

Meta-analysis and systematic review of teacher-delivered mental health interventions for internalizing disorders in adolescents

Article

Accepted Version

Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Shelemy, L., Harvey, K. and Waite, P. (2020) Meta-analysis and systematic review of teacher-delivered mental health interventions for internalizing disorders in adolescents. *Mental Health & Prevention*, 19. 200182. ISSN 2212-6570 doi: <https://doi.org/10.1016/j.mhp.2020.200182> Available at <https://centaur.reading.ac.uk/90128/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1016/j.mhp.2020.200182>

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Meta-analysis and systematic review of teacher-delivered mental health interventions for internalizing disorders in adolescents

Lucas Shelemy^{a*}, Dr Kate Harvey^a and Dr Polly Waite^{a, b},

^aSchool of Psychology and Clinical Language Sciences, University of Reading, Whiteknights, Reading, RG6 6AL, U.K.

^bDepartments of Experimental Psychology and Psychiatry, University of Oxford, Anna Watts Building, Oxford OX2 6GG, UK.

*Corresponding author: l.shelemy@pgr.reading.ac.uk; +44 (0)118 378 8523; ORCID Lucas Shelemy: 0000-0002-2831-3434

Declaration of Interest Statement

The PhD of the lead researcher (LS) is funded by The Charlie Waller Memorial Trust, a charity aiming to improve mental health in young people. The Trust were not involved in the research process, from study design to submission. PW is funded through a NIHR-funded postdoctoral fellowship (PDF-2016-09-092). The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article. The lead author may be contacted for further information or data from the study.

Acknowledgements

We would like to thank the undergraduate students Rosalind Jones, Sophia Humphries, Daniel Gilson and Katie Patterson who assisted with the screening and quality assessment of articles. The PhD of the lead researcher (LS) is funded by The Charlie Waller Memorial Trust, a charity aiming to improve mental health in young people. The Trust had no involvement in the research process, from study design to submission. PW is funded by a National Institute for Health Research (NIHR) Postdoctoral Research Fellowship (PDF-2016-09-092) for this project. This publication presents independent research. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article. The lead author may be contacted for further information or data from the study.

Funding

This work was supported by a studentship awarded to LS by The Charlie Waller Memorial Trust.

Abstract

A large proportion of emotional problems begin in adolescence and negatively impact quality of life into adulthood. There have been multiple teacher-delivered, classroom-based programs created to reduce symptoms of internalizing problems amongst young people. This meta-analysis and systematic review aims to examine the effectiveness of teacher-delivered interventions for depression, anxiety, post-traumatic stress disorder (PTSD) and obsessive-compulsive disorder (OCD) symptoms in adolescents, and a range of factors that may impact outcomes. Database searches were conducted from PsycInfo, Medline (PubMed), Scopus, the Cochrane Library and the British Educational Index (from database inception to January 2020). Quality assessment of studies was conducted using the EPHPP Quality Assessment Tool. Fifty-two intervention studies were identified that quantitatively assessed, via controlled design, intervention effects on internalizing disorder symptoms. Two meta-analyses found teacher-delivered interventions were significantly better than control conditions at improving depression ($g = -0.12$), anxiety ($g = -0.13$) and PTSD symptoms ($g = -0.66$) in students. Improvements were only maintained at follow-up for anxiety symptoms and no effect sizes reached a 'small' threshold. However, the effect sizes were 'moderate' within the context of universal prevention programs for young people. No interventions measured OCD outcomes. Overall, the findings suggest that teachers may not be the optimal deliverers of mental health interventions. Improved outcomes were associated with interventions that lasted up to 16 weeks, had program sessions of 45-90 minute duration, and included two or more days of training for teachers. Future studies should aim to improve reporting quality on number of sessions, teacher training and fidelity of intervention. Increased reporting of outcomes from adolescents with high versus low baseline mental health scores would enable a better understanding of for whom interventions are most effective.

Keywords: Teacher; mental health; meta-analysis; systematic review; intervention; school

1.1 Introduction

Approximately 20% of adolescents experience an emotional mental health problem in any given year (Merikangas et al., 2010), with 2.6% of children and young people experiencing depression and 6.5% experiencing an anxiety disorder (Polanczyk, Salum, Sugaya, Caye, & Rohde, 2015). Many of these mental health problems persist into adulthood, with 75% of lifetime mental health disorders first emerging before the age of 24 years (Kessler et al., 2007). Depression and anxiety disorders negatively impact adolescents' academic performance (Owens, Stevenson, Hadwin, & Norgate, 2012; Verboom, Sijtsma, Verhulst, Penninx, & Ormel, 2014), school attendance (Archambault, Janosz, Morizot, & Pagani, 2009), social relationships (Cook, Williams, Guerra, Kim, & Sadek, 2010), risk of suicidal behavior (Hetrick, Parker, Robinson, Hall, & Vance, 2011) and future employment (Butterworth, Leach, Pirkis, & Kelaher, 2012). The prevalence of mental health problems has led to a rise in the number of interventions attempting to address these difficulties, including *targeted* therapeutic approaches towards young people identified as being 'at-risk' in some way (e.g. students showing early signs of anxiety), and *universal* sessions delivered to a large group of young people (e.g. everyone in a class) (Mrazek & Haggerty, 1994).

Schools may be an ideal setting for the implementation of mental health interventions for adolescents. They are a practical environment for interventions due to the available space, resources and equipment for teaching and learning (Barrett & Pahl, 2006; Masia-Warner, Nangle, & Hansen, 2006). They may also be seen as a less stigmatizing setting to learn new skills and information about mental health than clinical environments (Rambaldo, Wilding, Goldman, McClure, & Friedberg, 2001). A recent meta-analysis of universal and targeted school-based mental health prevention programs delivered by both external (e.g., mental health professionals) and internal (e.g., teachers) facilitators found small effect sizes for improving depression (Hedges' $g = 0.23$) and anxiety outcomes ($g = 0.20$) at post-

intervention compared to control conditions, which were maintained at 12-month follow-up (Werner-Seidler, Perry, Callear, Newby, & Christensen, 2017). A review of school-based PTSD interventions found an overall effect size of ($d = 0.68$) at post-intervention compared to control conditions (Rolfesnes & Idsoe, 2011).

1.2 Teachers as deliverers of interventions

Many classroom-based programs have been created to be delivered by teachers rather than mental health professionals (Rones & Hoagwood, 2000; Dray et al., 2017; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011), with 40.8% of school-based mental health interventions involving teachers in some way throughout the delivery and up to 18.4% of interventions delivered solely by teachers (Franklin, Kim, Ryan, Kelly, & Montgomery, 2012). Although critics of teacher-delivered interventions note that many school staff lack knowledge of and training in mental health and intervention delivery (Frey, Lingo, & Nelson, 2011), there are other factors that suggest teachers may be effective in delivering mental health interventions in schools. Teachers have highly developed classroom management skills and a deep understanding of their pupils and how to best capture their attention (Hester et al., 2004; Leflot, van Lier, Onghena, & Colpin, 2010). Their pre-existing relationship with the students may also be beneficial when teaching about sensitive topics and core mental health concepts. In addition, their extensive contact with a large number of young people means that there are opportunities to disseminate mental health-related program content to the wider curriculum and school environment. Given effective training, preparation and supervision, it may be the case that teachers are effective and sustainable deliverers of mental health interventions (Easton & Erchul, 2011; Lane, Weisenbach, Little, Phillips, & Wehby, 2006).

Nevertheless, to date there is mixed evidence to suggest that teacher-led mental health interventions are effective for children and adolescents across broad age ranges (5-18 years).

Sub-analyses of previous reviews have found that interventions that were delivered or supported by school staff significantly improved depression ($n = 23$, $g = 0.17$, Werner-Seidler, Perry, Calear, Newby, & Christensen, 2017; $n = 6$, median $d = 0.39$, Calear & Christensen, 2010) and anxiety symptoms ($n = 19$, $g = 0.18$, Werner-Seidler et al., 2017; $n = 6$, median $d = 0.36$, Neil & Christensen, 2009), compared to control conditions. Effect sizes from these reviews ranged from below the ‘small’ threshold to ‘small’ (Cohen, 1988). However, while a sub-analysis conducted by Teubert & Piquart (2011) of teacher-delivered programs focusing on the prevention of anxiety across the same large age range found a similar effect size ($n = 16$, $g = 0.15$), the effectiveness of the programs was not significantly better than the control interventions at post-test and at follow up. There are no known analyses of teacher-delivered interventions for PTSD or OCD outcomes.

1.3 Focus on adolescence

To date, there have been no reviews that focus specifically on the effectiveness of teacher-delivered interventions for the adolescent age group aged 11-18. The increase in the prevalence and severity of mental health difficulties in adolescents compared to children (under 11) may impact on the effectiveness of school-based mental health programs (Green et al., 2005; Waite & Creswell, 2014). Significant differences between child and adolescent samples have been found in previous prevention and intervention meta-analyses in both directions (Reynolds, Wilson, Austin, & Hooper, 2012; Teubert & Piquart, 2011). It is feasible that teacher-delivered interventions targeted at adolescents have different outcomes compared to previous analyses that have combined data from both age groups (Calear & Christensen, 2010; Neil & Christensen, 2009; Teubert & Piquart, 2011; Werner-Seidler et al., 2017).

1.4 Factors that may moderate intervention outcomes

There has been little study of the teacher-as-deliverer factors that may impact the efficacy of such interventions (e.g., the content of the intervention, the number of sessions or length of the intervention, training and supervision given to teachers and fidelity with which the program is delivered), or the young people receiving the intervention (e.g., whether it is delivered to the whole class or those identified as being ‘at risk’ in some way).

In terms of the content of school-based interventions, previous studies have not found a significant difference in outcomes between cognitive-behavioral skills programs and other content types (e.g. social skills training or relaxation practice) on symptoms of depression and anxiety, when delivered by a range of professionals (Werner-Seidler et al., 2017). However, it is not known whether this is the case for interventions delivered by teachers.

The number of sessions and/or length of intervention may be linked to intervention efficacy. Two meta-analyses have demonstrated that a greater number of sessions in psychological treatment and prevention programs are associated with larger effect sizes for reducing anxiety symptoms in adolescents (Fisak, Richard, & Mann, 2011; Reynolds et al., 2012).

The fidelity of the delivery of the intervention appears to impact outcomes post-intervention and at long-term follow up (Frey et al., 2011), with greater clinician adherence to treatments and interventions manuals associated with improved outcomes for substance misuse and youth behavior problems (Derzon, Sale, Springer, & Brounstein, 2005; Schoenwald, Sheidow, & Letourneau, 2004). Training and ongoing supervision from professionals with mental health expertise might be assumed to be of critical importance on intervention fidelity and subsequent outcomes, given that many teachers lack specific knowledge and training around mental health (Roeser & Midgley, 1997). However, to date, the impact of teacher fidelity, training and supervision in relation to outcomes has not been evaluated (Franklin et al., 2017).

Finally, there may be differences in outcomes depending on the characteristics of the population receiving the intervention, i.e., whether participants are ‘at risk’ (e.g. with a high baseline symptom level or having parents with mental health difficulties) or from a universal sample. Whilst some studies have found that adolescents with higher baseline anxiety scores have shown greater improvements in wellbeing outcomes compared to those with low baseline anxiety following intervention (Blake et al., 2018), others have found that baseline positive and negative affect does not moderate the relationship between intervention and mental health outcomes in adolescence (Rash, Matsuba, & Prkachin, 2011; Wang et al., 2017).

1.5 The Present Study

This meta-analysis and systematic review aims to analyze and synthesize published controlled studies of teacher-led mental health interventions in schools targeting DSM-5 internalizing disorders (depression, anxiety disorders, post-traumatic stress disorder (PTSD) and obsessive-compulsive disorder (OCD)) that are delivered to adolescents aged 11-18 years (APA, 2013). It will examine:

- a) The effectiveness of teacher-delivered interventions in reducing symptoms or diagnoses of internalizing disorders in adolescents compared to control conditions, and
- b) Whether specific factors related to the content and delivery of interventions (intervention content, number of sessions, length of intervention, amount of training and supervision, fidelity of delivery, delivery to high risk groups) are associated with findings.

This will provide important information about whether teacher-delivered interventions are effective and what factors are associated with better outcomes, in order to improve the design and delivery of interventions (Adelman & Taylor, 2003; Fazel, Hoagwood, Stephan, & Ford, 2014; Han & Weiss, 2005a; Neil & Christensen, 2009).

2.1 Method

2.2 Protocol and registration

The meta-analysis and systematic review method follow the guidelines from the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) (Moher, Liberati, Tetzlaff, Altman, & PRISMA Group, 2009), and the AMSTAR checklist for systematic review quality (Shea et al., 2017) (both included in Appendix A). The search terms and criteria were developed under consultation with a research librarian. A research protocol for the review was devised on the basis of the existing literature and PROSPERO records prior to starting the review data search. The protocol and search strategy were registered on PROSPERO; registration number CRD42015027610.

2.3 Search Strategy

An electronic search was conducted in January 2016 and then updated in January 2020 using the following electronic databases: PsycINFO, Medline (PubMed), Scopus, Cochrane Library of Systematic Reviews and the British Educational Index (BEI). The search terms used were formed of three main terms, using the keywords (school* OR sixth-form) AND (mental OR depress* OR anxi* OR phobi* OR panic* OR mood) AND (prevent* OR intervent*). These search terms were based on multiple scoping searches. The aim of these terms was to generate a large database of interventions within schools, to ensure a comprehensive and inclusive list of possible articles. Secondary searches were also conducted via hand-searching, using the reference lists of included articles. Search terms were used to search databases that contained articles with these terms in the abstract or title. All eligible studies based on the abstract and title were added to an EndNote database and transferred into Microsoft Excel. These were de-duplicated automatically.

2.4 Screening criteria

Eligible studies in the present review had to fulfil the following inclusion criteria: (1) published within a peer reviewed journal; (2) published in the English language, due to a lack

of resources for accurate translation; (3) the intervention was part of either a randomized or non-randomized controlled trial (non-randomized studies were included to capture the breadth of the research literature and account for studies in which randomization was not possible or described); (4) the content was focused on improving mental and/or emotional health; (5) an intervention was delivered to adolescents between the mean age of 11-18 years (rounded up to the nearest year); (6) school teachers were explicitly described as the deliverers of all of the intervention content; (7) have control or comparison conditions that did not include content related to mental health or wellbeing; (8) there was recorded outcome data for at least one DSM-5 internalizing disorder (depression, anxiety and related disorders, PTSD and OCD) (APA, 2013) that took the form of either clinical diagnoses or relevant symptoms from a validated symptom rating scale. Any length of follow-up for longitudinal studies was included. There were no restrictions of publication date for inclusion. Studies in which teachers supervised participants use of an online program were not included as the teachers were not active deliverers of the intervention.

2.5 Screening procedure

A pilot screening was first conducted to ensure inter-coder reliability. Two researchers screened the same 500 studies. Outcomes were compared and Cohen's kappa (κ) statistics were calculated for categorical screening codes and found to be excellent ($\kappa=.96$; McHugh, 2012).

The first stage of selection involved the lead researcher independently screening the papers' abstracts and titles based on the inclusion criteria. Four undergraduate students acted as second raters during the screening stages. In the second stage the same individuals screened the full text of the remaining papers. Reliability between raters was checked and found to be good (κ range = .86 - .96). Discrepancies between researchers were resolved through discussion with the second and third authors.

The screening procedure and number of articles screened and accepted at each stage, can be seen in Figure 1.

2.6 Data extraction and review synthesis

The data were extracted from the studies by the lead researcher and an undergraduate student independently. A pilot data extraction was conducted with 10 studies and high inter-coder reliability was found ($\kappa=.84$; McHugh, 2012). Discrepancies were resolved via discussion with the co-authors. The data extraction form was based on the Cochrane EPOC checklist (EPOC, 2017) and recorded details about study characteristics, intervention design, the results and measures of outcomes. The standardized form for data extraction also contained data on (a) demographic information including study location, setting, design, participant ethnicity and socio-economic status; (b) number of participants; (c) mean age of participants and age range; (d) gender of participants; (e) follow-up time points in longitudinal studies; (f) number of teacher program leaders, (g) length of teacher training, (h) length of intervention, (i) name of intervention; (j) number of sessions delivered; (k) length of each session; (l) length of follow-up; (m) comparison groups used, (n) internalizing disorder/symptom measures used.

Several factors required additional interpretation for the data extraction. This included the content of interventions, the amount of ongoing and post-training supervision teachers received, and the measure and amount of fidelity to the program manual from teachers. The interventions were classified as ‘cognitive-behavioral’, ‘social, relationship and interpersonal’ or ‘relaxation and meditation’ based on authors’ explicit description of the intervention. These classification labels were used as they were the most common types of intervention delivered. As interventions were not being compared to each other, interventions could have more than one content classification. Studies in which the type of intervention content was not named were not included in content sub-analyses. Any information about

supervision and fidelity of the delivery of the intervention was copied into the data extraction form. Supervision, when reported, was classified as ‘regular’ when studies reported that face-to-face meetings took place throughout the intervention. Studies that described supervision in a way that implied it did not occur regularly (e.g. optional telephone calls) or did not provide details about supervision were classified as ‘not identified’. Studies that reported the intervention fidelity as either ‘high’ and/or ‘good’ were given a ‘high fidelity’ classification. Studies that reported fidelity otherwise (e.g. components were not delivered) were classified as ‘reduced fidelity’.

The following outcomes were also recorded: (o) mean and standard deviation outcomes for measures of internalizing disorders at pre, post and follow-up time points; p effect sizes; (q) notes about bias. Further data was requested from several authors of papers and this was included, if provided. The study characteristics and outcomes were extracted and synthesized as a review.

2.7 Risk of bias

All studies were systematically evaluated for risk of bias using the Effective Public Health Practice Project (EPHPP) Quality Assessment Tool for Quantitative Studies (EBHPP, 1998). This standardized assessment tool allows classification of studies based on quality of design and measures (1) selection bias, (2) study design, (3) confounders, (4) blinding, (5) data collection methods, (6) withdrawals and dropouts, (7) intervention integrity and (8) analyses. Total scores were generated, with 1 indicating the highest quality level, and 3 being the lowest. The assessment was conducted by the lead author and repeated by an undergraduate research assistant for reliability ($\kappa = .92$). Discrepancies were resolved in discussion with the co-authors.

2.8 Data Meta-Analysis

Primary meta-analyses were conducted separately for studies that measured outcomes for (i) depression symptoms/diagnoses (ii) anxiety symptoms/diagnoses, and (iii) PTSD symptoms. The significance level for these three analyses was set as 1.7% following a Bonferroni correction. No studies measured outcomes for OCD or other anxiety related disorders. Studies that measured both outcomes were included in both primary analyses. Seven papers were excluded from all analyses due to insufficient outcome data reported (Fitzpatrick et al., 2013; Karam et al., 2008; Lombas et al., 2019; Malgady, Rogler, & Costantino, 1990; Patton et al., 2006; Roberts et al., 2018; Williford et al., 2012). Two further papers were excluded from both analyses as the measure and outputs were an integrated internalizing disorder symptom score (Flynn, Joyce, Weihrauch, & Corcoran, 2018; Perry et al., 2014). As several studies used multiple measures for the same symptom, an a priori decision was made to use outcome data from one measure in each study. When multiple measures were used, the BDI (Beck Depression Inventory; Beck, Steer, & Brown, 1996) or CES-D (Centre for Epidemiological Studies Depression Scale; Radloff, 1977) for depressive symptoms and the SCAS (Spence Children's Anxiety Scale; Spence, 1998) for anxiety symptoms were selected for analysis as these were the most frequently used measures across the studies.

A random effects model was used due to the diversity in study design, populations and evaluation methods. The standardized mean difference effect size Hedges' g was calculated using the post-intervention mean and standard deviation values of both the intervention and control conditions. This outcome measure was used to determine the impact of the interventions on mental health outcomes. Missing standard deviation values were calculated from standard errors and sample sizes reported in the paper where possible. Negative effect sizes signified that participants in the intervention condition reported a greater reduction in symptoms than those in the control group. Positive values signified that

participants in the control group reported a greater reduction in symptoms than those in the intervention condition. The size of the Hedges' g was interpreted using Cohen's (1988) guidance as small (0.2), medium (0.5) and large (0.8). The data was analyzed using Comprehensive Meta-Analysis software (Biostat, 2018). The I^2 statistic was calculated to identify the heterogeneity of the studies in each of the analyses. The heterogeneity thresholds used were 0-29% for insubstantial heterogeneity, 30%-49% for moderate heterogeneity, 50-74% for substantial heterogeneity, and 75-100% for high heterogeneity (Borenstein, Hedges, Higgins, & Rothstein, 2009). Additional analyses were also conducted for studies that included follow-up data and for studies with selective targeted interventions.

Three separate sub-analyses were conducted for intervention content (studies that taught a) cognitive-behavioral skills, b) social, relationship and interpersonal skills, and c) relaxation and/or mindfulness skills). Sub-analyses were also conducted for length of intervention (<2 weeks, 2-16 weeks, >16 weeks), number of sessions (<8 sessions, 8-16 sessions, >16 sessions), session length (<45 minutes, 45-90 minutes, >90 minutes), and training length (<1 day, 2+ days) as moderator variables. The ranges within each category were decided by identifying the most common intervention characteristic across the studies. Sub-analyses of studies with regular supervision and fidelity (high vs low) were also conducted. Several studies had supervision that was 'not identified'. Studies also varied significantly in how high-risk adolescents were identified across the studies. The high heterogeneity among the studies where supervision was coded as 'not identified' and where studies involved high-risk groups meant that sub-analyses of these factors could not be conducted. A 5% significance level was used for the sub-group analyses, yet findings were interpreted with caution given the multiplicity of statistical tests.

A sensitivity analysis was conducted in which studies with a high risk of bias (a score of 3 identified using the EPHPP assessment tool) were excluded in order to examine whether

this changed the pattern of results. This was not necessary for analyses of PTSD outcomes as risk of bias for these studies was low. The ‘fail safe n’ was calculated to estimate the number of unpublished studies that would need to exist for the overall effect size to be considered non-significant (Rosenthal, 1979).

3.1 Results

3.2 Study characteristics

Fifty-two independent studies were included in the present review, including a total of 49,084 participants. The sample size in each study varied from 36 (Baker & Butler, 1984) to 7,741 participants (Williford et al., 2012). All but three of the included interventions were universal and delivered to a whole class of students. The two targeted selective interventions were delivered to students who were identified as having a high risk of developing mental health problems (O’Leary-Barrett et al., 2013) and additional educational needs (Martin, 2008). One indicative intervention was delivered to students that scored ‘borderline to abnormal’ on the Strengths and Difficulties Questionnaire (SDQ; Alampay et al., 2019). Further details about the factors relating to the content and delivery of the interventions can be found in Table 1.

The most common study location was Australia (n = 14, 27%), followed by the USA (n = 11, 21%), Israel (n = 4, 8%), Ireland (n = 4, 8%) and the UK (n = 3, 6%). The mean age of the participants across the studies ranged from 11 (Barrett & Turner, 2001) to 17 (Van der Gucht et al., 2017). Two interventions (4%) were delivered solely to male students (Gelkopf & Berger, 2008; Martin, 2008), and five interventions (10%) were delivered only to female students (Flynn et al., 2018; Garcia, Pintor, Vazquez, & Alvarez-Zumarraga, 2011; Green, Grant, & Rynsaardt, 2007; Harnett & Dadds, 2004; Pluess & Boniwell, 2015).

Twenty-three of the included studies (44%) used a no intervention control, 12 used a waiting list control (23%), and fifteen studies (29%) involved an active control including art

and physical health education (Garcia et al., 2011) and exercise and nutrition training (Melnyk et al., 2013). Two studies (4%) did not report the type of control condition used. All studies assessed outcomes using adolescent-reported measures of symptoms rather than diagnostic interviews.

Thirty-eight studies (73%) measured depression outcomes, 33 studies (63%) measured anxiety outcomes and four studies (9%) measured PTSD symptoms.

Twenty studies (38%) measured outcomes at follow-up, with time points ranging from 2 to 24 months.

3.3 Risk of bias

The methodological quality of the included studies varied considerably. Eighteen studies (35%) received the highest quality rating (1), whilst 13 studies (25%) received the lowest quality rating (3). Table 1 presents the quality ratings for each study. Forty-three studies (83%) failed to report adequate blinding procedures for participants, teachers or researchers. Whilst thirty-three studies (63%) reported statistically similar baseline characteristics between the intervention and control condition for age, sex and pre-intervention mental health outcomes, the remaining 19 studies (37%) either did not report on baseline characteristics or stated that there were significant differences between conditions at pre-test measurement. Four studies (8%) had high drop-out attrition rates of over 40% at post-intervention (Perry et al., 2014; Sawyer et al., 2010; Silbert & Berry, 1991; Williford et al., 2012). The Egger's test for study effects and funnel plot asymmetry suggested that there was no publication bias (Egger, Smith, Schneider, & Minder, 1997). The 'fail safe n' was high indicating that the effect estimates are reliable (for depression: $n = 335$, for anxiety: $n = 265$; Rosenthal, 1979).

3.4 Meta-Analyses

Forty-two papers (81%) were included in the two primary meta-analyses of universal interventions (29 in the depression outcome meta-analysis and 26 in the anxiety outcome meta-analysis). The primary analyses are presented in forest plots in Figure 2 for depression outcomes, Figure 3 for anxiety outcomes and Figure 4 for PTSD outcomes.

3.4.1 Outcomes post-intervention

Random-effects meta-analyses on the between-group difference for end-point scores of universal interventions demonstrated a significant effect in favor of the intervention group on measures of depression ($n = 29$, $g = -0.12$, 95% CI = -0.19 to -0.05, $p = .001$, $I^2 = 19\%$) and anxiety ($n = 26$, $g = -0.13$, 95% CI = -0.21 to -0.04, $p = .005$, $I^2 = 11\%$) with a low level of heterogeneity between studies (Higgins, Thompson, Deeks, & Altman, 2003). However, in both analyses Hedges' g did not meet the threshold of a small effect size (Cohen, 1988).

There was no overall significant effect of the selective or indicated interventions on depression or anxiety outcomes compared to control.

Four studies reported outcomes using measures of PTSD symptoms. Three of these studies found significant improvements in symptoms following intervention compared to control, with an overall medium effect size ($n = 4$, $g = -0.66$, 95% CI = -1.13 to -0.18, $p = .006$, $I^2 = 0\%$). Again, heterogeneity was low.

The results from the sensitivity analyses excluding studies with a high risk of bias did not change the pattern of results for either depression ($n = 24$, $g = -0.11$, 95% CI = -0.19 to -0.03, $p = .006$, $I^2 = 24\%$) or anxiety ($n = 19$, $g = -0.13$, 95% CI -0.24 to -0.03, $p = .014$, $I^2 = 5\%$).

3.4.2 Outcomes at follow-up

An analysis was conducted for when follow-up data was reported in full. No significant effect was found for depression ($n = 20$, $g = -0.04$, 95% CI = -0.10 to 0.02, $p = .23$, $I^2 = 0\%$) at follow-up. Whilst a significant effect in favor of the intervention group at follow-

up was found for anxiety ($n = 9$, $g = -0.08$, 95% CI = -0.16 to <0.00 , $p = .042$, $I^2 = 0\%$), this did not reach the threshold of a small effect size (Cohen, 1988). There was no moderating effect of the time in which follow-up data was recorded for both analyses. No follow-up data was reported for the PTSD studies.

3.5 Intervention Factors

Table 2 shows findings from the sub-analyses of intervention content, intervention and session length, training, supervision and program fidelity, including the sample size, Hedges' g statistic, confidence intervals, p values and heterogeneity statistics. Although there were a number of significant differences between intervention groups and controls, notably, none of the effect sizes reached the 'small' effect size threshold (Cohen, 1988).

3.5.1 Outcomes for intervention content

Twenty-seven studies (52%) used an intervention that taught cognitive-behavioral skills, 15 (29%) focused on social, relationship and interpersonal skills and 9 (17%) on relaxation/mindfulness. A small number of studies taught positive psychology (Shoshani & Steinmetz, 2014), Acceptance and Commitment Therapy skills (ACT; Van der Gucht et al., 2017), psychoeducation (Perry et al., 2014) and Dialectical Behavior Therapy skills (Flynn et al., 2018).

Interventions that taught cognitive-behavioral skills and relaxation/mindfulness skills indicated a significant effect in favor of the intervention condition for depression outcomes (and reached a small effect size for relaxation/mindfulness) but not for anxiety outcomes. For interventions that included social, relationship and interpersonal skills, there was not a significant effect for the intervention condition in comparison to control for either depression or anxiety.

3.5.2 Intervention and session length.

Reported individual session length ranged from a 3-12 minutes (e.g. a teacher-led mindfulness exercise: Britton et al., 2014) to 3 hours, with the majority lasting between 45 to 90 minutes (n = 30, 58%). The number of sessions ranged from 2 sessions in total to daily lessons for 5 months, with the majority of programs reporting 8 to 16 sessions (n = 31, 60%). Most intervention sessions occurred weekly or fortnightly (n = 35, 67%). There was high variability in the time the interventions ran for, with some occurring across two weeks, whilst others intermittently across an academic year.

Sub-analyses found significant improvements in favor of the intervention condition for depression and anxiety outcomes for interventions with 8 to 16 sessions. These findings were also shown for interventions with sessions lasting 45-90 minutes, and/or that lasted for up to 16 weeks. Intervention conditions with less than 8 sessions also demonstrated a significant improvement in anxiety outcomes compared to controls. There was no significant difference between the intervention and control condition for interventions with over 16 sessions, and/or with sessions lasting over 90 minutes or less than 45 minutes for either depression or anxiety outcomes.

3.5.3 Training.

Thirty-seven studies (71%) reported that training was provided for teachers prior or during the intervention. Training length ranged from 2 hours (Buttigieg et al., 2015) to 6 days (Gillham et al., 2012). Studies that included training that lasted 2 days or more found significant improvements favoring the intervention condition in both depression and anxiety outcomes. No significant effect was found for studies with 1 day or less of training for teachers for depression or anxiety outcomes.

3.5.4 Supervision.

Twenty-one studies (40%) reported that support was provided for teachers from the research team. Thirteen of these studies (25%) reported regular in-person supervision meetings. For

these studies, significant improvements in depression but not anxiety were found in favor of the intervention condition. Other forms of support included optional meetings or contact with the research team for advice or email.

3.5.5 Fidelity to intervention.

Twenty-one studies (40%) assessed and reported on teachers' fidelity or adherence to the intervention manual. Methods included videotaping of sessions, diary reports from teachers and random observations. There was high heterogeneity between fidelity measurements and subsequent interpretation. Eleven of the studies (21%) reported fidelity as 'high', 'good' or 'delivered fully'. The remaining 10 studies (19%) reported that several components or sessions were not fully taught, due to reasons such as teachers choosing to focus on key concepts rather than the whole curriculum (Harnett & Dadds, 2004) and not being able to fit all the sessions into the school year (Kindt, Kleinjan, Janssens, & Scholte, 2014).

Significant improvements for depression outcomes in the intervention condition compared to control was found for studies with high fidelity but not for anxiety. No significant differences were found between intervention and control condition for studies with lower reported fidelity for both depression or anxiety.

3.5.6 High-risk groups.

Six studies (12%) reported additional outcomes for high-risk groups in addition to the universal sample outcomes, reporting findings from participants with high baseline anxiety (Lowry-Webster, Barrett, & Dadds, 2001; Silbert & Berry, 1991), hopelessness (Gillham et al., 2012) and depression scores (Sheffield et al., 2006; Spence, Sheffield, & Donovan, 2003), as well as existing parent psychopathology (Kindt et al., 2014). An additional study analyzed whether baseline parental support for the participant (a factor associated with risk of developing mental health problems) moderated outcomes using the data from Sawyer et al.,

(2010) (Spence et al., 2014). High levels of heterogeneity between high-risk characteristics meant a sub-analysis was not conducted. Two studies (4%) did not report a significant difference between high-risk groups in the intervention and control conditions (Sheffield et al., 2006; Silbert & Berry, 1991). The remaining five studies (10%) all reported significant improvements in depression and/or anxiety outcomes in favor of the intervention compared to control for those with high risk of developing mental health problems (e.g. $d = 0.28$, Gillham et al., 2012; $d = 0.34$, Spence et al., 2014), with small effect sizes when reported.

4.1 Discussion

This study presents a meta-analysis of teacher-delivered mental health interventions to improve internalizing disorder outcomes in adolescents. We found that compared to (largely non-active) control conditions, teacher-delivered interventions significantly improved both depression ($g = -0.12$), anxiety ($g = -0.13$) and PTSD ($g = -0.66$) symptoms immediately post-intervention. However, with the exception of PTSD outcomes and the effect of relaxation/mindfulness for depression, none of the effect sizes found in the meta-analyses or sub-analyses reached the ‘small’ threshold described by Cohen (1988). Sensitivity analyses showed these findings to be robust regardless of the quality of studies. Studies that measured outcomes at follow-up time points showed that this remained significant for anxiety ($g = -0.08$), but not for depression outcomes. Interventions that were up to 16 weeks in length and/or with sessions lasting 45-90 minutes were associated with improved outcomes for both depression and anxiety outcomes. Likewise, studies with two or more days of training for teachers found improved depression and anxiety outcomes compared to control interventions. Significantly reduced depression outcomes compared to controls were found for programs in which the content was based on cognitive-behavioral or relaxation/meditation skills, levels of fidelity were recorded as high or where regular supervision was provided for teachers.

The overall meta-analysis effect sizes and follow-up results are relatively consistent with other sub-analyses of teacher-led interventions across wider age ranges (Calear & Christensen, 2010; Neil & Christensen, 2009; Werner-Seidler et al., 2017). There has been no previous examination of whether the amount of training teachers receive is associated with effectiveness. Studies with two or more days of training showed improved outcomes, suggesting that adequate time for teachers to become familiar with intervention concepts and materials is crucial to ensure program effectiveness. Future studies could explore whether amount of training is linked to intervention fidelity and subsequent outcomes.

The medium effect size found for studies measuring PTSD outcomes is based on just four studies, all with similar program content and shared study authorship. The effect size is consistent with a previous review of school-based PTSD interventions (Rolfesnes & Idsoe, 2011). The studies took place in Israel and Sri Lanka which were identified by the authors as having high levels of trauma following violent conflict and/or poverty (Berger & Gelkopf, 2009; Berger, Gelkopf, & Heineberg, 2012). Thus, future investigation should be conducted to investigate whether these findings can be extrapolated to other countries and contexts. Nevertheless, this finding is interesting given the increasing numbers of trauma-informed training programs for teachers, emphasizing the potential for school staff to effectively support students who have experienced trauma (McInerney & McKlindon, 2014; Steele & Malchiodi, 2012).

The present study found that cognitive-behavioral and relaxation/mindfulness interventions showed statistically significant improvements on depression outcomes in contrast to other types of program content. This differs to findings from a previous review in which no differences between intervention content types were found (Werner-Seidler et al., 2017). For anxiety symptoms, the lack of significant findings for relaxation is consistent with recent research that suggests that the use of relaxation is associated with limited impact in

treatment (Peris et al., 2015). It may also be that cognitive-behavioral interventions in these programs did not involve components that have been demonstrated to be associated with improved outcomes, such as the use of exposure (Ale, McCarthy, Rothschild, & Whiteside, 2015). Social, relationship and interpersonal skills interventions did not seem to improve mental health outcomes, in contrast with findings from one meta-analysis showing that social-interaction interventions in schools and community settings can improve depression and anxiety in young people (García-Carrión, Villarejo-Carballido, & Villardón-Gallego, 2019). Due to limited reporting from studies about how the intervention sessions and content were structured, in this review, studies which shared content types were classified under the same category, regardless of the proportions of content type delivered. It is feasible that interventions which solely delivered one content type may differ in efficacy compared to those with a blend of different content types.

The finding that regular supervision for teachers was related to intervention efficacy for depression outcomes was consistent with findings that teachers feel more confident teaching about and supporting mental health in students when they are receiving regular consultation and support (Han & Weiss, 2005; Shelemy, Harvey, Waite, 2019). To improve intervention outcomes, the level of teacher engagement in an intervention may be crucial. This may be achieved via increased supervision and an intervention that meets the needs of school staff and that accounts for their time and competing work demands. Consultation with teachers prior to intervention development may enhance intervention engagement and subsequent efficacy (Lynn, McKay, & Atkins, 2003; Rothì, Leavey, & Best, 2008).

Despite significant improvements found for both depression and anxiety outcomes in the intervention group compared to controls, the overall low effect sizes bring into question the value of teacher-delivered universal interventions that aim to improve mental health outcomes for adolescents, especially given the additional resource-cost and need for prior

training of teachers. Effect sizes were smaller than those reported by studies of clinician-delivered interventions, indicating that teacher-led programs may not be the optimal choice for schools (Calear & Christensen, 2010; Werner-Seidler et al., 2017). The null effect at follow-up for depression outcomes suggests that the interventions have limited preventative effects. In contrast, significant follow up effects for anxiety outcomes have potential to have a meaningful impact at a population level (Lakens, 2013). A better understanding is needed of whether such interventions are cost-effective given the low effect sizes, limited follow-up benefits and additional training for teachers.

It is important to consider that effect sizes should be interpreted when considering the research context of the specific area of study (Bloom, Hill, Black, & Lipsey, 2008; Funder & Ozer, 2019; Harris, 2009). One meta-analysis of 11 reviews of universal mental health preventative and promotion programs for young people found that the median effect sizes on internalizing behavior was $d = 0.12$, with the median average effect from all universal prevention programs being between 0.07 to 0.16 standard deviations (Tanner-Smith, Durlak, & Marx, 2018). The authors emphasize that Cohen's interpretation of effect sizes (e.g. 0.2, 0.5 and 0.8 suggesting small, medium and large effects; (Cohen, 1988) are not suitable for when examining effect sizes from universal prevention studies (Tanner-Smith et al., 2018). Large effect sizes of 0.50 or higher within the context of school mental health intervention programs are likely to be unattainable (Bloom et al., 2008; Tanner-Smith et al., 2018). Instead effect sizes should be interpreted relative to those from other meta-analyses in the same field of research (Hill, Bloom, Black, & Lipsey, 2008). When considering this model of interpretation, the current findings suggest 'moderate' effect sizes that are within the upper 50th percentile of mean effects obtained for universal internalizing mental health programs.

Only seven studies reported independent outcomes for high-risk adolescents, or for students who had high baseline levels of depression or anxiety. Significant differences were

found for five out of the seven studies in favor of the intervention condition and where effect sizes were reported, they were in the small range. However, different assessments of ‘high-risk’ and the small number of studies means that the size of the overall effect is uncertain. Among universal interventions, where there are significant improvements, it is currently not clear whether the improved outcomes are for a subset of students, or across larger groups of students in a school. To justify the continued use of interventions within schools to reduce symptoms of depression or anxiety, it would be important to demonstrate that there are improvements among those who may have poor mental health or have a high risk of developing difficulties in the future, rather than only among young people who show low levels of depression and anxiety symptoms at baseline.

4.2 Strengths and limitations

This review and meta-analysis were conducted with a high level of systematic rigor. The search for studies was exhaustive, with study selection, data extraction and quality assessment all completed with high inter-rater reliability. The method followed PRISMA and AMSTAR guidelines to ensure high quality reporting of findings. However, there are a number of limitations. Pre-test baseline data were not consistently reported across the studies and subsequently were not used in the meta-analyses of post-intervention data, which potentially may impact on study findings. Adjustments to the significance level for the sub-analyses was not made (unlike for the primary analyses). Given the high multiplicity of the sub-analyses, any interpretations made must be treated with caution.

All relevant outcome data from studies were used in the analyses, despite not always being the primary focus of change for each study. The results from studies with secondary measures of internalizing disorder outcomes may have confounded the current analyses. Future research should explore whether internalizing disorder symptoms as primary or secondary measure moderates meta-analysis outcomes.

The low quality of many studies has potential to impact on any conclusions made; however, our sensitivity analyses excluding studies with low quality of evidence did not alter outcomes, suggesting that the results are stable across different quality of study conditions. Due to the research setting, blinding of teachers and adolescents was not feasible, and many studies did not report appropriate research team blinding. As in previous studies (e.g. Weare & Nind, 2011), the limited reporting of study characteristics (e.g. training, supervision and fidelity) made it difficult to clearly establish the association between intervention factors and outcomes. The process of calculating intervention fidelity was often not described and subsequently could be a source of potential authorship bias. Going forward, we would suggest that all intervention studies should conform to pre-established guidelines and checklists to improve the quality of the research and reporting style of studies.

4.3 Recommendations

The present study suggests that teacher-led mental health interventions have limited effectiveness in helping reduce depression and anxiety in adolescents. Significant improvements found in depression and anxiety at post-intervention. The size of the effect can be interpreted within the context of universal prevention programs ('moderate'; Tanner-Smith et al., 2018) or by Cohen's criteria (below 'small'; Cohen, 1988). Over 2 days of training appears to be important to ensure improved outcomes for both depression and anxiety outcomes, whilst other factors, such as the type of intervention, regular supervision and intervention fidelity appear to be related to outcomes for symptoms of depression. A limitation of meta-analyses is that they often assess a heterogeneous group of programs with varying program quality. The effect sizes of universal teacher-delivered interventions are variable, with many included studies showing 'moderate' effect sizes. It is crucial that future research continues to investigate the factors that impact on a program's efficacy.

To better understand for whom interventions are most effective, it is crucial that future evaluations report outcomes for participants who are at high risk of developing mental health difficulties. Regression analyses would be useful to better understand the effectiveness of interventions based on different predictive factors about students.

Nevertheless, the interventions delivered may have led to other possible benefits beyond symptom outcomes that were not measured in the studies, such as knowledge and understanding of mental health difficulties or help seeking behavior. Qualitative studies in which students are invited to reflect on their experiences of receiving teacher-delivered interventions may enable us to better understand potential impacts of the interventions that may not be captured in current analyses of outcomes.

Finally, none of studies reported cost-effectiveness of the intervention. As the agenda of school-based interventions becomes more focused on sustainability and ease of dissemination, future interventions should detail the costs of intervention and share data that encourages cost-effectiveness analyses (Brunwasser & Garber, 2016).

4.4 Conclusion

Our review indicates that teacher-delivered interventions are more effective than control interventions at improving depression and anxiety outcomes in adolescents post-intervention yet benefits only persist for anxiety symptoms in the longer term. With the exception of PTSD outcomes and studies of high-risk groups, effect sizes were below the ‘small’ interpretation threshold, suggesting that teachers may not be the optimal intervention deliverers when attempting to improve mental health outcomes at a universal level. Training over two days is likely to be important to ensure higher intervention efficacy. Pre-test mental health group differences should be included in all future intervention analyses so that those for whom programs are most effective can be identified. Further research should be conducted to identify what factors impact on program quality and efficacy.

5.1 References

- Adelman, H. S., & Taylor, L. (2003). On Sustainability of Project Innovations as Systemic Change. *Journal of Educational and Psychological Consultation, 14*(1), 1–25.
https://doi.org/10.1207/S1532768XJEPC1401_01
- Alampay, L. P., Tan, L. J. T. G., Tuliao, A. P., Baranek, P., Ofreneo, M. A., Lopez, G. D., ... Angangco, T. (2019). A Pilot Randomized Controlled Trial of a Mindfulness Program for Filipino Children. *Mindfulness, 1–14*.
- Ale, C. M., McCarthy, D. M., Rothschild, L. M., & Whiteside, S. P. H. (2015). Components of cognitive behavioral therapy related to outcome in childhood anxiety disorders. *Clinical Child and Family Psychology Review, 18*(3), 240–251.
- American Psychological Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.
- Archambault, I., Janosz, M., Morizot, J., & Pagani, L. (2009). Adolescent behavioral, affective, and cognitive engagement in school: Relationship to dropout. *Journal of School Health, 79*(9), 408–415.
- Baker, S. B., & Butler, J. N. (1984). Effects of preventive cognitive self-instruction training on adolescent attitudes, experiences, and state anxiety. *Journal of Primary Prevention, 5*(1), 17–26.
- Barrett, P. M., & Pahl, K. M. (2006). School-Based Intervention: Examining a Universal Approach to Anxiety Management. *Australian Journal of Guidance and Counselling, 16*(01), 55–75. <https://doi.org/10.1375/ajgc.16.1.55>
- Barrett, P., & Turner, C. (2001). Prevention of anxiety symptoms in primary school children: Preliminary results from a universal school-based trial. *British Journal of Clinical Psychology, 40*(4), 399–410.
- Baum, N. L., Cardozo, B. L., Pat-Horenczyk, R., Ziv, Y., Blanton, C., Reza, A., ... Brom, D.

- (2013). Training teachers to build resilience in children in the aftermath of war: A cluster randomized trial. In *Child & Youth Care Forum* (Vol. 42, pp. 339–350). Springer.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Beck depression inventory-II. *San Antonio, 78*(2), 490–498.
- Berger, R., & Gelkopf, M. (2009). School-based intervention for the treatment of tsunami-related distress in children: a quasi-randomized controlled trial. *Psychotherapy and Psychosomatics, 78*(6), 364–371.
- Berger, R., Gelkopf, M., & Heineberg, Y. (2012). A teacher-delivered intervention for adolescents exposed to ongoing and intense traumatic war-related stress: A quasi-randomized controlled study. *Journal of Adolescent Health, 51*(5), 453–461.
- Biostat. (2018). Comprehensive Meta-Analysis. Biostat. Retrieved January 21, 2019, from <https://www.meta-analysis.com/>
- Blake, M. J., Blake, L. M., Schwartz, O., Raniti, M., Waloszek, J. M., Murray, G., ... McMakin, D. L. (2018). Who benefits from adolescent sleep interventions? Moderators of treatment efficacy in a randomized controlled trial of a cognitive-behavioral and mindfulness-based group sleep intervention for at-risk adolescents. *Journal of Child Psychology and Psychiatry, 59*(6), 637–649.
- Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness, 1*(4), 289–328.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons. Retrieved January 21, 2019, from <https://www.wiley.com/en-gb/Introduction+to+Meta+Analysis-p-9780470057247>
- Bradley, R. T., McCraty, R., Atkinson, M., Tomasino, D., Daugherty, A., & Arguelles, L.

- (2010). Emotion self-regulation, psychophysiological coherence, and test anxiety: results from an experiment using electrophysiological measures. *Applied Psychophysiology and Biofeedback*, 35(4), 261–283.
- Britton, W. B., Lepp, N. E., Niles, H. F., Rocha, T., Fisher, N. E., & Gold, J. S. (2014). A randomized controlled pilot trial of classroom-based mindfulness meditation compared to an active control condition in sixth-grade children. *Journal of School Psychology*, 52(3), 263–278.
- Brunwasser, S. M., & Garber, J. (2016). Programs for the Prevention of Youth Depression: Evaluation of Efficacy, Effectiveness, and Readiness for Dissemination HHS Public Access. *J Clin Child Adolesc Psychol*, 45(6), 763–783.
<https://doi.org/10.1080/15374416.2015.1020541>
- Butterworth, P., Leach, L. S., Pirkis, J., & Kelaheer, M. (2012). Poor mental health influences risk and duration of unemployment: a prospective study. *Social Psychiatry and Psychiatric Epidemiology*, 47(6), 1013–1021.
- Buttigieg, J. P., Shortt, A. L., Slaviero, T. M., Hutchinson, D., Kremer, P., & Toumbourou, J. W. (2015). A longitudinal evaluation of the Resilient Families randomized trial to prevent early adolescent depressive symptoms. *Journal of Adolescence*, 44, 204–213.
- Calcar, A. L., & Christensen, H. (2010). Systematic review of school-based prevention and early intervention programs for depression. *Journal of Adolescence*, 33(3), 429–438.
<https://doi.org/10.1016/J.ADOLESCENCE.2009.07.004>
- Challen, A. R., Machin, S. J., & Gillham, J. E. (2014). The UK Resilience Programme: A school-based universal nonrandomized pragmatic controlled trial. *Journal of Consulting and Clinical Psychology*, 82(1), 75.
- Clarke, G. N., Hawkins, W., Murphy, M., & Sheeber, L. (1993). School-based primary prevention of depressive symptomatology in adolescents: Findings from two studies.

Journal of Adolescent Research, 8(2), 183–204.

Cochrane Effective Practice and Organisation of Care (EPOC). Screening, data extraction and management. (2017). Retrieved January 21, 2019, from <http://epoc.cochrane.org/resources/epoc-resources-review-authors>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd edition. Routledge.

Cook, C. R., Williams, K. R., Guerra, N. G., Kim, T. E., & Sadek, S. (2010). Predictors of bullying and victimization in childhood and adolescence: A meta-analytic investigation. *School Psychology Quarterly*, 25(2), 65.

Derzon, J. H., Sale, E., Springer, J. F., & Brounstein, P. (2005). Estimating intervention effectiveness: Synthetic projection of field evaluation results. *Journal of Primary Prevention*, 26(4), 321–343.

Dowling, K., Simpkin, A. J., & Barry, M. M. (2019). A Cluster Randomized-Controlled Trial of the MindOut Social and Emotional Learning Program for Disadvantaged Post-Primary School Students. *Journal of Youth and Adolescence*, 1–19.

Dray, J., Bowman, J., Campbell, E., Freund, M., Wolfenden, L., Hodder, R. K., ... Wiggers, J. (2017). Systematic Review of Universal Resilience-Focused Interventions Targeting Child and Adolescent Mental Health in the School Setting. *Journal of the American Academy of Child & Adolescent Psychiatry*, 56(10), 813–824.
<https://doi.org/10.1016/j.jaac.2017.07.780>

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. *Child Development*, 82(1), 405–432.
<https://doi.org/10.1111/j.1467-8624.2010.01564.x>

Easton, J. E., & Erchul, W. P. (2011). An Exploration of Teacher Acceptability of Treatment

Plan Implementation: Monitoring and Feedback Methods. *Journal of Educational and Psychological Consultation*, 21(1), 56–77.

<https://doi.org/10.1080/10474412.2011.544949>

Effective Public Health Practice Project. (1998). Retrieved January 21, 2019, from

<https://merst.ca/ephpp/>

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109), 629–634.

Eggert, L. L., Thompson, E. A., Herting, J. R., & Nicholas, L. J. (1995). Reducing suicide potential among high-risk youth: Tests of a school-based prevention program. *Suicide and Life-Threatening Behavior*, 25(2), 276–296.

Fazel, M., Hoagwood, K., Stephan, S., & Ford, T. (2014). Mental health interventions in schools in high-income countries. *The Lancet Psychiatry*, 1(5), 377–387.

[https://doi.org/10.1016/S2215-0366\(14\)70312-8](https://doi.org/10.1016/S2215-0366(14)70312-8)

Fisak, B. J., Richard, D., & Mann, A. (2011). The Prevention of Child and Adolescent Anxiety: A Meta-analytic Review. *Prevention Science*, 12(3), 255–268.

<https://doi.org/10.1007/s11121-011-0210-0>

Fitzpatrick, C., Conlon, A., Cleary, D., Power, M., King, F., & Guerin, S. (2013). Enhancing the mental health promotion component of a health and personal development programme in Irish schools. *Advances in School Mental Health Promotion*, 6(2), 122–138.

Flynn, D., Joyce, M., Weihrauch, M., & Corcoran, P. (2018). Innovations in Practice:

Dialectical behaviour therapy–skills training for emotional problem solving for adolescents (DBT STEPS-A): evaluation of a pilot implementation in Irish post-primary schools. *Child and Adolescent Mental Health*, 23(4), 376–380.

Franklin, C. G. S., 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.
<https://doi.org/10.1016/J.CHILDYOUTH.2012.01.027>
- Franklin, C., Kim, J. S., Beretvas, T. S., Zhang, A., Guz, S., Park, S., ... Maynard, B. R. (2017). The Effectiveness of Psychosocial Interventions Delivered by Teachers in Schools: A Systematic Review and Meta-Analysis. *Clinical Child and Family Psychology Review*, *20*(3), 333–350. <https://doi.org/10.1007/s10567-017-0235-4>
- Frey, A., Lingo, A., & Nelson, C. M. (2011). Positive behavior support and response to intervention in elementary schools. *Interventions for Achievement and Behavior Problems: Preventive and Remedial Approaches*, 397–433.
- Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168.
- Garaigordobil, M. (2004). Effects of a Psychological Intervention on Factors of Emotional Development During Adolescence¹. *European Journal of Psychological Assessment*, *20*(1), 66–80.
- García-Carrión, R., Villarejo-Carballido, B., & Villardón-Gallego, L. (2019). Children and Adolescents Mental Health: A Systematic Review of Interaction-Based Interventions in Schools and Communities. *Frontiers in Psychology*, *10*.
- García, C., Pintor, J., Vazquez, G., & Alvarez-Zumarraga, E. (2011). Project Wings, a coping intervention for Latina adolescents: a pilot study. *Western Journal of Nursing Research*, *35*(4), 434–458.
- Gelkopf, M., & Berger, R. (2008). A school-based, teacher-mediated prevention program (ERASE-Stress) for reducing terror-related traumatic reactions in Israeli youth: A quasi-randomized controlled trial. *Journal of Child Psychology and Psychiatry*, *50*(8), 962–

971.

- Gillham, J. E., Reivich, K. J., Brunwasser, S. M., Freres, D. R., Chajon, N. D., Kash-MacDonald, V. M., ... Gallop, R. J. (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.
- Green, H., McGinnity, Á., Meltzer, H., Ford, T., & Goodman, R. (2005). Mental health of children and young people in Great Britain, 2004. Basingstoke: Palgrave Macmillan.
- Green, S., Grant, A., & Rynsaardt, J. (2007). Evidence-based life coaching for senior high school students: Building hardiness and hope. *International Coaching Psychology Review, 2*(1), 24–32.
- Han, S. S., & Weiss, B. (2005a). Sustainability of Teacher Implementation of School-Based Mental Health Programs. *Journal of Abnormal Child Psychology, 33*(6), 665–679.
<https://doi.org/10.1007/s10802-005-7646-2>
- Harnett, P. H., & Dadds, M. R. (2004). Training school personnel to implement a universal school-based prevention of depression program under real-world conditions. *Journal of School Psychology, 42*(5), 343–357.
- Harris, D. N. (2009). Toward policy-relevant benchmarks for interpreting effect sizes: Combining effects with costs. *Educational Evaluation and Policy Analysis, 31*(1), 3–29.
- Hester, P. P., Baltodano, H. M., Hendrickson, J. M., Tonelson, S. W., Conroy, M. A., & Gable, R. A. (2004). Lessons learned from research on early intervention: What teachers can do to prevent children's behavior problems. *Preventing School Failure: Alternative Education for Children and Youth, 49*(1), 5–10.
- Hetrick, S. E., Parker, A. G., Robinson, J., Hall, N., & Vance, A. (2012). Predicting suicidal risk in a cohort of depressed children and adolescents. *Crisis: The Journal of Crisis Intervention and Suicide Prevention, 33*(1), 13–20.

- Hiebert, B., Kirby, B., & Jaknavorian, A. (1989). School-based relaxation: Attempting primary prevention. *Canadian Journal of Counselling*.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Bmj*, *327*(7414), 557–560.
- Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*(3), 172–177.
- Karam, E. G., Fayyad, J., Karam, A. N., Tabet, C. C., Melhem, N., Mneimneh, Z., & Dimassi, H. (2008). Effectiveness and specificity of a classroom-based group intervention in children and adolescents exposed to war in Lebanon. *World Psychiatry*, *7*(2), 103–109.
- Kessler, R. C., Amminger, G. P., Aguilar-Gaxiola, S., Alonso, J., Lee, S., & Ustün, T. B. (2007). Age of onset of mental disorders: a review of recent literature. *Current Opinion in Psychiatry*, *20*(4), 359–364. <https://doi.org/10.1097/YCO.0b013e32816ebc8c>
- Kimber, B., Sandell, R., & Bremberg, S. (2008). Social and emotional training in Swedish classrooms for the promotion of mental health: results from an effectiveness study in Sweden. *Health Promotion International*, *23*(2), 134–143.
- Kindt, K., Kleinjan, M., Janssens, J. M. A. M., & Scholte, R. H. J. (2014). Evaluation of a school-based depression prevention program among adolescents from low-income areas: A randomized controlled effectiveness trial. *International Journal of Environmental Research and Public Health*, *11*(5), 5273–5293.
- Kumakech, E., Cantor-Graae, E., Maling, S., & Bajunirwe, F. (2009). Peer-group support intervention improves the psychosocial well-being of AIDS orphans: Cluster randomized trial. *Social Science & Medicine*, *68*(6), 1038–1043.
- Lai, E. S. Y., Kwok, C.-L., Wong, P. W. C., Fu, K.-W., Law, Y.-W., & Yip, P. S. F. (2016). The effectiveness and sustainability of a universal school-based programme for

- preventing depression in Chinese adolescents: A follow-up study using quasi-experimental design. *PloS One*, *11*(2), e0149854.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. <https://doi.org/10.3389/fpsyg.2013.00863>
- Lane, K. L., Weisenbach, J. L., Little, M. A., Phillips, A., & Wehby, J. (2006). Illustrations of function-based interventions implemented by general education teachers: Building capacity at the school site. *Education and Treatment of Children*, *549–571*.
- Leflot, G., van Lier, P. A. C., Onghena, P., & Colpin, H. (2010). The Role of Teacher Behavior Management in the Development of Disruptive Behaviors: An Intervention Study with the Good Behavior Game. *Journal of Abnormal Child Psychology*, *38*(6), 869–882. <https://doi.org/10.1007/s10802-010-9411-4>
- Lombas, A. S., Jiménez, T. I., Arguís-Rey, R., Hernández-Paniello, S., Valdivia-Salas, S., & Martín-Albo, J. (2019). Impact of the Happy Classrooms Programme on Psychological Well-being, School Aggression, and Classroom Climate. *Mindfulness*, 1–19.
- Lowry-Webster, H. M., Barrett, P. M., & Dadds, M. R. (2001). A universal prevention trial of anxiety and depressive symptomatology in childhood: Preliminary data from an Australian study. *Behaviour Change*, *18*(1), 36–50.
- Lynn, C. J., McKay, M. M., & Atkins, M. S. (2003). School Social Work: Meeting the Mental Health Needs of Students through Collaboration with Teachers. *Children & Schools*, *25*(4), 197–209. <https://doi.org/10.1093/cs/25.4.197>
- Malgady, R. G., Rogler, L. H., & Costantino, G. (1990). Hero/heroine modeling for Puerto Rican adolescents: a preventive mental health intervention. *Journal of Consulting and Clinical Psychology*, *58*(4), 469.
- Martin, A. J. (2008). Enhancing student motivation and engagement: The effects of a

- multidimensional intervention. *Contemporary Educational Psychology*, 33(2), 239–269.
- Masia-Warner, C., Nangle, D. W., & Hansen, D. J. (2006). Bringing evidence-based child mental health services to the schools: General issues and specific populations. *Education and Treatment of Children*, 29(2), 165–172.
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica: Biochemia Medica*, 22(3), 276–282.
- McInerney, M., & McKlindon, A. (2014). Unlocking the door to learning: Trauma-informed classrooms & transformational schools. *Education Law Center*, 1–24.
- Melnyk, B. M., Jacobson, D., Kelly, S., Belyea, M., Shaibi, G., Small, L., ... Marsiglia, F. F. (2013). Promoting healthy lifestyles in high school adolescents: A randomized controlled trial. *American Journal of Preventive Medicine*, 45(4), 407–415.
- Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., ... Swendsen, J. (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 & Adolescent Psychiatry*, 49(10), 980–989.
- Merry, S., McDowell, H., Wild, C. J., Bir, J., & Cunliffe, R. (2004). A randomized placebo-controlled trial of a school-based depression prevention program. *Journal of the American Academy of Child & Adolescent Psychiatry*, 43(5), 538–547.
<https://doi.org/10.1097/01.chi.0000117063.63530.68>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Mrazek, P. J., & Haggerty, R. J. (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington: National Academies Press.
- Neil, A. L., & Christensen, H. (2009). Efficacy and effectiveness of school-based prevention

- and early intervention programs for anxiety. *Clinical Psychology Review*, 29(3), 208–215. <https://doi.org/10.1016/j.cpr.2009.01.002>
- O’Leary-Barrett, M., Topper, L., Al-Khudhair, N., Pihl, R. O., Castellanos-Ryan, N., Mackie, C. J., & Conrod, P. J. (2013). Two-year impact of personality-targeted, teacher-delivered interventions on youth internalizing and externalizing problems: a cluster-randomized trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 52(9), 911–920.
- Olowokere, A. E., & Okanlawon, F. A. (2014). The effects of a school-based psychosocial intervention on resilience and health outcomes among vulnerable children. *The Journal of School Nursing*, 30(3), 206–215.
- Patton, G. C., Bond, L., Carlin, J. B., Thomas, L., Butler, H., Glover, S., ... Bowes, G. (2006). Promoting social inclusion in schools: a group-randomized trial of effects on student health risk behavior and well-being. *American Journal of Public Health*, 96(9), 1582–1587.
- Peris, T. S., Compton, S. N., Kendall, P. C., Birmaher, B., Sherrill, J., March, J., ... McCracken, J. T. (2015). Trajectories of change in youth anxiety during cognitive—behavior therapy. *Journal of Consulting and Clinical Psychology*, 83(2), 239.
- Perry, Y., Petrie, K., Buckley, H., Cavanagh, L., Clarke, D., Winslade, M., ... Christensen, H. (2014). Effects of a classroom-based educational resource on adolescent mental health literacy: A cluster randomised controlled trial. *Journal of Adolescence*, 37(7), 1143–1151.
- Pluess, M., & Boniwell, I. (2015). Sensory-processing sensitivity predicts treatment response to a school-based depression prevention program: Evidence of vantage sensitivity. *Personality and Individual Differences*, 82, 40–45.
- Polanczyk, G. V, Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual

- Research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry*, 56(3), 345–365.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1(3), 385–401.
- Rambaldo, L. R., Wilding, L. D., Goldman, M. L., McClure, J. M., & Friedberg, R. D. (2001). School-based interventions for anxious and depressed children. *Innovations in Clinical Practice: A Source Book*, 19, 347–358.
- Rash, J. A., Matsuba, M. K., & Prkachin, K. M. (2011). Gratitude and well-being: Who benefits the most from a gratitude intervention? *Applied Psychology: Health and Well-Being*, 3(3), 350–369.
- Reynolds, S., Wilson, C., Austin, J., & Hooper, L. (2012). Effects of psychotherapy for anxiety in children and adolescents: A meta-analytic review. *Clinical Psychology Review*, 32(4), 251–262.
- Rivet-Duval, E., Heriot, S., & Hunt, C. (2011). Preventing adolescent depression in Mauritius: A universal school-based program. *Child and Adolescent Mental Health*, 16(2), 86–91.
- Roberts, C. M., Kane, R., Bishop, B., Cross, D., Fenton, J., & Hart, B. (2010). The prevention of anxiety and depression in children from disadvantaged schools. *Behaviour Research and Therapy*, 48(1), 68–73.
- Roberts, C. M., Kane, R. T., Rooney, R. M., Pintabona, Y., Baughman, N., Hassan, S., ... Silburn, S. R. (2018). Efficacy of the Aussie Optimism Program: Promoting Pro-social Behavior and Preventing Suicidality in Primary School Students. A Randomised-Controlled Trial. *Frontiers in Psychology*, 8, 1392.
- Roeser, R. W., & Midgley, C. (1997). Teachers' Views of Issues Involving Students' Mental

- Health. *The Elementary School Journal*, 98(2), 115–133. <https://doi.org/10.1086/461887>
- Rohde, P., Stice, E., Shaw, H., & Gau, J. M. (2015). Effectiveness trial of an indicated cognitive–behavioral group adolescent depression prevention program versus bibliotherapy and brochure control at 1-and 2-year follow-up. *Journal of Consulting and Clinical Psychology*, 83(4), 736.
- Rolfsnes, E. S., & Idsoe, T. (2011). School-based intervention programs for PTSD symptoms: A review and meta-analysis. *Journal of Traumatic Stress*, 24(2), 155–165.
- Rones, M., & Hoagwood, K. (2000). School-based mental health services: A research review. *Clinical Child and Family Psychology Review*, 3(4), 223–241.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638.
- Rothì, D. M., Leavey, G., & Best, R. (2008). On the front-line: Teachers as active observers of pupils' mental health. *Teaching and Teacher Education*, 24(5), 1217–1231.
<https://doi.org/10.1016/J.TATE.2007.09.011>
- Ruttledge, R., Devitt, E., Greene, G., Mullany, M., Charles, E., Frehill, J., & Moriarty, M. (2016). A randomised controlled trial of the FRIENDS for Life emotional resilience programme delivered by teachers in Irish primary schools. *Educational & Child Psychology*, 33(2), 69–89.
- Sawyer, M. G., Pfeiffer, S., Spence, S. H., Bond, L., Graetz, B., Kay, D., ... Sheffield, J. (2010). School-based prevention of depression: a randomised controlled study of the beyondblue schools research initiative. *Journal of Child Psychology and Psychiatry*, 51(2), 199–209. <https://doi.org/10.1111/j.1469-7610.2009.02136.x>
- Schoenwald, S. K., Sheidow, A. J., & Letourneau, E. J. (2004). Toward effective quality assurance in evidence-based practice: Links between expert consultation, therapist fidelity, and child outcomes. *Journal of Clinical Child and Adolescent Psychology*,

33(1), 94–104.

Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., ... Kristjansson, E. (2017). AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *Bmj*, *358*, j4008.

Sheffield, J. K., Spence, S. H., Rapee, R. M., Kowalenko, N., Wignall, A., Davis, A., & McLoone, J. (2006). Evaluation of universal, indicated, and combined cognitive-behavioral approaches to the prevention of depression among adolescents. *Journal of Consulting and Clinical Psychology*, *74*(1), 66.

Shelemy, L., Harvey, K., & Waite, P. (2019). Supporting students' mental health in schools: what do teachers want and need? *Emotional and Behavioural Difficulties*, *24*(1), 100–116.

Shoshani, A., & Steinmetz, S. (201AD). Positive psychology at school: A school-based intervention to promote adolescents' mental health and well-being. *Journal of Happiness Studies*, *15*(6), 1289–1311.

Silbert, K. L., & Berry, G. L. (1991). Psychological effects of a suicide prevention unit on adolescents' levels of stress, anxiety and hopelessness: implications for counselling psychologists. *Counselling Psychology Quarterly*, *4*(1), 45–58.

Spence, S. H. (1998). A measure of anxiety symptoms among children. *Behaviour Research and Therapy*, *36*(5), 545–566.

Spence, S. H., Sawyer, M. G., Sheffield, J., Patton, G., Bond, L., Graetz, B., & Kay, D. (2014). Does the absence of a supportive family environment influence the outcome of a universal intervention for the prevention of depression? *International Journal of Environmental Research and Public Health*, *11*(5), 5113–5132.

Spence, S. H., Sheffield, J. K., & Donovan, C. L. (2003). Preventing adolescent depression:

- an evaluation of the problem solving for life program. *Journal of Consulting and Clinical Psychology*, 71(1), 3.
- Steele, W., & Malchiodi, C. A. (2012). *Trauma-informed practices with children and adolescents*. Routledge.
- Tanner-Smith, E. E., Durlak, J. A., & Marx, R. A. (2018). Empirically based mean effect size distributions for universal prevention programs targeting school-aged youth: A review of meta-analyses. *Prevention Science*, 19(8), 1091–1101.
- Teubert, D., & Pinquart, M. (2011). *A meta-analytic review on the prevention of symptoms of anxiety in children and adolescents*. *Journal of Anxiety Disorders* (Vol. 25).
<https://doi.org/10.1016/j.janxdis.2011.07.001>
- Van der Gucht, K., Griffith, J. W., Hellemans, R., Bockstaele, M., Pascal-Claes, F., & Raes, F. (2017). Acceptance and commitment therapy (ACT) for adolescents: outcomes of a large-sample, school-based, cluster-randomized controlled trial. *Mindfulness*, 8(2), 408–416.
- Verboom, C. E., Sijtsma, J. J., Verhulst, F. C., Penninx, B. W. J. H., & Ormel, J. (2014). Longitudinal associations between depressive problems, academic performance, and social functioning in adolescent boys and girls. *Developmental Psychology*, 50(1), 247.
- Wahl, M., Adelson, J., Patak, M., Pössel, P., & Hautzinger, M. (2014). Teachers or psychologists: who should facilitate depression prevention programs in schools? *International Journal of Environmental Research and Public Health*, 11(5), 5294–5316.
- Waite, P., & Creswell, C. (2014). Children and adolescents referred for treatment of anxiety disorders: Differences in clinical characteristics. *Journal of Affective Disorders*, 167, 326–332.
- Wang, R. A. H., Nelson-Coffey, S. K., Layous, K., Bao, K. J., Davis, O. S. P., & Haworth, C. M. A. (2017). Moderators of wellbeing interventions: Why do some people respond

more positively than others? *PloS One*, *12*(11), e0187601.

Weare, K., & Nind, M. (2011). Mental health promotion and problem prevention in schools:

what does the evidence say? *Health Promotion International*, *26*(suppl_1), i29–i69.

Werner-Seidler, A., Perry, Y., Cascar, A. L., Newby, J. M., & Christensen, H. (2017).

School-based depression and anxiety prevention programs for young people: A systematic review and meta-analysis. *Clinical Psychology Review*, *51*, 30–47.

<https://doi.org/10.1016/J.CPR.2016.10.005>

Williford, A., Boulton, A., Noland, B., Little, T. D., Kärnä, A., & Salmivalli, C. (2012).

Effects of the KiVa anti-bullying program on adolescents' depression, anxiety, and perception of peers. *Journal of Abnormal Child Psychology*, *40*(2), 289–300.

Wong, N., Kady, L., Mewton, L., Sunderland, M., & Andrews, G. (2014). Preventing anxiety

and depression in adolescents: a randomised controlled trial of two school based

Internet-delivered cognitive behavioural therapy programmes. *Internet Interventions*, *1*(2), 90–94.

World Health Organization. (2003). Caring for children and adolescents with mental

disorders: setting WHO directions.

Table 1: Characteristics of the 52 studies included in the systematic review and meta-analysis

Author and year	Location	Control	N	Baseline age mean	Outcome Measure	Intervention Name	Content	Delivery of Intervention	Training	Supervision	Fidelity	Quality Rating
Alampay et al., 2019	The Philippines	AC	186	11.9	Anx: STAI; Dep: MFQ	Kamalayan	Mindfulness, relaxation	10 90-minute weekly sessions	5 3-hour sessions, plus a 6 day intensive workshop	-	Sessions observed. 77% of participants attended 5 or more sessions.	1
Baker & Butler, 1984	USA	AC	36	-	Anx: STAI	Cognitive Self-Instruction	CB	8 45-minute lessons across 3 weeks	-	-	-	3
Barrett & Turner, 2001	Australia	NI	489	10.8	Anx: SCAS, RCMAS; Dep: CDI	FRIENDS	CB	10 75-minute weekly sessions plus optional parent session(s)	1 day	-	88-92% concordance with session and manual content	2
Baum et al., 2013	Israel	WL	563	11.1	Anx: SCARED; PTSD: UCLA-PTSD	Building Resilience Intervention	Resilience, relaxation, art	Variable – teacher has control over time spent on intervention	12 hours in 4 3-hour sessions	-	-	3
Berger & Gelkopf, 2009	Sri Lanka	WL	166	-	Dep: Brief BDI; PTSD: UCLA PTSD	ERASE Stress Sri Lanka	CB, relaxation, art	12 90-minute weekly sessions	3 days	2 3-hour sessions	Fidelity measured but not reported	2
Berger et al., 2012	Israel, Gaza	WL	154	12.8	Anx: SCARED; PTSD: UCLA PTSD	ERASE-Stress	CB, relaxation	16 90-minute weekly sessions plus optional parent session(s)	3 days	6 2-hour sessions	High fidelity and intervention quality reported	1
Bradley et al., 2010	USA	AC	136	15.3	Anx: TAI	TestEdge	-	2 lessons a week for one term	2 days	-	-	3
Britton et al., 2014	USA	AC	101	11.8	Anx: STAI-C; Dep and Anx: YSR	Mindfulness	Mindfulness, relaxation	Daily sessions for 3-12 minutes for 6 weeks	8-week mindfulness course	-	-	1

Author and year	Location	Control	N	Baseline age mean	Outcome Measure	Intervention Name	Content	Delivery of Intervention	Training	Supervision	Fidelity	Quality Rating
Buttigieg et al., 2015	Australia	NI	2027	12.3	Dep: CES-D	Resilient Families Intervention	Social skills	10 45-minute sessions held once a week	2 hours	-	Teacher-completed checklist indicated all components delivered	1
Challen et al., 2014	UK	NI	2844	11.5	Anx: RCMAS; Dep: CDI	UK Resilience Programme	CB, social skills	18 hours of intervention was delivered over 11-16 sessions weekly or fortnightly	10 days	Phone call support groups every 2-4 weeks	Facilitators were observed and rated using quality measures. High quality reported (9.85/12)	1
Clarke et al., 1993	USA	NI	622	15.4	Dep: CES-D	-	-	5 50-minute sessions	2 hours	-	-	3
Dowling et al., 2019	Ireland	AC	497	15.9	Anx: DASS Dep: DASS	MindOut	Social skills	13 weekly sessions	1 day	-	-	2
Eggert et al., 1995	USA	AC	105	16.0	Dep: CES-D	Personal Growth Class	Social skills	Daily 55-minute lessons for either 5 or 10 months	-	-	High fidelity ensured following daily monitoring	3
Fitzpatrick et al., 2013	Ireland	NI	1072	13.6	Emotional: SDQ	Working Things Out	-	One period per week over 8 months	1 day	-	-	2
Flynn et al., 2018	Ireland	NI	72	15.3	Emotional: Emotional Symptom Index	STEPS-A	Dialectical Behavioral skills	22 weekly sessions	-	-	-	2
Garaigordobil, 2004	Spain	NI	174	12.9	Anx: STAI-C	-	Social skills	2-hour session once a week across an academic year	-	-	-	3

Author and year	Location	Control	N	Baseline age mean	Outcome Measure	Intervention Name	Content	Delivery of Intervention	Training	Supervision	Fidelity	Quality Rating
Garcia et al., 2011	USA	AC	42	14.8	Dep: DASS	Project Wings Girls' Group	Social skills, relaxation	16 3-hour weekly sessions	30 hours	-	-	1
Gelkopf & Berger, 2008	Israel	WL	114	13.1	Dep: (Brief) BDI, PTSD: UCLA PTSD	ERASE-Stress	Relaxation	12 90-minute weekly lessons	7 3-hour sessions	3 2-hour sessions	High ratings of 5 to 6 out of 6 on all fidelity outcomes	1
Gillham et al., 2012	USA	NI	408	92% between 11-13	Anx: RCMAS; Dep: CDI, RADS-2	Penn Resilience Programme (PRP) for Adolescents	CB	10-12 90-minute sessions after school once a week plus 6 follow up booster sessions after 5 months	6 days (30 hours)	90-minute meetings every 2-3 weeks	68% of the intervention items were covered 'to some degree', and 47% 'satisfactorily'	1
Green et al., 2007	Australia	WL	56	16.1	Dep and Anx: DASS	Life Coaching Programme	CB	10 sessions over 28 weeks	2 ½ day workshops	-	-	3
Harnett & Dadds, 2004	Australia	NI	212	13.6	Anx: RCMAS; Dep: RADS	Resourceful Adolescent Programme	CB	11 45-minute sessions	1 day	11 90-minute supervision and planning meetings	A reduction in fidelity was reported throughout the intervention	2
Hiebert et al., 1989	Canada	AC	113	13-14	Anx: STAI	Progressive Relaxation	Relaxation	11 1-hour sessions 3 times a week	-	-	-	3
Karam et al., 2008	Lebanon	NI	194	11.8	Dep: DSM-III-R	-	CB	12 hour-long daily sessions over 12 consecutive school days	1 day	Supervised every 2-3 sessions	Rated using teacher diaries but not reported	2
Kimber et al., 2008	Sweden	NI	903	11-15	Anx: YSR	Social-Emotional	CB, social skills	A 45-minute session per week over the	Training provided	Monthly supervision	Moderate to high	2

Author and year	Location	Control	N	Baseline age mean	Outcome Measure	Intervention Name	Content	Delivery of Intervention	Training	Supervision	Fidelity	Quality Rating
						Training		school year			performance ratings	
Kindt et al., 2014	Netherlands	NI	1343	13.4	Dep: CDI	Op Volle Kracht (based on the PRP)	CB, social skills	16 weekly lessons were taught	4 days	-	95% and 65% of the first and last 8 lessons were taught respectively	1
Kumakech et al., 2009	Uganda	NI	298	11.8	Dep: BYI	Peer Group Support Intervention	Social skills	16 1-hour exercises over 10 weeks	Training provided	Weekly supervision	-	1
Lai et al., 2016	Hong Kong	NI	1359	15.1	Dep and Anx: DASS	The Little Prince is Depressed	CB	12 45-minute sessions	3 hours	On-going consultation provided	-	3
Lombas et al., 2019	Spain	NI	524	13.6	Dep: SDS (control outcomes not reported)	Happy Classrooms	Mindfulness, relaxation	2 5-minute activities a week for 18 weeks	16 hours	Two 2-hour sessions and online email contact	Teachers reported fidelity with high variation found.	2
Lowry-Webster et al., 2001	Australia	WL	594	15.2	Anx: SCAS, RCMAS; Dep: CDI	FRIENDS	CB	10 weekly 1-hour classes plus 2 booster sessions at 1 and 3 months after the intervention and optional parent session(s)	2 days	Regular meetings with program leader	Videotaped sessions. No significant problems found.	2
Malgady et al., 1990	USA	AC	90	13.7	Anx: STAI	Hero/Heroine Intervention	Role-modelling	18 90-minute weekly sessions	-	-	-	2
Martin, 2008	Australia	NI	53	15.0	Anx: MES-HS	Motivation and Engagement	CB, motivation	13 weekly modules, with each taking 20-30 minutes to complete	-	-	-	1

Author and year	Location	Control	N	Baseline age mean	Outcome Measure	Intervention Name	Content	Delivery of Intervention	Training	Supervision	Fidelity	Quality Rating
Melnyk et al., 2013	USA	AC	779	14.7	Dep and Anx: HLBS	COPE Healthy Lifestyles TEEN	CB, physical activity	15 20-minute weekly sessions	1 day	-	Decreased fidelity was reported in half of the observed sessions	2
Merry et al., 2004	New Zealand	AC	392	14.2	Dep: BDI-II, RADS-2	Resourceful Adolescent Programme – Kiwi	CB, social skills	11 weekly/twice-weekly sessions	2.5 days	-	Weekly integrity checklist completed but not reported	1
O’Leary-Barrett et al., 2013	UK	WL	1024	13.7	Dep and Anx: BSI	-	CB	2 90-minute sessions	3 days	4 hours of supervision with trainer	Fidelity measured but not reported	2
Olowokere & Okanlawon, 2014	Nigeria	NI	109	12.4	Anx: SCAS, Dep: CES-D	-	Resilience	6 2-hour weekly sessions	-	-	-	3
Patton et al., 2006	Australia	-	7594	13-14	Dep: MFQ	The Gatehouse Project	Social skills	20 45-minute lessons were delivered across 10 weeks	-	-	-	2
Perry et al., 2014	Australia	AC	380	14.8	Dep and Anx (combined outcome): DASS-21	HeadStrong	Psycho-education	10 hours of class time over 5-8 weeks	1 day	-	-	2
Pluess & Boniwell, 2015	UK	NI	363	11.4	Dep: CES-D	SPARK Resilience Programme	CB	12 1-hour sessions across 3-4 months	-	-	-	2
Rivet-Duval et al., 2011	Mauritius	WL	160	13.7	Dep: RADS-2	Resourceful Adolescent Programme – Adolescent	CB, social skills	11 1-hour weekly sessions	16 hours over 2 days	½ day support 6 months into intervention	-	2

Author and year	Location	Control	N	Baseline age mean	Outcome Measure	Intervention Name	Content	Delivery of Intervention	Training	Supervision	Fidelity	Quality Rating
Roberts et al., 2010	Australia	AC	496	12.0	Anx: RCMAS; Dep: CDI	Aussie Optimism Programme	CB, social skills	10 1-hour lessons that were delivered at times that suited the teachers	16 hours	8 hour-long sessions	Fidelity was measured from teacher logbooks and was reported as “good”	2
Roberts et al., 2018	Australia	NI	1471	11.1	Dep and Anx: DICA-IV (diagnostic tool)	Aussie Optimism Programme	CB, social skills	20 1-hour weekly sessions	8 hours	5 optional hours of coaching	Teachers completed logbooks and average 17 sessions were delivered in full	1
Rohde et al., 2015	USA	AC	378	15.5	Dep: K-SADS	Cognitive-behavioral Skills	CB	6 1-hour weekly sessions	-	Supervision provided	Fidelity was concluded as being “good”	1
Ruttledge et al., 2016	Ireland	WL	709	10.9	Anx: SCAS	FRIENDS	CB	10 weekly sessions plus optional parent session(s)	2 days	Unspecified number of meetings with supervisor	All key components were delivered	1
Sawyer et al., 2010	Australia	AC	5634	13.1	Dep: CES-D	BeyondBlue	Social skills, resilience	10 45-minute sessions across one school term for three years (30 sessions in total)	1 day	-	Teachers reported that they covered 70% of activities	1
Sheffield et al., 2006	Australia	NI	2479	14.3	Anx: SCAS; Dep: CDI, CES-D	-	CB	8 45-minute lessons weekly for a school term	6 hours	Infrequent use of optional supervision	-	1
Shoshani & Steinmetz, 2014	Israel	WL	1038	13.7	Dep and Anx: BSI	Maytiv	Positive psychology	15 fortnightly sessions	15 2-hour fortnightly	-	Full administration	1

Author and year	Location	Control	N	Baseline age mean	Outcome Measure	Intervention Name	Content	Delivery of Intervention	Training	Supervision	Fidelity	Quality Rating
Silbert & Berry, 1991	USA	WL	323	14-18	Anx: STAI	Suicide Prevention Unit	Psycho-education	2 50-minute class sessions	-	-	-	3
Spence et al., 2003	Australia	WL	1500	12.9	Dep: BDI	Problem Solving For Life	CB	8 45-minute weekly sessions	6 hours	No support provided	Full fidelity reported for 5 sessions, with half not completing 3 sessions.	2
Van der Gucht et al., 2017	Belgium	NI	616	17	Anx: YSR	Acceptance and Commitment Therapy (ACT)	ACT	4 2-hour weekly sessions	2 days	2 2-hour sessions	-	2
Wahl et al., 2014	Germany	NI	646	14.0	Dep: CES-D	LARS & LISA	CB	10 90-minute sessions delivered weekly	2 days	1.5 hour supervision was held biweekly	A high level of intervention fidelity and acceptance was found	2
Williford et al., 2012	Finland	-	7741	11.2	Anx: FNE and SAD (combined), Dep: Brief BDI	KiVa	Anti-bullying	20 hours of lessons were delivered	-	-	-	3
Wong et al., 2014	Australia	NI	265	15	Anx: GAD-7, Dep: PHQ-9	Thiswayup Schools	CB	6 or 7 40-minute weekly sessions	No training	The research team could be contacted for advice via phone	-	3

Author and year	Location	Control	N	Baseline age mean	Outcome Measure	Intervention Name	Content	Delivery of Intervention	Training	Supervision	Fidelity	Quality Rating
-----------------	----------	---------	---	----------------------	--------------------	----------------------	---------	-----------------------------	----------	-------------	----------	-------------------

and email

Note. Control group AC: Active Control; NI: No Intervention; WL: Waiting List. **Outcome measures** Anx: Anxiety; Dep: Depression; PTSD: Post-Traumatic Stress Disorder. **Quality Rating** 1 indicates highest quality level, 3 indicates lowest quality level. **Content** CB: Cognitive-behavioral skills; ACT: Acceptance and Commitment Therapy.

Table 2: Summary of effect sizes, confidence intervals and p values on depression and anxiety outcomes for intervention factors

Intervention Group	Depression Outcomes		Anxiety Outcomes	
	N of studies	Effect size (Hedges' <i>g</i>), 95% CI and p-value	N of studies	Effect size (Hedges' <i>g</i>), 95% CI and p-value
Content				
Cognitive-behavioral skills	20	-0.10 (-0.17 to -0.02; <i>p</i> = .015)	17	-0.08 (-0.18 to 0.02; <i>p</i> = .134)
Social skills	11	-0.06 (-0.15 to 0.03; <i>p</i> = .220)	6	-0.05 (-0.22 to 0.10; <i>p</i> = .485)
Relaxation/mindfulness	3	-0.32 (-0.60 to -0.04; <i>p</i> = .026)	3	-0.17 (-0.41 to 0.07; <i>p</i> = .180)
Session length				
<45 minutes	1	0.09 (-0.06 to 0.23; <i>p</i> = .235)	3	-0.10 (-0.41 to 0.20; <i>p</i> = .502)
45-90 minutes	22	-0.12 (-0.20 to -0.04; <i>p</i> = .003)	16	-0.12 (-0.24 to -0.01; <i>p</i> = .029)
90+ minutes	2	-0.18 (-0.79 to 0.43; <i>p</i> = .562)	3	-0.18 (-0.45 to 0.10; <i>p</i> = .206)
Number of sessions				
<8 sessions	2	-0.09 (-0.20 to 0.01; <i>p</i> = .090)	3	-0.16 (-0.27 to -0.06; <i>p</i> = .002)
8-16 sessions	26	-0.13 (-0.22 to -0.04; <i>p</i> = .002)	20	-0.13 (-0.24 to -0.02; <i>p</i> = .020)
16+ sessions	2	0.05 (-0.01 to 0.10; <i>p</i> = 0.08)	3	-0.10 (-0.55 to 0.34; <i>p</i> = .655)
Length of intervention				
0-16 weeks	20	-0.13 (-0.22 to -0.04; <i>p</i> = .005)	17	-0.17 (-0.28 to -0.06; <i>p</i> = .002)
16+ weeks	4	-0.10 (-0.35 to 0.14; <i>p</i> = .422)	4	-0.01 (-0.38 to 0.35; <i>p</i> = .941)
Training length				
1 day or less	9	-0.07 (-0.18 to 0.07; <i>p</i> = .285)	6	-0.07 (-0.19 to 0.07; <i>p</i> = .410)
2 days or more	14	-0.12 (-0.20 to -0.04; <i>p</i> = .004)	13	-0.14 (-0.26 to -0.01; <i>p</i> = .039)
Supervision				
Regular and in-	9	-0.18 (-0.30 to -0.06; <i>p</i> = .004)	9	-0.15 (-0.33 to 0.03; <i>p</i> = .106)

person				
Reported Fidelity				
High/Good	8	-0.14 (-0.25 to -0.02; $p = .021$)	6	-0.13 (-0.34 to 0.08; $p = .231$)
Lower fidelity	6	-0.11 (-0.27 to 0.05; $p = .184$)	5	-0.05 (-0.25 to 0.15; $p = .624$)