Published in final edited form as:

J Clin Child Adolesc Psychol. 2016 ; 45(6): 763–783. doi:10.1080/15374416.2015.1020541.

Programs for the Prevention of Youth Depression: Evaluation of Efficacy, Effectiveness, and Readiness for Dissemination

Steven M. Brunwasser and Judy Garber

Vanderbilt University

Abstract

Objective—To evaluate the current state of evidence of the effectiveness of depression prevention programs for youth, assess the degree to which current evidence supports broad implementation, and outline additional steps needed to close the gap between effectiveness and dissemination.

Method—We used the Society for Prevention Research’s Standards of Evidence (Flay et al., 2005) to evaluate the degree to which existing depression prevention programs have established intervention efficacy, effectiveness, and readiness for dissemination. We reviewed all depression prevention programs for youth that have been evaluated in at least two published, randomized controlled trials in which the intervention was compared to a no-intervention control group. A total of 37 studies evaluating 11 different programs were reviewed with regard to depressive symptoms and diagnoses post intervention and at follow-up (at least 6 months).

Results—Eight programs demonstrated significant main effects on depressive symptoms relative to controls in multiple RCTs; five programs had at least one trial with significant main effects present at least one year post-intervention. Two programs demonstrated efficacy for both depressive symptoms and depressive episodes across multiple independent trials. Regarding effectiveness, six programs had at least one study showing significant effects when delivered by endogenous service providers; four programs had significant effects in studies conducted independently of the program developers.

Conclusions—Several programs have demonstrated promise in terms of efficacy, but no depression prevention program for children or adolescents as yet has garnered sufficient evidence of effectiveness under real-world conditions to warrant widespread dissemination at this time.

Keywords

Depression; Prevention; Youth; Dissemination

Over the last decade, substantial progress has been made regarding research on the prevention of depression in children and adolescents. Indeed, recent meta-analyses have concluded that some depression prevention programs, particularly those conducted with

Corresponding Author: Judy Garber, Ph.D., Department of Psychology & Human Development, Vanderbilt University, 0552 Peabody, 230 Appleton Place, Nashville, TN 37203-5721, 615-343-8714 (office), 615-343-9494 (FAX), jgarber.vanderbilt@gmail.com.

SUPPLEMENTAL MATERIAL

Supplemental material for this manuscript can be accessed on the publisher’s website.

targeted samples, consistently yield small but significant effects in preventing increases in depressive symptoms or onsets of depressive episodes (e.g., Merry et al., 2011; Sandler et al., 2014). Nevertheless, researchers, clinicians, and policy makers continue to ask whether these efficacious programs are ready for more wide-spread implementation and dissemination (Merry, 2013; Spence, & Shortt, 2007).

The purpose of this review was to evaluate the current state of evidence of the effectiveness of depression prevention programs for youth, assess the degree to which current evidence supports broad implementation, and suggest additional steps needed to close the gap between effectiveness and dissemination. We used the Society for Prevention Research's Standards of Evidence (SPR-SOE; Flay et al., 2005) as the primary guide for evaluating the degree to which existing depression prevention programs have established intervention efficacy, effectiveness, and readiness for dissemination. We specifically concentrated on criteria from the guidelines that are most germane to the prevention of depression. Given our focus on the transition from effectiveness to dissemination, we emphasized the extent to which studies have met criteria related to external and ecological validity (e.g., replication by independent research teams) more than internal validity (e.g., degree to which the measures used were internally consistent). Clearly, internal validity is a prerequisite to establishing effectiveness, but a thorough review of these criteria is beyond the scope of this article. For a review of the degree to which universal, school-based depression prevention programs meet the SPR-SOE standards for internal validity see Spence and Shortt (2007).

When and if dissemination entities (e.g., schools, HMOs, pediatric clinics) decide to broadly implement depression prevention, they are likely to choose an established program with a coherent structure and theory. Therefore, we organized the evidence for depression prevention by type of program. Differences in program content, however, are just one of many potential sources of between-study heterogeneity in effects. The fact that some programs have performed better in controlled trials than others may be as much a reflection of the different contexts in which the programs were delivered as their different contents.

Method

We reviewed evidence of efficacy and effectiveness for all youth depression prevention programs that have been evaluated in multiple, published, randomized controlled trials (RCTs) in which the intervention condition was compared to some form of a no-intervention control condition such as assessment-only, waitlist, or usual care. To be included, studies had to explicitly describe the intervention as a program that specifically targeted the prevention of depression (either symptom severity or diagnoses) as the primary outcome. Studies were *excluded* if they (1) used non-random allocation procedures (e.g., Jaycox, Reivich, Gillham, & Seligman, 1994; Shochet et al., 2001), (2) lacked a "no intervention" or "usual care" control condition (e.g., Muris, Bogie, & Hoogsteder, 2001), or only compared the primary intervention to alternative interventions or placebo control conditions that were not naturally occurring non-interventions (e.g., Beardslee, Wright, Salt, & Drezner, 1997; Merry, McDowell, Wild, Bir, & Cunliffe, 2004; Van Voorhees et al., 2008), (3) did not measure depression outcomes explicitly (e.g., Peng et al., 2014), (4) were not published in a peer-reviewed, English-language journal (e.g., Gillham, 1995), (5) included primarily adult

participants (i.e., over 18 years old), (6) described the intervention as a treatment program or evaluated its effects among a sample likely meeting criteria for a current depressive episode at baseline, or (7) evaluated an intervention in a sample with past depressive episodes that had not fully remitted (i.e., relapse prevention). Finally, given that an important requirement for dissemination is replication, we excluded programs that had fewer than two RCTs meeting the above inclusion criteria (e.g., Compas et al., 2009). Nevertheless, we refer to studies of some excluded programs that demonstrate important points about effectiveness and dissemination.

Studies comparing interventions to alternative or placebo control groups provide an important test of whether the theory-driven program is superior to an alternative that is lacking some or all of the specific intervention components. The current review, however, excluded studies that compared theory-driven interventions *only* to alternative or placebo control interventions, because a natural control condition is needed to determine whether either or both interventions were better than receiving nothing at all. For example, a theory-driven intervention might appear to be efficacious by virtue of outperforming an alternative intervention, but we cannot know without also having a no-intervention control condition whether the group differences were due to a beneficial effect of the theory-driven intervention, an iatrogenic effect of the alternative intervention, or both. Given that our focus is first on the program's overall effectiveness rather than its comparative effectiveness, we excluded trials lacking some kind of no-intervention control and focused on comparisons with natural control conditions.

Articles were identified via a search of the PsycInfo database. The following search terms were used: (*Depression* OR "*Major Depression*" OR *dysthymia* OR *depress**) AND ("*Primary Mental Health Prevention*" OR "*Prevention*" OR "*Relapse Prevention*"). We limited the search to English-language, peer-reviewed articles with human participants between the ages of 2 and 19, inclusive. The search identified 614 articles.

In total, 11 different intervention programs met the above criteria and were included in this review: *Penn Resiliency Program* (PRP; Gillham, Jaycox, Reivich, Seligman, & Silver, 1990), *Coping with Stress Course* (CwSC; Clarke et al., 1995), *FRIENDS for Life* (FRIENDS; Barrett, 2005), *Aussie Optimism Program* (Roberts, 2006), *Blues Group* (Stice, Burton, Bearman, & Rohde, 2007), *Feeling Good* Bibliotherapy (Burns, 1989; Stice et al., 2007), *Supportive-Expressive Group Intervention* (Stice et al., 2007), Interpersonal Psychotherapy-Adolescent Skills Training (IPT-AST; Young, Mufson, & Davies, 2006), *Resourceful Adolescent Program* (RAP; Shochet, Holland, & Whitefield, 1997), *LARS & LISA* (Pössel, Horn, Groen, & Hautzinger, 2004), and *Problem-Solving for Life* (PSFL; Spence, Sheffield, & Donovan, 2003). A total of 37 studies evaluated these 11 programs.

Evaluating Evidence for Individual Programs

In accordance with the SPR-SOE guidelines specificity statement (Flay et al., 2005), we evaluated efficacy and effectiveness separately with regard to effects on continuous measures of symptom severity and on diagnostic outcomes. That is, a program could show high levels of efficacy in targeting depressive symptoms but have no evidence of efficacy in reducing depressive episodes, or the reverse.

The SPR-SOE criteria include the distinction between efficacy and effectiveness. We recognize, however, that this difference is somewhat arbitrary, and that ultimately it may make more sense to characterize interventions along a continuum with respect to the different aspects of clinical trials (e.g., endogenous service providers; real-world settings). Nevertheless, for the purpose of the current review of youth depression prevention programs and evaluation of their readiness for dissemination, we adopted the SPR-SOE criteria that clearly differentiate efficacy and effectiveness trials.

Evidence of efficacy—We evaluated programs with regard to the following criteria: (1) Was the program found to be superior to a no-intervention control condition in at least two separate RCTs? (2) If there were more than two trials, does the preponderance of evidence indicate that the program was superior to a no-intervention control condition? (3) Is there evidence that effects are present at a follow-up assessment of 6 months or longer?

Several programs have been tested in at least three independent RCTs that evaluated effects on depressive symptoms. Providing an overall appraisal of program effects across multiple trials is challenging due to considerable between-study differences in contexts, methodology, and program effects. Therefore, to aid in the overall appraisal of program effects, we calculated standardized mean difference scores (Hedges' g) for depressive symptom measures for all studies at post-intervention and at the first follow-up assessment of six months or longer (Hedges, 1981). When summary statistics (i.e., N s, means, SD s) were provided only for subgroups within the sample, we pooled the summary statistics across subgroups and then used the pooled data to calculate a single effect size (ES) for that specific study and assessment. When multiple measures of depressive symptoms were used in a single study, we calculated a pooled estimate of g so that studies provided only one ES per time point. Calculating a standard error for g estimates obtained by pooling across measures requires that the correlation between the two measures is known. When the correlation was not explicitly reported, we estimated it based on reports of correlations between the two instruments in the literature (Borenstein, Hedges, Higgins, & Rothstein, 2009).

Weighted average ESs (g_+) were calculated at both post and follow-up. Whereas g represented an individual study ES estimate, g_+ was an average estimate across multiple studies of the same program, with each study weighted by its inverse variance. We conducted separate fixed-effect meta-analyses for each program in the Metafor package (Viechtbauer, 2010) in R version 3.1.0 (R Core Team, 2014). Fixed effect meta-analysis makes the restrictive assumption that each study provides an estimate of a single, true ES and that deviations of individual study ESs from the true ES are attributable to random sampling variability. In other words, studies are merely replications with unimportant differences in methodology and implementation. In contrast, random effects meta-analysis assumes that studies differ in meaningful ways, resulting in a distribution of true ESs rather than a true ES common to all studies (Borenstein, Hedges, & Rothstein, 2007). The assumptions of the random effects meta-analysis approach are clearly more realistic in the context of youth depression prevention studies. However, given that no program provided more than 13 ES estimates at a given time point (and most provided 3 or fewer ESs), estimates of between-study variability (the central feature of random effects analyses) could

not be made with adequate precision to justify random-effects analyses. The consequence of using fixed-effect meta-analysis is that the g_+ estimates for each program are relevant only to studies with characteristics similar to those included in the meta-analysis; generalizations should not be made to the hypothetical population of heterogeneous studies evaluating the program of interest (Borenstein et al., 2007).

Cochrane's Q test was used to assess whether the degree of heterogeneity in ESs across studies exceeded chance expectation (Cochran, 1954). One program (PRP) provided a sufficient number of ESs ($k_{\text{post}} = 13$; $k_{\text{follow}} = 12$) to allow for meaningful subgroup analyses based on effectiveness criteria. We evaluated whether the g_+ estimates for this program were significant in the subgroup of studies in which intervention groups were led by endogenous group leaders (e.g., school teachers) and in subgroups of studies conducted independently of the program developers. Such analyses were not practical for the other programs, however, due to the small number of studies conducted.

Evidence of effectiveness—The following criteria were used to evaluate evidence of effectiveness: (1) Was the intervention delivered in a real-world setting where there was a high level of access to the population of interest? (2) Was the intervention facilitated by endogenous providers; that is, group leaders who are naturally present, and/or widely available, in the intervention setting (e.g., school teachers)? (3) Were there intervention effects across diverse samples and intervention sites? (4) Have intervention effects been replicated by an independent research team with limited involvement of the program developers in the implementation and evaluation of the program? If the program evaluators or their research team provided training or consultation in the delivery of the program but were not involved in the actual intervention implementation and evaluation, then the study was considered to be an independent evaluation.

Results

Evidence for Efficacy and Effectiveness of Individual Prevention Programs

Detailed characteristics of the intervention programs reviewed here are presented in Tables 1 and 2. Table 1 provides information about individual studies grouped within intervention programs. Table 2 summarizes intervention effects across studies for each program. Supplemental materials also are provided that include a Table that presents further details about the eleven youth depression prevention programs, and the forest plots of the program effects at post and follow-up for each of the eleven programs reviewed here.

Penn Resiliency Program (PRP)

PRP is a cognitive-behavioral (CB) group intervention for youth in late childhood and early adolescence that has been evaluated in 13 published RCTs with diverse samples of youth from four countries. The program has been implemented with universal (e.g., Pattison & Lynd-Stevenson, 2001), selective (e.g., Cardemil, Reivich, & Seligman, 2002), and indicated (e.g., Wijnhoven, Creemers, Vermulst, Scholte, & Engels, 2014) samples. Although typically a school-based intervention, PRP also has been delivered in primary care clinics (Gillham, Hamilton, Freres, Patton, & Gallop, 2006).

PRP Efficacy—Significant intervention effects on depressive symptoms have been reported for a subgroup of the study samples across eight trials. Main effects on symptoms have been found as long as 24 months post intervention (Cardemil, Reivich, Beevers, Seligman, & James, 2007) and subgroup effects have been found at 30 months post-intervention (Gillham et al., 2007). The mean ES across all studies was small but significant at both post ($g_+ = -0.08$, $k = 13$, 95% CI $[-0.15, -0.01]$) and follow-up ($g_+ = -0.19$, $k = 12$, 95% CI $[-0.27, -0.11]$).

The effects of PRP on depressive symptoms have been inconsistent both across and within trials. There is considerable between-study variability in effects sizes ($Q_{\text{Post}} = 21.25$, $p < .05$ and $Q_{\text{Follow}} = 26.24$, $p < .01$). Two well-powered studies reported no effects on depressive symptoms (Kindt, Kleinjan, Janssens, & Scholte, 2014; Roberts, Kane, Thomson, Bishop, & Hart, 2003) and others found beneficial effects in a subgroup of participants, but not in the overall sample (Gillham et al., 2007, 2012, 2006). Two studies found that the control condition reported either *greater* symptom reductions (Cardemil et al., 2002) or reduced risk for clinically relevant symptoms (Kindt et al., 2014) compared to PRP, thus suggesting iatrogenic effects. Few studies of PRP have evaluated effects on depression diagnoses, with little evidence of benefit in aggregate (Brunwasser, Gillham, & Kim, 2009); one study did find that PRP reduced risk for combined depression and anxiety diagnoses among participants with elevated baseline depressive symptoms (Gillham et al., 2006).

PRP Effectiveness—Six studies have evaluated PRP delivered entirely or predominantly by endogenous providers. Two of these studies reported significant effects in the overall sample (Chaplin et al., 2006; Yu & Seligman, 2002) and two other studies found significant effects in a subgroup of participants (Gillham et al., 2007, 2012). The mean ES across studies using endogenous providers was significant at follow-up ($g_+ = -0.15$, $k = 6$, 95% CI $[-0.23, -0.07]$) but not at the immediate post-intervention evaluation ($g_+ = -0.06$, $k = 7$, 95% CI $[-0.13, 0.02]$). PRP's effect was robust at both post ($g_+ = -0.24$, $k = 6$, 95% CI $[-0.44, -0.04]$) and follow-up ($g_+ = -0.45$, $k = 6$, 95% CI $[-0.66, -0.23]$) when delivered by research team members.

Five RCTs have been conducted with minimal involvement from the program developers. The mean ES among these studies was significant at follow-up ($g_+ = -0.19$, $k = 5$, 95% CI $[-0.30, -0.08]$), but not at post ($g_+ = -0.02$, $k = 5$, 95% CI $[-0.12, 0.08]$). Among studies conducted by the program developers or their research team, the mean ES was significant at both post $g_+ = -0.14$, $k = 8$, 95% CI $[-0.24, -0.04]$ and follow-up, $g_+ = -0.19$, $k = 7$, 95% CI $[-0.30, -0.08]$. Finally, two studies conducted by independent research groups that evaluated program effects with endogenous providers did not find significant benefit on depressive symptoms (Kindt et al., 2014; Roberts et al., 2003).

Coping with Stress Course (CwSC)

CwSC is a CB based group intervention for adolescents ages 13 to 18 (Clarke et al., 1995). The program has been evaluated in four RCTs in schools, health maintenance organizations (HMOs), and University clinics using universal, indicated, and selected + indicated samples. Two trials were conducted by investigators who were not the program developers.

CwSC Efficacy—All four RCTs found significant effects on at least one measure of depressive symptoms. Main effects on symptoms have endured for as long as 12 months post-intervention (Clarke et al., 2001) and subgroup effects have endured through 33 months post-randomization (Beardslee et al., 2013). The mean ES across trials was significant at both post ($g_+ = -0.33$, $k = 4$, 95% CI $[-0.47, -0.20]$) and follow-up ($g_+ = -0.18$, $k = 4$, 95% CI $[-0.32, -0.04]$). There is relatively little evidence of between-study heterogeneity in symptom-based ESs ($Q_{\text{Post}} = 0.47$, *ns* and $Q_{\text{Follow}} = 4.14$, *ns*), although the strength of the effects has varied within studies across initial symptom levels (Horowitz, Garber, Ciesla, Young, & Mufson, 2007) and current parent depression at baseline (Garber et al., 2009). All three studies evaluating diagnostic outcomes found reduced risk of depressive episodes in the overall sample: odds ratios (ORs) ranged from 0.49 to 0.77, lasting as long as 33 months (Beardslee et al., 2013).

CwSC Effectiveness—Two RCTs have evaluated the effectiveness of CwSC when delivered by endogenous providers (Clarke et al., 1995, 2001), with significant benefits of CwSC on depression symptoms and episodes reported in both studies. One trial conducted independently of the program developers found significant effects on two depressive symptom measures immediately post-intervention (Horowitz et al., 2007). No RCTs testing CwSC have been conducted by completely independent research groups using endogenous providers. In the multi-site trial (Garber et al., 2009), however, three of the four sites did not include program developers as either the group leaders or supervisors.

Aussie Optimism Program (AOP)

AOP is a CB-based intervention with two primary components focused on optimistic thinking skills (i.e., cognitive restructuring) and social life skills (e.g., assertiveness) (Roberts, 2006). One study tested AOP among youths ages 11-13 (Roberts et al., 2010), and two studies evaluated a version of the program (Aussie Optimism Program—Positive Thinking Skills; AOP-PTS) adapted to be developmentally appropriate for children ages 9-10 (Rooney et al., 2006; Rooney, Hassan, Kane, Roberts, & Nesa, 2013). In all three RCTs, the AOP programs were implemented as a whole-classroom, universal intervention in low-income Australian primary schools. In all cases, cluster randomization was employed, with schools randomized to either AOP or control after matching on key characteristics.

AOP Efficacy—In both studies evaluating AOP-PTS with children ages 9-10, there were significant benefits of the program relative to usual care control groups at post; however, there was no evidence of enduring effects on depressive symptoms (Rooney et al., 2006, 2013). There was no evidence of an effect of AOP on depressive symptoms in the trial with 11-13 year old children (Roberts et al., 2010). The mean ES across the three trials was not significant at either post ($g_+ = -0.09$, $k = 3$, 95% CI $[-0.19, 0.01]$) or follow-up ($g_+ = -0.03$, $k = 3$, 95% CI $[-0.13, 0.08]$). There was significant between-study heterogeneity at post ($Q = 12.49$, $p < .01$) but not at follow-up ($Q = 3.88$, *ns*). When evaluating just the two studies of AOP-PTS, the mean ES was significant at post ($g_+ = -0.19$, $k = 2$, 95% CI $[-0.31, -0.06]$) but not follow-up ($g_+ = -0.09$, $k = 2$, 95% CI $[-0.21, 0.04]$). In one of the two studies evaluating diagnostic outcomes, there was a significant reduction in the incidence of

depressive episodes through 9 months (Rooney et al., 2006), but there was no effect on diagnostic outcomes in a larger study (Rooney et al., 2013).

AOP Effectiveness—Two studies have evaluated AOP interventions as delivered by trained school teachers. One study found significant improvement in depressive symptoms relative to controls from pre- to post-intervention but no sustained effect (Rooney et al., 2013), and the second study found no effect on depressive symptoms (Roberts et al., 2010). No trials have been conducted independently of the program developers.

Blues Group

Three RCTs have evaluated the *Blues Group*, a CB-based group intervention, compared to both no-intervention control conditions and several alternative interventions that contain some but not all of the components believed to contribute to the *Blues Group*'s effects on depression outcomes (Rohde, Stice, Shaw, & Brière, 2014; Stice, Rohde, Seeley, & Gau, 2008; Stice et al., 2007). In all three studies, the *Blues Group* was compared to a CB bibliotherapy condition (Burns, 1989), which contained CB training but no group environment. In two of the three studies, the *Blues Group* and CB Bibliotherapy were compared to the Supportive-Expressive Group Intervention, which contained a supportive group environment but no CB training. We grouped these three interventions together below because they were all compared to the same no-intervention control conditions within the same RCTs.

Blues Group Efficacy—Significant effects on depressive symptoms have been reported in all three RCTs, with effects enduring as long as 12 months post-intervention (Stice et al., 2008). The mean ES across trials was significant at post ($g_+ = -0.45$, $k = 3$, 95% CI $[-0.63, -0.28]$) and at follow-up ($g_+ = -0.21$, $k = 3$, 95% CI $[-0.38, -0.03]$). There was little evidence of between-study heterogeneity at either post ($Q = 4.77$, *ns*) or 12-month follow-up ($Q = 4.12$, *ns*). Both studies examining the impact on depression diagnoses found significant benefits of the intervention relative to control, with reduced incidence of depressive episodes across 6-month (OR = 0.12; Rohde et al., 2014) and 24-month (OR = 0.53; Stice et al., 2010) follow-ups.

Blues Group Effectiveness—One study evaluated effects when school personnel led the intervention groups and found significant reductions on depressive symptoms immediately post-intervention ($d = 0.29$) and in the incidence of depressive episodes (OR = 0.12) through six-month follow-up compared to controls (Rohde et al., 2014). The *Blues Group* has not been evaluated by an independent research group.

Bibliotherapy (Feeling Good) Efficacy

Significant intervention effects on depressive symptoms compared to the no-intervention control were reported at post and through 6 months in two of three trials. The mean ES across the three trials was significant at both post ($g_+ = -0.18$, $k = 3$, 95% CI $[-0.36, -0.002]$) and follow-up ($g_+ = -0.25$, $k = 3$, 95% CI $[-0.43, -0.07]$), with non-significant between-study heterogeneity, respectively: ($Q_{\text{Post}} = 4.77$, *ns*; $Q_{\text{Follow}} = 4.12$, *ns*). One of the two trials evaluating diagnostic outcomes (Stice, Rohde, Gau, & Wade, 2010; Stice et al.,

2008) found that youths in the *Bibliotherapy* condition were at significantly reduced risk for depressive episodes relative to controls (OR = 0.17).

Bibliotherapy (Feeling Good) Effectiveness

The degree of provider endogeneity was less relevant for the bibliotherapy conditions in these studies given that youth completed the intervention with minimal guidance and support. There have been no evaluations of *Feeling Good* as a youth depression prevention program by other research groups, but *Feeling Good* has demonstrated efficacy in youth treatment trials (e.g., Ackerson, Scogin, McKendree-Smith, & Lyman, 1998). Of note, the research group conducting these prevention trials of bibliotherapy did not develop the *Feeling Good* materials, so these trials could be considered independent implementations.

Supportive-Expressive Group Intervention (SEGI)

SEGI Efficacy—The SEGI program was based on a non-directive treatment used as a comparison condition in an adolescent treatment outcome trial (Brent et al., 1997). Youth receiving SEGI reported significantly greater reductions in depressive symptoms relative to controls from baseline to post in one RCT (Stice et al., 2007) and from baseline through 6 months in another trial (Stice et al., 2010, 2008). The mean ES across the 2 trials was significant at both post ($g_+ = -0.39$, $k = 2$, 95% CI $[-0.64, -0.14]$) and follow-up ($g_+ = -0.36$, $k = 2$, 95% CI $[-0.62, -0.11]$). In the one trial evaluating diagnostic outcomes, SEGI significantly reduced risk for depressive episodes (OR = 0.49) through 6 months post-intervention compared to controls (Stice et al., 2008).

Supportive-Expressive Group Intervention Effectiveness—No depression prevention studies have evaluated SEGI as delivered by endogenous providers and there have been no evaluations by independent research groups.

Resourceful Adolescent Program (RAP)

RAP is a CB-based group program. There have been three RCTs of RAP in two countries with a diverse group of participants (Rivet-Duval, Heriot, & Hunt, 2011; Rose, Hawes, & Hunt, 2014; Stallard et al., 2012). Rose and colleagues compared two variations of the program to each other and to a control condition: *RAP* followed by a placebo intervention and *RAP* followed by an interpersonal intervention. We only evaluated effects from this study at post intervention as this was the only assessment conducted after completion of the *RAP* program but before commencement of the placebo and interpersonal interventions. Therefore, only the post assessment provided information about the effectiveness of the RAP program independent of other interventions. RAP also has been evaluated in two non-randomized controlled trials (Harnett & Dadds, 2004; Shochet et al., 2001) and an RCT with only a placebo comparison condition (Merry et al., 2004). Findings from these studies are mixed.

RAP Efficacy—Youth receiving RAP reported significant reductions in depressive symptoms at post-intervention in one of the three trials (Rivet-Duval et al., 2011). In one study, youth receiving *RAP* reported *higher* levels of depressive symptoms at the 12-month follow-up when using an instrumental variable analysis to estimate the complier average

causal effect (i.e., the intervention effect accounting for adherence to the intervention) (Stallard et al., 2012). Mean ESs across trials were small and not significant at either post ($g_+ = -0.05$, $k = 2$, 95% CI $[-0.25, 0.15]$) or follow-up ($g_+ = 0.12$, $k = 3$, 95% CI $[-0.004, 0.25]$). There was significant between-study heterogeneity at post ($Q = 4.89$, $p < .05$), but not follow-up ($Q = 3.42$, ns). None of these studies evaluated diagnostic outcomes.

RAP Effectiveness—One study, in which school-based *RAP* groups were led by facilitators with at least a bachelor's degree in a relevant discipline or experience working with children, found no evidence of an intervention effect (Stallard et al., 2012). Although these providers were not inherently present in the participating schools, they were easily available to implement the intervention. All three RCTs were conducted by independent research groups, with one finding a short-term benefit on depressive symptoms at post intervention (Rivet-Duval et al., 2011).

LARS&LISA (L&L)

[Lust An Realistischer Sicht & Leichtigkeit Im Sozialen Alltag Lust; translated as Desire for a Realistic View and Ease in Social Aspects of Everyday Life] is a group CB-based intervention that has been implemented with universal samples of youth in school settings in both Germany and the United States. There have been three RCTs with published English-language articles. Non-English articles (Groen, Pössel, Al-Wiswasi, & Petermann, 2003) evaluating L&L were not included in this review.

L&L Efficacy—Significant effects on depressive symptoms favoring L&L have been reported in all three studies for at least a subgroup of participants. Main effects on symptoms have been present at four months post-intervention (Pössel, Martin, Garber, & Hautzinger, 2013) and subgroup effects have been present at 12 months post-intervention (Pössel, Seemann, & Hautzinger, 2008). Mean ESs across trials were not significant at either post ($g_+ = 0.01$, $k = 3$, 95% CI $[-0.13, 0.15]$) or follow-up ($g_+ = -0.10$, $k = 3$, 95% CI $[-0.25, 0.04]$), and between-study heterogeneity was not significant at either post ($Q = 4.29$, ns) or follow-up ($Q = 3.64$, ns). Diagnostic outcomes were not assessed in these studies.

L&L Effectiveness—No studies have evaluated L&L delivered by endogenous providers, and there have been no evaluations by independent research groups published in English.

Problem-Solving for Life (PSFL)

PSFL is a CB-based group intervention that has been evaluated in two large effectiveness trials (Sheffield et al., 2006; Spence et al., 2003; Spence, Sheffield, & Donovan, 2005). PSFL has been delivered to universal and indicated samples. In one study, schools were randomly assigned to no-intervention control, or one of three delivery strategies for the PSFL program: universal, indicated, or universal + indicated. In the universal + indicated condition, PSFL was delivered to all youth regardless of symptom levels in a whole-classroom format (universal), and then a subsample of youths with elevated symptom levels received additional training in smaller groups (i.e., indicated) (Sheffield et al., 2006).

PSFL Efficacy—A significant short-term effect on depressive symptoms at post-intervention was reported in one (Spence et al., 2003) of the two trials. The mean ES across trials at post was significant but small ($g_+ = -0.19$, $k = 2$, 95% CI $[-0.28, -0.11]$), whereas the mean ES was not significant at follow-up ($g_+ = 0.03$, $k = 2$, 95% CI $[-0.06, -0.12]$). Both studies conducted diagnostic evaluations but neither found a significant intervention effect.

PSFL Effectiveness—In both trials, endogenous providers (school personnel) led the intervention groups. One study found a significant effect on depressive symptoms at post intervention (Spence et al., 2003). No evaluations by independent research teams have been reported.

Interpersonal Psychotherapy-Adolescent Skills Training (IPT-AST)

IPT-AST is based on Interpersonal Psychotherapy for Adolescents (Mufson, Dorta, Moreau, & Weissman, 2004) and contains both individual and group components. The program has been evaluated using both universal (Horowitz et al., 2007) and indicated samples (Young et al., 2006; Young, Mufson, & Gallop, 2010) in three RCTs, two of which were conducted in schools with predominantly low-income and Hispanic youth (Young et al., 2006, 2010).

IPT-AST Efficacy—Youths receiving IPT-AST reported significantly greater reductions in depressive symptoms compared to controls in all three RCTs, with effects persisting through 6 months in one trial (Young et al., 2006). The mean ES across trials was significant at both post ($g_+ = -0.49$, $k = 3$, 95% CI $[-0.71, -0.28]$) and follow-up ($g_+ = -0.24$, $k = 3$, 95% CI $[-0.46, -0.01]$), with significant between-study heterogeneity at both time points: $Q_{\text{Post}} = 11.38$, $p < .01$ and $Q_{\text{Follow}} = 9.12$, $p < .05$. Two studies evaluated diagnostic outcomes. In the first, there was a nonsignificant trend ($p = .08$) whereby IPT-AST tended to reduce risk for depressive episodes relative to control (OR = .10) (Young et al., 2006). In the second, youth in IPT-AST were at reduced risk for depressive episodes through 6 months of follow-up (OR = 0.00; no youth meeting criteria in IPT-AST and 19.1% in control) (Young et al., 2010).

IPT-AST Effectiveness—In the study by Horowitz et al. (2007), the group leaders were trained and supervised by the program developer, but the leaders had not been involved in the program creation. No other studies have evaluated IPT-AST delivered by endogenous providers.

FRIENDS for Life

The effect of the *FRIENDS* program on depression outcomes has been evaluated in four school-based RCTs in the Brisbane region of Queensland, Australia. *FRIENDS* is primarily an anxiety prevention program but depression outcomes have been reported in several studies.

FRIENDS Efficacy—Two RCTs have shown significant benefits of *FRIENDS* on depressive symptoms, with the main effects on symptoms enduring for 12 months (Lock & Barrett, 2003; Lowry-Webster, Barrett, & Dadds, 2001; Lowry-Webster, Barrett, & Lock, 2003). The mean ES across studies for depressive symptoms was significant, favoring

FRIENDS relative to controls at follow-up ($g_+ = -0.24$, $k=3$, 95% CI $[-0.34, -0.14]$), but not at post ($g_+ = -0.04$, $k=4$, 95% CI $[-0.14, 0.05]$). There was significant between-study variability at post ($Q = 12.61$, $p < .01$), but not at follow-up ($Q=1.23$, ns). Barrett and Turner (2001) reported a greater *increase* in depressive symptoms in the *FRIENDS* condition relative to controls from baseline to post in the teacher-led subgroup and no effect of a clinical psychologist-led subgroup relative to controls. Diagnostic assessments were conducted in two studies, but few cases of depression were diagnosed and therefore no statistical analyses of effects on depression diagnoses were reported in either study.

FRIENDS Effectiveness—Significant main effects on depressive symptoms favoring *FRIENDS* were reported in two studies with endogenous providers (Lock & Barrett, 2003; Lowry-Webster et al., 2001, 2003). In both studies, intervention effects were evaluated across multiple schools with a diverse pool of participants. No RCTs conducted independently of the program developers that evaluated depression outcomes were found.

How Have Depression Prevention Programs Fared in Large Effectiveness Trials?

With few exceptions, depression prevention programs in youth have not fared well in large effectiveness trials. Next we summarize findings from seven RCTs that are notable for having large, diverse samples ($N > 1,000$), spanning multiple intervention sites, and implemented with a high degree of ecological validity. Perhaps the most promising evidence of broad effectiveness comes from an Australian cluster randomized trial of MoodGYM, an online teacher-supported CB program. In adolescents from 30 Australian schools ($N=1477$), the intervention reduced depressive symptoms through 6 months follow-up for males, although not for females (Calear, Christensen, Mackinnon, Giffiths, & O’Kearney, 2009).

Two large effectiveness studies (Sheffield et al., 2006; Spence et al., 2003, 2005) have evaluated the *Problem Solving for Life (PSFL)* intervention across multiple sites implemented by endogenous providers (i.e., school teachers). In the first study, 16 schools in the Brisbane region of Queensland, Australia were matched on demographic characteristics and then randomized to PSFL or an assessment-only control condition ($N = 1500$). Among youth with elevated baseline symptoms, those in the *PSFL* condition reported significantly greater decreases in depressive symptoms, in the short-term. No further significant group differences were found, however, at any subsequent assessments spanning four years of follow-up.

Sheffield and colleagues (2006) conducted a large cluster RCT evaluating the effectiveness of using universal, indicated, and universal plus indicated delivery approaches. Thirty-four schools in Queensland and New South Wales, Australia ($N = 2479$) participated; after being matched on demographic variables, students were randomly assigned to universal, indicated, or universal plus indicated intervention, or an assessment-only control. No intervention effects were found on levels of depressive symptoms, rates of depressive episodes among high-risk participants, or in levels of targeted mediators over a 12-month follow-up.

The *beyondblue* program was evaluated in a large ($N = 5634$) effectiveness trial conducted in 50 Australian secondary schools (Sawyer et al., 2010). *Beyondblue* is comprised of four components designed to target established individual-level risk factors through classroom

training, improve social interactions at the school-level, facilitate access to professional services, and impart information needed to improve identification of problems and help-seeking. Schools were matched into pairs and then randomized to either *beyondblue* or a comparison condition. Over a five-year follow-up, no evidence was found of intervention effects on depressive symptoms or on the four individual-level hypothesized mediators (Sawyer et al., 2010).

The *Resourceful Adolescent Program (RAP)* was tested in a cluster RCT in eight schools (out of 66 invited) drawn from the East Midlands and southwest of England ($N = 5030$; Stallard et al., 2012). Participating schools were randomized to *RAP*, an attention-control condition, or usual care. Although the program was delivered universally in schools, the investigators specifically focused analyses of effects among students with high baseline symptoms (i.e., indicated sample). At the 12-month follow-up, the intervention group reported significantly *higher* levels of depressive symptoms compared to the usual care condition.

The effectiveness of the *YoPiensoSientoActuo* (“I Think, Feel, and Act”) program, a CB-based classroom intervention, was evaluated in a cluster-randomized trial with 22 schools (from a potential sampling of 85 schools) in socioeconomically poor regions in Chile (Araya et al., 2013). Trios of mental health workers who had received training and ongoing supervision from senior clinicians led the intervention groups. Despite a relatively high attendance rate at intervention sessions, youths in intervention classrooms evidenced no benefits on depressive symptoms relative to those receiving the standard school curriculum over a 12-month follow-up.

Finally, there have been two large effectiveness studies of the Penn Resiliency Program (PRP). Challen and colleagues (2014) evaluated the effectiveness of an adapted version of PRP (the U.K. Resilience Program or UKRP) when delivered under highly realistic conditions in 16 U.K. secondary schools ($N = 2844$). There was a small beneficial effect of UKRP at post-intervention ($d = 0.09$) that did not persist at the one and two-year follow-up assessments. Intervention effects were stronger among students in classrooms rated as having high-quality intervention implementation. A notable limitation of this study, however, was the lack of random assignment to study conditions.

Kindt and colleagues (2014) randomly assigned classrooms within 13 participating secondary schools in the Netherlands to the Op Volle Kracht intervention (OVK; a Dutch adaption of PRP), as delivered by school teachers, as compared to an assessment-only control group. Results indicated that there were no main effects on depressive symptoms. Moreover, OVK participants were *more* likely to report clinically significant levels of symptoms at the 12-month follow-up as compared to controls. There was a significant moderated effect such that among youth whose parents had psychopathology, those in the OVK condition were less likely to report symptoms in the clinical range at the 12-month follow-up relative to their counterparts in the control condition. Finally, another large, school-based implementation trial is currently underway in the Netherlands (Tak & Engel, 2013; Tak et al., 2012), but outcomes from this study have not yet been published.

In summary, evidence from these large pragmatic effectiveness trials raises considerable doubt about the viability of classroom-based depression prevention efforts when delivered to scale (Merry, 2013). It is unlikely that the null findings in these studies were due to low power or inadequate dosage. In fact, attendance at group sessions in these studies tended to be higher than is typically seen in the depression prevention literature, and most studies reported high levels of intervention fidelity. Thus, whereas meta-analytic reviews of the overall youth depression prevention literature (Horowitz & Garber, 2006; Merry et al., 2011; Stice et al., 2009) have concluded that some depression prevention efforts have been efficacious, results from these large scale depression prevention dissemination trials have been less encouraging.

Discussion

Overall Evaluation

This review indicates that some depression prevention programs have demonstrated efficacy. For depressive symptoms, eight programs have revealed main effects relative to a control, in multiple RCTs: PRP, CwSC, AOP, Blues Group, Bibliotherapy, IPT-AST, LARS&LISA, and FRIENDS. All but two programs (LARS&LISA and RAP) had significant mean ESs at either post or follow-up, or both. Five of the 11 programs had at least one trial that demonstrated main effects on depressive symptoms that were present at least one year post-intervention (PRP, CwSC, Blues Group, Bibliotherapy, and FRIENDS). For six programs (PRP, CwSC, Blues Group, Bibliotherapy, SEGI, and IPT-AST), the mean ES across studies was significant at both post and follow-up; some of these programs (e.g., PRP) had more power than most to detect significant effects. Only the CwSC and Blues Group programs have demonstrated efficacy in targeting both depressive symptoms and depressive episodes across multiple independent trials.

With regard to effectiveness, nearly all studies evaluated program effects in realistic settings with broad access to the populations of interest (e.g., schools, HMOs, clinics). Six programs had at least one study showing effects on depression outcomes when delivered by endogenous service providers (PRP, CwSC, AOP, FRIENDS, Blues Group, and PSFL). Three programs have had effects on depression outcomes in studies conducted independently of the program developers (PRP, CwSC, and RAP); the effects of CwSC and RAP in these studies did not endure beyond the immediate post-intervention evaluation, however (Horowitz et al., 2007; Rivet-Duval et al., 2011). PRP is the only intervention to demonstrate efficacy with regard to depressive symptoms across multiple trials conducted by independent research groups, although effects have been inconsistent. No study of any program has been conducted independently of the program developers and has been shown to be effective with endogenous providers. Bibliotherapy, however, a self-administered intervention, has shown efficacy in studies conducted independently of the intervention developer.

Despite the fact that several programs have demonstrated promise in terms of efficacy, no youth depression prevention program has yet garnered sufficient evidence of effectiveness under realistic conditions to warrant widespread dissemination at this time. This conclusion is consistent with that of a previous review evaluating whether or not depression prevention

programs had satisfied the SPR-SOE criteria for dissemination (Spence & Shortt, 2007). The CwSC, Blues Group, and IPT-AST are notable for showing consistent beneficial effects across trials. Conducting large scale evaluations of these programs under increasingly realistic conditions and with decreasing involvement from the program developers seems like a logical next step on the path toward dissemination.

Additional Questions Related to Efficacy and Effectiveness

Durability of effects—The SPR-SOE guidelines for efficacy mandate that an intervention demonstrates “effects for at least one long-term follow-up ... e.g., at least 6 months” (p. 5), although the guidelines acknowledge that the appropriate duration will vary depending upon the intervention and the disorder. All but two (RAP and PSFL) of the programs reviewed here have demonstrated effects on depression outcomes in at least a subgroup of participants for as long as 6 months. In fact, follow-up g_+ estimates (i.e., the mean ES across trials at the first follow-up of 6 months or greater) were significant for seven programs (PRP, CwSC, FRIENDS, Blues Group, Bibliotherapy, SEGI, and IPT-AST). It is not clear, however, what duration of effects on depression outcomes is needed in order for prevention programs to have a public health impact that is meaningful enough to warrant widespread dissemination.

Is it worthwhile to broadly disseminate programs whose effects on depression diminish or disappear after 6 months? Even brief improvements in depression severity or a delay in the onset of depressive disorders might lead to other benefits (e.g., academic and social functioning). Moreover, how long can we realistically expect these relatively brief interventions to have sustained effects? Stressors continue to occur across the lifespan; it might be helpful for individuals to have “boosters” that remind them of how to cope with negative life events as they occur over the course of development. An analogy can be made to the process of losing weight and maintaining it, which require making healthy food choices as a continuous life style and not just when “dieting.” Dealing with stressors in a healthy way takes continued application of the skills learned in these various depression prevention programs, otherwise individuals may fall back into their “automatic” and often dysfunctional patterns of thinking and acting. Therefore, providing some kind of boosters on an “as needed” basis, either in person or through the internet might help youths sustain the benefits accrued in the short-term from these depression prevention programs. The efficacy of such longer-term “reminders” and the best method(s) for providing them should be the focus of future systematic investigations.

Few youth mental health prevention programs have evaluated long-term outcomes into adulthood. One noteworthy exception is the New Beginnings Program, a family-based intervention that targeted a range of outcomes, not just depression, reduced risk for internalizing disorders 15 years after the intervention (Wolchik et al., 2013). These findings are encouraging; this long-term follow-up approach should be emulated in future depression prevention studies.

Practical significance—Within the efficacy criteria, the SPR-SOE guidelines state that interventions must “demonstrate practical significance in terms of public health impact” (p. 5). The term “practical significance” is rather vague, however; there is no consensus as to

what constitutes a “meaningful” effect. The threshold is particularly nebulous when considering effects on dimensional measures of symptom severity. For example, PRP reduces depression symptom severity by approximately a fifth of a *SD* on average at the 6- to 8-month follow-up evaluations. Whether a reduction of this magnitude translates into practical benefits for youth, and how many adolescents would have to receive the intervention in order for there to be a measurable public health benefit are still unknown (Brunwasser et al., 2009).

The practical implication of effects on diagnostic outcomes is somewhat clearer given that clinically significant distress or functional impairment is a requirement for depression diagnoses (American Psychiatric Association, 2013). Thus, reductions in risk for diagnostic outcomes reflect a meaningful benefit that can be converted to easily interpretable metrics, like the number needed to treat (NNT), which conveys how many youth would have to receive an intervention in order to prevent one case with a depression diagnosis (Kraemer & Kupfer, 2006). For example, the CwSC program would need to be delivered to approximately 9 (Beardslee et al., 2013; Clarke et al., 1995; Garber et al., 2009) adolescents to prevent one case of depression. Although both PRP and CwSC reduce depressive symptoms, CwSC clearly results in meaningful benefits on diagnoses, whereas PRP has not established efficacy in preventing diagnosed depression disorders.

Diagnostic outcomes have drawbacks, however, in that intervention benefits are only captured if they reduce the likelihood of crossing the (somewhat arbitrary) diagnostic threshold. That is, no distinction is made between youth with no symptoms and those who just narrowly fall short of the diagnostic threshold, or between those whose symptoms just barely meet the diagnostic threshold and those whose symptoms are far in excess of it. An increasing number of studies are evaluating functional outcomes in addition to diagnostic or symptom-severity measures (e.g., Brunwasser et al., 2014; Young, Kranzler, Gallop, & Mufson, 2012). This focus on functioning should become part of the criteria defining practical and meaningful benefits that are a prerequisite to widespread dissemination.

Cost-effectiveness and cost-offsets—Showing that programs effectively prevent depressive symptoms and disorders and improve functioning is not the only hurdle to broad implementation. Researchers also must demonstrate that the program provides good value relative to the cost of implementation. Depression is one of the leading causes of disability worldwide, carrying an immense fiscal toll due to lost productivity and increased healthcare expenditures (Ferrari et al., 2013). Prevention efforts that reduce the prevalence, duration, or severity of depression should result in substantial cost offsets, making implementation efforts a good investment.

Studies evaluating the costs and savings associated with depression prevention trials are still relatively rare. Thus far, however, economic analyses have been encouraging (Mihalopoulos & Vos, 2013). For example, when delivered to a targeted sample in primary care clinics, the incremental cost-effectiveness ratio (ICER) of CwSC was found to be good by traditional standards: \$9275 per quality-adjusted life year (Lynch et al., 2005). Another examination of cost-effectiveness aggregated data from eight indicated trials that had evaluated depression diagnostic outcomes, and found that the ICER (ranging from \$5400 to \$14000 Australian

per disability-life year averted depending on model assumptions) was below a threshold commonly used to indicate cost-effectiveness (Mihalopoulos, Vos, Pirkis, & Carter, 2012).

As depression prevention research becomes increasingly focused on demonstrating effectiveness and readiness for dissemination, documenting the costs of intervention implementation should become part of the standard criteria for evaluating readiness for dissemination. Additionally, researchers should collect measures that optimally lend themselves to cost-effectiveness analyses (e.g., depression-free days). Many complications accompany cost-effectiveness evaluations such as accounting for declining implementation costs as the intervention becomes increasingly integrated into the settings of interest. Moreover, because many prevention researchers are relatively unfamiliar with this line of research, additional training and external consultation will be needed.

Evidence of intervention superiority over displaced activities—If broadly implemented, some depression prevention programs would occur in place of something else. For example, school-based interventions might serve as an alternative health curriculum, thereby displacing the regular health class. Such comparisons provide an ecologically sound test of whether the program of interest is better at preventing depression than the usual activities. If the displaced activity and the prevention program were designed to achieve the same goal (e.g., prevent depression), then this would be a fair comparison; otherwise, it is not. That is, showing that a prevention program is superior at reducing risk for depression as compared to a health class designed to promote healthy behaviors (e.g., eating habits), is not evidence that the depression program is necessarily overall “better” than the standard curriculum. It doesn’t make sense to judge the value of the standard curriculum by determining its relative impact on an outcome it was not designed to affect. Rather, to establish the superiority of a specific prevention program, researchers should explicitly measure the targeted outcomes of both the displaced activities (e.g., eating habits) and the specific intervention (e.g., depression). Evaluating both programs in tandem not only can provide information about their relative value, but it also can test whether existing programs are achieving their goals.

Evaluating the relative benefits over “usual” activities is more complicated when there is no inherent structure for the control condition. For example, when evaluating clinic-based prevention programs and self-guided interventions (e.g., web-based programs), there generally is no standard activity for all control participants. Thus, there will be individual variability in the types of displaced activities, making it difficult to gauge their relative value. As prevention moves toward dissemination, researchers should consider the relative benefits among a number of desirable outcomes rather than only about the outcomes in which they are most invested. This will require working closely with representatives of dissemination sites to learn more about the goals of existing (and potentially replaceable) activities. Additionally, information about such displaced activities should be described in published research articles.

Evidence of intervention superiority over simpler alternatives—Dismantling studies test whether subcomponents of a multifaceted intervention are sufficient to produce beneficial results that are at least comparable to those achieved with the full intervention. A

good example of this in the depression prevention literature is the work of Stice, Rohde, and colleagues (Rohde et al., 2014; Stice et al., 2008; Stice et al., 2007). In their studies, a multi-component CB program (the *Blues Group*) was compared to several more parsimonious interventions, each containing components of the full program. These studies were designed to determine which of the various interventions best balanced potency and feasibility, which is critical for moving toward dissemination given the logistical challenges of broad implementation in real-world settings. Brief programs do not necessarily achieve their aims better than longer ones, however (although see a meta-analysis by Stice et al., 2009). If more intensive programs produce stronger or more sustained effects, then efforts should be made to secure the necessary resources. All things being equal, preference typically will be given to simpler and less costly programs, as long as they are truly effective.

Web-based and self-guided interventions—An intriguing and burgeoning approach to prevention involves interventions that are not dependent upon trained leaders and supportive dissemination sites. Web-based interventions, like CATCH-IT (Van Voorhees, Ellis, Stuart, Fogel, & Ford, 2005) and SPARX (Merry et al., 2012), and self-guided interventions such as *Feeling Good* Bibliotherapy (Burns, 1989; Stice et al., 2007) can more easily go to scale and avoid many of the practical challenges that accompany in-person interventions (e.g., leader training, travel burden, intervention fidelity and quality). On the other hand, web-based programs have some difficulty recruiting individuals to actually participate and also have a problem with attrition (Andersson & Cuijpers, 2009).

Some web-based interventions have been effective in treatment RCTs for adult depression (Andersson & Cuijpers, 2009) and have shown promising effects in youth depression prevention trials (Merry et al., 2012; Van Voorhees et al., 2008). CB-based bibliotherapy has been shown to be efficacious in reducing depressive symptoms (Stice et al., 2007), although a recent trial found that a group-based CB intervention was significantly more effective than bibliotherapy at reducing the incidence of depressive episodes (Rohde et al., 2014).

Additional research is needed to determine whether self-guided interventions are sufficient to prevent depression or whether some degree of in-person support also is needed. If effective, these interventions would offer a more easily scalable and potentially more cost-effective alternative to in-person interventions, as long as problems with recruitment and retention can be addressed.

Next Steps for Depression Prevention

Depression prevention research has reached an important crossroad. On the one hand, evidence emerging from efficacy and small-scale effectiveness trials is promising, particularly with targeted samples. On the other hand, large-scale school-based effectiveness trials generally have been disappointing. Some programs (CwSC, Blues Group, IPT-AST) have produced significant effects on both depressive symptoms and episodes. The CwSC and Blues Group programs have performed well in trials with endogenous providers (Clarke et al., 1995, 2001; Rohde et al., 2014), and CwSC demonstrated short-term efficacy in targeting depressive symptoms in a trial conducted independently of the program developers (Horowitz et al., 2007). IPT-AST has not yet demonstrated effects when delivered by

endogenous providers or completely independent of program developers, which is an important future direction.

Would these programs (CwSC, Blues Group, IPT-AST) fare any better in large-scale effectiveness trials than PRP, RAP, or PSFL? What factors might be responsible for the apparent drop-off in intervention effects under increasingly realistic conditions? Perhaps the implementation of depression prevention programs requires greater scaffolding than can be provided when delivered on a broad scale. For example, the ability of program developers to provide support and training likely diminishes in larger trials. Smaller effectiveness RCTs are more likely to select providers (e.g., teachers) based on willingness and enthusiasm for implementing the protocol. In large effectiveness trials, often all teachers in specific grades teach the course as part of their standard curriculum, which may be perceived as an unwanted burden. Finally, when interventions are incorporated into the regular class curriculum, they may seem like just another course and thus less special or appealing to students.

Can we justify additional pragmatic implementation studies in schools for programs that have shown promising effects in smaller-scale studies, or should we focus more on other settings (e.g., clinics) or methods of delivery (e.g., web-based programs)? Perhaps the large-scale implementation trials conducted thus far were premature. PSFL and RAP were broadly implemented without strong support from RCTs. PRP had a stronger evidence base in terms of improving depressive symptoms prior to large-scale implementations, but the sources of inconsistency in PRP's effects are not well understood (Brunwasser et al., 2009). Ideally, we would have known more about the contexts in which PRP is most and least effective before evaluating it on such a large scale.

Some may argue that incremental approaches to intervention evaluation, like those endorsed in the SPR-SOE guidelines, are too slow and inefficient, because program efficacy could be established in a series of costly RCTs but still fail when implemented under real world circumstances. Might a similar conclusion have been reached using a single, pragmatic effectiveness trial that saved time and resources? On the other hand, important insights into the conditions under which the program is most effective may result from efficacy trials, which then can inform the design and implementation of pragmatic effectiveness trials, and subsequently provide a test of the program's utility under realistic circumstances. Thus, both incremental and pragmatic approaches have advantages and disadvantages. The precise benefits and drawbacks of each approach to the prevention of youth depression, however, remain to be identified and tested.

The schism between efficacy and effectiveness research is neither new nor peculiar to depression prevention (Glasgow, Magid, Beck, Ritzwoller, & Estabrooks, 2005). A potential way to bridge this gap is by combining desired features of traditional efficacy and effectiveness studies; that is, balancing internal and ecological validity in efficacy trials (e.g., Clarke, 1995; Glasgow et al., 2005). For example, rather than initially conducting an RCT exclusively with providers who are trained and monitored closely by the program developers and then conducting a separate RCT with endogenous providers, a single, multiple-cohort design could accommodate both approaches. In the first cohort, non-endogenous providers

(e.g., research team members) would lead intervention groups with endogenous providers (e.g., school teachers) being aids. In the second cohort, the same teachers could take the lead role in facilitating the intervention with support from the research team. Finally, the trained teachers could lead the intervention independently with no input from the research team. Cohort number then could be used as a moderator in outcome analyses to determine whether effects attenuate with diminishing support from the research team. Not only does this provide a method of expediting the efficacy-to-effectiveness transition, but it also might be a particularly efficient and effective method of training endogenous providers. The extent to which the endogenous providers still need to be supervised and monitored for fidelity and quality also should be examined.

One other important issue for dissemination concerns *where* depression prevention programs should be implemented. Although schools are where the children are, there may be iatrogenic effects of conducting the programs in the school setting. First, some children may be ostracized if they reveal distress in front of peers, making it more difficult for them to fully participate and thereby potentially benefit from the program. Second, as noted earlier, providing a depression prevention program in school may displace other activities from which students also may benefit.

Where, then, should youth depression prevention programs be implemented? Given the high rates of depression in offspring of depressed parents (Beardslee, Gladstone, & O'Connor, 2011) and evidence of effective preventive interventions with such at-risk youth (e.g., Clarke et al., 2001; Compas et al., 2009; Garber et al., 2009), providing services to children of parents being treated for depression within the same clinical setting might be particularly appealing to families. This would require having the staff and space to accommodate such interventions with the children, but these seem like manageable obstacles. Pediatricians' offices are another logical setting for possible dissemination efforts.

Finally, in addition to conducting new RCTs, another approach to advancing depression prevention research would be to make full-use of existing data to better understand the contexts in which interventions are most (and least) effective. Single RCTs often lack the statistical power needed to detect moderation and mediation effects. Meta-analysis is helpful in summarizing effects across studies, but does not provide adequate information about the contexts in which interventions are most effective, because moderators only can be measured at the study level (Brown, Wang, & Sandler, 2008). A synthesis study currently is underway in which investigators are harmonizing subject-level data from many depression prevention RCTs in order to maximize the likelihood of detecting moderation and mediation (Perrino et al., 2013). The findings from this data synthesis project may greatly improve our understanding of when and why prevention programs are effective, and thereby inform the design of future dissemination trials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

FUNDING

This work was supported in part from National Institute of Mental Health grants (R01MH64735; R01MH100258) and training grant (T32 MH018921), and the National Center for Advancing Translational Sciences, Grant 2 UL1 TR000445-06.

References

- Ackerson J, Scogin F, McKendree-Smith N, Lyman RD. Cognitive bibliotherapy for mild and moderate adolescent depressive symptomatology. *Journal of Consulting and Clinical Psychology*. 1998; 66(4):685. doi:10.1037/0022-006X.66.4.685. [PubMed: 9735587]
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5TM. 5th. American Psychiatric Publishing, Inc; Arlington, VA: 2013.
- Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cognitive Behaviour Therapy*. 2009; 38(4):196–205. [PubMed: 20183695]
- Araya R, Fritsch R, Spears M, Rojas G, Martinez V, Barroilhet S, Montgomery AA. School intervention to improve mental health of students in Santiago, Chile: a randomized clinical trial. *JAMA Pediatrics*. 2013; 167(11):1004–1010. doi:10.1001/jamapediatrics.2013.2361. [PubMed: 23999656]
- Barrett, PM. FRIENDS for Life: Group leaders' manual for children. Barrett Research Resources Pty Ltd; 2005.
- Barrett PM, Turner C. Prevention of anxiety symptoms in primary school children: Preliminary results from a universal school-based trial. *British Journal of Clinical Psychology*. 2001; 40(4):399–410. doi:10.1348/014466501163887. [PubMed: 11760616]
- Beardslee WR, Brent DA, Weersing VR, Clarke GN, Porta G, Hollon SD, Garber J. Prevention of Depression in At-Risk Adolescents: Longer-term Effects. *JAMA Psychiatry*. 2013; 70(11):1161–1170. doi:10.1001/jamapsychiatry.2013.295. [PubMed: 24005242]
- Beardslee WR, Wright EJ, Salt P, Drezner K. Examination of children's responses to two preventive intervention strategies over time. *Journal of the American Academy of Child & Adolescent Psychiatry*. 1997; 36(2):196–204. doi:10.1097/00004583-199702000-00010. [PubMed: 9031572]
- Borenstein, M.; Hedges, L.; Rothstein, H. Meta-analysis: fixed effect vs. random effects. 2007. Retrieved from <http://www.meta-analysis.com/downloads/Meta-analysis%20fixed%20effect%20vs%20random%20effects.pdf>
- Borenstein, M.; Hedges, LV.; Higgins, JPT.; Rothstein, HR. Introduction to meta-analysis. John Wiley & Sons; 2009.
- Brent DA, Holder D, Kolko D, Birmaher B, Baugher M, Roth C, Johnson BA. A clinical psychotherapy trial for adolescent depression comparing cognitive, family, and supportive therapy. *Archives of General Psychiatry*. 1997; 54(9):877–885. doi:10.1001/archpsyc.1997.01830210125017. [PubMed: 9294380]
- Brown CH, Wang W, Sandler IN. Examining how context changes intervention impact: The use of effect sizes in multilevel mixture meta-analysis. *Child Development Perspectives*. 2008; 2(3):198–205. doi:10.1111/j.1750-8606.2008.00065.x. [PubMed: 20585469]
- Brunwasser SM, Gillham JE, Kim ES. A meta-analytic review of the Penn Resiliency Program's effect on depressive symptoms. *Journal of Consulting and Clinical Psychology*. 2009; 77(6):1042–1054. doi:10.1037/a0017671. [PubMed: 19968381]
- Burns, DD. The feeling good handbook: Using the new mood therapy in everyday life. William Morrow & Co; 1989.
- Cardemil EV, Reivich KJ, Beevers CG, Seligman MEP, James J. The prevention of depressive symptoms in low-income, minority children: Two-year follow-up. *Behaviour Research and Therapy*. 2007; 45(2):313–327. [PubMed: 16643843]

- Cardemil EV, Reivich KJ, Seligman MEP. The prevention of depressive symptoms in low-income minority middle school students. *Prevention & Treatment*. 2002; 5(1) doi: 10.1037/1522-3736.5.1.58a.
- Chaplin TM, Gillham JE, Reivich K, Elkon AGL, Samuels B, Freres DR, Seligman MEP. Depression prevention for early adolescent girls: A pilot study of all girls versus co-ed groups. *The Journal of Early Adolescence*. 2006; 26(1):110–126. [PubMed: 26139955]
- Clarke GN. Improving the transition from basic efficacy research to effectiveness studies: methodological issues and procedures. *Journal of Consulting and Clinical Psychology*. 1995; 63(5):718. [PubMed: 7593864]
- Clarke GN, Hawkins W, Murphy M, Sheeber LB, Lewinsohn PM, Seeley JR. Targeted prevention of unipolar depressive disorder in an at-risk sample of high school adolescents: A randomized trial of group cognitive intervention. *Journal of the American Academy of Child & Adolescent Psychiatry*. 1995; 34(3):312–321. doi:10.1097/00004583-199503000-00016. [PubMed: 7896672]
- Clarke GN, Hornbrook C, Lynch F, Polen M, Gale J, Beardslee WR, Seeley JR. A randomized trial of a group cognitive intervention for preventing depression in adolescent offspring of depressed parents. *Archives of General Psychiatry*. 2001; 58(12):1127–1134. doi:10.1001/archpsyc.58.12.1127. [PubMed: 11735841]
- Cochran WG. The combination of estimates from different experiments. *Biometrics*. 1954; 10(1):101–129.
- Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJL, Whiteford HA. Burden of depressive disorders by country, sex, age, and year: Findings from the Global Burden of Disease Study 2010. *PLoS Medicine*. 2013; 10(11):e1001547. doi:10.1371/journal.pmed.1001547. [PubMed: 24223526]
- Flay BR, Biglan A, Boruch RF, Castro FG, Gottfredson D, Kellam S, Ji P. Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science*. 2005; 6(3):151–175. doi:10.1007/s11121-005-5553-y. [PubMed: 16365954]
- Garber J, Clarke GN, Weersing VR, Beardslee WR, Brent DA, Gladstone TRG, Iyengar S. Prevention of depression in at-risk adolescents: A randomized controlled trial. *JAMA: Journal of the American Medical Association*. 2009; 301(21)(3):2215–2224. [PubMed: 19491183]
- Gillham JE. Preventing depressive symptoms in school children. *Dissertation Abstracts International: Section B: The Sciences and Engineering*. 1995; 55(9-B):4119.
- Gillham JE, Hamilton J, Freres DR, Patton K, Gallop R. Preventing depression among early adolescents in the primary care setting: A randomized controlled study of the Penn Resiliency Program. *Journal of Abnormal Child Psychology*. 2006; 34(2):195–211. doi:10.1007/s10802-005-9014-7.
- Gillham, JE.; Jaycox, LH.; Reivich, KJ.; Seligman, MEP.; Silver, T. The Penn Resiliency Program. University of Pennsylvania; Philadelphia: 1990. Unpublished Manual
- Gillham JE, Reivich KJ, Brunwasser SM, Freres DR, Chajon ND, Kash-MacDonald VM, Seligman MEP. Evaluation of a group cognitive-behavioral depression prevention program for young adolescents: A randomized effectiveness trial. *Journal of Clinical Child and Adolescent Psychology*. 2012; 41(5):621–639. doi:10.1080/15374416.2012.706517. [PubMed: 22889296]
- Gillham JE, Reivich KJ, Freres DR, Chaplin TM, Shatté AJ, Samuels B, Seligman MEP. School-based prevention of depressive symptoms: A randomized controlled study of the effectiveness and specificity of the Penn Resiliency Program. *Journal of Consulting & Clinical Psychology*. 2007; 75(1):9–19. doi:10.1037/0022-006X.75.1.9. [PubMed: 17295559]
- Glasgow RE, Magid DJ, Beck A, Ritzwoller D, Estabrooks PA. Practical clinical trials for translating research to practice: design and measurement recommendations. *Medical Care*. 2005; 43(6):551–557. [PubMed: 15908849]
- Groen G, Pössel P, Al-Wiswasi S, Petermann F. Universelle, schulbasierte Prävention der Depression im Jugendalter: Ergebnisse einer Follow-Up-Studie. *Kindheit und Entwicklung*. 2003; 12(3):164–174. doi:10.1026//0942-5403.12.3.164.
- Harnett PH, Dadds MR. Training school personnel to implement a universal school-based prevention of depression program under real-world conditions. *Journal of School Psychology*. 2004; 42(5):343–357. doi:10.1016/j.jsp.2004.06.004.

- Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*. 1981; 6(2):107–128. doi:10.2307/1164588.
- Horowitz JL, Garber J, Ciesla JA, Young JF, Mufson L. Prevention of depressive symptoms in adolescents: A randomized trial of cognitive-behavioral and interpersonal prevention programs. *Journal of Consulting and Clinical Psychology*. 2007; 75(5):693–706. doi:10.1037/0022-006X.75.5.693. [PubMed: 17907851]
- Jaycox LH, Reivich KJ, Gillham JE, Seligman MEP. Prevention of depressive symptoms in school children. *Behaviour Research and Therapy*. 1994; 32(8):801–816. doi:10.1016/0005-7967(94)90160-0. [PubMed: 7993324]
- Kindt K, Kleinjan M, Janssens JMAM, Scholte RHJ. 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*. 2014; 11(5):5273–5293. [PubMed: 24837666]
- Kraemer HC, Kupfer DJ. Size of Treatment Effects and Their Importance to Clinical Research and Practice. *Biological Psychiatry*. 2006; 59(11):990–996. doi:10.1016/j.biopsych.2005.09.014. [PubMed: 16368078]
- Lock S, Barrett PM. A Longitudinal Study of Developmental Differences in Universal Preventive Intervention for Child Anxiety. *Behaviour Change*. 2003; 20(4):183–199. doi:10.1375/bech.20.4.183.29383.
- Lowry-Webster HM, Barrett PM, Dadds MR. A universal prevention trial of anxiety and depressive symptomatology in childhood: Preliminary data from an Australian study. *Behaviour Change*. 2001; 18(1):36–50. doi:10.1375/bech.18.1.36.
- Lowry-Webster HM, Barrett PM, Lock S. A universal prevention trial of anxiety symptomatology during childhood: Results at 1-year follow-up. *Behaviour Change*. 2003; 20(1):25–43. doi:10.1375/bech.20.1.25.24843.
- Lynch FL, Hornbrook M, Clarke GN, Perrin N, Polen MR, O'Connor E, Dickerson J. Cost-effectiveness of an Intervention to Prevent Depression in At-Risk Teens. *Archives of General Psychiatry*. 2005; 62(11):1241–1248. doi:10.1001/archpsyc.62.11.1241. [PubMed: 16275811]
- Merry SN. Preventing depression in adolescents: time for a new approach? *JAMA Pediatrics*. 2013; 167(11):994–995. doi:10.1001/jamapediatrics.2013.2862. [PubMed: 24000104]
- Merry SN, McDowell H, Wild CJ, Bir J, Cunliffe R. A randomized placebo-controlled trial of a school-based depression prevention program. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2004; 43(5):538–547. doi:10.1097/00004583-200405000-00007. [PubMed: 15100560]
- Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MFG. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. *BMJ: British Medical Journal*. 2012:344.
- Mihalopoulos C, Vos T. Cost-effectiveness of preventive interventions for depressive disorders: an overview. *Expert Rev. Pharmacoecon. Outcomes Res*. 2013; 13(2):237–242. doi:10.1586/ERP.13.5. [PubMed: 23570434]
- Mihalopoulos C, Vos T, Pirkis J, Carter R. The population cost-effectiveness of interventions designed to prevent childhood depression. *Pediatrics*. 2012; 129(3):e723–e730. doi:10.1542/peds.2011-1823. [PubMed: 22312000]
- Mufson L, Dorta KP, Moreau D, Weissman MM. Interpersonal psychotherapy for depressed adolescents (2nd). 2004
- Muris P, Bogie N, Hoogsteder A. Effects of an early intervention group program for anxious and depressed adolescents: A pilot study. *Psychological Reports*. 2001; 88(2):481–482. doi:10.2466/PRO.88.2.481-482. [PubMed: 11351893]
- Pattison C, Lynd-Stevenson RM. The prevention of depressive symptoms in children: The immediate and long-term outcomes of a school based program. *Behaviour Change*. 2001; 18(2):92–102.
- Peng L, Li M, Zuo X, Miao Y, Chen L, Yu Y, Wang T. Application of the Pennsylvania resilience training program on medical students. *Personality and Individual Differences*. 2014; 61-62:47–51. doi:10.1016/j.paid.2014.01.006.

- Perrino T, Howe G, Sperling A, Beardslee W, Sandler IN, Shern D, Brown CH. Advancing science through collaborative data sharing and synthesis. *Perspectives on Psychological Science*. 2013; 8(4):433–444. doi:10.1177/1745691613491579. [PubMed: 24244216]
- Pössel P, Horn AB, Groen G, Hautzinger M. School-Based Prevention of Depressive Symptoms in Adolescents: A 6-Month Follow-up. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2004; 43(8):1003–1010. [PubMed: 15266195]
- Pössel P, Martin NC, Garber J, Hautzinger M. A randomized controlled trial of a cognitive-behavioral program for the prevention of depression in adolescents compared with nonspecific and no-intervention control conditions. *Journal of Counseling Psychology*. 2013; 60(3):432–438. doi:10.1037/a0032308. [PubMed: 23506514]
- Pössel P, Seemann S, Hautzinger M. Impact of comorbidity in prevention of adolescent depressive symptoms. *Journal of Counseling Psychology*. 2008; 55(1):106–117.
- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2014. Retrieved from <http://www.R-project.org/>
- Rivet-Duval E, Heriot S, Hunt C. Preventing adolescent depression in Mauritius: A universal school-based program. *Child and Adolescent Mental Health*. 2011; 16(2):86–91. doi:10.1111/j.1475-3588.2010.00584.x.
- Roberts CM. Embedding Mental Health Promotion Programs in School Contexts: The Aussie Optimism Program. *International Society for the Study of Behavior Newsletter*. 2006; 2(50):1–4.
- Roberts CM, Kane R, Bishop B, Cross D, Fenton J, Hart B. The prevention of anxiety and depression in children from disadvantaged schools. *Behaviour Research and Therapy*. 2010; 48(1):68–73. doi:10.1016/j.brat.2009.09.002. [PubMed: 19786274]
- Roberts CM, Kane R, Thomson H, Bishop B, Hart B. The prevention of depressive symptoms in rural school children: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*. 2003; 71(3):622–628. doi:10.1037/0022-006X.71.3.622. [PubMed: 12795585]
- Rohde P, Stice E, Shaw H, Brière FN. Indicated cognitive behavioral group depression prevention compared to bibliotherapy and brochure control: Acute effects of an effectiveness trial with adolescents. *Journal of Consulting and Clinical Psychology*. 2014; 82(1):65–74. doi:10.1037/a0034640. [PubMed: 24099432]
- Rooney RM, Hassan S, Kane R, Roberts CM, Nesa M. Reducing depression in 9–10 year old children in low SES schools: A longitudinal universal randomized controlled trial. *Behaviour Research and Therapy*. 2013; 51(12):845–854. doi:10.1016/j.brat.2013.09.005. [PubMed: 24185214]
- Rooney RM, Roberts CM, Kane R, Pike L, Winsor A, White J, Brown A. The prevention of depression in 8-to 9-year-old children: a pilot study. *Australian Journal of Guidance and Counselling*. 2006; 16(01):76–90. doi:10.1375/ajgc.16.1.76.
- Rose K, Hawes DJ, Hunt CJ. Randomized controlled trial of a friendship skills intervention on adolescent depressive symptoms. *Journal of Consulting and Clinical Psychology*. 2014; 82(3):510–520. doi:10.1037/a0035827. [PubMed: 24491074]
- Sheffield JK, Spence SH, Rapee RM, Kowalenko N, Wignall A, Davis A, McLoone J. Evaluation of universal, indicated, and combined cognitive-behavioral approaches to the prevention of depression among adolescents. *Journal of Consulting and Clinical Psychology*. 2006; 74(1):66–79. doi:10.1037/0022-006X.74.1.66. [PubMed: 16551144]
- Shochet IM, Dadds MR, Holland D, Whitefield K, Harnett PH, Osgarby SM. The efficacy of a universal school-based program to prevent adolescent depression. *Journal of Clinical Child Psychology*. 2001; 30(3):303–315. doi:10.1207/S15374424JCCP3003_3. [PubMed: 11501248]
- Shochet, IM.; Holland, D.; Whitefield, K. Resourceful Adolescent Program: Group leader's manual. Queensland University of Technology; 1997.
- Spence SH, Sheffield JK, Donovan CL. Preventing adolescent depression: An evaluation of the Problem Solving For Life program. *Journal of Consulting and Clinical Psychology*. 2003; 71(1):3–13. doi:10.1037/0022-006X.71.1.3. [PubMed: 12602420]
- Spence SH, Sheffield JK, Donovan CL. Long-Term Outcome of a School-Based, Universal Approach to Prevention of Depression in Adolescents. *Journal of Consulting and Clinical Psychology*. 2005; 73(1):160–167. doi:10.1037/0022-006X.73.1.160. [PubMed: 15709843]

- Stallard P, Sayal K, Phillips R, Taylor JA, Spears M, Anderson R, Montgomery AA. Classroom based cognitive behavioural therapy in reducing symptoms of depression in high risk adolescents: pragmatic cluster randomised controlled trial. *BMJ: British Medical Journal*. 2012; 345
- Stice E, Burton E, Bearman SK, Rohde P. Randomized trial of a brief depression prevention program: An elusive search for a psychosocial placebo control condition. *Behaviour Research and Therapy*. 2007; 45(5):863–876. [PubMed: 17007812]
- Stice E, Rohde P, Gau JM, Wade E. Efficacy trial of a brief cognitive-behavioral depression prevention program for high-risk adolescents: Effects at 1- and 2-year follow-up. *Journal of Consulting and Clinical Psychology*. 2010; 78(6):856–867. doi:10.1037/a0020544. [PubMed: 20873893]
- Stice E, Rohde P, Seeley JR, Gau JM. Brief cognitive-behavioral depression prevention program for high-risk adolescents outperforms two alternative interventions: A randomized efficacy trial. *Journal of Consulting and Clinical Psychology*. 2008; 76(4):595–606. doi:10.1037/a0012645. [PubMed: 18665688]
- Van Voorhees BW, Ellis J, Stuart S, Fogel J, Ford DE. Pilot Study of a Primary Care Internet-Based Depression Prevention Intervention for Late Adolescents. *Canadian Child and Adolescent Psychiatry Review*. 2005; 14(2):40–43. [PubMed: 19030514]
- Van Voorhees BW, Vanderplough-Booth K, Fogel J, Gladstone T, Bell C, Stuart S, Reinecke MA. Integrative Internet-based depression prevention for adolescents: A randomized clinical trial in primary care for vulnerability and protective factors. *Journal of the Canadian Academy of Child and Adolescent Psychiatry / Journal de l'Académie Canadienne de Psychiatrie de L'enfant et de L'adolescent*. 2008; 17(4):184–196. 2008.
- Viechtbauer W. Metafor: meta-analysis package for R. R Package Version. 20102010:1–0.
- Wijnhoven LAMW, Creemers DHM, Vermulst AA, Scholte RHJ, Engels RCME. Randomized controlled trial testing the effectiveness of a depression prevention program (“Op Volle Kracht”) among adolescent girls with elevated depressive symptoms. *Journal of Abnormal Child Psychology*. 2014; 42(2):217–228. doi:10.1007/s10802-013-9773-5. [PubMed: 23893066]
- Wolchik SA, Sandler IN, Tein J-Y, Mahrer NE, Millsap RE, Winslow E, Reed A. Fifteen-year follow-up of a randomized trial of a preventive intervention for divorced families: Effects on mental health and substance use outcomes in young adulthood. *Journal of Consulting and Clinical Psychology*. 2013; 81(4):660–673. doi:10.1037/a0033235. [PubMed: 23750466]
- Young JF, Kranzler A, Gallop RJ, Mufson L. Interpersonal Psychotherapy-Adolescent Skills Training: Effects on school and social functioning. *School Mental Health*. 2012; 4(4):254–264. doi:10.1007/s12310-012-9078-9. [PubMed: 23393545]
- Young JF, Mufson L, Davies M. Efficacy of Interpersonal Psychotherapy-Adolescent Skills Training: An indicated preventive intervention for depression. *Journal of Child Psychology and Psychiatry*. 2006; 47(12):1254–1262. doi:10.1111/j.1469-7610.2006.01667.x. [PubMed: 17176380]
- Young JF, Mufson L, Gallop RJ. Preventing depression: A randomized trial of interpersonal psychotherapy-adolescent skills training. *Depression and Anxiety*. 2010; 27(5):426–433. doi: 10.1002/da.20664. [PubMed: 20112246]
- Yu DL, Seligman MEP. Preventing depressive symptoms in Chinese children. *Prevention & Treatment*. 2002; 5(1) doi:10.1037/1522-3736.5.1.59a.

Table 1

Study Results and Characteristics by Intervention Program.

		Efficacy Evidence				Participants and Intervention Delivery		
		Depressive Symptoms		Depression Diagnoses				
	Study Length	Measures	Months	Measure	Months	Intervention Delivery Site	Sample Size & Breadth of Recruitment	Intervention Providers
Coping with Stress Course (CwSC)								
Clarke 1995	12mos	CES-D HDRS	Post <i>ns</i>	K-SADS/LIFE	12mos	High schools	N=172; pool of 1652 students; 3 suburban HS	School personnel
Clarke 2001	24mos	CES-D HDRS CBCL-D	12 <i>ns</i> ^a <i>ns</i>	K-SADS	24mos	HMO clinics	N=94; pool of 410,000-member HMO database; 3374 potentially eligible	MS level clinicians with clinical experience
Horowitz 2007	6mos	CDI CES-D	Post Post	None	NA	During school (Wellness class)	N=380; from pool of 600; 3 suburban/rural schools	MS or PhD level clinicians with therapy experience
Garber 2009; Beardslee 2013	9mos 33mos	CES-D CDRS-R	9 (M), 33 (S) <i>ns</i>	K-SADS/LIFE	33mos	HMO & University clinics	N=316; 2494 assessed for eligibility; recruited from 4 regions of US	MS or PhD level clinicians with therapy experience
Penn Resiliency Program (PRP)								
Quayle 2001	6mos	CDI	6	None	NA	In schools, during school day	N=47; pool of 70; all-girls private school; Australia	Post-graduate clinical psychology students trained by program developers
Pattison 2001	8mos	CDI	<i>ns</i>	None	NA	In schools, during school day	N=66; of 150; rural school; Australia	Mental health professionals; trained by program developers
Cardemil 2002								
<i>Latino Sample</i>	24mos	CDI	24	None	NA	In an urban middle school during school day	N=65; pool of 977; middle school (mostly low income and Latino)	MS level in psychology, counseling, or education trained by program developer
<i>African American Sample</i>	24mos	CDI	3 (I)	None	NA		N=106; pool of 828; mid school (most low income & African American)	Program developer and a MS level graduate student
Yu 2002	6mos	CDI	6	None	NA	In school, outside of school hours	N=220; pool of 1416 (355 invited); 1 middle school, 1 HS; China	Teachers affiliated with participating schools

		Efficacy Evidence				Participants and Intervention Delivery		
		Depressive Symptoms		Depression Diagnoses				
	Study Length	Measures	Months	Measure	Months	Intervention Delivery Site	Sample Size & Breadth of Recruitment	Intervention Providers
Roberts 2003	30mos	CDI	<i>ns</i>	None	NA	In school, during school day	N=189; from pool of 720 (208 invited); 18 rural primary schools	School psychologists & nurses trained by program developer or research staff
Gillham, Reivich, 2006	12mos	CDI	12	None	NA	After-school sessions in school	N=44; pool of 470; suburban middle schools	Trained research assistants & 1 program developer
Gillham, Hamilton, 2006	24mos	CDI	12 (S)	HMO records	<i>ns</i>	Primary care clinics	N=271; pool of 6,000 HMO members	Child mental health clinicians in HMO
Chaplin, 2006	12mos	CDI	Post	None	NA	After-school; middle schools	N=208; pool of 1500; 2 middle schools	School personnel and research assistants
Gillham, 2007	36mos	CDI & CDRS-R	30 (S) <i>ns</i>	None	NA	After-school; middle schools	N=697; pool of 4000 (710 eligible); 3 suburban schools	School personnel trained by research team; some groups led by research assistants
Gillham, 2012						After-school sessions; 5 middle schools	N=408; pool of 8000 (655 eligible); 5 suburban middle schools	School personnel trained by research team; some groups led by research assistants
<i>PRP-A</i>	6mos	CDI & RADS	Post <i>ns</i>	DISC-IV	<i>ns</i>			
<i>PRP-AP</i>	6mos	CDI & RADS	<i>ns ns</i>	DISC-IV	<i>ns</i>			
Wijnhoven, 2014	6mos	CDI & CES-D	6 6	None.	NA	After school; 3 secondary schools	N=102; 800 screened (118 eligible); 3 schools; Netherlands	Experienced group therapist
Kindt, 2014	12mos	CDI	<i>ns</i>	None	NA	Secondary schools; part of curriculum	N=1343; pool of 1440; 12 low-income schools; Netherlands	Teacher trained by members of research team
<i>Blues Group</i>								
Stice, 2007	6mos	BDI	1	None.	NA	Unclear	N=225; 2 high schools & one college	Clinical psych grad student; undergrads; trained by staff
Stice 2008	24mos	K-SADS BDI	12 6	K-SADS	24mos	Sessions held at schools	N=341; pool of 698 (366 eligible); 6 high schools	Clinical psych grad students & undergrads; trained by staff
Rohde 2014	6mos	K-SADS	Post	K-SADS	6mos	After-school sessions at school	N=378; pool of 8020 (388 eligible); 5 HSs	School personnel trained by program developers
Bibliotherapy: <i>Feeling Good</i> [comparison condition in same studies as <i>Blues Group</i>]								
Stice 2007	6mos	BDI	6	None	NA	Completed on their own	N=225; 2 HSs, 1 college	Self-administered

		Efficacy Evidence				Participants and Intervention Delivery		
		Depressive Symptoms		Depression Diagnoses				
	Study Length	Measures	Months	Measure	Months	Intervention Delivery Site	Sample Size & Breadth of Recruitment	Intervention Providers
Stice 2008	24mos	K-SADS BDI	6 <i>ns</i>	K-SADS	24mos	Completed on their own	N=341; pool of 698 (366 eligible); 6 HSs	Self-administered
Rohde 2014	6mos	K-SADS	<i>ns</i>	K-SADS	<i>ns</i>	Completed on their own	N=378; pool of 8020 (388 eligible); 5 HSs	Self-administered
Supportive-Expressive Group Intervention [comparison condition in same studies as <i>Blues Group</i>]								
Stice 2007	6mos	BDI	Post	None	NA	Unclear	N=225; 2 HSs, 1 college	Clinical grad student & undergrads; trained by staff
Stice 2008	24mos	K-SADS BDI	6 6	K-SADS	6mos	Sessions held at schools	N=341; pool of 698 (366 eligible); 6 HSs	Grad & undergrad students; trained by program developers
Resourceful Adolescent Program (RAP)								
Rivet-Duval 2011	6mos	RADS	Post	None	NA	Delivered in schools	N=160; 2 single-sex secondary schools in Mauritius	Research team member; certified trainer
Stallard 2012	12mos	SMFQ	12 (I)	None	NA	Delivered in school, part of curriculum	N=5030; pool of 5503; 8 schools (66 invited) representative of UK population	BA level in appropriate profession; experience with youth; teachers worked with certified facilitators
Rose 2014	Post ^b	RADS CDI	<i>ns</i> <i>ns</i>	DISCAP	NA ^c	Delivered in school; part of curriculum	N=210; 4 single-sex secondary schools, Sydney, Australia	Graduate students; trained by certified RAP trainer
LARS&LISA								
Pössel 2004	6mos	CES-D	3 (S)	None	NA	Delivered during school; part of curriculum	N=324; 6 middle schools in Tübingen, Germany	MS/ graduate level psychologists trained by program developer
Pössel 2008						Delivered during school; part of curriculum	N=301; 4 middle schools (out of 6 invited) Germany	MS/ graduate level psychologists trained by program developer
<i>Comorbid: Anxiety</i>	12mos	SBB-DES	Post (M) 6 (S)	None	NA			
<i>Externalizing</i>	12mos	SBB-DES	6 (S)	None	NA			
<i>Overall Sample</i>	12mos	SBB-DES	12 (S)	None.	NA			
Pössel 2013	12mos	CDI	4	None	NA	In school; during Wellness Class	N=518; HS in southern USA	MS or PhD psychologists, trained by program developer
Problem Solving For Life (PSL)								
Spence 2003	48mos	BDI	Post	ADIS-C	<i>ns</i>	Part of school curriculum	N=1500; 16 HSs; Brisbane, Australia	Teachers, completed 6 hours of training

		Efficacy Evidence				Participants and Intervention Delivery		
		Depressive Symptoms		Depression Diagnoses				
	Study Length	Measures	Months	Measure	Months	Intervention Delivery Site	Sample Size & Breadth of Recruitment	Intervention Providers
Sheffield 2006								
<i>Universal</i>	12mos	CDI CES-D	<i>ns</i> <i>ns</i>	ADIS-C	<i>ns</i>	Part of curriculum to whole class	N=2479; 34 schools Queensland and New South Wales, Australia; >50% consent	Teachers, completed 6 hours of training
<i>Indicated</i>	12mos	CDI CES-D	<i>ns</i> <i>ns</i>	ADIS-C	<i>ns</i>	Delivered during school hours		School counselors, health providers; 1 day training
<i>Universal + Indicated</i>	12mos	CDI CES-D	<i>ns</i> <i>ns</i>	ADIS-C	<i>ns</i>	Delivered during school hours		School counselors, health providers; 1 day training
Interpersonal Psychotherapy - Adolescent Skills Training (IPT-AST)								
Young 2006	12mos	CES-D	6	None	NA	After-school sessions in school	N=41; pool of 558 (41 eligible); 3 Catholic schools (most low income & Hispanic)	Program developer & MS psychologists or MSWs; trained by program developer
Horowitz 2007	6mos	CDI & CESD	Post <i>ns</i> Post (S)	None	NA	During school (Wellness classes)	N=380;pool of 600; 3 suburban/ rural schools	MS or PhD level clinicians trained by program developer
Young 2010	18mos	CES-D CDRS-R	Post Post	LIFE	6mos	After-school sessions in school	N=57; pool of 1117 (215 eligible); 3 single-sex schools (most Hispanic)	Program developer & MS or PhD level mental health professionals
FRIENDS								
Barrett 2001								
<i>Psychologists</i>	Post	CDI	<i>ns</i>	None	NA	Delivered in schools; part of curriculum	N=489; 10 schools (12 schools invited); 89% recruitment rate	Clinical psychologists
<i>Teacher Leaders</i>	Post	CDI	Post (I)	None	NA			Teachers completed a full-day training
Lowry-Webster 2001	12mos	CDI	12	ADIS-C	NA ^d	Part of the regular curriculum; given to whole class	N=594; 7 Catholic schools in Brisbane, Australia; 97.2% consent	Teachers completed a 2-day workshop by postgraduate psychology students
Lock 2003	24mos	CDI	12	ADIS-C	NA ^d	In schools; part of curriculum	N=977; 7 schools, economically diverse, Australia; 78% consent	MS level psychologists or doctoral candidates
Barrett 2005	12mos	CDI	<i>ns</i>	None	NA	In schools; part of curriculum	N=693; 7 schools, Australia; 79% consent	MS level psychologists or doctoral candidates

		Efficacy Evidence				Participants and Intervention Delivery		
		Depressive Symptoms		Depression Diagnoses				
	Study Length	Measures	Months	Measure	Months	Intervention Delivery Site	Sample Size & Breadth of Recruitment	Intervention Providers
Aussie Optimism Program (AOP)								
Rooney et al. (2006)	18 mos	CDI	Post	DICA	9	In schools; part of curriculum	N=136; 4 state primary schools; low income; Australia	BS level psychologists; 8 hr of training with program developers
Roberts et al. (2010)	18 mos	CDI	ns	None	NA	In schools; part of curriculum	N=496; 12 low-income primary schools; Australia; >85% consent	Teachers; 16 hr of training
Rooney et al. (2013)	54 mos	CDI	Post	DICA	ns	Delivered to whole classes during school hours	N=910; 22 low-income primary schools in West Australia; 89% consent	Teachers; 8 hr of training with program developers

Note. *ns* = non-significant effects; CES-D = Center for Epidemiological Studies-Depression Scale; KSADS = Schedule for Affective Disorders and Schizophrenia for School-Age Children; LIFE = Longitudinal Interval Follow-up Evaluation; HDRS = Hamilton Depression Rating Scale; CBCL-D = Children's Behavior Checklist depression items; CDI = Children's Depression Inventory; CDRS-R = Children's Depression Rating Scale-Revised; RADS = Reynold's Adolescent Depression Scale; DISC-IV = Diagnostic Interview Schedule for Children for DSM-IV; BDI = Beck Depression Inventory; SBB-DES = Self-Report Questionnaire-Depression; SMFQ = Short Mood & Feelings Questionnaire; DISCAP = Diagnostic Interview Schedule for Children, Adolescents, and Parents; ADIS-C = Anxiety Disorders Interview Schedule for Children. MS = Master's level; MSW = Masters in Social Work; HS = high school; mos = months; grad = graduate; (M) = significant main effect in the overall sample; (S) = effect only significant in a subgroup of the sample; (I) = intervention group's mean score was significantly worse than the control group's mean score;

^aRate of decrease in HDRS was significantly greater in the intervention group relative to control; group differences were not significant at any follow-ups

^bThis study compared two intervention conditions to control: RAP followed by a placebo program or RAP followed by the Peer Interpersonal Relatedness (PIR) program. Post assessment was after both intervention conditions completed the RAP program and prior to completing the placebo and PIR programs.

^cNumber of individuals with diagnoses was too small for meaningful formal analyses.

^dAnalyses evaluated effects on anxiety and depressive disorders combined, not on depression diagnoses specifically.

Table 2

Summary of Program Effects across Randomized Controlled Trials.

	Post Average Symptom Effect: g + [95% CI] k , Q	Follow-up ^a Avg Symptom Effect: g + [95% CI] k , Q	Longest Duration Effect on Symptoms, in months	Studies Showing Effect on Depression Diagnosis	Studies Showing Effects with Endogenous Providers	Effects Independent of Program Developers ^b
Penn Resiliency Program	-0.08 [-0.15, -0.01] 13, 21.25*	-0.19 [-0.27, -0.11] 12, 26.24**	30 (S) Gillham 2007	None	Chaplin 2006: (M); school personnel Gillham 2007: (S); school personnel Gillham 2012: (S) school personnel	Yes
Coping with Stress Course	-0.33 [-0.47, -0.20] 4, 0.47	-0.18 [-0.32, -0.04] 4, 4.14	33 (M) Beardslee 2013	Clarke 1995 (M) Clarke 2001 (M) Garber 2009 (M)	Clarke 1995: (M) school personnel Clarke 2001: (M) HMO clinicians	Yes
Blues Group	-0.45 [-0.63, -0.28] 3, 3.95	-0.21 [-0.38, -0.03] 3, 3.74	24 (M) Stice 2008	Stice 2008 (M) Rohde 2014 (M)	Rohde 2014: (M); teachers	No
Bibliotherapy <i>Feeling Good</i>	-0.18 [-0.36, -0.002] 3, 4.77	-0.25 [-0.43, -0.07] 3, 4.12	24 (M) Stice 2008	Stice 2008 (M)	NA: Providers not needed	No
Supportive- Expressive Group	-0.39 [-0.64, -0.14] 2, 7.35**	-0.36 [-0.62, -0.11] 2, 0.01	6 (M) Stice 2008	Stice 2008 (M)	None	No
Resourceful Adolescent Program	-0.05 [-0.25, 0.15] 2, 4.89*	0.12 [-0.004, 0.25] 3, 3.42	Post (M) Rivet-Duval 2011	None	None	Yes
LARS&LISA	0.01 [-0.13, 0.15] 3, 4.29	-0.10 [-0.25, 0.04] 3, 3.64	12 (S) Pössel 2008	None	None	No
Problem-Solving For Life	-0.19 [-0.28, -0.11] 2, 4.07*	0.03 [-0.06, 0.12] 2, 0.01	Post (M) Spence 2003	None	Spence 2003: (M); school personnel	No
IPT-AST	-0.49 [-0.71, -0.28] 3, 11.38**	-0.24 [-0.46, -0.01] 3, 9.12*	6 (M) Young 2006, 2010	Young 2010 (M)	None	No
FRIENDS for Life	-0.04 [-0.14, 0.05] 4, 12.61**	-0.24 [-0.34, -0.14] 3, 1.23	12 (M) Lowry-Webster 2001; Lock 2003	None	Lowry-Webster 2001: (M) teachers Lock 2003: (M)teachers	No
Aussie Optimism Program (AOP)	-0.09 [-0.19, 0.01] 3, 12.49*	-0.03 [-0.13, 0.08] 3, 3.88	Post (M) Rooney 2006; Rooney, 2013	Rooney 2006 (M)	Rooney 2013: (M) teachers	No

Note. g + = mean effect size (Hedges' g); k = number of studies contributing effect sizes; P = percentage of variability attributable to differences across studies; (M) = significant main effect in the overall sample; (S) = effect significant only in a subgroup of the sample

*
 $p < .05$,

**
 $p < .01$

^aReflects the weighted averaged effect across studies at the first follow-up assessment of 6 months or later (range 6-15 months).

^bThis column indicates if there have been any published randomized controlled trials in which the program was compared to an assessment-only, waitlist, or treatment-as-usual control condition and the program developers were not involved in carrying out the research.