 2007). Direct effects were defined as social and/or behavioral benefits (e.g., increased time on-task, improved social interactions) observed as a result of nonacademic outcomes being a primary goal of peer tutoring. These effects were examined in studies described in one of two ways. First, peer tutoring was used solely to train peers as tutors of behavioral and/or social skills (e.g., Lo & Cartledge, 2004; Locke & Fuchs, 1995). Second, peer tutoring was implemented as an academic intervention, but the purpose of the study was to observe academic as well as behavioral and/or social outcomes with equal emphasis (e.g., Maheady & Sainato, 1985; Sideridis et al., 1997). Collateral effects were investigated for studies in which the primary focus was academic achievement. Cook et al. (2012) defined collateral effects as “effects that are secondary outcomes, or byproducts, of an intervention”; they are “behavioral improvements that result from the implementation of [an academic] intervention” (p. 48). Peer tutoring has had a positive collateral effect on a range of problem behaviors, social interactions, and academic engagement. For example, DuPaul et al. (1998) reported a decrease in off-task behaviors and fidgeting as a collateral effect for students with ADHD. McDonnell, Mathot-Buckner, Thorson, and Fister (2001) found a reduction in noncompliant and

aggressive behaviors, whereas Brady (2007, 1997) and Maheady and Sainato (1985) reported collateral increases in the frequency of positive social interactions between peers. Finally, increased levels of academic responding (academic engagement) were also observed as collateral benefits (Ginsburg-Block & Fantuzzo, 1997; McDonnell et al., 2001).

### EXISTING PEER TUTORING META-ANALYSES

Six peer tutoring meta-analyses addressing academic, behavioral, and/or social outcomes were found in a search of the literature (Bowman-Perrott et al., 2013a; Cohen et al., 1982; Cook, Scruggs, Mastropieri, & Casto, 1985; Ginsburg-Block, Rohrbeck, & Fantuzzo, 2006; Jun, Ramirez, & Cumming, 2010; Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003). Five of the meta-analyses included only group design studies and reported Cohen's (1988) *d* effect sizes, or effect sizes calculated from *t* and *F* statistics. Effect sizes ranged from 0.28 across content areas (Rohrbeck et al., 2003) to 1.08 for literacy outcomes (Jun et al., 2010). One meta-analysis (Bowman-Perrott et al., 2013a) examined single-case research design (SCR) studies across subject areas, reporting an overall TauU effect size of 0.75. Whereas effect sizes for academic achievement were reported in all six studies, only three addressed the effects of peer tutoring on behavioral and social outcomes (Cohen et al., 1982; Cook et al., 1985; Ginsburg-Block et al., 2006). A brief review of the latter meta-analyses follows.

First, Cohen et al. (1982) examined school tutoring programs across 65 studies with students in grades 1 through 12. Effect sizes for academic achievement were reported for tutors and tutees as 0.33 and 0.40, respectively; students who participated in peer tutoring outperformed students in control groups on content area tests. The authors reported that elementary school students benefitted more than their middle and high school peers, and that cross-age tutoring produced greater academic gains than same-age tutoring ( $ES = 0.35$  and  $ES = 0.28$ , respectively). They also

found that peer tutoring had positive effects on students' attitudes ( $ES = 0.29$  for tutees,  $ES = 0.42$  for tutors) but not self-esteem ( $ES = 0.09$  for tutees,  $ES = 0.18$  for tutors). Second, Cook et al. (1985) examined 19 studies of peer tutoring arrangements in which elementary and secondary school students with disabilities served as tutors and tutees, the majority of whom were students with EBD and LD. Both tutors and tutees made academic gains as a result of participating in peer tutoring ( $ES = 0.59$  and  $ES = 0.65$ , respectively). However, no effect was found for tutors on improvements in self-esteem ( $ES = -0.06$ ) and a small effect was found for tutees ( $ES = 0.12$ ). Mixed findings were reported for attitudes toward the subject areas being tutored ( $ES = 0.25$  for tutors,  $ES = 0.86$  for tutees) and for behavior rating measures ( $ES = 0.89$  for tutors,  $ES = 0.10$  for tutees). Third, Ginsburg-Block et al. (2006) examined peer-assisted learning interventions across 36 studies focused on elementary school students, reporting a mean effect size of 0.48 for academic achievement. An effect size of 0.40 was noted for self-concept (15 studies), 0.52 for social skills (30 studies), and 0.65 for behavior (12 studies). Moreover, the use of interdependent group reward contingencies was found to moderate self-concept and social outcomes but not behaviors.

These three meta-analyses provide valuable information on the direct and collateral impact of peer tutoring on students' behavioral and social outcomes. However, there are some limitations. First, Cohen et al. (1982) reported effect sizes for attitude and self-concept but failed to examine both social skills and behavioral measures. Second, Cook et al. (1985) examined attitude, self-concept, and behaviors among students with disabilities. However, data were not disaggregated by disability category, and descriptions of behaviors, behavior rating scales, and behavior checklists were not provided. In addition, social skills measures were not examined. Third, Ginsburg-Block et al. (2006) examined self-concept, social skills, and behavioral outcomes. However, the types of behaviors examined were limited to "behaviors necessary for classroom learning"

(p. 743). Furthermore, the analyses were also limited to elementary school students. Fourth, of these three meta-analyses, only one (Ginsburg-Block et al., 2006) provided confidence intervals for the reported effect sizes. Fifth, none of the three meta-analyses examined disability, a reciprocal tutoring format, or direct and collateral effects as potential moderators. Sixth, and most germane to the current study, none of the existing meta-analyses of social and behavioral benefits of peer tutoring included SCR studies.

### NEED FOR CURRENT SINGLE-CASE META-ANALYSIS

Single-case research methods can provide a rigorous evaluation of applied interventions and can be used to evaluate the efficacy of interventions used in schools (Horner et al., 2005; Kratochwill et al., 2010). Additionally, meta-analytic methods for summarizing SCR studies are becoming more widely used (Burns, Zaslofsky, Kanive, & Parker, 2012; Methe, Kilgus, Neiman, & Riley-Tillman, 2012; Vannest, Davis, Davis, Mason, & Burke, 2010). However, studies using SCR designs have often been excluded from meta-analyses (Allison & Gorman, 1993; Cumming, 2012), most likely because of an absence of standards for quality SCR and the calculation of effect sizes for these types of data. This omission is unfortunate because many peer tutoring studies have used SCR designs (e.g., Greenwood, 1991; Kamps et al., 1999; Mahedy et al., 1988). However, recent developments in articulating standards for high-quality SCR designs (Cooper, 2011; Horner et al., 2005; Horner & Kratochwill, 2012; Kratochwill et al., 2010) and methodology for standardizing, aggregating, and analyzing SCR results (Parker, Vannest, Davis, & Sauber, 2011) make it possible to determine the efficacy of evidence-based interventions using these data.

In addition to reporting estimates of the size of the effect, the American Psychological Association (2010) advocates for including confidence intervals when reporting research findings (Wilkinson & APA Task Force on Statistical Inference, 1999). Moreover, an ef-

fect size with confidence intervals provides a standardized metric that allows for aggregating results across studies. Such aggregation is important for SCR because the methodology depends on direct and systematic replication to establish the efficacy of an intervention (Horner et al., 2005).

### PURPOSE AND RESEARCH QUESTIONS

Peer tutoring meta-analyses conducted to date on behavioral and social measures have only synthesized research from group design studies, to the exclusion of the direct and collateral effects of peer tutoring from the SCR literature. The goal of the current meta-analysis was to begin to fill this void by synthesizing the contributions of the SCR peer tutoring literature on students' behavioral and social outcomes. Given the link between achievement and behavioral and social variables, academic data will also be investigated.

The current meta-analysis extends the existing peer tutoring SCR meta-analysis (Bowman-Perrott et al., 2013a) by investigating nonacademic outcomes associated with peer tutoring, as well as the relation between nonacademic and academic outcomes. In addition, it contributes to the peer tutoring literature in the following ways. This study is (a) the first meta-analysis of peer tutoring studies in nearly 30 years to report behavioral and social outcomes for both elementary and secondary school students; (b) the first nonacademic peer tutoring meta-analysis in nearly a decade to examine the relation between behavioral and social outcomes and academic achievement; (c) the first nonacademic peer tutoring meta-analysis in nearly a decade to provide support for peer tutoring based on American Psychological Association standards of reporting effect sizes with confidence intervals; (d) the only peer tutoring meta-analysis to examine disability, academic engagement, the effect of a reciprocal format, and collateral and direct effects as potential moderators; and (e) the first peer tutoring meta-analysis to investigate the contribution of SCR studies that examine social and behavioral out-

comes. This meta-analysis addressed the following questions:

1. What is the overall effect size for behavioral and social outcomes across SCR studies?
2. What is the effect of potential moderators on behavioral and social outcomes?
3. What is the correlation between academic achievement and behavioral and social measures?

## METHOD

### Literature Search and Inclusion Criteria

A search of the literature was conducted to identify relevant studies published in peer-reviewed journals using the following databases: (a) Education Full Text, (b) Educational Resources Information Center (ERIC), and (c) PsycINFO. Search terms included *peer tutoring*, *reciprocal peer tutoring*, *classwide peer tutoring*, *peers as tutors*, *peer-mediated instruction*, and *peer-assisted learning*. A total pool of 1,758 articles was found, which was reduced by limiting the search to include the terms *behavior* and *social*. The revised search resulted in 240 studies. Abstracts were reviewed to determine whether behavioral and social variables were investigated alone or in addition to academic achievement. Finally, an ancestral search was conducted for studies included in published peer tutoring literature reviews.

After the initial search, the following were excluded from the group of 240 studies: (a) duplicate studies; (b) studies involving tutoring between college students; (c) studies in which college students, parents, or other adults served as tutors; (d) studies that focused on peer tutoring in an academic domain that did not include a measure of social or behavioral outcomes; (e) reviews of the literature; and (f) descriptive reviews or evaluations of peer tutoring programs. Selection criteria for inclusion required (a) publication in peer-reviewed journals between 1966 and 2012, (b) an SCR design, (c) peer tutoring as an intervention (with or without an academic component) to address social or behavioral outcomes, (d) students in prekindergarten (pre-K) through

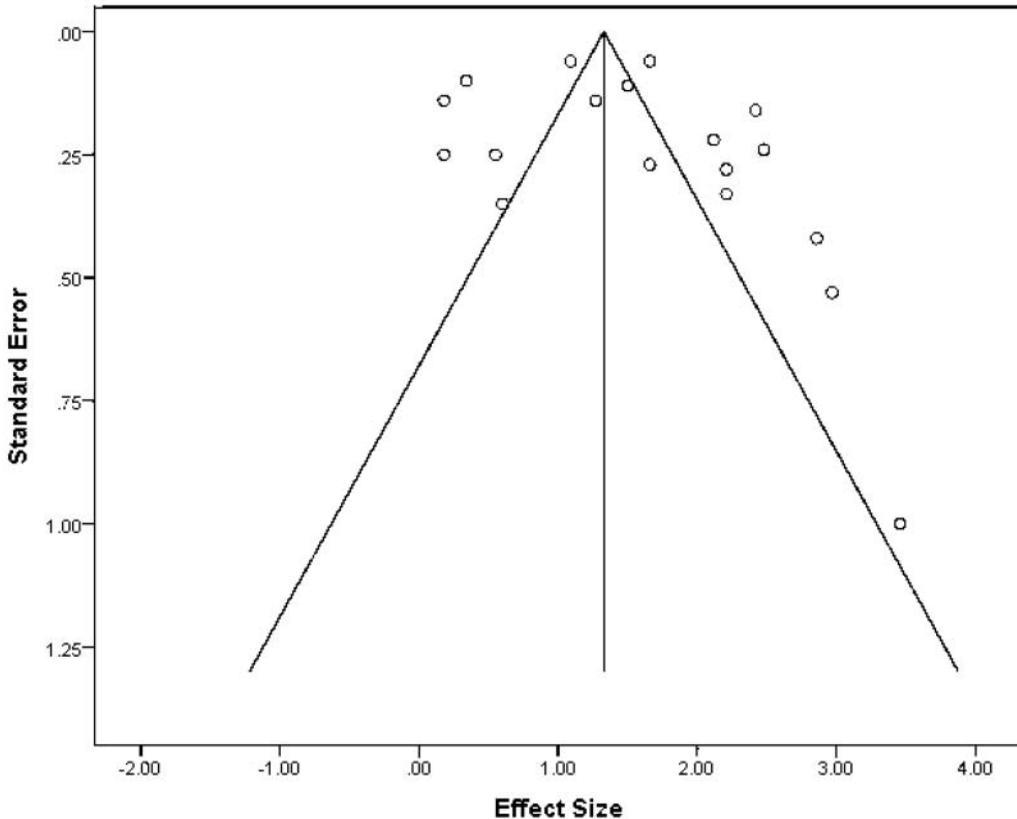
grade 12 serving as tutors, and (e) baseline conditions that did not involve some form of peer tutoring. A pool of 24 studies remained. Limiting the search to peer-reviewed journal articles helped to ensure some minimal standard for methodologic rigor for the included studies.

The 24 single-case studies were evaluated using design standards to demonstrate experimental control (Horner et al., 2005; Horner, Swaminathan, Sugai, & Smolkowski, 2012; Kratochwill et al., 2010). The standards were as follows: (a) the peer tutoring intervention had to be systematically manipulated; (b) studies had to demonstrate experimental control with at least three demonstrations of the effect of peer tutoring at three points in time or with three different phase changes; (c) studies using reversal designs had to include a minimum of four phases with at least five data points per phase to meet standards or with three data points per phase to meet standards with reservations; (d) multiple-baseline designs had to consist of at least six phases with at least five data points per phase to meet standards or with three data points per phase to meet standards with reservations; and (e) alternating-treatments designs had to include at least three alternating treatments with baseline. Studies with designs that did not meet these criteria (e.g., AB, ABA designs) were excluded. A total of 20 SCR studies remained for inclusion in the analyses.

### Fixed-Effects Model

Although neither a fixed-effects model nor a random-effects model is an “exact fit” for SCR data, a fixed-effects model seemed to be a better fit because the number of cases was small for reasonable estimation of variances under the random-effects model (Greenhouse & Iyengar, 2009). Thus, the TauU effect size was calculated within a fixed-effects model (see Parker et al., 2011) using WinPepi (Abramson, 2011). Rather than regarding the studies in this meta-analysis as random samples, they were all regarded as estimates of an unknown “true” effect size. Variations in the true effect size were sought through moderator

Figure 1. Funnel Plot to Detect Publication Bias



Note. Effect sizes represent Cohen's  $d$  values transformed from TauU effect size values.

analysis (Hedges & Olkin, 1985). On the basis of a fixed-effects model, inferences were made about the similarity of the procedures in the peer tutoring interventions implemented across studies. However, some degree of generalization can be assumed because of the inclusion of moderator-variable analyses (Hedges & Olkin, 1985).

### Publication Bias

*Publication bias* refers to the tendency for studies yielding favorable (statistically significant) results to be published (see Rosenthal & DiMatteo, 2001). Publication bias was visually inspected by creating a funnel plot (see Figure 1), a scatter plot depicting treatment effects based on each study's effect size and standard error. The funnel plot was statistically tested for skewness (asymmetry) in Win-

Pepi using the Egger's test (Egger, Smith, Schneider, & Minder, 1997). Whereas the funnel plot showed slight skewness, suggesting publication bias, the intercept for Egger's test (0.98, 90% CI [-1.50, 3.46],  $p = .60$ ) indicated that this result was not significant.

Heterogeneity was measured using the  $H$  and  $I^2$  statistics of Higgins and Thompson (2002), where  $H = 3.2$  (95% CI [2.7, 3.8]) and  $I^2 = 90.2\%$  (95% CI [86.2, 93.0]). Although these results indicate evidence of considerable heterogeneity, two cautions must be considered. First, "the [ $H$ ] test has poor power with few studies . . . and it can therefore be difficult to decide either whether heterogeneity is present or whether it is clinically important" (Higgins & Thompson, 2002, p. 1552). Second, with regard to statistical heterogeneity, "there may be situations when the fixed-effects anal-

ysis is appropriate even when there is substantial heterogeneity of results (e.g., when the question is specifically about the particular set of studies that have already been conducted)” (Hedges & Vevea, 1998, p. 487). The  $I^2$  statistic indicated that most of the variance observed is reliable and not just a result of sampling error. Furthermore, sensitivity analyses indicated that no single study had an undue impact on the findings.

## Effect Size Estimation

### *TauU*

*TauU*, also referred to in this article as *Tau*, is a weighted effect size measure based on nonoverlap between A and B phases (Parker et al., 2011). It is derived from Kendall’s rank correlation and the Mann-Whitney *U* test between groups. Kendall’s rank correlation is an analysis algorithm of time and score, comparing ordered scores and all possible pairs of data. Each pairwise comparison represents an improved score, a score that has not improved, or a tie. Because it is the percentage of all data pairs that show improvement, it measures the tendency for scores to improve over time and calculates monotonic trendedness (the tendency for scores to improve over time in the absence of an intervention). The Mann-Whitney *U* test assesses for nonoverlap between two groups, but with SCR, it is applied to phases rather than groups. Scores from two phases are combined for a cross-group ranking. The separate rankings are then statistically compared for mean differences. The Mann-Whitney *U* algorithm uses two continuous variables: score and time. By replacing the time variable with a dummy code that serves as a numerical stand-in (0/1) to represent Phases A and B, an identical result is produced. This produces the proportion of pairwise comparisons that improve from Phase A to Phase B, also known as the percentage of nonoverlapping data (Scruggs, Mastropieri, & Casto, 1987).

*TauU* does not have to meet the assumptions (e.g., homoscedasticity) required of Cohen’s *d*, Hedges’s *g*, or  $R^2$  (Parker et al., 2011). There are several other advantages to

using *TauU*. Namely, *TauU* (a) has more statistical power than many other nonoverlap or nonparametric techniques; (b) is distribution-free and is suitable for ordinal and interval scaled scores; (c) is complementary of visual analysis; (d) avoids ceiling effects; (e) addresses autocorrelation with the “runs test” as an add-on; (f) controls for level and trend and controls for preexisting trend in Phase A; and (g) provides an overall effect size for summarizing the effects of an intervention. *TauU* values, equivalent to nonoverlap of all pairs (NAP), can be tentatively interpreted as small effects, 0 to 0.65; medium to high effects, 0.66 to 0.92; and large effects, 0.93 to 1.0 (Parker & Vannest, 2009).

### *TauU* Phase Contrasts

*TauU* was calculated for each phase contrast. In the present meta-analysis, data from all A and B phases that met design standards were coded and analyzed. For example, an effect size was calculated for the A1/B1 contrast and a separate effect size was calculated for the A2/B2 contrast in each study. Each effect size was then entered into WinPepi (Abramson, 2011, 2012) and combined using the meta-analysis function to aggregate the data and arrive at a single effect size for a given study. Specifically, multiple A/B phase contrasts (strata) were entered into WinPepi and combined to obtain an effect size for each study (Abramson, 2012; Parker et al., 2011). Statistical significance for *TauU* values was then determined using a 95% confidence interval. When determining whether change is reliable, a 90% to 95% confidence interval is standard (Nunnally & Bernstein, 1994), indicating a reasonable change of 5% to 10% likelihood of error. Because one article contained two studies (Blake, Wang, Cartledge, & Gardner, 2000), the number of articles and the number of studies were not equal in all analyses. Data for maintenance phases were also analyzed for studies that reported them ( $n = 8$ ).

### Cohen’s *d* Transformation

Raw scores or graphed data were not available for one study (DuPaul et al., 1998), which instead reported means and standard

deviations. For this study, Cohen's  $d$  was first calculated and then transformed to TauU (Rosenthal, 1994) to aggregate TauU values across studies to calculate an overall effect size. A three-step process was used. First, a Cohen's  $d$  effect size was calculated by hand using the following formula:

$$\text{Cohen's } d = M_1 - M_2 / \sigma_{\text{pooled}}$$

where  $\sigma_{\text{pooled}} = \sqrt{[(\sigma_1^2 + \sigma_2^2)/2]}$ . The Cohen's  $d$  effect size was also obtained from WinPepi along with the Cohen's  $d$  standard error (Abramson, 2011). Second, Cohen's  $d$  was transformed to Tau using the following formula (Acion, Peterson, Temple, & Amdt, 2006; Parker & Vannest, 2009):

$$\text{Tau} = 1 - (1 - d/3.464)^2$$

Third, the standard error of Tau ( $SE_{\text{Tau}}$ ) was calculated from the Cohen's  $d$  standard error by dividing Cohen's  $d$  by Tau and then dividing the Cohen's  $d$  value by the quotient. For example, dividing a Cohen's  $d$  of 1.39 by a Tau of 0.64 yields a quotient of 2. Dividing the Cohen's  $d$  standard error of 0.57 by 2 yields a quotient of 0.28, the  $SE_{\text{Tau}}$ . Transformed Cohen's  $d$  values were entered into WinPepi to arrive at an effect size for the DuPaul et al. (1998) study.

## Coding

Eighteen study variables were coded including behavior type (social interactions/social skills, disruptive/off-task, and academic engagement behaviors) and disability status. The first author coded each of the studies. Two graduate students were trained on the codes and independently coded each study across all potential moderators in separate Excel spreadsheets. The formula used for interobserver agreement was the sum of agreement/total number of agreements plus disagreements times 100, which is an appropriate reliability calculation for noncategorical data. Initial agreement across variables ranged from 81% to 100%. Disagreements were resolved after the first author and graduate students reread

and discussed the articles, resulting in 100% final agreement across all codes. The coding guide is available from the first author on request.

## Potential Moderators

Articles were coded to examine six potential moderator variables hypothesized to have an impact on social and behavioral outcomes. They were selected because they (a) had been investigated in previous meta-analyses (e.g., grade level, the use of reward) or (b) had not yet been investigated in nonacademic systematic reviews of the peer tutoring literature (e.g., disability status, reciprocal format) but represent theoretically relevant aspects of peer tutoring interventions or student characteristics. They were as follows: use of reward versus no reward, reciprocal versus nonreciprocal, same-age versus cross-age, behavior type, disability status, and direct versus collateral effects.

A reliable difference, or a difference that is large enough so that it cannot be accounted for solely chance, given the number of participants and data points, was calculated for levels of each of the potential moderators, with the exception of behavior type and disability group. The reliable difference formula below was used to determine whether different levels of a given moderator (e.g., elementary (L1) and secondary (L2) grade levels) differed statistically from one another.

$$(L1-L2)/\sqrt{[(SE_{\text{Tau}1})^2 + (SE_{\text{Tau}2})^2]}$$

The statistical test ( $p < .05$ ) indicated whether such findings were credible; that is, whether the results would change substantially over several retestings. A 95% confidence interval ( $\alpha = .05$ ) was set for each effect size. Reliable difference  $z$  test scores and  $p$  values are reported in the Results section.

## Reward Versus No Reward

Studies that included some type of reward structure (e.g., independent, interdependent) and reward type (viz., tangible, intangible) were considered to use a reward. Levels

included the use of rewards, coded *reward*, and no rewards, coded *no rewards*.

### **Reciprocal Versus Nonreciprocal**

Two levels were used in coding this potential moderator. Peer tutoring arrangements in which students in a dyad or triad reversed roles as tutor and tutee during the peer tutoring session were coded *reciprocal*. Those that did not use reciprocal tutoring were coded *nonreciprocal*.

### **Same Age Versus Cross Age**

This variable was assigned two levels. Studies that included students in the same room serving as tutors and tutees (e.g., Class-Wide Peer Tutoring) and involved students of the same age or grade level were coded *same age*. Studies that reported tutors and tutees across different ages or grade levels were coded *cross age*.

### **Behavior Type**

Behaviors described in each study were assigned to one of three categories: social interactions/social skills (coded *SIS*), disruptive/off-task (coded *DOT*), and academic engagement (coded *AE*). Only one study focused on social skills (Prater, Serna, & Nakamura, 1999). It was grouped with studies that examined related social interaction behaviors. Behaviors in the social interactions/social skills category included positive and negative peer interactions, positive and negative statements, and giving and receiving positive feedback. Disruptive/off-task behaviors included non-compliance, arguing, hitting, making faces, and making gestures intended to provoke others. Academic engagement included observable academic-related tasks such as reading aloud, writing, reading silently, talking about academic tasks, and attending to assigned tasks.

### **Disability Status**

Studies were coded based on students' reported primary disability. That is, studies that included students with ADHD or attention deficit disorder (ADD) under other health impairment (OHI) were coded together. Studies

involving students with EBD, LD, autism, MR/ID, and S/L were coded as such.

### **Direct Versus Collateral Effects**

If the primary purpose of a study was to examine academic outcomes with behavioral and/or social variables investigated secondarily, it was coded *collateral*. If the primary purpose was to examine the effects of peer tutoring on students' social and/or emotional outcomes or if academic and nonacademic variables were both examined, it was coded *direct*.

### **Dummy Coding**

Dummy coding assigns an artificial variable (e.g., 0 or 1) to represent the overlap or nonoverlap of data points between A and B phases. Dummy coding for A and B phases (e.g., A1/B1, A2/B2) was calculated by hand for each study. Dummy coding involved visually inspecting graphed data in adjacent A and B phases to assign dummy codes based on the overlap or nonoverlap of data points (see Parker et al., 2011). Dummy codes for each A and B phase were then entered into the TauU calculator (Vannest, Parker, & Gonen, 2011) to obtain TauU values and the  $SE_{\text{TauU}}$  for each phase contrast. Next, TauU and  $SE_{\text{TauU}}$  values were entered separately into WinPepi for each study and "combined" to arrive at an effect size and confidence interval.

Reliability was calculated for 20% ( $n = 43$ ) of all of the A/B phase contrasts across the 19 studies for which graphed or raw data were provided. The formula used for inter-coder reliability was the sum of agreement/total number of agreements plus disagreements times 100 (House, House, & Campbell, 1981). Reliability was 100%.

### **Correlations With Academic Achievement**

Effect sizes for behavioral, social, and academic outcomes across the nine studies reporting these data were entered into an Excel spreadsheet in separate cells, and correlations were calculated using the CORREL function. Because of the varied nature of the behavior

types (e.g., disruptive and off-task behaviors are not favorable, academic engagement and social interactions/social skills are favorable), separate correlation coefficients were calculated for (a) academic achievement and disruptive/off-task behaviors, (b) academic achievement and social interactions/social skills, and (c) academic achievement and academic engagement.

## RESULTS

### Descriptive Data Summary

An overview of individual study characteristics is provided in Table 1. The 20 studies examined in this meta-analysis included 128 participants in pre-K through grade 12 and included 234 phase contrasts. Two studies included pre-K students, nine focused on elementary school-aged students (kindergarten through grade 5), and six included secondary school-aged students (grades 6 through 12). Three studies included both elementary and secondary school students. Treatment fidelity data were collected for 13 of the studies, with a mean rating of 90%. Social validity data were collected in 12 studies from teachers and students; favorable ratings were reported by both. That is, teachers reported being satisfied with peer tutoring procedures and improvements in student behaviors; students reported that they liked peer tutoring procedures and enjoyed working with peer partners.

Average interobserver agreement for all behaviors across the 20 studies was 90.5% (range = 67% to 100%). The most common research design was a multiple-baseline design ( $n = 13$ ). One study used a multiple-probe design, meeting the same criteria as the multiple-baseline design studies. Five studies used an ABAB design; one study used an alternating-treatment design, comparing a peer tutoring condition with a “no peer tutoring” condition. Of the 18 studies that provided graphed data, 10 met SCR design quality standards; 8 met standards with reservations (Kratowill et al., 2010). Eight studies collected maintenance data; phases lasted an average of seven sessions. All studies used direct observation to collect behavioral and social outcome data.

One study (Franca, Kerr, Reitz, & Lambert, 1990) also used peer nominations to assess social status (a measure of peer acceptance or rejection). The two studies by Blake et al. (2000) measured different behavioral outcomes for students in two different schools.

With regard to participant characteristics, gender was reported in all but one study. Of participants, 64 were male and 26 were female. The ethnicity of participants in studies reporting these data were as follows: Caucasian ( $n = 39$ ), African American ( $n = 21$ ), and Hispanic/Latino ( $n = 6$ ). None of the studies reported the inclusion of English language learners. Participants’ problem behaviors included fighting, off-task behaviors, class disruption, noncompliance, profanity, being disrespectful, fidgeting, and not following directions.

### Overall Effect, Potential Moderators, and Correlations With Academic Achievement

The overall effect of peer tutoring on participants’ social and behavioral outcomes was examined across all 20 studies (representing 20 effect sizes) with a weighted mean effect size of 0.62 ( $SE = 0.02$ , 95% CI [0.58, 0.66]). The finding for the first research question was that peer tutoring had a small to moderate effect on social and behavioral measures. Figure 2 illustrates the range of effect sizes and confidence intervals across all of the studies at a 95% confidence level. Moderators examined were (a) reward versus no reward, (b) reciprocal versus nonreciprocal, (c) same age versus cross age, (d) behavior type, (e) disability status, and (f) direct versus collateral. All four moderators for which reliable difference values could be calculated were confirmed as significant (see Table 2). The correlation between behavioral and social outcomes and academic achievement was examined as well. The second research question is answered across the next six sections. The third research question is addressed in the following section. Obtained TauUs were transformed to Cohen’s  $d$  values (Rosenthal, 1994) using the formula that follows. They are

**Table 1. Summary of Study Characteristics**

Study	Behavior Type	Primary Disability <sup>a</sup>	Reciprocal	Reward	Same Age	Direct Versus Collateral	Dependent Measure(s)/Target Behavior(s)	Summary of Social/Behavioral Findings	Research Design	Grade Level	N <sup>c</sup>
Barton-Arwood, Webby, & Falk (2005) <sup>b</sup>	DOT, AE	EBD, ADD, LD	Yes	Yes	Yes	Collateral	Negative talk, aggression, AE	Generally, low rates of negative talk and aggressive behavior and higher rates of engagement were observed for students participating in a reading intervention. However, engagement tended to be high at baseline as well. Observed changes in student behavior may have been the result of other factors (e.g., teacher behaviors) that were not measured.	MBD across participants	ES	6
Blake, Wang, Cartledge, & Gardner (2000): Study 1	SIS	EBD	No	No	No	Direct	Positive and negative behaviors (initiations and terminations of a game), positive and negative statements	Positive behaviors and communication began to increase before the intervention. Similarly, negative behaviors and communication began to decrease before peer-mediated social skills instruction began. Both trends continued during intervention and maintenance phases.	MBD across behaviors	MS	3
Blake, Wang, Cartledge, & Gardner (2000): Study 2	SIS	EBD	Yes	No	No	Direct	Verbally and nonverbally supportive and abusive behaviors	Verbally and nonverbally supportive behaviors increased during the intervention; verbally and nonverbally abusive behaviors decreased. For all participants, behaviors remained at intervention levels or improved during the maintenance phase.	MBD across participants	ES, MS	6
Bowman-Perrott (2009) <sup>b</sup>	DOT	EBD	Yes	No	No	Collateral	On- and off-task behaviors	Peer tutoring resulted in an increased amount of time spent on-task and a reduction in off-task behaviors.	Alternating treatment	MS, HS	8

(Table 1 continues)

**Table 1. Continued**

Study	Behavior Type	Primary Disability <sup>a</sup>	Reciprocal	Reward	Same Age	Direct Versus Collateral	Dependent Measure(s)/Target Behavior(s)	Summary of Social/Behavioral Findings	Research Design	Grade Level	N <sup>c</sup>
Brady (1997) <sup>b</sup>	SIS	S/L	Yes	Yes	Yes	Collateral	Positive and negative initiations, positive and negative responses, length of social interaction	The majority of the initiations and responses were positive. The length of social interaction with peers increased for 4 students during peer tutoring; it did not increase for the remaining 4 students. Results may reflect social interactions taking place in a "contrived" versus a more natural setting.	MBD across behaviors	PS	8
Christensen, Young, & Marchant (2007)	DOT	LD, at-risk EBD	No	Yes	Yes	Direct	Socially appropriate classroom behaviors	A peer partner who modeled appropriate classroom behaviors was part of the target student's behavior intervention planning. Although components of the intervention were not analyzed separately, the use of a peer to facilitate behavior change was identified as an effective intervention component.	ABAB	ES	1
DuPaul, Ervin, Hook, & McGoey (1998)	DOT	ADHD	Yes	No	Yes	Direct	Active on-task, passive on-task, off-task, fidgeting	Peer tutoring resulted in increased active on-task behaviors in students with ADHD. A reduction in passive on-task, off-task, and fidgeting behaviors during tutoring sessions was also observed.	ABAB	ES	18
Franca, Kerr, Reitz, & Lambert (1990) <sup>b</sup>	SIS	EBD	No	No	Yes	Collateral	Positive and negative social interactions, attitudes toward math, self-concept, social status (peer acceptance or rejection)	Peer tutoring resulted in increased positive social interactions and a reduction in negative social interactions between tutors and tutees. Improved attitudes toward math were observed; minimal improvements were noted for self-concept and social status.	MBD across tutor-tutee dyads	MS	8

(Table 1 continues)

**Table 1. Continued**

Study	Behavior Type	Primary Disability <sup>a</sup>	Reciprocal	Reward	Same Age	Direct Versus Collateral	Dependent Measure(s)/Target Behavior(s)	Summary of Social/Behavioral Findings	Research Design	Grade Level	N <sup>c</sup>
Gumpel & Frank (1999)	SIS	None	No	No	No	Direct	No social interactions, positive social interactions	Results indicated that cross-age peer tutoring increased the frequency of positive social interactions for tutors and tutees. Instances of "no social interactions" were reduced.	MBD across participants	K, MS	4
Kamps, Dugan, Potucek, & Collins (1999): Study 1	SIS	Autism	No	No	No	Direct	Social interactions	An increase in social interactions was observed for the student with autism who served as a tutor.	ABAB	ES	1
Lo & Cartledge (2004) <sup>b</sup>	DOT	LD (1 student)	Yes	Yes	Yes	Collateral	Off-task behaviors	Results indicated a reduction in off-task behaviors among the majority of the participants during peer tutoring, as well as peer tutoring with a group contingency. Only 1 student benefitted more from peer tutoring with the group contingency than peer tutoring alone.	MBD with reversal	ES	8
Locke & Fuchs (1995)	DOT, SIS	EBD, ADHD	Yes	Yes	Yes	Direct	On-task behaviors, positive social interactions	Peer-mediated reading instruction resulted in improvements in on-task behaviors and positive social interactions for all participants.	ABAB	ES	3
Maheady & Sainato (1985)	SIS	LD (3 students)	No	Yes	Yes	Direct	Positive and negative initiations	An increase in positive initiations and a decrease in negative initiations were reported for all participants.	ABAB	ES	6
McDonnell, Mathot-Buckner, Thorson, & Fister (2001) <sup>b</sup>	DOT, AE	MR/ID (3 students)	Yes	No	No	Direct	Competing behaviors <sup>d</sup>	Peer tutoring, combined with a multilevel curriculum and accommodations, resulted in improved levels of academic responding and decreased levels of competing behaviors for students with and without disabilities.	MPD across participants	MS	6

(Table 1 continues)

Table 1. Continued

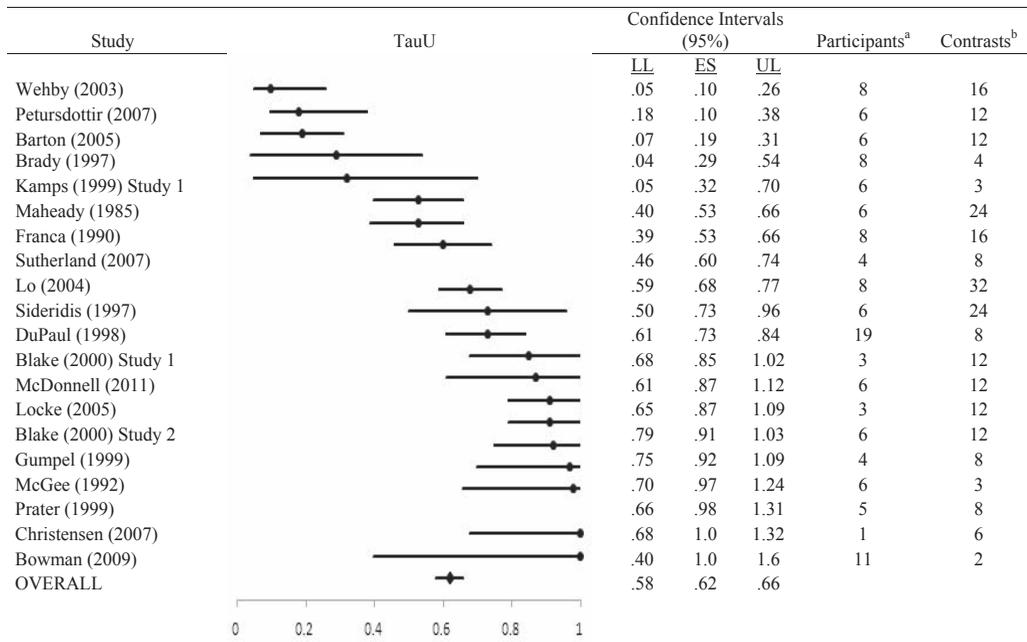
Study	Behavior Type	Primary Disability <sup>a</sup>	Reciprocal	Reward	Same Age	Direct Versus Collateral	Dependent Measure(s)/Target Behavior(s)	Summary of Social/Behavioral Findings	Research Design	Grade Level	N <sup>c</sup>
McGee, Almeida, Sulzer-Azaroff, & Feldman (1992)	SIS	Autism	No	No	No	Direct	Social interactions	Peer incidental teaching promoted reciprocal interactions among 3 children with autism and their peers without disabilities. During fading, 1 student increased initiating social interactions with peers, 1 did not, and the third showed "inconsistent" increases. Results suggest that adding play-related common stimuli to academic peer tutoring may increase social interactions between students with autism and their peers without disabilities.	MBD across participants	PS	3
Petursdottir, McComas, & McMaster (2007)	SIS	Autism (1 student)	No	No	Yes	Direct	Social interactions	Results indicated that adolescents with LD can be trained to teach social skills to their peers without disabilities. Students trained by their peers acquired the social skills of giving positive feedback and contributing to discussion more quickly than when taught by their teachers. However, students taught by their peers did not maintain those skills. Students trained by their teachers did maintain those skills. Both groups made less improvement in accepting negative feedback.	MBD across participants	K	4
Prater, Serna, & Nakamura (1999)	SIS	LD	No	No	Yes	Direct	Giving positive feedback, contributing to discussions, accepting negative feedback	Results indicated that adolescents with LD can be trained to teach social skills to their peers without disabilities. Students trained by their peers acquired the social skills of giving positive feedback and contributing to discussion more quickly than when taught by their teachers. However, students taught by their peers did not maintain those skills. Students trained by their teachers did maintain those skills. Both groups made less improvement in accepting negative feedback.	MBD across skills	MS	17
Sideridis et al. (1997) <sup>b</sup>	DOT, AE	LD (2 students), MR/ID (1 student)	Yes	Yes	Yes	Direct	Academic responding, social interactions, competing behaviors <sup>a</sup>	Peer tutoring significantly increased levels of academic responding and positive social interactions. A decrease in competing behaviors was also reported for all participants.	ABAB	MS	6

(Table 1 continues)

Table 1. Continued

Study	Behavior Type	Primary Disability <sup>a</sup>	Reciprocal	Reward	Same Age	Direct Versus Collateral	Dependent Measure(s)/Target Behavior(s)	Summary of Social/Behavioral Findings	Research Design	Grade Level	N <sup>c</sup>
Sutherland & Snyder (2007) <sup>b</sup>	DOT, AE	EBD	Yes	Yes	Yes	Direct	Active responding, disruptive behaviors	Peer tutoring and self-graphing resulted in a decrease in disruptive behaviors and an increase in active responding for all 4 students with EBD. Follow-up data showed that students generally maintained the desired behaviors.	MBD across participants	MS	4
Wehby, Falk, Barton-Arwood, Lane, & Cooley (2003) <sup>b</sup>	DOT, AE	EBD, ADHD, LD, MR/ID	Yes	No	No	Collateral	Attending (engagement), inappropriate behaviors	Results showed that peer tutoring combined with a reading intervention resulted in slight improvement in the average percentage of attending time during instruction. No improvements were made in rates of inappropriate behaviors.	MBD across groups	K, ES	8

*Note.* DOT = disruptive/off-task behavior; SIS = social interaction/social skills; AE = academic engagement; EBD = emotional/behavioral disorder; S/L = speech/language disorders; LD = learning disabilities; ADHD = attention deficit hyperactivity disorder; ADD = attention deficit disorder; MR/ID = mental retardation/intellectual disability; PS = preschool; K = kindergarten; ES = elementary school; MS = middle school; HS = high school; MBD = multiple baseline design; MPD = multiple probe design; None = no disability reported.  
<sup>a</sup>Disability refers to participants' reported primary disability or students identified as behaviorally at risk.  
<sup>b</sup>Studies reported academic as well as social and/or behavioral data.  
<sup>c</sup>N refers to the number of participants for whom data were reported (often identified in studies as target students).  
<sup>d</sup>Competing behaviors include refusal to follow class rules or teacher instructions.

**Figure 2. Forest Plot for Effects of Peer Tutoring on Students' Social and Behavioral Outcomes**

Note. Circles represent effect sizes from individual studies. They are proportionate to their weight in the overall effect size. Circles and confidence interval bars represent the precision of each study effect size. The diamond represents the overall effect size across all 20 studies. The number of participants (a) and the number of phase contrasts (b) are reported. ES = effect size; LL = lower limit; UL = upper limit.

presented in the Discussion section to help interpret the findings alongside the effect sizes reported in the meta-analyses of group design studies.

$$d = 3.464 \times [1 - \sqrt{(1 - \text{TauU})}]$$

### Reward Versus No Reward

Studies that did not use rewards as a part of the peer tutoring program yielded a larger effect size ( $ES = 0.68$ ,  $SE = 0.03$ , 95% CI [0.63, 0.74]) than those that did ( $ES = 0.56$ ,  $SE = 0.03$ , 95% CI [0.50, 0.61]). The values obtained from the reliable difference formula were  $z = 2.86$ ,  $p = .01$ . The effect size for students with LD ( $ES = 0.75$ ,  $SE = 0.35$ , 95% CI [0.06, 1.00]) participating in peer tutoring interventions using rewards was higher than the effect size for students with EBD ( $ES = 0.45$ ,  $SE = 0.04$ , 95% CI [0.37, 0.53]). Although these data could not be analyzed for students representing the other disability

groups because there were only one or two studies representing each category, a summary of their findings is provided. Only one of the three studies involving MR/ID used rewards (Sideridis et al., 1997); none of the studies that involved students with autism or ADHD/ADD used them. Brady (1997) included the use of reward for students with S/L difficulties.

The effect size for studies of disruptive/off-task behaviors using rewards ( $ES = 0.65$ ,  $SE = 0.04$ , 95% CI [0.58, 0.72]) was higher than that in those not using rewards ( $ES = 0.53$ ,  $SE = 0.05$ , 95% CI [0.44, 0.62]). The opposite was true for studies addressing social skills/social interactions. Studies of social skills/social interactions that did not use rewards had an effect size of 0.76 ( $SE = 0.03$ , 95% CI [0.69, 0.82]), whereas studies that did use rewards had an effect size of 0.57 ( $SE = 0.05$ , 95% CI [0.46, 0.68]). The effect size for the use of rewards in studies examin-

**Table 2. Summary of Effect Size Results for Moderator Variables**

Moderator Variable	No. of Studies	No. of Participants	ES	SE	95% CI		$z^a$	$p^a$
					LL	UL		
Reward	9	28	0.56	0.03	0.50	0.61		
No reward	11	100	0.68	0.03	0.63	0.74	2.86	.01
Reciprocal	11	79	0.60	0.02	0.56	0.65		
Nonreciprocal	9	49	0.68	0.03	0.62	0.74	2.22	.02
Same age	12	83	0.52	0.02	0.48	0.56		
Cross age	8	45	0.88	0.04	0.81	0.95	8.00	.01
Behavior type <sup>b</sup>								
SIS	11	56	0.69	0.03	0.63	0.75		
DOT	10	72	0.60	0.03	0.55	0.66		
AE	6	30	0.38	0.05	0.29	0.48		
Disability								
OHI (ADHD, ADD) <sup>c</sup>	4	23	0.63	0.06	0.52	0.74		
EBD	8	55	0.61	0.03	0.55	0.67		
LD	6	24	0.57	0.05	0.47	0.68		
Autism	3	5	0.49	0.09	0.32	0.66		
SL	1	8	0.29	0.13	0.04	0.54		
MR/ID	3	5	0.93	0.15	0.63	1.00		
Direct	14	79	0.75	0.02	0.70	0.80		
Collateral	6	49	0.43	0.03	0.37	0.49	8.88	.01

*Note.* SIS = social interaction/social skills; DOT = disruptive/off-task behavior; AE = academic engagement; OHI = other health impaired; ADHD = attention deficit hyperactivity disorder; ADD = attention deficit disorder; EBD = emotional and behavioral disorders; LD = learning disabilities; SL = speech/language disability; MR/ID = mental retardation/intellectual disability; ES = effect size; SE = standard error; CI = confidence interval; LL = lower level; UL = upper level.

<sup>a</sup>Reliable difference  $z$  test scores and corresponding  $p$  values are reported for the four moderators for which levels were able to be analyzed (e.g., reciprocal tutoring levels were coded as *yes reciprocal* or *no nonreciprocal*).

<sup>b</sup>Six studies measured 2 of the 3 behavior types, so the reported numbers do not equal the total number of studies.

<sup>c</sup>Only 1 of the 23 students in this disability category was identified as having ADD.

ing academic engagement (ES = 0.52, SE = 0.06, 95% CI [0.39, 0.64]) was much larger than that in those that did not use rewards (ES = 0.21, SE = 0.07, 95% CI [0.06, 0.35]). In addition, the confidence intervals did not overlap for the two levels of this potential moderator.

### **Reciprocal Versus Nonreciprocal**

Studies that did not use reciprocal tutoring as an intervention component had a higher effect size (0.68, SE = 0.03, 95% CI [0.62, 0.74]) than those that did (ES = 0.60, SE = 0.02, 95% CI [0.56, 0.65]). The reliable difference values were  $z = 2.22$ ,  $p = .02$ . Studies using a nonreciprocal tutoring format were more likely to involve cross-age tutoring.

Conversely, those using a reciprocal arrangement were more likely to involve same-age tutoring. Students with EBD, ADHD, MR/ID, and S/L were more likely to be a part of peer tutoring interventions that used reciprocal tutoring, whereas students with LD were more likely to be a part of nonreciprocal tutoring arrangements. All three studies involving students with autism used a nonreciprocal format. By behavior type, studies focusing on social interactions and social skills that used a reciprocal format had a higher effect size (0.81, SE = 0.05, 95% CI [0.72, 0.91]) than those that did not (ES = 0.65, SE = 0.03, 95% CI [0.58, 0.72]). Analyses were not possible for disruptive/off-task behaviors because all but

one disruptive/off-task behavior study used a reciprocal format. These data were not available for academic engagement because all seven of the studies used reciprocal peer tutoring.

### **Same Age Versus Cross Age**

Studies that involved cross-age tutoring yielded a larger effect size ( $0.73$ ,  $SE = 0.03$ , 95% CI [0.66, 0.80]) than those using same-age tutoring ( $ES = 0.57$ ,  $SE = 0.02$ , 95% CI [0.53, 0.62]). The reliable difference values were  $z = 8.00$ ,  $p = .01$ . Students with LD ( $ES = 0.72$ ,  $SE = 0.02$ , 95% CI [0.67, 0.77]) benefitted more from same-age tutoring than students with EBD ( $ES = 0.46$ ,  $SE = 0.04$ , 95% CI [0.39, 0.53]), for whom cross-age tutoring yielded a higher effect size ( $0.68$ ,  $SE = 0.04$ , 95% CI [0.59, 0.76]). No studies involving students with LD used cross-age tutoring. Students with S/L participated in same-age peer tutoring. Students with autism, MR/ID, and ADHD participated in both same- and cross-age tutoring. Further examination of the data by behavior type showed that disruptive/off-task behavior studies using a same-age format had a higher ES ( $0.67$ ,  $SE = 0.03$ , 95% CI [0.61, 0.73]) than those that did not ( $ES = 0.29$ ,  $SE = 0.07$ , 95% CI [0.15, 0.43]). Nonoverlapping confidence intervals were observed between studies examining disruptive/off-task behaviors across these formats. Studies of social skills/social interactions that used cross-age tutoring had a higher ES ( $0.87$ ,  $SE = 0.04$ , 95% CI [0.80, 0.95]) than those that used same-age tutoring ( $ES = 0.54$ ,  $SE = 0.04$ , 95% CI [0.46, 0.62]); the confidence intervals did not overlap for these variables. Academic engagement studies that used same-age tutoring had a larger ES ( $0.57$ ,  $SE = 0.06$ , 95% CI [0.39, 0.64]) than those that used cross-age tutoring ( $ES = 0.21$ ,  $SE = 0.07$ , 95% CI [0.06, 0.35]).

### **Behavior Type**

The majority of the studies ( $n = 11$ ) focused on social interaction/social skills, disruptive/off-task behaviors made up 10 studies, and academic engagement ( $n = 6$ ) comprised the smallest group of studies. Because six

studies measured two of the three behavior types, the number of studies across behavior types does not equal the total number of studies. Social interactions and social skills ( $ES = 0.69$ ,  $SE = 0.03$ , 95% CI [0.63, 0.75]) included positive and negative peer interactions, positive and negative statements, and giving and receiving positive feedback. Disruptive/off-task behaviors ( $ES = 0.60$ ,  $SE = 0.03$ , 95% CI [0.55, 0.66]) yielded a similar effect to social interactions and social skills. Academic engagement ( $ES = 0.38$ ,  $SE = 0.05$ , 95% CI [0.29, 0.48]) produced the smallest effect.

### **Disability Status**

All but one study included students with disabilities; the remaining study identified participants as “socially rejected” and “isolated” (Gumpel & Frank, 1999, p. 115). The effect of peer tutoring across behavior types showed the following: Students with MR/ID and ADHD/ADD benefitted most from peer tutoring to reduce disruptive and off-task behaviors ( $ES = 0.66$  and  $ES = 0.62$ , respectively). Students with LD ( $ES = 0.55$ ) and EBD ( $ES = 0.33$ ) were not as likely to experience reductions in these behaviors. Students with ADHD and ADD and students with EBD benefitted most from social skills and social interactions ( $ES = 0.84$  and  $ES = 0.78$ , respectively). Participants with S/L disabilities and students with autism were less likely to show improvements in these areas. Students with MR/ID showed increased academic engagement ( $ES = 0.57$ ) compared with their peers with ADHD/ADD ( $ES = 0.40$ ), LD ( $ES = 0.36$ ), and EBD ( $ES = 0.20$ ). Students without disabilities showed positive behavioral and social gains ( $ES = 0.92$ ,  $SE = 0.09$ , 95% CI [0.75, 1.00]).

### **Direct Versus Collateral Effects**

Fourteen studies (70%) examined social and behavioral outcomes as the primary purpose of the peer tutoring intervention (investigating direct effects). Six studies investigated the collateral effects of peer tutoring on these outcomes. Studies that investigated the direct effect of peer tutoring on social and

behavioral outcomes yielded an effect size of 0.75 ( $SE = 0.02$ , 95% CI [0.70, 0.80]), whereas those examining collateral effects resulted in an effect size of 0.43 ( $SE = 0.03$ , 95% CI [0.37, 0.49]). The reliable difference results were  $z = 8.88$ ,  $p = .01$ .

### **Correlations With Academic Achievement**

The third research question was answered by calculating Pearson's  $r$  correlations for students' behavioral, social, and academic outcomes. Only 9 of the 20 studies reported data for these analyses. An overall correlation of  $r = 0.57$  was obtained for all behavioral and social variables. Because of the variation across behavior types, three additional correlations were calculated: one for academic achievement and disruptive/off-task behaviors, one for academic achievement and social skills/social interactions, and one for academic achievement and academic engagement. First, the correlation between academic achievement and disruptive/off-task behaviors was  $r = 0.31$ . This correlation represents six studies, 34 phase contrasts, and 64 participants. Second, the correlation for academic achievement and social skills/social interactions was  $r = -0.18$ , representing two studies, 19 phase contrasts, and 16 participants. Third, the correlation between academic achievement and academic engagement was  $r = 0.52$ . This finding represents five studies, 37 phase contrasts, and 30 students. An additional effect size was calculated for academic achievement for the nine studies that reported these data (0.61,  $SE = 0.02$ , 95% CI [0.56, 0.66]).

### **Maintenance**

Most of the studies ( $n = 12$ ) did not report maintenance data. However, for the eight studies that did, the effect size was relatively large (0.79,  $SE = 0.05$ , 95% CI [0.70, 0.88]), representing 43 participants with EBD, LD, and autism. Individual study TauU values ranged from 0.51 to 1.00. Seven of the eight studies that collected maintenance data

addressed social skills and social behaviors; one focused on disruptive behaviors.

## **DISCUSSION**

This meta-analysis examined the direct and collateral effects of peer tutoring on behavioral and social outcomes for elementary and secondary school students. Whereas numerous studies support the efficacy of peer tutoring with regard to academic outcomes, fewer have investigated its impact on nonacademic measures. This is the first peer tutoring meta-analysis to investigate nonacademic outcomes associated with peer tutoring for studies using SCR designs. The overall effect (ES = 0.62) means that students engaged in peer tutoring achieved some benefit in displaying appropriate behaviors, improved social skills, increased positive social interactions with peers, and increased academic engagement.

### **Moderator Variables**

Findings from this meta-analysis support the use of rewards as a moderator of behavioral and social outcomes. Students with LD and secondary students, in particular, derived greater benefit from peer tutoring arrangements that included rewards. Students displaying disruptive and off-task behaviors also benefitted from rewards, regardless of disability status. This finding may prove helpful not only in the implementation of peer tutoring programs, but also interventions that help reduce disciplinary exclusion because disruptive behaviors are all too often met with suspension (Skiba, Peterson, & Williams, 1997). In addition, academic engagement was more likely to improve in studies that used rewards compared with those that did not use rewards. This finding suggests that regardless of age and disability status, students are more motivated to work toward academic goals with some kind of reward structure in place.

Social skills and social interactions yielded the largest effect size among the target behaviors. This is logical given the nature of peer-mediated instruction. The effect sizes obtained for social skills/social interactions (0.69

with a transformed Cohen's *d* value of 1.55) and disruptive/off-task behaviors (0.60 with a transformed Cohen's *d* value of 1.28) were larger than those reported by Ginsburg-Block et al. (2006), who found small effect sizes for social skills (0.28, 95% CI [0.25, 0.32]) and moderate effect sizes for behavior (e.g., on-task; 0.45, 95% CI [0.39, 0.50]). Cook et al. (1985) reported behavior rating measures of 0.89 for tutors and 0.10 for tutees. However, it is unclear whether Cook et al.'s behavior rating measures included a range of problem behaviors or whether social skills were included. Only one study in the current meta-analysis examined self-concept. None examined attitudes toward subject matter.

In the present study, the effect size for social skills and social interactions was not larger with the use of rewards, which is the opposite of what Ginsburg-Block et al. (2006) reported. Perhaps positive social interactions with peers serve as their own reward. Future research can further investigate this relation. Rohrbeck et al. (2003) indicated that rewards are "used to foster social motivation" (p. 252). They also noted that the use of rewards provides an avenue for students to work toward a common goal, as well as "the establishment of norms emphasizing academic achievement" (p. 243). The same may be true for establishing social and behavioral norms.

Reciprocal peer tutoring was more frequently used with elementary school students, with students with EBD, and in studies in which academic engagement was examined. Students with LD and autism were more likely to participate in nonreciprocal peer tutoring arrangements. Because peer tutoring studies that used a reciprocal format yielded a larger effect size for social interactions and social skills, more studies should include students with LD and autism in this type of tutoring arrangement.

Cross-age tutoring had the largest effect on students' behavioral and social outcomes among all of the moderator variables (ES = 0.73 with a transformed Cohen's *d* of 1.69). Perhaps this is related to the finding that cross-age peer tutoring yielded a higher effect size than same-age tutoring for students with EBD, who (by definition) experience significant be-

havioral challenges that interfere with everyday functioning. It might be that students with EBD benefitted more socially from cross-age peer tutoring because of the use of older peers serving as effective models for appropriate behaviors and patterns of social interactions. Cohen et al. (1982) also found that cross-age tutoring had a larger effect size (0.49) than studies that did not use cross-age tutoring (ES = 0.29). Same-age tutoring yielded a larger effect for students with LD, for students displaying disruptive/off-task behaviors, and in studies that examined academic engagement.

With regard to overall gains, students with MR/ID showed the largest nonacademic gains from peer tutoring, reflecting decreases in disruptive/off-task behaviors and improvements in academic engagement. Although this effect size only represents three studies and five students, the social benefit participants gained from peer tutoring is encouraging. Moderate overall benefits were found for students with ADHD/ADD and EBD. Students with ADHD/ADD, EBD, and LD showed their greatest gains in improved social skills and social interactions with peers. Students with autism also benefitted socially, achieving small to moderate benefits in increased social interactions with peers.

Moderator analyses showed that studies examining the direct effect of peer tutoring yielded larger effect sizes (e.g., Blake et al., 2000; Prater et al., 1999). This makes sense in light of reports of the increasing numbers of children and youth who experience social-emotional difficulties (National Academy of Sciences, 2009). One possible implication for future research is to investigate whether features of studies that examine direct effects (e.g., instrumentation) differ from those found in studies of collateral effects. Because this is the first meta-analysis to examine this moderator, its effect should be investigated further.

### Academic Achievement and Maintenance

The effect size for academic achievement, calculated from the few studies that reported these data, was 0.61 (with a trans-

formed Cohen's  $d$  of 1.31). This finding is most similar to the effect sizes for achievement reported by Cook et al. (1985), with 0.59 for tutors and 0.65 for tutees. The effect size obtained for this study was larger than the effect sizes reported by Cohen et al. (1982) (0.33 for tutors, 0.40 for tutees), as well as Ginsburg-Block et al. (2006) (0.48).

Maintenance data suggest promising findings: Students maintained the behavioral and social gains derived from their involvement with peer tutoring. For example, Franca et al. (1990) reported that, "tutees showed slight decreases in the occurrence of positive interactions during follow-up, although they did maintain 0% level of negative social interactions" (p. 122). Gumpel and Frank (1999) noted positive social interactions (compared with baseline) for students identified as socially rejected during maintenance. Furthermore, Kamps et al. (1999) indicated that the duration of social interactions increased during follow-up for the participants with autism and their peers without disabilities. Similarly, McGee, Almeida, Sulzer-Azaroff, and Feldman (1992) observed increased levels of initiation and increased responses to peer initiations during fading for students with autism. Sutherland and Snyder (2007) concluded that active responding occurred more frequently than at baseline and that disruptive behaviors decreased for three of the four participants with EBD. Blake et al. (2000) reported positive behaviors and positive verbal exchanges with peers in both studies. Finally, Maheady and Sainato (1985) noted partial maintenance of observed behavioral changes.

### Correlations With Academic Achievement

The overall correlation between behavioral and social outcomes and academic achievement was  $r = 0.57$ , indicating a moderately strong relationship. That is, as improvements were noted in students' behavioral and social outcomes, progress was also observed in their academic achievement. This finding is similar to the  $r = 0.59$  reported by Ginsburg-Block et al. (2006) for the correla-

tion between social and academic achievement. It also falls within the range reported by Malecki and Elliott (2002), based on teacher ratings of students' social skills and problem behaviors when correlated with standardized reading, language, and math scores ( $r = -0.80$  to  $-0.50$ ).

More detailed correlational analyses by behavior type showed a relatively small correlation representing the strength of the relationship between academic achievement and disruptive/off-task behaviors ( $r = 0.31$ ). Thus, it appears that improvements in disruptive and off-task behaviors were associated with a minimal increase in academic achievement. Perhaps further examining this behavior type would help to understand the relationship better (e.g., calculating separate correlations for achievement and aggressive behaviors and achievement and out-of-seat behaviors). A correlation of  $r = -0.18$  was obtained when academic achievement was correlated with social skills/social interactions. This finding shows that, at least among this group of studies, social skills and academic achievement are not related. It also suggests that improved social skills and positive, frequent social interactions with peers do not increase as a function of improved academic achievement. The relation between academic achievement and academic engagement was also examined. The correlation between the two was  $r = 0.52$ , suggesting a moderately strong relation. That is, academic engagement increased as academic achievement improved.

### Limitations and Future Research

The findings of this meta-analysis should be considered in light of the following limitations and need for further research. First, not all studies published in the SCR literature on peer tutoring included the same variables. This precluded the examination of some potential moderators or a full exploration of differences between moderators. For example, there were limited numbers of studies involving some disability categories (e.g., autism, S/L), limiting comparisons and generalizability of some analyses. Future research should

examine the impact of peer tutoring on other types of behaviors for students across the autism spectrum, as well as for students identified with S/L difficulties. Second, variability may have been introduced by the way behavioral and social gains were measured across studies (e.g., teacher report, standardized assessment, direct observation). Third, the majority of the studies focused on elementary school students. Although early intervention efforts are critical, more research is needed on the effect of peer tutoring for middle and high school students. Finally, caution should be used in interpreting the TauU and Cohen's *d* conversions because such transformations are approximations. Despite these limitations, a significant strength of this meta-analysis is the application of design quality standards (Kratochwill et al., 2010) to the studies included. Furthermore, a nonsignificant Egger's test result provides some assurance that the studies were not notably biased.

Several additional factors may serve as potential moderators of student outcomes in future analyses. Gender could be a potential moderator variable because boys are more likely than girls to display oppositional and hyperactive, aggressive, and disruptive behaviors (Loeber, Green, Keenan, & Lahey, 1995; Loeber & Hay, 1994; McIntosh et al., 2008). Boys are also more likely to be suspended and to be identified as having conduct disorder (Darney et al., 2013). In addition to gender, ethnicity may be a potential moderator of effects because it is related to other types of behavioral outcomes (e.g., suspension and expulsion; Achilles, McLaughlin, & Croninger, 2007; Bowman-Perrott et al., 2013b). Other considerations for future peer tutoring research are how behavioral and social data are collected (direct observation versus teacher ratings) and how academic achievement is measured (standardized tests, curriculum-based measurement, teacher ratings of student academic competence and achievement). These could potentially affect individual study findings as well as summaries of aggregated data across studies. The reliability and validity of instruments used to measure students' outcomes could potentially moderate outcomes. Furthermore, grade level

(e.g., first versus fourth grade) could possibly act as a moderator of peer tutoring effects on social and behavioral variables (Bowman-Perrott et al., 2013a; Cohen et al., 1982). Lastly, investigating the potential impact of dosage (intensity, duration, and number of sessions) of peer tutoring studies might be beneficial as well.

### Implications for Practice

It is critical that school personnel identify and implement approaches to improve social and behavioral outcomes for their students. Approaches that also encourage academic gains are strongly recommended (Council for Exceptional Children, 2008; National Association of School Psychologists, 2010). The results of this meta-analysis, as well as others (e.g., Ginsburg-Block et al., 2006), highlight the social and behavioral benefits associated with peer tutoring. As a viable strategy for students with and without disabilities, peer tutoring can be implemented in general education, special education, and alternative education settings. Although the results of this study show that cross-age peer tutoring yielded larger benefits, the application of peer tutoring with same-age/same-grade students is well supported by the findings and in the peer tutoring literature. Findings also suggest that regardless of age and disability status, students are motivated to work toward academic goals with a reward structure in place. Furthermore, a higher rate of academic engagement, a key feature of peer tutoring, was shown to be related to improvements in academic achievement.

The results of this meta-analysis demonstrate how students with MR/ID, ADHD, EBD, LD, and autism can benefit from participation in peer tutoring. This is important because students with ADHD, EBD, and LD often have comorbid behavior and learning problems and are most at risk for disciplinary exclusion (Achilles et al., 2007; Bowman-Perrott et al., 2013b). Thus, they require comprehensive academic and behavioral supports to be successful in school. This finding also has important implications and application for students with MR/ID and autism because social skills are related to problem solving, self-

determination (Agran, Blanchard, Wehmeyer, & Hughes, 2002), and self-management (Koegel, Koegel, Hurley, & Frea, 1992). Finally, students without disabilities identified as “at risk” behaviorally also benefitted from peer tutoring.

Overall, peer tutoring yields behavioral and social supports to students and can produce simultaneous academic benefits as well. Furthermore, maintenance of behavioral and social improvements as a result of participating in peer tutoring is encouraging because such gains help contribute to students’ overall success.

### FOOTNOTES

\*References marked with an asterisk indicate studies included in the meta-analysis.

### REFERENCES

- Abramson, J. H. (2011). WINPEPI updated: Computer programs for epidemiologists, and their teaching potential. *Epidemiologic Perspectives & Innovations*, 8(1). Retrieved from <http://www.epiperspectives.com/content/8/1/1>
- Abramson, J. H. (2012). *WinPepi programs Compare 2 manual (version 2.69)*. Retrieved from <http://www.biomedcentral.com/1742-5573/content/1/1/6>
- Achilles, G. M., McLaughlin, M. J., & Croninger, R. G. (2007). Sociocultural correlates of disciplinary exclusion among students with emotional, behavioral, and learning disabilities in the SEELS national database. *Journal of Emotional and Behavioral Disorders*, 15, 33–45.
- Acion, L., Peterson, J. J., Temple, S., & Amtdt, S. (2006). Probabilistic index: An intuitive non-parametric approach to measuring the size of treatment effects. *Statistics in Medicine*, 25, 591–602.
- Agran, M., Blanchard, C., Wehmeyer, M., & Hughes, C. (2002). Increasing problem-solving skills of students with developmental disabilities participating in general education. *Remedial and Special Education*, 23, 279–288.
- Algozzine, B., Wang, C., & Violette, A. S. (2011). Reexamining the relationship between academic achievement and social behavior. *Journal of Positive Behavior Interventions*, 13, 3–16.
- Allison, D. B., & Gorman, B. S. (1993). Calculating effect sizes for meta-analysis: The case of the single case. *Behaviour Research and Therapy*, 31, 621–631.
- American Psychological Association. (2010). *The publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.
- Barriga, A. Q., Doran, J. W., Newell, S. B., Morrison, E. M., Barbetti, V., & Robbins, B. D. (2002). Relationships between problem behaviors and academic achievement in adolescents: The unique role of attention problems. *Journal of Emotional and Behavioral Disorders*, 10, 233–240.
- \*Barton-Arwood, S. M., Wehby, J. H., & Falk, K. B. (2005). Reading instruction for elementary age students with emotional and behavioral disorders: Academic and behavioral outcomes. *Exceptional Children*, 72, 7–27.
- Becker, B. E., & Luthar, S. S. (2002). Social-emotional factors affecting achievement outcomes among disadvantaged students: Closing the achievement gap. *Educational Psychology*, 37, 197–214.
- Benner, G. J., Beaudoin, K., Kinder, D., & Mooney, P. (2005). The relationship between the beginning reading skills and social adjustment of a general sample of elementary aged children. *Education and Treatment of Children*, 28, 250–264.
- Blackorby, J., Knokey, A., Wagner, M., Levine, P., Schiller, E., & Sumi, C. (2007). *SEELS: What makes a difference: Influences on outcomes for students with disabilities*. Retrieved from <http://www.seels.net/info-product.htm>
- \*Blake, C., Wang, W., Cartledge, G., & Gardner, R. (2000). Middle school students with serious emotional disturbances serve as social skills trainers and reinforcers for peers with SED. *Behavioral Disorders*, 25, 280–298.
- \*Bowman-Perrott, L. (2009). ClassWide Peer Tutoring: An effective strategy for students with emotional and behavioral disorders. *Intervention in School and Clinic*, 44(5), 259–267.
- Bowman-Perrott, L., Benz, M. R., Hsu, H., Kwok, O., Eisterhold, L., & Zhang, D. (2013b). Patterns and predictors of disciplinary exclusion over time: An analysis of the SEELS national dataset. *Journal of Emotional and Behavioral Disorders*, 21, 83–96.
- Bowman-Perrott, L., Davis, H., Vannest, K. J., Williams, L., Greenwood, C. R., & Parker, R. (2013a). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review*, 42(1), 39–55.
- Bowman-Perrott, L. J., Greenwood, C. R., & Tapia, Y. (2007). The efficacy of peer tutoring used in secondary alternative school classrooms with small teacher/pupil ratios and students with emotional and behavioral disorders. *Education and Treatment of Children*, 30, 65–87.
- \*Brady, N. (1997). The teaching game: A reciprocal peer tutoring program for preschool children. *Education and Treatment of Children*, 20, 123–149.
- Burke, M. D., Hagan-Burke, S., & Sugai, G. (2003). The efficacy of function-based interventions for students with learning disabilities who exhibit escape-maintained problem behaviors: Preliminary results from a single-case experiment. *Learning Disability Quarterly*, 26, 15–25.
- Burns, M. K., Zaslofsky, A. F., Kanive, R., & Parker, D. C. (2012). Meta-analysis of incremental rehearsal using phi coefficients to compare single-case and group designs. *Journal of Behavioral Education*, 21, 185–202.
- \*Christensen, L., Young, K. R., & Marchant, M. (2007). Behavioral intervention planning: Increasing appropriate behavior of a socially withdrawn student. *Education and Treatment of Children*, 30, 81–103.
- Cleary, T. J., & Zimmerman, B. J. (2004). Self-regulation empowerment program: A school based program to enhance self-regulated and self-motivated cycles of

- student learning. *Psychology in the Schools*, 41, 537–550.
- Cloward, R. D. (1967). Studies in tutoring. *Journal of Experimental Education*, 36, 14–25.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta analysis of findings. *American Educational Research Journal*, 19, 237–248.
- Cook, C. R., Dart, E., Collins, T., Restori, A., Daikos, C., & Delport, J. (2012). Preliminary study of the confined, collateral, and combined effects of reading and behavioral interventions: Evidence for a transactional relationship. *Behavioral Disorders*, 38, 38–56.
- Cook, S. B., Scruggs, T. E., Mastropieri, M. A., & Casto, G. C. (1985). Handicapped students as tutors. *The Journal of Special Education*, 19, 483–492.
- Cooper, H. (2011). *Reporting research in psychology: How to meet journal article reporting standards* (6th ed.). Baltimore: American Psychological Association.
- Council for Exceptional Children. (2008). *Classifying the state of evidence for special education professional practices: CEC practice study manual*. Arlington, VA: Author.
- Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York: Routledge.
- Darney, D., Reinke, W. M., Herman, K. C., Stormont, M., & Ialongo, N. S. (2013). Children with co-occurring academic and behavior problems in first grade: Distal outcomes in twelfth grade. *Journal of School Psychology*, 51, 117–128.
- Delquadri, J. C., Greenwood, C. R., Stretton, K., & Hall, R. V. (1983). The peer tutoring spelling game: A classroom procedure for increasing opportunity to respond and spelling performance. *Education and Treatment of Children*, 6, 225–239.
- \*DuPaul, G. J., Ervin, R. A., Hook, C. L., & McGoey, K. E. (1998). Peer tutoring for children with attention deficit hyperactivity disorder: Effects on classroom behavior and academic performance. *Journal of Applied Behavior Analysis*, 31, 579–592.
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315, 629–634.
- Fantuzzo, J. W., King, J. A., & Heller, L. R. (1992). Effects of reciprocal peer tutoring on mathematics and school adjustment: A component analysis. *Journal of Educational Psychology*, 84, 331–339.
- \*Franca, V. M., Kerr, M. M., Reitz, A. L., & Lambert, D. (1990). Peer tutoring among behaviorally disordered students: Academic and social benefits to tutor and tutee. *Education and Treatment of Children*, 13, 109–128.
- Fuchs, D., & Fuchs, L. S. (2005). Peer-assisted learning strategies promoting word recognition, fluency, and reading comprehension in young children. *The Journal of Special Education*, 39, 34–44.
- Fuchs, L. S., Fuchs, D., & Karns, K. (2001). Enhancing kindergartners' mathematical development: Effects of peer-assisted learning strategies. *Elementary School Journal*, 101, 495–510.
- Fuchs, L. S., Fuchs, D., Phillips, N. B., Hamlett, C. L., & Karns, K. (1995). Acquisition and transfer effects of classwide peer-assisted learning strategies in mathematics for students with varying learning histories. *School Psychology Review*, 24, 604–620.
- Garcia-Vazquez, E., & Ehly, S. W. (1992). Peer tutoring effects on students who are perceived as not socially accepted. *Psychology in the Schools*, 29, 256–266.
- Ginsburg-Block, M., & Fantuzzo, J. (1997). Reciprocal peer tutoring: An analysis of “teacher” and “student” interactions as a function of training and experience. *School Psychology Quarterly*, 12, 134–149.
- Ginsburg-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. W. (2006). A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology*, 98, 732–749.
- Gooding, L. F. (2011). The effect of a music therapy social skills training program on improving social competence in children and adolescents with social skills deficits. *Journal of Music Therapy*, 48, 440–462.
- Greenhouse, J. N., & Iyengar, S. (2009). Sensitivity analysis and diagnostics. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 417–433). New York: Russell Sage Foundation.
- Greenwood, C. R. (1991). Longitudinal analysis of time, engagement, and achievement in at-risk versus non-risk students. *Exceptional Children*, 57, 521–535.
- Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1989). Longitudinal effects of classwide peer tutoring. *Journal of Educational Psychology*, 81, 371–383.
- \*Gumpel, T. P., & Frank, R. (1999). An expansion of the peer-tutoring paradigm: Cross-age peer tutoring of social skills among socially rejected boys. *Journal of Applied Behavior Analysis*, 32, 115–118.
- Hawkins, R. O., Musti-Rao, S., Hughes, C., Berry, L., & McGuire, S. (2009). Applying a randomized interdependent group contingency component to classwide peer tutoring for multiplication fact fluency. *Journal of Behavioral Education*, 18, 300–318.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, 3, 486–504.
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21, 1539–1558.
- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S. L., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practices in special education. *Exceptional Children*, 71, 165–179.
- Horner, R. H., & Kratochwill, T. R. (2012). Synthesizing single-case research to identify evidence-based practices: Some brief reflections. *Journal of Behavioral Education*, 21, 266–272.
- Horner, R. H., Swaminathan, H., Sugai, G., & Smolkowski, K. (2012). Considerations for the systematic analysis and use of single-case research. *Education and Treatment of Children*, 35, 269–290.
- House, A. E., House, B. J., & Campbell, M. B. (1981). Measures of interobserver agreement: Calculation formulas and distribution effects. *Journal of Behavioral Assessment*, 3, 37–57.
- Jun, S. W., Ramirez, G., & Cumming, A. (2010). Tutoring adolescents in literacy: A meta-analysis. *Journal of Education*, 45, 219–238.

- \*Kamps, D. M., Dugan, E., Potucek, J., & Collins, A. (1999). Effects of cross-age peer tutoring networks among students with autism and general education students. *Journal of Behavioral Education, 9*, 97–115.
- Koegel, L. K., Koegel, R. I., Hurlley, C., & Frea, W. D. (1992). Improving social skills and disruptive behavior in children with autism through self-management. *Journal of Applied Behavior Analysis, 25*, 341–353.
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation*. Retrieved from [http://ies.ed.gov/ncee/wwc/pdf/wwc\\_scd.pdf](http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf)
- Lassen, S. R., Steele, M. M., & Sailor, W. (2006). The relationship of school-wide positive behavior support to academic achievement in an urban middle school. *Psychology in the Schools, 43*, 701–712.
- \*Lo, Y., & Cartledge, G. (2004). Total class peer tutoring and interdependent group oriented contingency: Improving the academic and task related behaviors of fourth-grade urban students. *Education and Treatment of Children, 27*, 235–262.
- \*Locke, W. R., & Fuchs, L. S. (1995). Effects of peer-mediated reading instruction on the on task behavior and social interaction of children with behavior disorders. *Journal of Emotional and Behavioral Disorders, 3*, 92–99.
- Loeber, R., Green, S. M., Keenan, K., & Lahey, B. B. (1995). Which boys will fare worse? Early predictors of the onset of conduct disorder in a six-year longitudinal study. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*, 499–509.
- Loeber, R., & Hay, D. F. (1994). Developmental approaches to aggression and conduct problems. In M. Rutter & D. F. Hay (Eds.), *Development through life: A handbook for clinicians* (pp. 448–515). London: Blackwell.
- Maheady, L., Harper, G. F., & Sacca, K. (1988). A class-wide peer tutoring system in a secondary, resource room program for the mildly handicapped. *Journal of Research and Development in Education, 21*, 76–83.
- \*Maheady, L., & Sainato, D. M. (1985). The effects of peer tutoring upon the social status and social interaction patterns of high and low status elementary school students. *Education and Treatment of Children, 8*, 51–65.
- Malecki, C. K., & Elliott, S. N. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly, 17*, 1–23.
- Mastropieri, M. A., Spencer, V., Scruggs, T. E., & Talbott, E. (2001). Students with disabilities as tutors: An updated research synthesis. *Advances in Learning and Behavioral Disabilities, 15*, 247–279.
- \*McDonnell, J., Mathot-Buckner, C., Thorson, N., & Fister, S. (2001). Supporting the inclusion of students with moderate and severe disabilities in junior high school general education classes: The effects of class-wide peer tutoring, multi-element curriculum, and accommodations. *Education and Treatment of Children, 24*, 141–160.
- McEvoy, A., & Welker, R. (2000). Antisocial behavior, academic failure, and school climate: A critical review. *Journal of Emotional and Behavioral Disorders, 8*, 130–140.
- \*McGee, G. G., Almeida, C., Sulzer-Azaroff, B., & Feldman, R. S. (1992). Promoting reciprocal interactions via peer incidental teaching. *Journal of Applied Behavior Analysis, 1*, 117–126.
- McIntosh, K., Horner, R. H., Chard, D. J., Dickey, C. R., & Braun, D. H. (2008). Reading skills and function of problem behavior in typical school settings. *Journal of Special Education, 42*, 131–147.
- Methe, S. A., Kilgus, S. P., Neiman, C., & Riley-Tillman, T. C. (2012). Meta-analysis of interventions for basic mathematics computation in single-case research. *Journal of Behavioral Education, 21*, 230–253.
- Miles, S. B., & Stipek, D. (2006). Contemporaneous and longitudinal associations between social behavior and literacy achievement in a sample of low-income elementary school children. *Child Development, 77*, 103–117.
- National Academy of Sciences. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and abilities*. Washington, DC: National Academies Press. Retrieved from [http://www.nap.edu/openbook.php?record\\_id=12480](http://www.nap.edu/openbook.php?record_id=12480)
- National Association of School Psychologists. (2010). *Model for comprehensive and integrated school psychological services*. Retrieved from [http://www.nasponline.org/standards/2010standards/2\\_PracticeModel.pdf](http://www.nasponline.org/standards/2010standards/2_PracticeModel.pdf)
- Nelson, J. R., Benner, G. J., Lane, K., & Smith, B. W. (2004). Academic achievement of K-12 students with emotional and behavioral disorders. *Exceptional Children, 71*, 59–73.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw Hill.
- Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*, 357–367.
- Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy, 35*, 202–322.
- \*Petursdottir, A., McComas, J., & McMaster, K. (2007). The effects of scripted peer tutoring and programming common stimuli on social interactions of a student with autism spectrum disorder. *Journal of Applied Behavior Analysis, 40*, 353–357.
- \*Prater, M. A., Serna, L., & Nakamura, K. K. (1999). Impact of peer teaching on the acquisition of social skills by adolescents with learning disabilities. *Education and Treatment of Children, 22*, 19–35.
- Reinke, W. M., Heman, K. C., Petras, H., & Ialongo, N. S. (2008). Empirically derived subtypes of child academic and behavior problems: Co-occurrence and distal outcomes. *Journal of Abnormal Child Psychology, 36*, 759–770.
- Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology, 95*, 240–257.
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 232–243). New York: Russell Sage Foundation.
- Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52*, 59–82.

- Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). the quantitative synthesis of single subject research: Methodology and validation. *Remedial and Special Education, 8*(2), 24–33.
- \*Sideridis, G. D., Utley, C., Greenwood, C. R., Delquadri, J., Dawson, H., Palmer, P., & Reddy, S. (1997). Class-wide peer tutoring: Effects on the spelling performance and social interactions of students with mild disabilities and their typical peers in an integrated instructional setting. *Journal of Behavioral Education, 7*, 435–462.
- Skiba, R. J., Peterson, R. L., & Williams, T. (1997). Office referrals and suspension: Disciplinary intervention in middle schools. *Education and Treatment of Children, 20*, 1–21.
- Stainback, W. C., & Stainback, S. B. (1972). Effects of student to student tutoring on arithmetic achievement and personal social adjustment of low achieving tutees and high achieving tutors. *Education and Training of the Mentally Retarded, 7*, 169–172.
- \*Sutherland, K. S., & Snyder, A. (2007). Effects of reciprocal peer tutoring and self-graphing on reading fluency and classroom behavior of middle school students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 15*, 103–118.
- Topping, K. J., Peter, C., Stephen, P., & Whale, M. (2004). Cross-age tutoring of science in the primary school: Influence on scientific language and thinking. *Educational Psychology, 24*, 57–75.
- Utley, C. A., & Mortweet, S. L. (1997). Peer-mediated instruction and interventions. *Focus on Exceptional Children, 29*, 1–23.
- Vannest, K. J., Davis, J. L., Davis, C. R., Mason, B. A., & Burke, M. D. (2010). Effective intervention and measurement with a daily behavior report card: A meta-analysis. *School Psychology Review, 39*, 654–672.
- Vannest, K. J., Parker, R. I., & Gonen, O. (2011). *Single-case research: Web-based calculators for SCR analysis, version 1.0* [Web-based application]. College Station, TX: Texas A&M University. Retrieved from [www.singlecaseresearch.org](http://www.singlecaseresearch.org).
- \*Wehby, J. H., Falk, K. B., Barton-Arwood, S., Lane, K. L., & Cooley, C. (2003). The impact of comprehensive reading instruction on the academic and social behavior of students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 11*, 225–238.
- Wilkinson, L., & APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594–604.

Date Received: February 19, 2013

Date Accepted: September 5, 2013

Associate Editor: John Hitchcock ■

Lisa Bowman-Perrott, PhD, is an Assistant Professor in the Special Education program in the Department of Educational Psychology at Texas A&M University. Her research interests are academic and behavioral interventions for students with or at risk for emotional and behavioral disorders and disciplinary exclusion.

Mack D. Burke, PhD, is an associate professor in the Special Education program in the Department of Educational Psychology at Texas A&M University. His areas of interest focus on emotional and behavioral disorders and positive behavioral supports.

Nan Zhang, MS, is a doctoral student in the Special Education program at Texas A&M University. Her primary research interest is intervention efficacy for children and youth with internalizing behavior disorders.

Samar Zaini, MS, is a doctoral student in the Special Education program at Texas A&M University. Her research interests are positive behavioral interventions and support for young children, as well as interventions for children with attention deficit hyperactivity disorder.