Clin Child Fam Psychol Rev
DOI 10.1007/s10567-017-0235-4

The Effectiveness of Psychosocial Interventions Delivered by Teachers in Schools: A Systematic Review and Meta-Analysis

Cynthia Franklin[1] · Johnny S. Kim[2] · Tasha S. Beretvas[3] · Anao Zhang[4] · Samantha Guz[4] · Sunyoung Park[5] · Katherine Montgomery[6] · Saras Chung[7] · Brandy R. Maynard[8]

Abstract The growing mental health needs of students within schools have resulted in teachers increasing their involvement in the delivery of school-based, psychosocial interventions. Current research reports mixed findings concerning the effectiveness of psychosocial interventions delivered by teachers for mental health outcomes. This article presents a systematic review and meta-analysis that examined the effectiveness of school-based psychosocial interventions delivered by teachers on internalizing and externalizing outcomes and the moderating factors that influence treatment effects on these outcomes. Nine electronic databases, major journals, and gray literature (e.g., websites, conference abstract) were searched and field experts were contacted to locate additional studies. Twenty-four studies that met the study inclusion criteria were coded into internalizing or externalizing outcomes and further analyzed using robust variance estimation in meta-regression. Both publication and risk of bias of studies were further assessed. The results showed statistically significant reductions in students' internalizing outcomes ($d = .133$, 95% CI [.002, .263]) and no statistical significant effect for externalizing outcomes ($d = .15$, 95% CI [−.037, .066]). Moderator analysis with meta-regression revealed that gender (%male, $b = -.017$, $p < .05$), race (% Caucasian, $b = .002$, $p < .05$), and the tier of intervention ($b = .299$, $p = .06$) affected intervention effectiveness. This study builds on existing literature that shows that teacher-delivered Tier 1 interventions are effective interventions but also adds to this literature by showing that interventions are more effective with internalizing outcomes than on the externalizing outcomes. Moderator analysis also revealed treatments were more effective with female students for internalizing outcomes and more

✉ Cynthia Franklin
cfranklin@austin.utexas.edu

Johnny S. Kim
johnny.kim@du.edu

Tasha S. Beretvas
tberetvas@austin.utexas.edu

Anao Zhang
zhanganao@utexas.edu

Samantha Guz
samantharguz@utexas.edu

Sunyoung Park
sunyoung@utexas.edu

Saras Chung
saraschung@wustl.edu

Brandy R. Maynard
bmaynar1@slu.edu

[1] School of Social Work, University of Texas at Austin, 1925 San Jacinto Blvd D3500, RM 2.228, Austin, TX 78712-0358, USA

[2] Graduate School of Social Work, University of Denver, Denver, CO, USA

[3] College of Education, University of Texas, Austin, TX, USA

[4] School of Social Work, University of Texas at Austin, Austin, TX, USA

[5] Quantitative Method Program, College of Education, University of Texas at Austin, Austin, TX, USA

[6] Little Sunshine School in St. Louis, St. Louis, MO, USA

[7] Social System Design Lab, Washington University's Brown School of Social Work, St. Louis, MO, USA

[8] School of Social Work, Saint Louis University, St. Louis, MO, USA

Published online: 10 May 2017

effective with Caucasian students for externalizing outcomes.

Keywords School mental health · Teacher interventions · Systematic review · Meta-analysis · Response to intervention

Introduction

Many mental health challenges begin in childhood with approximately 1 in 5 adolescents meeting the diagnostic criteria for a mental disorder (Merikangas et al. 2010), yet most do not receive services or seek help due to a variety of perceived barriers (Gulliver et al. 2010). The high prevalence rates and low service utilization of mental health services by youth indicates a need for mental health services that can be easily accessible by a relatively large proportion of youth. Schools are well positioned to fill this gap in mental health service provision and have been increasingly providing a range of school-based psychosocial interventions. Rones and Hoagwood (2000) defined school-based psychosocial interventions as any program or intervention delivered in a school setting aimed at improving students behavioral, emotional, or social functioning. A more recent definition of psychosocial intervention was developed by the National Academy of Medicine in relationship to the treatment of mental and substance use disorders.

> *Psychosocial interventions for mental health and substance use disorders are interpersonal or informational activities, techniques, or strategies that target biological, behavioral, cognitive, emotional, interpersonal, social, or environmental factors with the aim of improving health functioning and well-being* (England, Butler, & Gonzales, 2015, p. 31).

Children and adolescents' mental disorders that are addressed by school-based, psychosocial interventions may be clustered into either internalizing or externalizing groups (American Psychiatric Association 2013, p. 13). When combined, these two groups of disorders are estimated to cost society 247 billion dollars annually (Perou et al. 2013), and put children and adolescents at significant educational risk. According to The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), "It was demonstrated that clustering of disorders according to what has been termed internalizing and externalizing factors represents an empirically supported framework. The internalizing group represents disorders with prominent anxiety, depressive, and somatic symptoms, and the externalizing group represents disorders with prominent impulsive, disruptive conduct, and substance use symptoms" (APA 2013, p. 13).

Low socioeconomic conditions, such as poverty, may exacerbate the internalizing condition and/or externalizing problem behaviors of students in school, which is further complicated by limited access to family and community resources (Bain and Diallo 2016). When mental disorders are not addressed, children and adolescents have significantly lower academic performance (Hoagwood et al. 2007), experience greater developmental difficulties (Bain and Diallo 2016), and are at higher risk of comorbid conditions such as substance use, suicide, school dropout, and incarceration (Rajaleid et al. 2016; Stagman and Cooper 2010). This may explain why over half of the psychosocial interventions that are delivered in schools are offered in low-income communities and aim to lessen the disparities in mental health care (Amaral et al. 2011). Many school-based psychosocial interventions are currently delivered by or involve teachers who are not traditionally trained to provide mental health care (Franklin et al. 2012; Han and Weiss 2005), and thus there is a particular need to examine relevant factors that may influence the effectiveness of teacher-delivered interventions on internalizing and externalizing outcomes (Durlak et al. 2011; Ojio et al. 2015).

School-Based Psychosocial Interventions Delivered by Teachers

Previous literature has focused on three major areas in relationship to understanding the effectiveness of school-based, psychosocial interventions delivered by teachers: (1) The content and nature of the mental health services provided, such as what types of intervention in what format is most effective (Paulus et al. 2016); (2) The role of teachers as service providers, like teachers' own perceptions of their involvement in supporting students' mental health (Reinke et al. 2011); and (3) Qualification, training, and supervision needed for teachers to be effective in mental health service provision, like the teacher training and/or supervision process for them to be more effective (Han and Weiss 2005).

The Role of Teachers

The Response to Intervention (RTI) framework explains teachers' involvement within school mental health services within the three Tiers (Bender and Shores 2007; Fuchs and Fuchs 2006; Greenfield et al. 2010). The Response to Intervention is a multi-tiered approach for identifying the needs of students. Beginning with universal screening in classroom, RTI aims to ensure that students are being appropriately supported along a three tiered system (Lenski 2011). Tier 1 interventions are school-wide approaches targeting all students in the school that may be delivered in the classroom (Kearney 2016; Paulus et al. 2016). Given

the nature of this framework approach, teachers are most likely to be involved with Tier 1 interventions. Tier 2 interventions provide student support beyond Tier 1 interventions and are more indicative of existing problems. Because Tier 1 and 2 interventions may overlap and may also scaffold on one another (Franklin et al. 2012), teachers may become involved in Tier 2 interventions. Tier 2 interventions, for example, may be delivered in small groups of students that target deficits in academic, psychosocial, or behavioral performance (Frey et al. 2011). Previous literature (e.g., Franklin et al. 2012; Kelly 2008; Simonsen et al. 2008) reported about 5 percent of students requiring Tier 3 interventions that are intensive and therapeutic and may also involve special education services. It should be noted, however, that this proportion guideline has been increasingly questioned (e.g., Ferri 2012; Kavale et al. 2005). Because RTI original came from the special education field, this guideline may be less appropriate now because of the diversity in school students' and inclusion criteria that are used to select students for multi-tier services. Although teachers play a role in the delivery of services across tiers, their roles vary from sole provider of the service to various levels of involvement in collaborative teams with other professionals in the RTI framework, such as mental health professionals, counselors, and other colleagues (Ringwalt et al. 2010; Wolmer et al. 2011).

Recent literature indicates that teachers have the ability to deliver quality evidence-based behaviorally oriented services on a Tier 1 and/or Tier 2 level (Anderson-Butcher and Ashton 2004; Berzin et al. 2011; Frey et al. 2011). This success could be due to teachers' existing familiarity with classroom and behavioral management. However, some literature also criticizes the ability of teachers to deliver psychosocial interventions (Chitiyo and Wheeler 2009; Han and Weiss 2005; Hawkins and Heflin 2010; Ringwalt et al. 2010; Tillery et al. 2010) across all the tiers of intervention. Overall, there are mixed perspectives on teachers' ability to effectively deliver a range of psychosocial interventions; however, there are existing reviews favoring the accumulative evidence that teachers are more effective with Tier 1 interventions (e.g., Durlak et al. 2011; Franklin et al. 2012; Han and Weiss 2005).

Content, Nature, and Format of Interventions Provided

Psychosocial interventions delivered by teachers typically take place in classrooms and include curricula that provide activities, strategies, and techniques that are aimed at improving mental health outcomes, functional impairment, and well-being (Durlak et al. 2011; Franklin et al. 2012). A meta-analysis of 213 school-based studies by Durlak et al. (2011) found that teachers delivered Tier 1 interventions in a classroom (53%) because delivering interventions in this way was believed to reduce stigmatization, limit preparation time done by teachers, and utilize teachers' existing skills. Examples of Tier 1 interventions that are often delivered by teachers include classroom and instructional-based social skill training interventions. Tier 2 interventions include small cognitive groups, while Tier 3 interventions are one-on-one sessions between the teacher and student.

Training, Qualification, and Supervision for Teachers

It is not known how much professional training and supervision that teachers need to be able to effectively deliver psychosocial interventions across the different tiers of intervention. Frey et al. (2011) argued that teachers may need extensive training and ongoing supervision to ensure the quality of their service delivery. Recent reviews of school mental health literature note the importance of teacher–student relationships (Paulus et al. 2016) and the fact that teachers commonly correct the majority of challenging behaviors in the classroom (Barnes et al. 2014). Behavioral management and relationship factors associated with effective teaching may also provide an advantage in delivery of psychosocial interventions that are aimed at a student's academic, behavioral, mental, and emotional growth (Helker et al. 2007). For this reason, Franklin et al. (2012) suggest that it is possible that the role of teachers in schools may be expanded to the delivery of a broader range of mental health interventions if initial training and ongoing supervision is provided.

Reviews of Psychosocial Interventions Delivered by Teachers

Franklin et al. (2012) investigated 49 school mental health studies on teacher's involvement and collaboration with other professionals in service delivery. Of the studies examined, teachers were involved in the delivery of 40.8% of the interventions reviewed and most of those interventions were at the Tier 1 level. This 10-year review came short of synthesizing the results of different studies or moderators of outcome but instead examined individual effect sizes across studies discovering mixed results. The authors concluded that further research is needed to determine the efficacy of psychosocial interventions delivered by teachers.

A recent meta-analysis of cognitive-behavioral oriented interventions found an overall small, positive treatment effect ($d = .23$) when implemented in school settings (Barnes et al. 2014); however, these interventions were not necessarily provided by teachers. Another systematic

review of school-based services found unclear evidence for the effects of social skills training in schools with 75% of the interventions showing no effect (Vreeman and Carroll 2007). Even though school-based psychosocial interventions have been shown to achieve different treatment effects, the studies have not targeted the different roles of teachers in relationship to their effectiveness with internalizing and externalizing outcomes.

Another review conducted by Durlak et al. (2011) examined the effectiveness of 213 school-based universal interventions on six types of outcomes associated with socio-emotional learning. Fifty-three percent of the interventions were classroom interventions conducted by teachers. Results indicated that teachers are effective as interventionists, with reduction in conduct and internalizing problems observed. This review, however, did not examine the full spectrum of teacher-delivered interventions across different Tiers. Authors in this review also were not able to investigate important demographic moderators that could impact the effectiveness of interventions delivered by teachers. Durlak et al. (2011) recommended, for example, that future reviews of school interventions should examine moderating effects of ethnicity and gender to determine how these factors may impact the effectiveness of interventions delivered by teachers and other school professionals. Other studies have also indicated that factors such as race and income are important to consider in the evaluation and implementation of school-based, psychosocial interventions (Eiraldi et al. 2016; Garcia et al. 2016; Paulus et al. 2016) because these interventions are frequently delivered to underrepresented ethnic minority students.

Baskin et al. (2010) completed a meta-analysis of 83 counseling and psychotherapy studies in schools and examined service provider moderators, grouping teachers into a paraprofessional category along with parents. The studies in this review did not group interventions into specific tiers of intervention; however, because of the focus of the study on psychotherapy, it is likely that many of the interventions fell somewhere along the Tier 2 and Tier 3 categories. Twenty-three of the 107 studies analyzed were implemented by paraprofessionals including both teachers and parents. In this study, licensed mental health professionals were shown to achieve better results ($d = .67$) than the paraprofessionals ($d = .45$).

In a meta-analysis on school-based programs for aggressive and disruptive behavior, results suggested that "interventions were generally more effective when implemented well and relatively intense, used one-on-one formats and were administered by teachers" (Wilson and Lipsey 2007, p. 148). A wide range of interventions, such as social competence training, behavioral interventions, therapy and counseling, multimodal programs, and peer mediation, were included in the review. This study, however, did not address specific tiers of interventions or other factors that may contribute to teacher effectiveness.

Finally, some researchers have specifically reviewed teacher implementation factors and sustainment in the delivery of evidence-based, psychosocial interventions (e.g., Han and Weiss 2005; Jennings and Greenberg 2009). Han and Weiss (2005), for example, developed a conceptual framework for effective implementation (e.g., training, feedback, fidelity) for psychosocial curriculum in classroom settings. This review did not specifically evaluate the effectiveness of implementation factors such as training and supervision in relationship to how these factors may influence distinct mental health outcomes.

Aims of Study

Although there is some evidence that teacher-delivered psychosocial interventions in schools may be effective, the current body of literature is inconclusive at best and a number of gaps remain. Prior reviews have not examined the effectiveness of teachers across all the three tiers of intervention, and differentiation between internalizing and externalizing outcomes has been neglected. There is also an incomplete understanding of the demographic, intervention type, and supervisory factors that may moderate outcomes. In response, our study presents the results from a systematic review and meta-analysis of randomized controlled trials (RCTs) that examined the effectiveness of psychosocial interventions delivered by teachers on internalizing and externalizing outcomes. In addition, we also examined factors that may moderate effects of interventions, including Tier of interventions, type of interventions delivered, the effects of manualized interventions and the effects of supervision, and the specific demographic factors such as grade level, age, race/ethnicity and gender.

Methods

Search Procedures

Following the Cochrane Collaboration Guidelines (Higgins and Green 2011), we searched multiple sources for eligible published and unpublished studies. The search included nine electronic databases, 19 intervention websites (intervention listed in Table 1), and contacted six experts for studies published from 2000 to September 2016. The six subject experts were consulted in the initial stages of the study to help conceptualize the study and develop the codebook. The experts were school researchers and with practice experience. These experts had advanced degrees

and worked with school professionals implementing interventions in school settings. We searched nine electronic databases including: PsycINFO, Academic Search Complete, CINAHL Plus, Education Full Text, ERIC, Professional Development Collection, PsycARTICLES, Psychology and Behavioral Sciences Collection, and Teacher Reference Center. Four primary search terms were used: "school*," "intervention*," "random*," and "teacher*" searched in all text. Because many studies do not mention teacher as being part of the intervention in the title or abstract even though teachers were involved in the intervention, we searched key terms in full text in all nine electronic databases. The search concluded in September 2016.

Inclusion and Exclusion Criteria

To be eligible for inclusion, a study needed to examine effects of a teacher-delivered psychosocial intervention in a school setting on internalizing or externalizing outcomes using a randomized controlled trial study design. The internalizing group was operationalized as behaviors with prominent anxiety, depressive, and somatic symptoms and other outcome measures of similar constructs. The externalizing group was operationalized as behaviors with prominent impulsive, disruptive conduct, and substance use symptoms and other outcome measures of similar constructs. For the purposes of this review, teacher was defined as a professional educator holding primary responsibility for one or more instructional areas and whose main role is to be in the classroom. Most of the studies reviewed listed teachers as the sole primary provider of services. A code was created to mark the few studies where the role of the teacher was unclear, so that these studies could be separated in the analysis. Furthermore, other professionals were also involved in implementation of interventions and therefore, a code was created to identify these professionals. Eligible studies must report sufficient data to calculate an effect size. We did not limit inclusion based on publication status, but studies must have been published in English. Studies were excluded if: (1) a study did not involve teachers being part of the intervention delivery; (2) a study did not report statistical information that could be used to calculate treatment effect size estimates; (3) study samples were not independent because a study reported a subset of the population that had been reported by other studies that were included in the review (e.g., if a study reported the effect of school-based CBT for children of all races and another study reported the effect of school-based CBT using the same dataset but for children who were Hispanic then results from the second study would be excluded; (4) a study did not measure an internalizing or externalizing behavior problem.

Data Extraction

A group of experienced school researchers and practitioners developed a coding sheet for this review (available upon request from the corresponding author). The four researchers who developed the coding sheet were all tenure-track faculty members affiliated with US research-intensive universities and one researcher had training and appointment with the Campbell Collaboration. Initially five studies were coded to pilot the codebook. Because of the complexity discovered in coding the statistical data needed for the calculation of the effect sizes, five additional studies were coded to make sure each coder could accurately code statistical data. Individual studies were then coded for participant and provider characteristics, intervention characteristics, and effect size data. Participant and provider characteristics included student's age, grade level, race/ethnicity, sex, socioeconomic status, disabilities, criteria used to screen participants, teacher's - years of experience, and the teacher's highest level of education. Intervention characteristics included type of intervention (i.e., behavioral, cognitive, social skills, counseling, peer mediation, and other), the tier of the intervention, treatment fidelity (whether teachers used a manual or written guide, if the teachers were trained to use the program, if the teachers received follow-up training or supervision, if other professionals in addition to a teacher delivered the intervention), duration of the intervention, and frequency of contact between students and interventionists. Other descriptors included role of the researcher and whether a control/comparison group received services after completion of the study. This study adhered to the internalizing and externalizing group framework (APA 2013) and coded outcome measures into either internalizing or externalizing outcomes based on the measures used in primary studies. Examples of internalizing outcome measures included: Children's Depression Inventory, Hopelessness Scale for Children, Children's Pessimistic Explanatory Style; and examples of externalizing outcome measures included: intentions to use substance, peer-reported aggressive behavior, and self-reported behaviors hurting himself/herself or others.

We used Cochrane Collaboration's tool for assessing the risk of bias in randomized trials (Higgins et al. 2011) to evaluate study bias in each study. In conducting evaluation for each study's risk of bias, this paper followed criteria specified in the Cochrane Handbook for Systematic Reviews of Interventions, Table 8.5.d: Criteria for judging risk of bias in the Risk of Bias Assessment Tool (Higgins et al. 2011).

Inter-Rater Agreement

Two coders coded the total sample of articles included in the meta-analysis. Then a third coder independently coded

the included studies in order to calculate inter-rater agreement. Any disagreements were settled with a fourth coder. Inter-rater reliability was calculated based on the agreements between all the data collected from the coding process including percentage codes and binary codes. The inter-rater reliability was determined by a percent agreement model, dividing the number of agreements over the number of possible agreements. Out of the 24 coded studies, there was a 93% agreement rate between the three primary coders. Additionally, two doctoral level coders with mental health clinical experience and familiarity with DSM-5 categorized outcome measures into internalizing or externalizing group and conducted confirmation checks on the grouping of outcome variables based on the measures in the studies.

Data Analysis

Data analysis proceeded in four stages using both SPSS (IBM, 2015) to calculate individual effect size estimates and R software (R Development Core Team 2008) to conduct robust variance estimation in meta-regression for moderator analysis: (1) descriptive statistics of study characteristics; (2) calculating individual effect size estimates (in SPSS) and synthesizing effect size estimates (in R); (3) assessing publication bias; and (4) moderator analysis using meta-regression (in R).

Effect Size Calculation and Adjustment

Reviewed outcomes in primary studies all used continuous measures, and their effect size estimates were calculated using Hedges's *g* effect size (Cooper et al. 2009) which represents the standardized mean difference that captures study findings when different measures or scales were used in studies (Cohen 1988; Glass 1976). All effect sizes (*g* statistic) were adjusted using Hedges's small sample size correction for unbiased estimates (Hedges and Olkin 1985) and noted as *d* in this review.

Synthesizing Effect Size Estimates and Moderator Analysis

Because internalizing and externalizing outcomes are theoretically and empirically different from each other in terms of etiology and treatment effectiveness, we synthesized the treatment effect size estimates for these outcomes separately. We used robust variance estimation (RVE) in meta-regression to synthesize the treatment effect size estimates and to conduct moderator analysis (Hedges et al. 2010; Tanner-Smith et al. 2016).

Several studies in this review reported multiple effect sizes using related outcomes for the same sample which introduces dependence into the resulting effect sizes. Compared to other statistical procedures that better handle within-study dependence [e.g., generalized least squares estimation (Gleser and Olkin 2009) and multilevel meta-analysis model (Van den Noortgate et al. 2013)], RVE fits better with our existing data because it makes no assumptions about effect sizes' sampling distributions and can estimate the covariance structure of the dependent effect sizes without actually knowing it (Hedges et al. 2010; Tanner-Smith and Tipton 2014). More importantly, simulation studies have suggested that RVE may yield accurate results with as few as 10 studies for estimating an average effect size (Tanner-Smith and Tipton 2014), and 20–40 studies for moderator analysis (e.g., Hedges et al. 2010; Tipton 2013). Methodological research recommended an ideal sample size of 5 effect sizes per primary study and about 40 primary studies for RVE to produce reasonably accurate results (Hedges et al. 2010; Tipton and Pustejovsky 2015). Because this review did not meet the above criteria, 24 primary studies and 123 effect size estimates, we used small sample size correction (Tanner-Smith and Tipton 2014; Tipton 2015) running RVE to control for a possibly inflated Type I error of test statistics and confidence intervals (Tipton and Pustejovsky 2015).

Having identified variability in the effect size estimates, we also conducted moderator analyses to explain some of that variability using meta-regression models with RVE that included continuous and categorical predictor (moderator) variables (Cooper et al. 2009; Konstantopoulos et al. 2009). For example, we explored whether intervention effect sizes were significantly greater for Tier 1 than for Tier 3 interventions.

Interpretation of the results was based on an alpha level of .05 with one caveat. Methodological studies (Tanner-Smith et al. 2016; Tipton 2015) reported the parameter estimation using RVE with small sample size correction is trustworthy as long as the degrees of freedom associated with the moderators are greater than or equal to four. Tanner-Smith et al. (2016) state, "If the degrees of freedom are very small [below four], a lower p value should be used; … if $p < .05$ is used as a threshold elsewhere, for these cases $p < .01$ should be used instead" (p. 94). Therefore, this study adopted an alpha level of .05 for all results but annotated any results that were significant at .05 level, not .01 level, for moderators that had degrees of freedom lower than 4.

Publication Bias

Publication bias occurs when the research that appears in the published literature is systematically unrepresentative

of all the research that has been done in an area (Rothstein et al. 2006). Studies with null results are less likely to be published than studies with statistically significant effects, thus including only published studies in a review may introduce an upward bias into the effect sizes (Cooper 2016). To assess publication bias, we used a funnel plot of the effect size estimates graphed against their standard errors to visually assess the potential for publication bias. We also used Vevea and Woods (2005) weight function model to conduct a sensitivity analysis again to assess the possibility of publication bias.

Results

Search Results

Figure 1 presents detailed steps and results from the literature search. The initial pool of eligible studies began with 36,823 studies for initial screening after duplicates were removed. Four separate coders excluded 36,497 studies based on reviewing titles and abstracts and then excluded 199 studies based on full-text review, resulting in 127 studies for initial coding. During the coding process, 75 studies were eliminated for conceptual reasons including study methodology and lack of focus on teacher implementation. Fifty-two studies remained and examined for statistical results necessary to calculate at least one effect size estimate. An additional 28 studies were eliminated because there was not enough statistical information to calculate an effect size estimate. As a result, 24 primary studies were included and reported in this systematic review and meta-analysis.

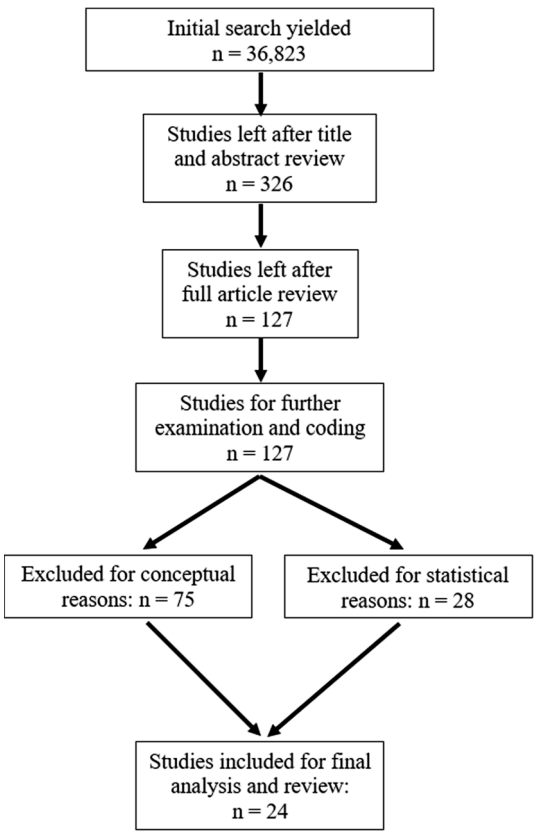

Fig. 1 Flowchart of inclusion and exclusion procedure

Study Characteristics

Table 1 presents study characteristics of all 24 studies, with a total sample size of $N = 32{,}985$ included in this review. Five studies reported internalizing disorders and 19 studies reported externalizing disorders. The reported mean age of students across the 24 studies was 11.35 years old, and over one-third ($n = 9$, 37.5%) of the studies used a mixture of grade levels, meaning that data were gathered from a combination of elementary, middle, and high school settings. The average percentage of male participants was 51.40% across all studies. Of studies that reported their samples' race/ethnicity ($n = 23$), the percentages of the participants were as follows: 34.90% White, 25.32% African-American, 36.07% Hispanic, and 18.28% other including Asian American and Native American.

Regarding treatment modalities, 75% of the studies included interventions with multiple components ($n = 18$) using a combination of behavioral, cognitive, social skills training, talk therapy, and peer mediation. More specifically, out of the 24 included studies, 58.33% of the studies used a behavioral strategy ($n = 14$), 66.66% used a cognitively oriented program ($n = 16$), 91.66% used a social skills program ($n = 22$), none of the studies reported using counseling programs, and 12.5% used peer mediation ($n = 3$). Twenty studies (91.66%) used teacher as the primary interventionist, and two studies (8.3%) used teachers who were primary but assisted by a mental health care provider not affiliated with the school. Less than half the studies ($n = 10$, 41.66%) used a Tier 1 format to treat students, leaving nine studies (37.5%) that treated students in a Tier 3 format and five studies (20.83%) that treated students in a Tier 2 format. With regard to treatment fidelity, 91.70% of the studies ($n = 22$) used manualized interventions or interventions with written guide. Ninety-two percent of the studies ($n = 22$) provided some form of training and/or ongoing supervision to teachers.

Meta-Analytic Results

The overall mean treatment effect sizes for both internalizing and externalizing outcomes, using an intercept only meta-regression model with RVE, are presented in Table 2. The treatment effect size for internalizing outcomes, with

Table 1 Characteristics of studies included in the review ($n = 24$)

Studies	Sample characteristics (% male and grades)	Intervention	Primary provider	Other provider(s)	Intervention modalities[β]	Measurement[γ]
Benner et al. (2012)	Males: 80.8–84.1% Mixed grade levels	Behavior Intervention	Teachers	NA	Behavioral	SOS
Botvin et al. (2001)	Males: 40% High school grades	Life Skills Training (LST)	Teachers	NA	Behavioral Cognitively Oriented Social Skills	Questionnaire*
Botvin et al. (2006)	Males: 51% Middle school grade levels	Life Skills Training (LST)	Teachers	NA	Behavioral Cognitively Oriented Social Skills	Items adopted*
Brown et al. (2001)	Males: 50.5% Mixed grade levels	Steps to Respect	Teachers	NA	Behavioral Cognitively Oriented Social Skills	SES TASB CTBPISS
Cappella et al. (2012)	Males: 57% Elementary grade levels	BRIDGE	Teachers	NA	Behavioral Social Skills	Classroom observation* BRIBRIEF SBEQ
Chaplin et al. (2006)	Males: 50.48% Mixed grade levels	Pennsylvania Resiliency Program for Adolescents	Teachers	NA	Behavioral Cognitively Oriented Social Skills	CDI HSC CASQ
Crean et al. (2013)	Males: 43% Mixed grade levels	Promoting Alternative Thinking Strategies	Teachers	NA	Social Skills	TRS, TCRS BASC-2, CRAS CRFDBS, CRVS CRGBNBAAS ASPS, HABAINS
Eron et al. (2002) MACSRG Study	Males: 53% Middle school grade levels	Early Yes I Can	Teachers	NA	Social Skills	PNI TRF CBCL-AS
Fonagy et al. (2009)	Males: 53.2% Elementary grade levels	School Psychiatric Consultation (SPC); Creating a Peaceful School Learning Environment (CAPSLE)	Teachers	NA	Behavioral Cognitively Oriented	
Frey et al. (2005)	Males: 51.8% Mixed grade levels	Second Step: Student Success Through Prevention	Teachers	NA	Behavioral Cognitively Oriented Social Skills	PPSB SESS-WSILM
Gillham et al. (2007)	Males: 54% Middle school grade levels	Penn Enhancement Program; and Penn Resiliency Program Adolescent	Teachers	NA	Behavioral Cognitively Oriented Social Skills	CDI

Table 1 continued

Studies	Sample characteristics (% male and grades)	Intervention	Primary provider	Other provider(s)	Intervention modalities[β]	Measurement[γ]
Gillham et al. (2012)	Males: 52% Middle school grade levels	Penn Resiliency Program Adolescent	Teachers	School-based service provider	Behavioral Cognitively Oriented Social Skills	CDI RADS-2 RCMAS NIMH DISC-IV CASQ, HSC CCSC, RCDI
Hecht et al. (2003)	Males: 50% Middle school grade levels	Keepin' it REAL	Teachers	NA	Social Skills	Questionnaire*
Holt et al. (2008)	Males: 47% High school grades	Achievement Mentoring Program (AMP)	Unknown	Teachers, and Mental Health Professional	Cognitively Oriented	PSSM CMSES Attendance records Discipline referrals
Iovannone et al. (2009)	Males: 82% Mixed grade levels	Prevent-Teach-Reinforce Model	Teachers	NA	Behavioral Peer Mediation	SSSR AET
Leff et al. (2009)	Males: 51.5% Elementary grade levels	Early Intervention Program (F2F)	Teachers	NA	Behavioral Cognitively Oriented Social Skills	Peer nomination of items adopted CSBQ, HAB AWLS, CDI
Leff et al. (2010)	Males: 51.5% Elementary grade levels	Modified Early Intervention Program (F2F)	Service Provider from community	Teachers	Cognitively Oriented Social Skills	Peer nomination of items adopted CSBQ HAB AWLS CDI
MACSRG Study (2007)	Males: NA Elementary grade levels	Class social cognitive intervention	Teachers	NA	Behavioral Cognitively Oriented	Items adopted** CFI NBAAS
Metz et al. (2006)	Males: 52% Mixed grade levels	Project Toward No Tobacco Use	Service Provider from community	Teachers	Cognitively Oriented Social Skills	USCSS
Murray et al. (2005)	Males: 75% High school grades	Teacher-student relationship program	Teachers	NA	Social Skills	CBCL Attendance record
Oneil et al. (2011)	Males: 54% Mixed grade levels	The Michigan Model for Health (MMH)	Teachers	NA	Social Skills	Items from SCASS-HEAP
Simonsen et al. (2011)	Males: 77.8% Mixed grade levels	Behavioral Education Program (BEP): CICO	School-based provider	Teachers	Behavioral	FACTS SOD SSRS
Spoth et al. (2005)	Males: 53% Middle school grade levels	Life Skills Training (LST)	Teachers	NA	Social Skills	Questionnaire* RAU

Table 1 continued

Studies	Sample characteristics (% male and grades)	Intervention	Primary provider	Other provider(s)	Intervention modalities[β]	Measurement[γ]
Spoth et al. (2008)	Males: 53% Middle school grade levels	Life Skills Training (LST) Strengthening Families Program (SFP)	Teachers	NA	Social Skills	Questionnaire* MPU APU

[β] Treatment modalities are defined as follow: *Behavioral Strategies*: Interventions involve the use of various behavioral techniques, such as rewards, token economies, contingency contracts, and the like to modify or reduce inappropriate behavior. *Cognitive-Oriented Programs:* Interventions focus on changing thinking processes or cognitive skills; programs focus on solving social problems, controlling anger, inhibiting hostile attributions, etc. *Social Skills Programs:* Interventions are designed to help youth better understand social behavior and learn appropriate social skills. Children learn communication skills, fighting avoidance skills, group entry skills, eye contract, "I" statements, etc. *Peer Mediation:* Student mediators are trained to offer mediation services for peers who experience interpersonal conflicts. Training generally focuses on a series of conflict resolution steps

[γ] *AET* Academic Engaged Time. *AUP* Advanced poly-substance use. *ASPS* Aggressive Social Problem Solving. *AWLS* Asher and Wheeler Loneliness Scale. *BASC-2-AO,TV* The Behavior Assessment Scale for Children-2 Acting Out, Teacher Version. *BASC-2-AS,TV* The Behavior Assessment Scale for Children-2 Aggression Subscale, Teacher Version. BASC-2-PS,TV: The Behavior Assessment Scale for Children-2 Conduct Problems Subscale, Teacher Version. *BRIBRIEF* The Behavioral Regulation Index of the Behavior Rating Inventory of Executive Function. *CASQ* Children's Attributional Style Questionnaire. *CASQ* Children's Attributional Style Questionnaire. *CBCL-AS* Child Behavior Checklist. *CCSC* Children's Coping Strategies Checklist. *CDI* Children's Depression Inventory. *CDRSR* Children's Depression Rating Scale – Revised. *CFI* The Children's Fantasy Inventory. *Classroom observation** classrooms were observed by a single coder blind to intervention condition. Observers were trained following standard procedures published before. *CMSES* Children's Multidimensional Self-Efficacy Scales. *CRAS* Child Report Aggression Scale. *CRFDBS* Child Report Frequency of Delinquent Behavior Survey. *CRGBNBAAS* Child Report General Beliefs subscale from the Normative Beliefs About Aggression Scale. *CRVS* Child Report Victimization Scale. *CSBQ* The Children's Social Behavior Questionnaire. *CTBPISS* Colorado Trust's Bullying Prevention Initiative Student Survey. *FACTS* Functional Assessment Checklist for Teachers and Staff. *HAB* Hostile Attributional Bias Measure. *HABAINS* Hostile Attribution Bias and Aggressive Interpersonal Negotiation Strategies. *HSC* Hopelessness Scale of Children. *Items adopted** Items that are similar to and adopted from previous published studies to measure verbal aggression, physical aggression, fighting, and delinquent behaviors. *Items adopted*** A measure derived from previous study was used to measure the child's intent to use aggressive responses. *Items from SCASS-HEAP* Items developed from the State Collaborative on Assessment and Student Standards-Health Education Assessment Project. MPU: Monthly poly-substance use. *NBAAS* The Normative Beliefs About Aggression Scale. *Peer nomination of items adopted* Peer nomination items included the standard five relational and three physical aggression items derived from the peer nomination survey designed. *PEQ* The Peer Experiences Questionnaire. *NIMH DISC-IV* NIMH Diagnostic Interview Schedule for Children version IV. *PNAVB* Peer Nominations of Aggression, Victimization and Bystanding. *PNI* Peer Nomination Inventory. *PPSB* Peer-Preferred Social Behavior subscale of the Walker-McConnel Scale of Social Competence and School Adjustment, Elementary Version. *PSSM* Psychological Sense of School Membership Scale. *Questionnaire** Questionnaire that included items assessing current alcohol and drug use and a series of scales measuring cognitive, attitudinal, and skills variables believed to be associated with the initiation of alcohol and drug use. *SES* School Environment Survey. *RADS-2* Reynolds Adolescent Depression Scale – 2. *RAU* Regular Alcohol Use. *RCDI* Region Child Depression Inventory. *RCMAS* Revised Children's Manifest Anxiety Scale. *SESS-WSILM* The Student Experience Survey: What School Is Like for Me. *SBEQ* Social Behavior and Experience Questionnaire. *SOS* Stage Observation System. *SOD* Structured Direct Observations. *SSRS* Social Skills Rating System. *TASB* Teacher Assessment of Student Behavior. *TCRS* Teacher-Child Rating Scales. *TRS* Teacher Report on Students. *TRF* Teacher's Report Form. *USCSS* University of Southern California Student Survey. *WWID* What Would I Do?

Table 2 Overall effect size for internalizing and externalizing disorders and between group differences

	K[a]	d[b]	95% CI[c]
Internalizing	27	.133*	[.002, .263]
Externalizing	96	.015	[−.037, .066]
Internalizing versus Externalizing[d]	123	.118*	[.034, .202]

* $p < .05$

[a] K = number of effect size estimates

[b] d = (small sample size corrected hedges' g) effect size

[c] CI = confidence intervals

[d] The internalizing versus externalizing row reports the differences in overall effect size between studies reporting internalizing and externalizing disorders

27 effect sizes from 5 studies, was $d = .133$ with a 95% confidence interval [.002, .263] which supported an overall positive and statistically significant effect of teacher-delivered psychosocial interventions for students' internalizing outcomes ($p < .05$). The overall treatment effect size for externalizing outcomes, from 96 effect sizes reported in 19 studies, was $d = .015$ with a 95% confidence interval [−.037, .066]. The effect size was not statistically significant ($p > .05$) thereby indicating no treatment effects for the interventions for externalizing related problems.

A further investigation compared the difference between the mean treatment effect size between internalizing and externalizing outcomes revealed that the two effect sizes

differed significantly by .118 ($p < .05$) with a 95% confidence interval for the difference of [.034, .202]. The positive difference in treatment effect size for internalizing versus externalizing outcomes indicates that teacher-delivered school-based psychosocial interventions are significantly more effective for internalizing than for externalizing outcomes.

Analysis of Within-Group [Internalizing Versus Externalizing Outcomes] Moderator Effects

Participant, intervention, and study characteristics were investigated in moderator analyses for internalizing and externalizing outcomes separately. Variables tested for moderating effects related to participant characteristics included age, gender (% Male), and race (% Caucasian, Black, Hispanic, and Others) and are presented in Table 3. For internalizing outcomes, the proportion of males in the sample was negatively associated with treatment effect size estimates and the association was statistically significant, $b = -.017$, $p < .05$, indicating treatments are more effective for female students. For externalizing outcomes, effect sizes were significantly positively associated with the proportion of Caucasian students in the sample ($b = .002$, $p < .05$), which indicates that these treatments are more effective for Caucasian students. It should be noted that % of male as a moderator for internalizing outcomes was significant at .05 level but had degrees of freedom smaller than 4. Therefore, this moderator should be interpreted with caution.

Moderating effects of treatment characteristics that were investigated included treatment mode and Tier of intervention (presented in Table 4). The analysis revealed that effect size estimates did not differ significantly ($p > .05$) as a function of the use of unimodal versus multimodal interventions ($b = .126$ and $b = -.034$ for internalizing and externalizing outcomes, respectively). Tiers of the intervention showed a trend toward significance ($p = .06$) when comparing the treatment effects of Tier 1 interventions with Tier 2 and Tier 3 interventions ($b = .299$, $p = .06$). Further subgroup analysis revealed that the overall treatment effect for Tier 1 interventions was $d = .211$, $p < .05$ while the overall treatment effect for Tier 2 + Tier 3 interventions was not statistically significant, $d = -.078$, $p > .05$.

This study intended to investigate the moderating effects of fidelity-related descriptors including whether an intervention is manualized, has written guides, and if training and/or supervision was provided. The low variability of these fidelity-related descriptors (reported previously) prevented us from conducting moderator analysis for these potential moderators.

This study also examined the type of comparison group as a moderator for effect size estimates (in Table 4). No significant differences ($p > .05$) in effect sizes were found as a function of whether a study used a treatment-as-usual control group versus studies that used nothing or waitlist comparison groups.

Table 3 Bivariate meta-regression results for participant characteristics

	Internalizing outcomes			Externalizing outcomes		
	K	*b*	SE	*K*	*b*	SE
Model 1						
Age	27	−.133	.080	94	.001	.080
Model 2						
Gender						
% Male	27	***−.017****	.001	94	−.001	.002
Model 3						
Race						
% Caucasian	27	.002	.001	94	.002*	.001
% Black	27	−.002	.001	94	−.001	.001
% Hispanic	27	−.003	.005	94	−.001	.001
% Others	27	.008	.012	94	.000	.002

Variables were entered separately but presented in the same table. *Italicized and bolded coefficient* should be interpreted with caution: the coefficient is significant at an alpha level of .05 but a degrees of freedom lower than 4. For these coefficients, a more stringent alpha level of .01 is recommended

K = number of effect size estimates; *b* = coefficient of the meta-regression analysis; *SE* = standard errors

* $p < .05$

Analysis of Publication Bias and Risk of Bias

Figure 2 presents the funnel plots for internalizing and externalizing outcomes. Both funnel plots are reasonably symmetric, indicating that publication bias does not appear to be a source of bias in this review. The conclusion was further confirmed by results of the sensitivity analysis using Vevea and Woods' weight function model (also presented in Fig. 2) with the thin line being the unadjusted effect size estimate and the heavy line representing the effect size estimate adjusted for publication bias. For internalizing outcomes, the unadjusted effect size estimate was $d = .137$ and the adjusted effect size estimate was $d = .146$. For externalizing outcomes, the unadjusted effect size estimate was $d = .019$ and the adjusted effect size estimate was $d = .042$. For both types of outcomes, the analyses provided greater alternative effect size estimate for the funnel plot to be symmetric supporting a likely lack of publication bias in these effect size estimates.

Risk of bias of studies (Table 5) indicated low risk of bias across studies in random sequence generation (100%

Table 4 Bivariate meta-regression results for intervention and study characteristics

	Internalizing outcomes			Externalizing outcomes		
	K	*b*	SE	*K*	*b*	SE
Treatment mode	27	.126	.118	93	−.034	.036
Tier of intervention[a]	27	.299[†]	.125	93	.176	.201
Comparison group	27	.039	.136	93	.024	.058

* $p < .05$; K = number of effect size estimates; SE = standard errors

[†] Marginally significant at $p = .06$ level, overall treatment effect (d) for Tier 1 intervention = .211, $p < .05$. Overall treatment effect for Tier 2 + Tier 3 = −.078, $p > .05$

[a] This moderator compares Tier 1 interventions with Tier 2 + Tier 3 interventions

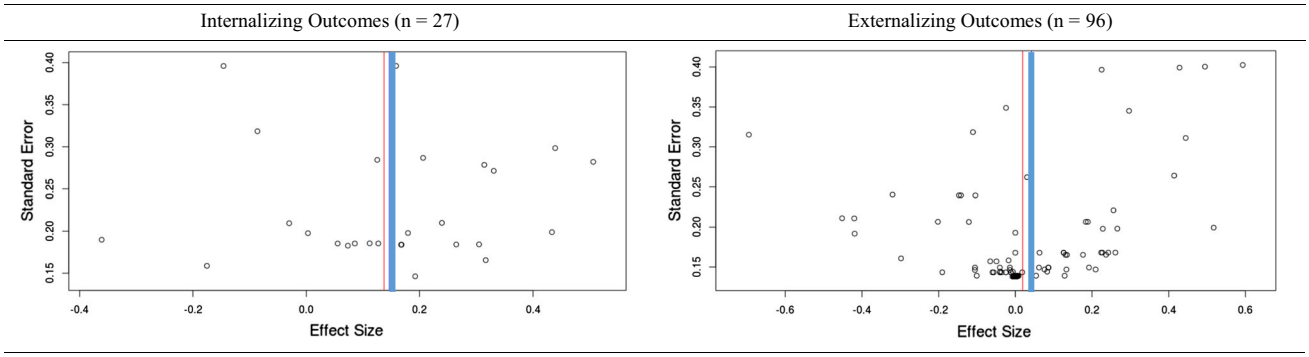

Fig. 2 Funnel plot for internalizing and externalizing outcomes. **Thin line* is unadjusted effect size estimate; *thick line* is effect size estimate adjusted for publication bias

met criteria) and selective reporting (96% met criteria). As shown in Fig. 2, studies reported unclear risk of bias in allocation concealment (63% met criteria) and incomplete outcome data (54% met criteria), and reported high risk of bias in blinding of participants and personnel and outcome assessment (8% and 21% met criteria, respectively).

Discussion

This systematic review and meta-analysis examined the effectiveness of school-based psychosocial interventions delivered by teachers on internalizing and externalizing outcomes and the factors that may moderate effects of those interventions. Overall, the interventions included in this review demonstrated a statistically significant positive effect on internalizing outcomes, but not on externalizing outcomes. In addition, teacher-provided interventions were statistically significantly more effective with internalizing outcomes in comparison to externalizing outcomes.

The findings from this study differs from previous reviews that concluded that teacher interventions are effective with both internalizing and externalizing outcomes (Durlak et al. 2011), and additional reviews that found structured interventions delivered by teachers were particularly effective with externalizing behavior such as aggression (Wilson et al. 2003). Prior reviews, however, did not include the same criteria for study inclusion, which may have accounted for the differing results across studies. The current study included only RCT's, for example, and was very specific to teacher interventions, while other studies investigated school mental health interventions across a greater variation of study designs and service providers. It is worth noting that the findings of the current study regarding greater effectiveness with internalizing outcomes is similar to past reviews on empirically supported psychotherapy studies for children and adolescents. Weisz, Hawley, and Doss (2004) reported higher numbers of significant treatment effects for internalizing outcomes than those for externalizing outcomes. Similarly, in a meta-analysis of psychotherapy for depression in children and adolescents, Weisz and McCarty (2006) did not find the treatment effects were transferrable to externalizing outcomes. This review and the other psychotherapy reviews included many of the same types of intervention modalities such as cognitive-behavioral interventions, and for this reason, it is recommended that program planners and researchers may want to re-examine specific interventions in relationship to their overall effectiveness for externalizing outcomes.

Durlak et al. (2011) recommended that reviews on school-based interventions examine gender and ethnicity to determine how these factors may impact the effectiveness

Table 5 Results of the Cochrane Collaboration's tool for assessing risk of bias

	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other source of bias
Benner et al. (2012)	+	+	+	−	+	+	+
Brown et al. (2001)	+	+	+	−	+	+	?
Botvin et al. (2001)	+	?	?	?	+	+	+
Botvin et al. (2006)	+	?	?	−	?	+	+
Cappella et al. (2012)	+	+	−	−	+	+	+
Chaplin et al. (2006)	+	+	?	−	?	+	+
Crean and Johnson (2013)	+	+	−	−	+	+	+
Eron et al. (2002)	+	?	−	−	+	+	+
Frey et al. (2005)	+	−	−	−	?	+	?
Fonagy et al. (2009)	+	+	−	+	+	+	+
Gillham et al. (2012)	+	+	?	−	+	+	?
Gillham et al. (2007)	+	+	?	+	?	+	+
Hecht et al. (2003)	+	?	?	−	+	+	+
Holt et al. (2008)	+	?	?	?	?	+	?
Iovannone et al. (2009)	+	+	−	+	?	+	+
Leff et al. (2010)	+	?	−	−	−	+	−
Leff et al. (2009)	+	?	−	−	−	+	−
Metz et al. (2006)	+	?	?	?	−	?	?
MACSRG (2007)	+	+	−	−	+	+	+
Murray and Malmgren (2005)	+	+	−	−	−	+	+
O'Neill et al. (2010)	+	+	−	?	+	+	+
Simonsen et al. (2011)	+	+	−	−	−	+	+
Spoth et al. (2005)	+	+	−	+	+	+	+
Spoth et al. (2008)	+	+	−	+	+	+	+
Number of "+"s	24	15	2	5	13	23	17

"+" = criteria were met in primary studies, thus no bias present; "?" = whether or not criteria met unclear from reading of primary studies; and "−" = criteria were not met in primary studies, and thus bias was present

of interventions delivered by teachers and other school professionals. The findings from this review suggest that interventions were more effective for female students with internalizing outcomes than males, and that interventions were more effective for Caucasian as compared to other ethnic minority students with externalizing outcomes. The moderating differences of gender are important to consider for future program planning and studies and may have

some clinical implications because females are frequently diagnosed with depression. This could suggest that the effects of teacher-delivered interventions may be helpful in alleviating internalizing symptoms for the female students. On the other hand, there is some prior evidence that male adolescents receive less positive intervention outcomes than female adolescents (Westwood and Pinzon 2008) and this suggests that males may need different types of interventions. One possible explanation to the gender differences could be that male students were more impacted by the interaction between internalizing and externalizing behaviors, resulting in poorer/smaller treatment effect in comparison to their female counterparts. This type of interaction has been found in other studies (e.g., Marmorstein 2007).

The gender differences need to be investigated in future studies on teacher-delivered mental health interventions and it may also be important to consider the gender of the teacher in interaction with the male students because many teachers are female. In addition, both race and gender may interact in moderating treatment effects and this was not investigated in this study but needs to be analyzed in future studies. Additionally, ethnic minority male adolescents (Hispanic and African-American) may be more disadvantaged than Caucasians because they may face additional stresses of racism and structural and cultural barriers to learning that exist in schools (Colins et al. 2010). The differences in the ways race (Caucasian) moderated outcomes further supports differences and is an important finding for school mental health practice because over half of school mental health interventions have been reported to be provided in low-income schools (Amaral et al. 2011) with underserved, ethnic minority students. Over 60% of the current reported sample was from African-American and Hispanic students. Educational disparities are prevalent in public schools, and this study suggests that the differences in outcomes achieved based on race need to be further addressed in future school interventions (Garcia et al. 2016). The majority of interventions from primary studies purported to be effective with different ethnic groups and some interventions were developed and tested with Hispanics (i.e., Hecht et al. 2003). All interventions were not adapted in the same way to be culturally competent with different minority populations and this could possibly account for the differences found in race and gender. A past systematic review showed that culturally adapted, evidence-based psychosocial interventions are generally effective for ethnic minority youths (Robles et al. 2016), but other reviews demonstrated that there is little evidence that cultural adaptations make evidence-based interventions significantly more effective than non-adapted interventions (Huey and Polo 2008). Outcomes may be improved for ethnic minority populations especially males when future studies further examine the cultural competence of school-based psychosocial interventions delivered by teachers and their effectiveness with ethnic minority groups.

Teachers are most effective in the delivery of Tier 1 level interventions where they can utilize their existing skills in the classroom. These findings support other reviews that indicate teachers are effective with Tier 1 interventions (e.g., Durlak et al. 2011; Franklin et al. 2012; Paulus et al. 2016; Stormont et al. 2011). The use of intervention manuals may play a significant part in the effectiveness of the teacher-delivered psychosocial interventions as most Tier 1 interventions provide teachers with structured curriculums. Overall, in these reviews, 22 studies (91.7%) were from a manualized program and only 2 studies (8.30%) were not from a manualized program. Past literature has also emphasized the importance of studying implementation factors such as training and supervision because they are critical to the success of teachers (Barnes et al. 2014; Frey et al. 2011). Ongoing supervision may be especially relevant for the fidelity of studies asking teachers to implement Tier 2 or Tier 3 interventions as opposed to Tier 1 interventions because teachers utilize existing professional skills in Tier 1 interventions but are developing new skills in Tier 2 and Tier 3 interventions. Most of the effect size estimates (thus most studies) in this review reported training and supervision across all three Tiers and this may be related to the fact that all studies were RCT's but these implementation practices may not reflect practice as usual in a school setting. In this review, 22 studies (91.7%) reported intervention under supervision with only 1 study (4.2%) where a study did not report this information and an additional 1 study (4.2%) from providers that reported not receiving supervision. The high rates of supervision provide more confidence in the quality of teachers' practices in this study; however, the articles did not provide enough information to evaluate the differences in the quality of the training and/or supervision being provided so that the overall quality of these implementation factors could be assessed. Future studies need to further evaluate distinct types of training and supervision and the quality of the supervision provided in relationship to teacher effectiveness, and especially for those teachers who are involved with Tier 2 and Tier 3 interventions. Future studies may also want to examine implementation of training/supervision under real world conditions to understand what is feasible in the day-to-day practices of teachers.

Limitations

First, despite our effort of being exhaustive in our literature search, there's no way for us to know if we included all the RCT's studies that met the criteria for our review. Second, the choice to include only RCT studies limits the type of evidence that is considered in this study. This decision made

it possible for us to examine more rigorous studies on psychosocial interventions delivered by teachers increasing our confidence in the effects of teacher-delivered interventions. The meta-analysis method used also made it necessary to limit several studies because it was not possible to calculate effect sizes based on the statistical techniques used in the individual studies. Reviews that include studies based on different study designs however, could potentially impact or even change the results of this current study. Third, our findings concerning the effects of teacher interventions on internalizing and externalizing outcomes are limited by the measures used in the primary studies that were frequently not direct measures of these constructs but were instead measures that assessed cognitive attributes and behaviors that are associated with internalizing and externalizing outcomes (e.g., hostile cognitive attributions, aggression, depression, substance use). It is recommended that in future studies that researchers analyze other moderators that may influence outcomes. Moderators such as the interactions between SES, race and gender, and implementation factors may be important areas to pursue. Fourth, some analyses used a small number of effect sizes from a relatively small number of studies, thus some caution must be used when interpreting these findings, especially for internalizing outcomes which were only included in 5 studies. On the positive side, this decision gave this reviewer greater generalizability to all studies that met our inclusion criteria. Additionally, we used robust variance estimation, a method handling the dependence between effect size estimates in the same study. This method significantly increased the sample size of this review from 24 studies to 121 effect size estimates, which compensates the relatively small number of studies in this review. Finally, the results of this meta-analysis could be confounded by the fact that the pooled effect size estimate came from various measures of internalizing and externalizing behaviors. For example, one intervention could have had a great impact on two specific outcome measures but not on two others and still have a significant overall treatment effect. This review tried to investigate the treatment effect size by specific measure but was unsuccessful because of the very low number of effect size per measure, resulting in too small degrees of freedom (<4) for the analytic result to be trustworthy (Tipton 2013). As a result, we chose the analytic strategy presented in this review as a starting point and encourage further analysis when more studies are available in the future.

Conclusion

This study builds on and extends existing school mental health literature that suggests that teacher-delivered Tier 1 interventions are effective. We were able to examine differential effects for internalizing and externalizing outcomes and found that Tier 1 interventions delivered by teachers are more effective with internalizing outcomes than externalizing outcomes. The overall treatment effects were found to be moderated by both race and gender; these interventions are more effective with female for internalizing outcomes and with Caucasians for externalizing outcomes. Future studies need to examine further the gender and cultural competence of existing school-based, psychosocial interventions and to improve upon these interventions for externalizing outcomes, male students, and ethnic minorities. School practitioners may also want to carefully evaluate Tier 1, teacher-delivered interventions to examine whether they are working well for all their students including males and students of color and supplement other classroom approaches when needed. This review also points to the need to give more attention to intervention implementation and report on the quality of teacher supervision to help gain a better understanding for the implementation factors that may influence outcomes. This may be especially true for Tier 2 and 3 interventions that may require greater effort by teachers and the learning of new skills. The training and supervision associated with teacher-delivered, psychosocial interventions may also benefit from further evaluation to determine feasibility and outcomes in the everyday practices of teachers.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval This article does not contain any studies with human participants performed by any of the authors.

References

*Indicates articles included in this review

Amaral, G., Geierstanger, S., Soleimanpour, S., & Brindis, C. (2011). Mental health characteristics and health-seeking behaviors of adolescent school-based health center users and nonusers. *Journal of School Health, 81*(3), 138–145.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Anderson-Butcher, D., & Ashton, D. (2004). Innovative models of collaboration to serve children, youths, families, and communities. *Children & Schools, 26*(1), 39–53.

Bain, R. M., & Diallo, A. F. (2016). Mental health services in school-based health centers: Systematic review. *The Journal of School Nursing: The Official Publication of the National Association of School Nurses, 32*(1), 8.

Barnes, T., Smith, S., & Miller, M. (2014). School-based cognitive-behavioral interventions in the treatment of aggression in the United States: A meta-analysis. *Aggression and Violent Behavior, 19*(4), 311–321.

Baskin, T. W., Slaten, C. D., Sorenson, C., Glover-Russell, J., & Merson, D. N. (2010). Does youth psychotherapy improve academically related outcomes? A meta-analysis. *Journal of Counseling Psychology, 57*(3), 290.

Bender, W. N., & Shores, C. (Eds.). (2007). *Response to intervention: A practical guide for every teacher*. New York: Corwin Press.

*Benner, G. J., Nelson, J. R., Sanders, E. A., & Ralston, N. C. (2012). Behavior intervention for students with externalizing behavior problems: Primary-level standard protocol. *Exceptional Children, 78*(2), 181–198.

Berzin, S. C., O'Brien, K. H. M., Frey, A., Kelly, M. S., Alvarez, M. E., & Shaffer, G. L. (2011). Meeting the social and behavioral health needs of students: Rethinking the relationship between teachers and school social workers. *Journal of School Health, 81*(8), 493–501.

*Botvin, G. J., & Griffin, K. W. (2006). Preventing youth violence and delinquency through a universal school-based prevention approach. *Prevention Science, 7*(1), 403–408.

*Botvin, G. J., Griffin, K. W., Diaz, T., & Ifill-Williams, M. (2001). Preventing binge drinking during early adolescence: One- and two-year follow-up of a school-based preventive intervention. *Psychology of Addictive Behaviors, 15*(4), 360–365.

*Brown, E. C., Low, S., Smith, B. H., & Haggerty, K. P. (2001). Outcomes from a school-randomized controlled trial of steps to respect: A bullying prevention program. *School Psychology Review, 40*(3), 423–443.

*Cappella, E., Hamre, B., Kim, H. Y., Henry, D. B., Frazier, S. L., & Atkins, M. S. (2012). Teacher consultation and coaching within mental health practice: Classroom and child effects in urban elementary schools. *Journal of Consulting and Clinical Psychology, 80*(4), 597–610.

*Chaplin, T. M., Gillham, J. E., Reivich, K., Elkon, A. G. L., Samuels, B., Freres, D. R., et al. (2006). Depression prevention for early adolescent girls: A pilot study of all girls versus co-ed groups. *Journal of Early Adolescence, 26*(1), 110–126.

Chitiyo, M., & Wheeler, J. J. (2009). Challenges faced by school teachers in implementing positive behavior support in their school systems. *Remedial and Special Education, 30*(1), 58–63.

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

Colins, O., Vermeiren, R., Vreugdenhil, C., van den Brink, W., Doreleijers, T., & Broekaert, E. (2010). Psychiatric disorders in detained male adolescents: A systematic literature review. *The Canadian Journal of Psychiatry, 55*(4), 255–263.

Cooper, H. (2016). *Research synthesis and meta-analysis: A step-by-step approach*. New York: SAGE Publication.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Russell Sage Foundation.

*Crean, H. F., & Johnson, D. B. (2013). Promoting alternative thinking strategies (PATHS) and elementary school aged children's aggression: Results from a cluster randomized trial. *American Journal of Community Psychology, 52*(1), 56–72.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82,* 474–501.

Eiraldi, R., Power, T. J., Schwartz, B. S., Keiffer, J. N., McCurdy, B. L., Mathen, M., et al. (2016). Examining effectiveness of group cognitive-behavioral therapy for externalizing and internalizing disorders in urban schools. *Behavior Modification, 40*(4), 611–639.

England, M. J., Butler, A. S., & Gonzales, M. L. (2015). *Psychosocial interventions for mental and substance use disorders: A framework for establishing evidence-based standards*. Washington, DC: The National Academy Press.

*Eron, L., Huesmann, R., Spindler, Guerra, N., Henry, D., Tolan, P., et al. (2002). A cognitive-ecological approach to preventing aggression in urban settings: Initial outcomes for high-risk children. *Journal of Consulting and Clinical Psychology, 70*(1), 179–194.

Ferri, B. A. (2012). Undermining inclusion? A critical reading of response to intervention (RTI). *International Journal of Inclusive Education, 16*(8), 863–880.

*Fonagy, P., Twelow, S. W., Vernberg, E. M., Nelson, J. M., Dill, E. J., Little, T. D., et al. (2009). A cluster randomized controlled trial of child-focused psychiatric consultation and a school systems-focused intervention to reduce aggression. *Journal of Child Psychology and Psychiatry, 50*(5), 607–616.

Franklin, C. G., Kim, J. S., Ryan, T. N., Kelly, M. S., & Montgomery, K. L. (2012). Teacher involvement in school mental health interventions: A systematic review. *Children and Youth Services Review, 34*(5), 973–982.

*Frey, K. S., Hirschstein, M. K., Snell, J. L., Edstrom, L. V. S., MacKenzie, E. P., & Broderick, C. J. (2005). Reducing playground bullying and supporting beliefs: An experimental trial of the steps to respect program. *Developmental Psychology, 41*(3), 479–491.

Frey, A., Lingo, A., & Nelson, C. M. (2011). Positive behavior support and response to intervention in elementary schools. In H. Walker & M. K. Shinn (Eds.), *Interventions for achievement and behavior problems: Preventive and remedial approaches* (pp. 397–433). National Association for School Psychologists: Bethesda.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*(1), 93–99.

Garcia, A. R., Kim, M., & DeNard, C. (2016). Context matters: The state of racial disparities in mental health services among youth reported to child welfare in 1999 and 2009. *Children and Youth Services Review, 66,* 101–108.

*Gillham, J. E., Reivich, K. J., Brunwasser, S. M., Freres, D. R., Chajon, N. D., Kash-MacDonald, V. M., et al. (2012). Evaluation of a group cognitive-behavioral depression prevention program for young adolescents: A randomized effectiveness trial. *Journal of Clinical Child & Adolescent Psychology, 41*(5), 621–639.

*Gillham, J. E., Reivich, K. J., Freres, D. R., Chaplin, T. M., Shatte, A., Samuels, B., et al. (2007). School-based prevention of depressive symptoms: A randomized controlled study of the effectiveness and specificity of the Penn Resiliency Program. *Journal of Consulting and Clinical Psychology, 75*(1), 9–19.

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

Gleser, L. J., & Olkin, I. (2009). Stochastically dependent effect sizes. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis*. New York, NY: Russell Sage Foundation.

Greenfield, R., Rinaldi, C., Proctor, C. P., & Cardarelli, A. (2010). Teachers' perceptions of a response to intervention (RTI) reform effort in an urban elementary school: A consensual qualitative analysis. *Journal of Disability Policy Studies, 21*(1), 47–63.

Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review. *BMC Psychiatry, 10*(1), 1.

Han, S. S., & Weiss, B. (2005). Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology, 33*(6), 665–679.

Hawkins, S. M., & Heflin, L. J. (2010). Increasing secondary teachers' behavior-specific praise using a video self-modeling and visual performance feedback intervention. *Journal of Positive Behavior Interventions, 13*(2), 97–108.

*Hecht, M. L., Marsiglia, F. F., Elek, E., Wagstaff, D. A., Kulis, S., Dustman, P., et al. (2003). Culturally grounded substance use prevention: An evaluation of the keepin' it R.E.A.L. Curriculum. *Prevention Science, 4*(4), 233–248.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*(1), 39–65.

Helker, W. P., Schottelkorb, A. A., & Ray, D. (2007). Helping students and teachers CONNECT: An intervention model for school counselors. *Journal of Professional Counseling, Practice, Theory, & Research, 35*(2), 31–45.

Higgins, J. P. T., & Green S. (Eds.). (2011) Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Retrieved from: www.handbook.cochrane.org.

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., et al. (2011). The Cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ, 343,* d5928.

Hoagwood, K. E., Olin, S. S., Kerker, B. D., Kratochwill, T. R., Crowe, M., & Saka, N. (2007). Empirically based school interventions targeted at academic and mental health functioning. *Journal of Emotional and Behavioral Disorders, 15*(2), 66–92.

*Holt, L. J., Bry, B. H., & Johnson, V. L. (2008). Enhancing school engagement in at-risk, urban minority adolescents through a school-based, adult mentoring intervention. *Child & Family Behavior Therapy, 30*(4), 297–318.

Huey, S. J., & Polo, A. J. (2008). Evidence-based psychosocial treatments for ethnic minority youths. *Journal of Clinical Child & Adolescent Psychology, 371,* 262–301.

IBM SPSS Statistics [computer program]. (2015). Version 23.0. Chicago, Illinois: SPSS Inc.

*Iovannone, R., Greenbaum, P. E., Wang, W., Kincaid, D., Dunlap, G., & Strain, P. (2009). Randomized controlled trial of the prevent-teach-reinforce (PTR) tertiary intervention for students with problem behaviors: Preliminary outcomes. *Journal of Emotional and Behavioral Disorders, 17*(4), 213–225.

Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research, 79*(1), 491–525.

Kavale, K. A., Holdnack, J. A., & Mostert, M. P. (2005). Responsiveness to intervention and the identification of specific learning disability: A critique and alternative proposal. *Learning Disability Quarterly, 28*(1), 2–16.

Kearney, C. A. (2016). *Managing school absenteeism at multiple tiers: An evidence-based and practical guide for professionals*. New York: Oxford University Press.

Kelly, M. S. (2008). *The domains and demands of school social work practice: A guide to working effectively with students, families and schools*. New York: Oxford University Press.

Konstantopoulos, S., & Hedges, L. V. (2009). Analyzing effect sizes: Fixed-effect models. In H. Cooper, L. V. Hedges, & J. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 279–293). New York: Russell Sage Foundation.

*Leff, S. S., Gullan, R. L., Paskewich, B. S., Abdul-Kabir, S., Jawad, A. F., Grossman, M., et al. (2009). An initial evaluation of a culturally adapted social problem-solving and relational aggression prevention program for urban African-American relationally aggressive girls. *Journal of Prevention & Intervention in the Community, 37*(4), 260–274.

*Leff, S. S., Waasdorp, T. E., Raskewich, B., Bullan, R. L., Jawad, A. F., MacEvoy, J. P., et al. (2010). The preventing relational aggression in schools everyday program: A preliminary evaluation of acceptability and impact. *School Psychology Review, 39*(4), 569–587.

Lenski, S. (2011). What RTI means for content area teachers. *Journal of Adolescent & Adult Literacy, 55*(4), 276–282.

MAGSRG Study. The Metropolitan Area Child Study Research Group. (2007). Changing the way children "think" about aggression: Social-cognitive effects of a preventive intervention. *Journal of Consulting and Clinical Psychology, 75*(1), 160–167.

Marmorstein, N. R. (2007). Relationships between anxiety and externalizing disorders in youth: The influences of age and gender. *Journal of Anxiety Disorder, 21*(3), 420–432.

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., et al. (2010). Lifetime prevalence of mental disorders in US adolescents: Results from the National Comorbidity Survey Replication-Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(10), 980–989.

*Metz, A. E., Fuemmeler, B. F., & Brown, R. T. (2006). Implementation and assessment of an empirically validated intervention program to prevent tobacco use among African-American middle-school youth. *Journal of Clinical Psychology in Medical Settings, 13*(3), 229–238.

*Murray, C., & Malmgren, K. (2005). Implementing a teacher-student relationship program in a high-poverty urban school: Effects on social, emotional, and academic adjustment and lessons learned. *Journal of School Psychology, 43,* 137–152.

*O'Neill, J. M., Clark, J. K., & Jones, J. A. (2010). Promoting mental health and preventing substance abuse and violence in elementary students: A randomized control study of the Michigan model for health. *Journal of School Health, 81*(6), 320–330.

Ojio, Y., Yonehara, H., Taneichi, S., Yamasaki, S., Ando, S., Togo, F., et al. (2015). Effects of school-based mental health literacy education for secondary school students to be delivered by school teachers: A preliminary study. *Psychiatry and Clinical Neurosciences, 69*(9), 572–579.

Paulus, F. W., Ohmann, S., & Popow, C. (2016). Practitioner review: School-based interventions in child mental health. *Journal of Child Psychology and Psychiatry*. doi:10.1111/jcpp.12584.

Perou, R., Bitsko, R. H., Blumberg, S. J., Pastor, P., Ghandour, R. M., Gfroerer, J. C., & Huang, L. N. (2013). Mental health surveillance among children – United States, 2005 – 2011. CDC Morbidity and Mortality Weekly Report. Retrieved from: http://www.cdc.gov/mmwr/preview/mmwrhtml/su6202a1.htm?s_cid=su6202a1_w.

R Development Core Team (2008). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rajaleid, K., Nummi, T., Westerlund, H., Virtanen, P., Gustafsson, P. E., & Hammarström, A. (2016). Social adversities in adolescence predict unfavorable trajectories of internalized mental health symptoms until middle age: Results from the Northern Swedish Cohort. *The European Journal of Public Health, 26*(1), 23–29.

Reinke, W. M., Stormont, M., Herman, K. C., Puri, R., & Goel, N. (2011). Supporting children's mental health in schools: Teacher perceptions of needs, roles, and barriers. *School Psychology Quarterly, 26*(1), 1–13.

Ringwalt, C. L., Pankratz, M. M., Jackson-Newsom, J., Gottfredson, N. C., Hansen, W. B., Giles, S. M., et al. (2010). Three-year trajectory of teachers' fidelity to a drug prevention curriculum. *Prevention Science, 11*(1), 67–76.

Robles, E. H., Maynard, B. R., Salas-Wright, C. P., & Todic, J. (2016). Effects of culturally adapted substance use interventions for Latino adolescents: A systematic review and meta-analysis. Published advance online, *Research on Social Work Practice*.

Rones, M., & Hoagwood, K. (2000). School-based mental health services: A research review. *Clinical Child and Family Psychology Review, 3*(4), 223–241.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.). (2006). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. West Sussex: Wiley.

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children, 31*(3), 351–380.

*Simonsen, B., Myers, D., & Briere, D. E., III. (2011). Comparing a behavioral Check-In/Check-Out (CICO) intervention to standard practice in an urban middle school setting using an experimental group design. *Journal of Positive Behavior Interventions, 13*(1), 31–48.

*Spoth, R., Randall, G. K., Shin, C., & Redmond, C. (2005). Randomized study of combined universal family and school preventive interventions: Patterns of long-term effects on initiation, regular use, and weekly drunkenness. *Psychology of Addictive Behaviors, 19*(4), 372–381.

*Spoth, R. L., Randall, G. K., Trudeau, L., Shin, C., & Redmond, C. (2008). Substance use outcomes 5½ years past baseline for partnership-based, family-school preventive interventions. *Drug and Alcohol Dependence, 96*(1), 57–68.

Stagman, S., & Cooper, J. L. (2010). Children's mental health; what every policymaker should know. (Policy brief). Mailman School of Public Health, National Center for Children in Poverty, Columbia University. Retrieved from: http://www.lacdcfs.org/katiea/practices/docs/Childrens_MH.pdf.

Stormont, M., Reinke, W., & Herman, K. (2011). Teachers' knowledge of evidence-based interventions and available school resources for children with emotional and behavioral problems. *Journal of Behavioral Education, 20*(2), 138–147.

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods, 5*(1), 13–30.

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental Life Course Criminology, 2,* 85–112.

Tillery, A. D., Varjas, K., Meyers, J., & Collins, A. S. (2010). General education teachers' perceptions of behavior management and intervention strategies. *Journal of Positive Behavior Interventions, 12*(2), 86–102.

Tipton, E. (2013). Robust variance estimation in meta-regression with binary dependent effects. *Research Synthesis Methods, 4*(2), 169–187.

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods, 20*(3), 375–393.

Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics, 40*(6), 604–634.

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods, 45*(2), 576–594.

Vevea, J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods, 10*(4), 428–443.

Vreeman, R. C., & Carroll, A. E. (2007). A systematic review of school-based interventions to prevent bullying. *Archives of Pediatrics and Adolescent Medicine, 161*(1), 78–88.

Weisz, J. R., Hawley, K. M., & Doss, A. J. (2004). Empirically tested psychotherapies for youth internalizing and externalizing problems and disorders. *Child and Adolescent Psychiatric Clinics of North America, 13*(4), 729–815.

Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin, 132*(1), 132–149.

Westwood, M., & Pinzon, J. (2008). Adolescent male health. *Paediatrics & Child Health, 13*(1), 31.

Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine, 33*(2), S130–S143.

Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *Journal of Consulting and Clinical Psychology, 71*(1), 136–149.

Wolmer, L., Hamiel, D., & Laor, N. (2011). Preventing children's posttraumatic stress after disaster with teacher-based intervention: A controlled study. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*(4), 340–348.