Journal of Consulting and Clinical Psychology
2009, Vol. 77, No. 3, 486–503

0022-006X/09/$12.00 DOI: 10.1037/a0015168

A Meta-Analytic Review of Depression Prevention Programs for Children and Adolescents: Factors That Predict Magnitude of Intervention Effects

Eric Stice
Oregon Research Institute

Heather Shaw
University of Texas at Austin

Cara Bohon
University of Oregon

C. Nathan Marti
University of Texas at Austin

Paul Rohde
Oregon Research Institute

In this meta-analytic review, the authors summarized the effects of depression prevention programs for youth as well as investigated participant, intervention, provider, and research design features associated with larger effects. They identified 47 trials that evaluated 32 prevention programs, producing 60 intervention effect sizes. The average effect for depressive symptoms from pre-to-posttreatment ($r = .15$) and pretreatment to-follow-up ($r = .11$) were small, but 13 (41%) prevention programs produced significant reductions in depressive symptoms and 4 (13%) produced significant reductions in risk for future depressive disorder onset relative to control groups. Larger effects emerged for programs targeting high-risk individuals, samples with more females, samples with older adolescents, programs with a shorter duration and with homework assignments, and programs delivered by professional interventionists. Intervention content (e.g., a focus on problem-solving training or reducing negative cognitions) and design features (e.g., use of random assignment and structured interviews) were unrelated to effect sizes. Results suggest that depression prevention efforts produce a higher yield if they incorporate factors associated with larger intervention effects (e.g., selective programs with a shorter duration that include homework).

Keywords: depression prevention, adolescents, meta-analytic review

Major depression is one of the most common psychiatric problems faced by adolescents, is marked by a recurrent course and elevated psychiatric comorbidity, and increases risk for future suicide attempts, academic failure, interpersonal problems, unemployment, and legal problems (Klein, Torpey, Bufferd, & Dyson, 2008). Thus, numerous researchers have designed and evaluated depression prevention programs. Most prevention programs have targeted factors that have been found to increase risk for future onset of depression or increases in depressive symptoms that have emerged from prospective studies, including negative cognitions, infrequent pleasant activities, social skill deficits, and problem-solving skill deficits (e.g., Clarke et al., 1992; Hankin, Abramson, & Siler, 2001; Lewinsohn et al., 1994; Nolen-Hoeksema, Girgus, & Seligman, 1992; Warner, Weissman, Fendrich, Wickramaratne, & Moreau, 1992).

Although numerous trials of depression prevention programs have been conducted, the results of the findings have not been comprehensively reviewed and analyzed with meta-analytic procedures. In a recent meta-analytic review, Horowitz and Garber (2006) synthesized this literature and included effect sizes from 29 depression prevention programs from 29 trials. However, our review identified 60 effect sizes for 32 prevention programs evaluated in 47 trials. In addition, Horowitz and Garber (2006) examined only five effect size moderators; they did not investigate several potentially relevant moderators, including the content of the interventions and methodological features such as use of random assignment and structured diagnostic interviews. Further, they did not use multiple coders and test for intercoder agreement, which is usual practice for meta-analytic reviews (Cooper & Hedges, 1994), so it is unclear whether the moderators were reliably coded. More generally, it is important to test whether results from a meta-analytic review replicate when an independent research group abstracts information from studies, synthesizes this information, and tests for effect size moderators. Thus, our objective in the present review was to extend the Horowitz and Garber review by including 31 new effect sizes from 18 recently completed depression prevention trials, by investigating 15 potential

Eric Stice and Paul Rohde, Oregon Research Institute; Heather Shaw and C. Nathan Marti, Department of Psychology, University of Texas at Austin; Cara Bohon, Department of Psychology, University of Oregon.

Heather Shaw is now at the Oregon Research Institute.

Preparation of this article was supported by a research grant MH 67183 from the National Institutes of Health. We thank Jane Gillham for her insightful and thoughtful comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Eric Stice, Oregon Research Institute, 1715 Franklin Blvd., Eugene, Oregon, 97403. E-mail: estice@ori.org

moderators of program effectiveness, and by conducting a formal evaluation of interrater agreement for abstracted information.

Putative Moderators of Intervention Effects

Examining moderators that predict magnitude of prevention program effects may identify aspects of the participants, interventions, providers, and research design associated with stronger effects. This information should increase the yield of future prevention efforts by identification of the conditions under which optimal prevention effects occur and the subgroups of individuals for whom alternative depression prevention programs need to be developed. These analyses may also advance theories regarding effective routes to reduce risk for depressive episodes and enhance the methodological rigor of trials. Thus, we investigated several potential moderators of intervention effects that were selected on the basis of theory, prior findings, and past literature reviews.

Participant Features

Participant risk status. Meta-analytic reviews have found that prevention programs often produce significantly stronger effects when interventions are offered to high-risk participants (selective and indicated prevention programs) versus all individuals in a population (universal prevention programs) for various outcomes, including depression (Horowitz & Garber, 2006), eating pathology (Stice & Shaw, 2004), and obesity (Stice, Shaw, & Marti, 2006). In addition, prevention programs for depression (Clarke et al., 1995), anxiety (Lowry-Webster, Barrett, & Dadds, 2001), eating pathology (McVey, Tweed, & Blackmore, 2007), behavior problems (Stoolmiller, Eddy, & Reid, 2000), and substance abuse (Murphy et al., 2001) have produced stronger effects for high-risk subsamples than for the full sample of individuals enrolled in universal prevention programs. In the depression prevention field, selective and indicated programs have targeted various groups at high risk for major depression, including children and adolescents with elevated depressive symptoms, a pessimistic explanatory style, parental mood disorders, and family conflict. Theoretically, high-risk youth are more motivated to engage in the prevention program content and have a greater opportunity to show symptom reduction (Stice & Shaw, 2004). Thus, we hypothesized that intervention effects would be larger for selective and indicated versus universal programs. Because the key distinction between these types of programs is that the former are offered to high-risk individuals, we use the term *participant risk status* to refer to this moderator.

Participant gender. We hypothesized that the effects for depression prevention programs would be larger for female versus male youth on the basis of evidence that adolescent girls report greater depressive symptoms and higher rates of major depression than adolescent boys (Hankin et al., 1998; Lewinsohn et al., 1994), which would make it easier to demonstrate prevention effects for the former. However, prior trials that have tested whether gender-moderated intervention effects generated mixed findings: several trials found that intervention effects for depressive symptoms were significantly larger for girls than for boys (Gillham, Hamilton, Freres, Patton, & Gallop, 2006; Petersen, Leffert, Graham, Alwin, & Ding, 1997; Shatte & Seligman, 1997), but other trials found that gender was unrelated to effect sizes (Horowitz, Garber, Ciesla, Young, & Mufson, 2007; Jaycox, Reivich, Gilham, & Seligman, 1994; Lock & Barrett, 2003; Reivich, 1996).

Participant ethnicity. We hypothesized that depression prevention programs would produce larger effects for samples containing greater proportions of ethnic minority youth, as there is evidence that ethnic minority youth report more depressive symptoms than White youth (Cuffe, Waller, Cuccaro, & Pumariega, 1995; Roberts, Chen, & Solovitz, 1995; Siegel, Aneshensel, Taub, Cantwell, & Driscoll, 1998), which might suggest that prevention programs would produce larger effects for these high-risk subgroups. Alternatively, it is possible that prevention programs that were largely developed by European American researchers and evaluated with European American samples may be culturally incongruent with ethnic minority populations or may not adequately address the life circumstances faced by minority youth. Although no studies have tested whether ethnicity moderates the effects of depression prevention programs, cognitive–behavior therapy (CBT) programs have been found to be effective for Latino but not African American youth (Cardemil, Reivich, Beevers, Seligman, & James, 2007; Cardemil, Reivich, & Seligman, 2002).

Participant age. We theorized that children and early adolescent youth may find it more difficult to grasp the concepts and skills taught in the interventions than older adolescents (Stice & Shaw, 2004). Meta-analytic reviews have found support for this hypothesis for depression (Horowitz & Garber, 2006) and eating disorder prevention programs (Stice & Shaw, 2004). We hypothesized that depression prevention programs would produce larger effects for older youth.

Intervention Features

Program content. Intervention content should influence whether a program produces effects (Stice, Shaw, & Marti, 2007). Theoretically, interventions that seek to change established risk factors for a particular psychiatric disorder should be more effective than those that focus on other factors. On the basis of content of extant depression prevention programs, we coded interventions as focusing on (a) reducing negative cognitions (cognitive change content), (b) encouraging engagement in pleasant activities (behavioral activation content), (c) promoting problem-solving skills (problem-solving content), and (d) promoting social skill development (social skills content). Because etiologic studies have provided support for each of these content areas (e.g., Lewinsohn et al., 1994; Nolen-Hoeksema et al., 1992; Warner et al., 1992), we hypothesized that programs that included these content areas would produce larger effects.

Intervention duration. Meta-analyses of prevention programs for other problems revealed that longer interventions produced superior effects compared with very brief interventions (Rooney & Murray, 1996; Stice & Shaw, 2004). Theoretically, longer interventions afford a greater opportunity for presentation of information concerning attitudinal and behavioral change skills, allow participants to reflect on intervention material between sessions, and give participants more opportunities to practice new skills and then return to the group for trouble-shooting advice. However, extremely long programs may not appeal to youth, resulting in greater attrition and smaller intervention effects. Given that there were few very brief interventions but several that were very long,

we hypothesized that smaller effects would emerge for longer interventions.

Homework. Theoretically, prevention programs that include homework exercises relevant to the principles taught in the program should produce larger intervention effects than programs without homework. Clinicians have similarly posited that homework strengthens the impact of treatment for depression (Burns & Spangler, 2000). Thus, we hypothesized that prevention programs with homework would produce larger intervention effects than programs without.

Provider Features: Professional Interventionists

Researchers have suggested that prevention programs are more effective when delivered by dedicated professional interventionists versus classroom teachers (Baranowski, Cullen, Nicklas, Thompson, & Baranowski, 2002). Teachers are not able to devote as much time to providing interventions due to classroom responsibilities and typically receive less training and supervision relative to professional interventionists. Further, professional interventionists are often able to repeatedly deliver the intervention, allowing them to refine their presentation strategies. In support of this supposition, eating disorder prevention programs delivered by professional interventionists have been shown to produce larger effects than those provided by school staff (Stice, Shaw, & Marti, 2007). Thus, we hypothesized that intervention effects would be significantly larger for programs delivered by dedicated interventionists versus classroom teachers.

Design Features

Random assignment. Trials in which participants are randomly assigned to condition should produce larger intervention effects than trials in which alternative approaches are used to allocate participants to condition (e.g., matching) because it is the best approach to generating groups that are equivalent on potential confounds at baseline (with sufficiently large sample sizes), which should minimize the odds that any of these confounds are correlated with treatment condition and maximize the ability to detect intervention effects. Accordingly, we hypothesized that intervention effects may be greater for trials that used random assignment relative to other allocation approaches. However, because the proper analysis of intervention effects involves tests of differential change across conditions, which adjusts for any initial differences at baseline on the outcome, we suspected that this effect might not emerge. Indeed, random assignment did not emerge as a moderator of effects sizes in meta-analytic reviews of eating disorder (Stice & Shaw, 2004) or obesity (Stice et al., 2006) prevention programs.

Interview assessment. We hypothesized that depression prevention programs that were evaluated in trials using diagnostic interviews to assess depressive symptoms would produce larger intervention effects than programs that were evaluated in trials using self-report surveys. Evidence suggests that diagnostic interviews are more sensitive measures of depressive symptoms than are self-report surveys (Roberts, Lewinsohn, & Seeley, 1991), presumably because interviewers can clarify ambiguous questions and probe for details that clarify whether a particular experience reflects depression or some other circumstance (i.e., illness).

Publication status. Numerous meta-analytic reviews have documented a *file-drawer phenomena* (Cooper & Hedges, 1994), in which studies showing significant effects are more likely to be published than those that show nonsignificant effects. This is concerning because meta-analytic reviews that focus solely on published articles may misrepresent the true population effect size. Accordingly, we sought to include both published and unpublished studies and tested whether publication status was related to the magnitude of intervention effects.

Incorrect unit of analysis. In many prevention trials, the classrooms or schools are the unit of random assignment to condition, but the data are analyzed as if the individual was the unit of randomization. This practice increases the risk for a false-positive finding because it artificially reduces the error term and increases the between-condition effect. The degrees of freedom for the test statistics are also artificially inflated, and the assumption of independent errors is violated. Therefore, we tested the hypothesis that trials in which the unit of random assignment was not equivalent to the unit of analysis would produce larger intervention effects than trials in which the unit of randomization and analyses matched.

Follow-up duration. Effect sizes for prevention programs are typically strongest at posttest and become smaller at each subsequent follow-up assessment (Stice, Shaw, & Marti, 2007). Thus, we coded the length of follow-up so that we could test whether this factor moderated intervention effects at follow-up and controlled for this potential confound as necessary.

We were interested in additional moderators but were unable to include them for various reasons. We wanted to test whether effect sizes would be larger for programs that involved more extensive interventionist training and programs with higher session attendance and smaller for programs evaluated using blinded assessors, but reports did not contain sufficient detail for coding. Other moderators were not coded because they did not have sufficient variability, including whether (a) the intervention modality was individual or group (all were group), (b) the intervention had psychoeducational content (almost all included this content), (c) booster sessions were used (almost none used such sessions), (d) an intervention was interactive or didactic (almost all were interactive), and (e) the study outcome was assessed with validated measures (all included validated measures).

Method

Sample of Studies

Five procedures were used to retrieve published and unpublished trials of depression prevention programs. First, a computer search was performed on PsychInfo, MedLine, and Dissertation Abstracts databases for the years 1980–2008 with the following keywords: *depression*, *depressive*, *prevention*, *preventive*, and *intervention*. Two research assistants and a librarian performed independent searches. Eric Stice reviewed the products of all three searches to identify pertinent articles. Second, the tables of content for journals that commonly publish articles in this area were reviewed for this same period (e.g., *Journal of Clinical and Consulting Psychology*). Third, we consulted narrative reviews and prior meta-analytic reviews of the depression prevention field to search for additional citations. Fourth, the reference sections of all identified articles were examined. Finally, established depression prevention researchers were asked for copies of unpublished articles (under review or in press) describing prevention trials.

Inclusion and Exclusion Criteria

We focused exclusively on studies that included a continuous measure of depressive symptoms or conducted interviews assessing criteria for major depression. We also focused exclusively on trials that were conceptualized as depression prevention programs and did not include trials in which depressive symptoms were treated as a secondary outcome. If multiple reports of the same trial were published, we recorded effect sizes from all available follow-ups. We focused on effect sizes testing for differential change in depressive symptoms because only nine trials tested whether the prevention program reduced the risk for onset of depression disorder among intervention participants relative to control participants.

We included trials in which participants were randomly assigned to a depression prevention program or to an attention control condition, an assessment-only control condition, or a wait-list control condition. We also included trials in which some other relevant comparison group was used (e.g., matched controls) in a quasi-experimental design.

We focused exclusively on studies that tested whether the change in the outcomes over time was significantly greater in the intervention group versus the control group. This could take the form of a Time × Condition interaction in a repeated-measures analysis of variance (ANOVA) model, an analysis of covariance (ANCOVA) model that controlled for initial levels of the outcome variable, or growth curve model that controlled for initial levels of the outcome. We also included trials that used logistic regression or survival models to test whether the incidence of major depression onset was significantly lower in the intervention condition versus a control condition, provided initially depressed participants were excluded from the analyses.

We restricted our focus to trials that targeted children and adolescents because of our interest in determining whether effective interventions have been designed for this developmental period. We believe that depression prevention programs should be implemented before most individuals are expected to show onset of their first major depression episode. We used a broad view of adolescence and included trials with a mean age of participants up to age 22 because this captured college-based depression prevention programs. Many developmental psychologists consider adolescence to span from approximately age 12 through age 24 (Arnett, 2000).

Effect Size Estimation Procedures

We calculated effect sizes for tests of differential change in depressive symptoms across the intervention and control conditions. However, if only the effect size for differential risk for onset of major depression across the conditions was available, that was used as the effect size. The correlation coefficient (r) was used as the index of effect size because of its similar interpretation across different combinations of interval, ordinal, and nominal variables (Pearson's r, Spearman's rho, and point biserial; Rosenthal, 1991) and because this effect size preserved the valence of the effects. Cohen's (1988) criteria for small ($r = .10$), medium ($r = .30$), and large ($r = .50$) effects were used. If effect sizes were reported in Cohen's (1988) *d,* we converted them to r with the formula provided on Page 20 of Rosenthal (1991). If effects were reported as odds ratios (OR), they were converted to r with the formula provided on Page 194 of Lipsey and Wilson (2001). If no effect sizes were reported, we generated them directly by calculating Cohen's d with the means and standard deviations (from the control group at baseline) reported in the article, which we then converted to r using the Rosenthal formula, or we reconstituted the data using weighted probability values to estimate a chi-square test that provided an OR, which was then converted to r using the Lipsey and Wilson formula. If none of these options were possible, we estimated effect sizes from the exact p values reported by the authors using the formula provided on Page 19 of Rosenthal (1991). If exact p values were not reported, they were generated from the test statistics (e.g., F) and degrees of freedom using Microsoft Excel (2004; Microsoft Corp., Redmond, WA). If none of these options worked, we contacted the authors and requested effect sizes. Effect sizes reflect analyses performed on the entire samples used in these studies. With these methods, we calculated effect sizes for posttest and then for all available follow-up points for all trials (e.g., 6-month, 12-month, and 24-month follow-ups). We averaged the follow-up effect sizes that were available for each trial.

Operationalization and Coding of Effect Size Moderators

Table 1 lists the numeric values, the operationalization, and descriptive statistics of each of the moderators. There were four categories of moderators that were coded for this study: (a) *participant features*: risk status (selective or universal), gender (percentage of females), ethnicity (percentage of Whites), and mean age; (b) *intervention features*: intervention content (reducing negative cognitions, behavior activation, problem-solving skills training, social skills training), intervention duration (in hours), and whether the intervention included homework; (c) *provider features:* the type of facilitator (professional interventionist or endogenous provider, such as teacher, nurse, or school counselor), and (d) *design features*: whether participants or other units of analyses were randomly assigned to condition, whether the assessment method for the main outcome (depressive symptoms) was a diagnostic interview or self-report, whether the study was published in a peer-reviewed outlet, whether the unit of analysis correctly matched the unit of randomization, and the length of the follow-up (in months).

An iterative approach was taken to ensure reliable abstraction of moderators from the reports. First, Heather Shaw and Cara Bohon generated a coding system for the moderators on an a priori basis. Second, they coded a sample of 10 studies and then discussed and resolved all discrepancies, refining the coding system as necessary. Third, the remaining studies were then coded independently and reliability coefficients calculated. Finally, Heather Shaw and Cara Bohon held consensus meetings to resolve any remaining disagreements with regard to the coding of moderators. This final corrected data set was used for all analyses.

Results

Descriptive Statistics

The literature search identified 46 trials that met the inclusion criteria, in which 32 different depression prevention programs

Table 1
Operationalization and Descriptive Statistics for Moderators

Moderator	Values	Coding description and criteria	Descriptive statistics
Participant features			
Risk status of participants	1 = *Selected/indicated*; 0 = *Universal*	Dichotomous variable representing whether the study was universally implemented or whether study participants were selected/indicated because they were from a group at increased risk for depression or had elevated depressive symptoms.	*Selected/indicated* (*n*) = 35; *Universal* (*n*) = 25
Participant gender	Percentage of females	Continuous variable representing the percentage of the sample that was female.	*M* = 57.06, *SD* = 19.44
Participant ethnicity	Percentage of Whites	Continuous variable representing the percentage of the sample that was White.	*M* = 63.22, *SD* = 29.36
Participant age	Age in years	Continuous variable representing the mean age of the sample.	*M* = 14.02, *SD* = 2.90
Intervention features			
Cognitive change content	1 = *Yes*; 0 = *No*	Dichotomous variable representing whether the intervention included changing negative cognitions thought to lead to depression.	*Yes* (*n*) = 45; *No* (*n*) = *15*
Behavioral activation content	1 = *Yes*; 0 = *No*	Dichotomous variable representing whether the intervention included increased engagement in pleasant activities.	*Yes* (*n*) = 14; *No* (*n*) = 46
Problem-solving content	1 = *Yes*; 0 = *No*	Dichotomous variable representing whether the intervention included improving problem-solving abilities.	*Yes* (*n*) = 32; *No* (*n*) = 28
Social skills content	1 = *Yes*; 0 = *No*	Dichotomous variable representing whether the intervention included improving social skills.	*Yes* (*n*) = 29; *No* (*n*) = 31
Intervention duration	No. of hours	Continuous variable representing the number of intervention hours.	*M* = 12.84, *SD* = 6.82
Homework	1 = *Yes*; 0 = *No*	Dichotomous variable representing whether the intervention included homework or practice assignments.	*Yes* (*n*) = 41; *No* (*n*) = 18
Provider features: Professional interventionist	1 = *Yes*; 0 = *No*	Dichotomous variable representing whether facilitator was a professional interventionist or an endogenous provider, such as a teacher, school nurse, or counselor.	*Yes* (*n*) = 46; *No* (*n*) = 12
Design features			
Randomization	1 = *Yes*; 0 = *No*	Dichotomous variable representing whether participants were randomly assigned to intervention and control conditions.	*Yes* (*n*) = 52; *No* (*n*) = 6
Interview assessment	1 = *Yes*; 0 = *No*	Dichotomous variable representing whether diagnostic interviews were used to assess depression.	*Yes* (*n*) = 10; *No* (*n*) = 50
Publication status	1 = *Yes*; 0 = *No*	Dichotomous variable representing whether paper coded for meta-analysis was published in a peer-reviewed journal.	*Yes* (*n*) = 50; *No* (*n*) = 10
Incorrect unit of analysis	1 = *Yes (Incorrect)*; 0 = *No (Correct)*	Dichotomous variable representing whether the unit of randomization was an incorrect match for the unit of analysis.	*Yes* (*n*) = 7; *No* (*n*) = 53
Follow-up duration	Length of follow-up in months	Continuous variable representing the length of follow-up.	*M* = 11.91, *SD* = 11.50

were evaluated (11 trials evaluated more than 1 program, and 9 programs were evaluated in 2–8 trials), resulting in a total of 60 effect sizes. Table 2 lists prevention programs, describes the samples, characterizes the interventions evaluated, and summarizes the main findings. Of the 32 prevention programs evaluated in these trials, 13 programs (41%) produced significant reductions in depressive symptoms, and 4 (13%) produced significant reductions in risk for future depressive disorder relative to control groups in at least 1 trial. Of these 32 prevention programs, 11 were universal, 19 were selective or indicated, and 2 programs were evaluated in both universal and selective samples. The average age of participants ranged from 10 to 19 years. The majority focused on both males and females ($n = 25$), but 7 focused solely on females.

We calculated interrater agreement between the two moderator coders for all trials included in this review (see Table 3). We used kappa (κ) coefficients for nominal variables and interclass correlation coefficients (ICC) for continuous variables; raters were treated as a random effect (Shrout & Fleiss, 1979). The ICC coefficients ranged from .95 to 1.0. The κ coefficients ranged from .74 to 1.00. These analyses indicate that there was high interrater agreement. Again, following their independent coding, the two raters held a consensus meeting to resolve coding differences, and we used this consensus-corrected data set for all analyses. Table 4 reports the magnitude of effect sizes for universal and selective programs, respectively, and coding for potential moderators of intervention effects.

Average Effect Size and Effect Size Heterogeneity

A Statistical Analysis System (SAS Institute, Cary, NC) macro that computed inverse variance-weighted average effect sizes for

random effects models was used to compute all mean values (Lipsey & Wilson, 2001). For all means and random effects regression models reported herein, Pearson's r values were converted to z scores for analysis, as recommended by Hedges and Olkin (1985). The average posttest effect size across all studies (M r = .15) was significantly larger than zero (z = 4.96, p < .001). The r values for posttest effect sizes ranged from −.47 to .68. There was significant heterogeneity in effect sizes at posttest (Q = 528.76, p < .001), indicating variability across effect sizes. The average follow-up effect size across all studies (M r = .11) was significantly larger than zero (z = 6.40, p < .001). The r values for follow-up effect sizes ranged from −.18 to .76. There was also significant heterogeneity in effect sizes at follow-up (Q = 145.69, p < .001).

Relations of Moderators to Observed Effects Sizes

Moderator analyses were conducted using inverse variance-weighted random-effects regression models. Random-effects models separate the overall variability in observed effect sizes from the within-intervention variance. If studies are treated as a source of random variability, random effects models can be generalized to a broader set of studies or potential studies. Regression models with maximum likelihood estimation were conducted using a SAS macro written for meta-analysis (Lipsey & Wilson, 2001).

Moderators were examined individually in regression models to investigate the univariate relations between moderators and effect sizes. Although some meta-analyses have used multivariate approaches that test whether each moderator shows a unique relation to effect sizes statistically controlling for the other moderators (Perepletchikov, Treat, & Kazdin, 2007; Weisz, Han, Granger, & Morton, 1995), others have used univariate approaches (Cooper & Hedges, 1994; Horowitz & Garber, 2006; Stice et al., 2006). We chose the latter approach because many of the correlations between the moderators are logical (Table 5). For instance, intervention duration was positively correlated with problem-solving content and social skills content, which seems reasonable because it takes many session hours to cover these complex topics. Cognitive change content was correlated with use of homework, which would be expected given that a hallmark of CBT interventions is the use of homework. Participant age was correlated with intervention duration, which seems logical given that it would take more sessions to convey concepts and skills to children versus adolescents.

The four continuous moderators—percentage of females, percentage of Whites, average age, and intervention duration—were standardized in a z score format. We tested for linear and quadratic effects for the continuous moderators to decrease the risk of model misspecification (Hosmer & Lemeshow, 2000). In the event of a nonsignificant quadratic effect, the quadratic term was removed. We included average length of follow-up in models for follow-up effect sizes when this factor produced a significant effect. In the event of a significant effect for average length of follow-up, we tested the linearity assumption by including the Moderator × Average Length of Follow-Up interaction. If this interaction effect was significant, this interaction was retained in the model. To probe the form of significant linear effects, we calculated average intervention effects for studies above and below the median split. To probe the form of significant quadratic effects, we calculated the average intervention effects for the three tertiles of the moderator.

Results for all univariate models are presented in Table 6. All four participant features moderated the magnitude of intervention effects. Significantly larger effects were observed in selective trials involving high-risk participants versus universal trials. The average effect for studies involving high-risk participants was moderate and significantly different from zero (M r = .23, p < .001, n = 34), whereas the average effect for universally implemented programs was trivial and not significantly different from zero (M r = .04, p = *ns,* n = 25).[1] Risk status of participants was also a significant predictor of effect sizes from follow-up assessments: selective trials exhibited a moderate average effect size (M r = .14, p < .001, n = 28), but universally implemented programs exhibited a small average effect size (M r = .06, p < .001, n = 21), though both effects differed significantly from zero. The percentage of the participants who were female in the trials was significantly related to effects sizes.[2] At posttest, interventions below the median (≥ 53% females) exhibited a small nonsignificant average effect size (M r = .05, p = *ns*, n = 26), whereas the average effect for interventions at or above the median was moderate and significant (M r = .22, p < .001, n = 32). A similar effect was observed with effect sizes from follow-ups: interventions below the median exhibited a small average effect size that was significant (M r = .09, p < .001, n = 21) and interventions at or above the median showed larger effects (M r = .12, p < .001, n = 27). Percentage of White participants exhibited a quadratic effect at posttest. Probing this pattern with tertile splits revealed that effects were similar for the lowest tertile, which was less than 55% Whites (M r = .24, p < .001, n = 11), and the middle tertile, which was between 55% and 83% Whites (M r = .25, p < .001, n = 13), but effect sizes were trivial and nonsignificant for interventions containing greater than 83% White participants (M r = .04, p = *ns*, n = 11). Participant age was a significant predictor of effect size at posttest;

[1] We also compared selective versus indicated programs to ensure that it was reasonable to combine these two types of programs. There were no differences between selective and indicated programs at posttest (z = −.69, p = .49) or at follow-up (z = 1.60, p = .11).

[2] Horowitz and Garber (2006) found that the impact of participant gender on effect sizes for depression prevention programs became nonsignificant when college student samples were excluded from the analyses. This pattern of findings implies that participant age may interact with participant gender to predict prevention program effect size. We therefore conducted a direct test of this hypothesis. At posttest, the main effect for age (z = 4.19, p < .001) and the Age × Percentage of Females interaction (z = 2.61, p = .009) were significant, whereas the main effect for percentage of females was not (z = −0.43, p = .67). We probed this interaction by examining mean effect size above and below the median for age (13.5 years) and the median for percentage of female participants (53%). The mean r was 0.07 (p = .02, n = 20) in cases in which age and percentage of females were below their respective medians; the mean r was 0.01 (p = .89, n = 7) in cases in which age was below the age median and percentage of females was above the percentage of females median; the mean r was 0.04 (p = .50, n = 7) in cases in which age was above the age median and percentage of females was below the percentage of females median; and the mean was r = 0.31 (p < .001, n = 29) in cases in which age and percentage of females were above their respective medians. Thus, the largest effects are clearly associated with studies involving older samples that were predominantly female. The Age × Percentage of Females interaction was not significant when we examined follow-up effect sizes.

Table 2
Descriptions of the Sample, Intervention Content, and Findings From Depression Prevention Trials

Study	Sample	Intervention	Findings
Barrett et al., 2006	669 girls and boys	Efficacy trial of a universal school-based CBT intervention designed to prevent child anxiety by teaching children coping and problem-solving skills.	No significant effects for depressive symptoms (CDI) compared with an assessment-only control group.
Beardslee et al., 2003	121 girls and boys	Efficacy trial of selective psychoeducational intervention targeting children of depressed parents that presented information on mood disorders, risk, and resilience, and how to facilitate relationships.	No significant effects for depressive symptoms (SADS–L) at 1-, 2-, and 4.5-year follow-ups compared with an attention control group.
Bearman et al., 2003	74 girls	Efficacy trial of selective CBT intervention targeting adolescent girls with elevated body dissatisfaction.	Significant effects for depressive symptoms (BDI) at posttest but not 6-month follow-up compared with a waitlist control group.
Burton et al., 2007	145 young women	Efficacy trial of selective CBT intervention targeting women with elevated depressive symptoms.	Significant effects for depressive symptoms (BDI) at posttest, 3-month, and 6-month follow-ups compared with control group.
Cardemil et al., 2007	168 girls and boys	Two-year follow-up of efficacy trial of a universal school-based CBT intervention that taught cognitive and social problem-solving skills.	Significant effects for depressive symptoms (CDI) compared with an assessment-only control group.
Chaplin et al., 2006	208 girls and boys	Efficacy trial of a girls-only and co-ed version of a universal CBT and social and problem-solving intervention.	Significant effects for depressive symptoms (CDI) compared with an assessment-only control group for both girls-only or co-ed groups.
Clarke et al., 1993			
Study 1	513 girls and boys	Efficacy trial of universal school-based psychoeducational intervention that provided information on the symptoms, causes, and treatments for depression.	No significant effects for depressive symptoms (CES–D) compared with an assessment-only control group.
Study 2	300 girls and boys	Efficacy trial of universal school-based behavioral skills training intervention that encouraged participants to engage in pleasant activities.	No significant effects for depressive symptoms (CES–D) compared with an assessment-only control group.
Clarke et al., 1995	125 girls and boys	Efficacy trial of selective school-based cognitive intervention targeting children with elevated depressive symptoms that taught cognitive techniques to identify and challenge negative or irrational thoughts.	Significant effects for depressive symptoms (CES–D) at posttest compared with an assessment-only control group. Significantly reduced risk for depression onset for CBT group versus controls though 18-month follow-up.
Clarke et al., 2001	94 girls and boys	Efficacy trial of a shortened version of a selective cognitive treatment program targeting adolescents with a depressed parent in which participants were taught cognitive restructuring techniques.	Significant effects for depressive symptoms (CES–D) at posttest and 1-year follow-up compared with assessment-only control group. Significantly reduced risk for depression onset for CBT group versus controls though 1-year follow-up.
Forsyth, 2000	59 college women (97%) and men (3%)	Efficacy trial of a selective interpersonal therapy-based program targeting college undergraduates with both elevated depressive symptoms and at least one other risk factor for depression (e.g., negative life events or low social support). The intervention emphasized role transitions, interpersonal disputes, problem solving and social skills.	Significant effects for depressive symptoms (BDI) at posttest and 3-month follow-up compared with a waitlist control group.
Garber et al., 2008	316 girls and boys	Efficacy trial of CBT program for preventing depression in at-risk adolescents across four sites (replication of Clarke et al., 2001).	Significantly reduced risk for onset of depression in CBT group compared with assessment-only control group through 8-month follow-up.
Gillham, 1994	108 girls and boys	Efficacy trial comparing a child-only to a child-and-parent condition of a selective school-based CBT intervention.	Significant effects for depressive symptoms (CDI) for child-only version at posttest but not 6-month follow-up compared with an assessment-only control group; no effects for child–parent version compared with assessment-only control group.

(*table continues*)

Table 2 (*continued*)

Study	Sample	Intervention	Findings
Gillham & Reivich, 1999	118 girls and boys	Follow-ups 2.5 and 3 years after a selective intervention (Gillham et al., 1995) that taught cognitive and social problem-solving skills to children at risk for depression.	No significant effects for depressive symptoms (CDI) at 2.5- or 3-year follow-ups compared with assessment-only control group.
Gillham, Hamilton, et al., 2006	271 girls and boys	Effectiveness trial of selective school-based CBT intervention that focused on problem-solving and social skills training delivered by therapists in a primary care setting.	No significant effects for depressive symptoms (CDI) at posttest, 6-month, 1-year, 18-month, or 2-year follow-ups compared with an assessment-only control group.
Gillham, Reivich, et al., 2006	40 girls and boys	Pilot study examining selective efficacy of school-based CBT intervention that included problem-solving and social skills training when combined with a parent component.	Significant effects for depressive symptoms (CDI) at 6-month and 1-year follow-ups, but not at posttest, compared with an assessment-only control group.
Gillham et al., 2007	697 girls and boys	Efficacy trial of selective school-based CBT intervention that included problem-solving and social skills training.	No significant effects for depressive symptoms (CDI) at posttest, 6-month, 1-year, 18-month, 2-year, 2.5-year, or 3-year follow-ups compared with assessment-only control group.
Gwynn & Brantley, 1987	60 girls and boys	Study investigating the effects on depressive symptoms of a selective educational support group targeting children of divorce.	Significant effects for depressive symptoms (CDI) at posttest compared with an assessment-only control group.
Hains & Ellman, 1994	21 girls and boys	Efficacy trial of a universal stress inoculation training intervention that included cognitive coping skills and relaxation skills to reduce the incidence of negative emotional arousal.	No significant effects for depressive symptoms (RADS) at posttest compared with assessment-only control group.
Horowitz et al., 2007	380 girls and boys	Efficacy trial of a universal school-based CBT intervention and a school-based interpersonal therapy intervention.	Significant effects for depressive symptoms (CES–D) for both interventions compared with an assessment-only control group at posttest; no effects at 6-month follow-up.
Johnson, 2000	100 girls and boys	An efficacy trial of a universal intervention based on a social/interpersonal and cognitive behavioral model.	No significant effects for depressive symptoms (RCDS) compared with assessment-only control group through 1-year follow-up.
Kellam, et al., 1994	685 girls and boys	Universal intervention that compared an enriched curriculum aimed at improving reading achievement with a classroom behavior management strategy designed to reduce aggressive behavior.	No significant effects for depressive symptoms (CDI) at posttest compared with an assessment-only control group.
Lamb et al., 1998	222 girls and boys	Efficacy trial of selective coping and problem-solving skills intervention.	No significant effects for depressive symptoms (RADS) at posttest compared with assessment-only control group.
Lowry-Webster et al., 2003	584 girls and boys	One-year follow-up of universal effectiveness trial of a CBT-based intervention (Lowry-Webster et al., 2001).	No significant effects for depressive symptoms (CDI) at 1-year follow-up compared with an assessment-only control group.
Merry et al., 2004	392 girls and boys	Effectiveness trial of a universal school-based CBT and interpersonal therapy intervention delivered by teachers.	Significant effects for depressive symptoms (RADS) at posttest but not at 18-month follow-up compared with attention control group.
Miller, 1999	56 boys and girls	Selective efficacy trial of CBT intervention targeted to kids at a juvenile detention camp.	No significant effects for depressive symptoms (CDI) at posttest compared with assessment-only control group.
Pattison & Lynd-Stevenson, 2001	66 girls and boys	Effectiveness trial comparing universal school-based CBT-based intervention with an active control group that switched the order of topics; both programs delivered by community mental health providers.	No significant effects for depressive symptoms (CDI) at posttest or 6-month follow-up compared with active control and assessment-only control groups.

(*table continues*)

Table 2 (*continued*)

Study	Sample	Intervention	Findings
Peden et al., 2001	92 college women	Efficacy trial of a selective CBT-based intervention targeting women with elevated depressive symptoms that focused on the identification and reduction of negative thinking.	Significant effects for depressive symptoms (CDI) at posttest and 6-month follow-up, but not at 18-month follow-up, compared with assessment-only control group.
Peterson et al., 1997	237 girls and boys	Efficacy trial of a universal school-based psychoeducational intervention that taught adolescents adaptive emotional, cognitive, and behavioral responses to stressors or challenges.	Significant effects for depressive symptoms (DISC) at posttest compared with assessment-only control group.
Possel et al., 2004	324 girls and boys	Efficacy trial of a universal school-based CBT intervention focused on cognitive and social factors.	No significant effects for depressive symptoms (CES–D) at posttest, 3-month, or 6-month follow-up compared with assessment-only control group.
Quayle et al., 2001	47 girls	Efficacy trial of a universal school-based CBT intervention.	Significant effects for depressive symptoms (CDI) at 6-month follow-up, but not at posttest compared with assessment-only control group.
Roberts et al., 2003	189 girls and boys	Effectiveness trial of selective school-based version of a CBT intervention delivered by school staff.	No effects for depressive symptoms (CDI) at posttest or 6-month follow-up compared with assessment-only control group.
Roosa et al., 1989	81 girls and boys	Efficacy trial of a selective intervention that consisted of a school-based curriculum which taught information on alcoholism, self-esteem enhancement, and coping strategies to children from alcoholic families.	No significant effects for depressive symptoms (CDI) at posttest compared with assessment-only control group.
Sandler et al., 1992	72 girls and boys	Efficacy trial of selective intervention that consisted of a family grief workshop and a family advisor program targeting children who experienced the death of a parent.	No significant effects for depressive symptoms (CDI) at posttest compared with a waitlist control group.
Sawyer et al., 2008	5,634 girls and boys	Effectiveness trial of a universal school-based intervention that sought to improve problem-solving and social skills, resilient thinking style, and coping strategies.	No significant effects for depressive symptoms (CES–D) at posttest compared with assessment-only control group.
Seligman et al., 1999	231 college women and men	Efficacy trial of a selective CBT program targeting college students with negative attributional style.	Significant effects for depressive symptoms (BDI) at 1-, 2-, and 3-year follow-ups compared with an assessment-only control group.
Seligman et al., 2007	240 college women and men	Efficacy trial of a selective CBT program with ongoing Web-based materials and e-mail coaching.	No significant effects for depressive symptoms (BDI) or episodes (SCID) at posttest or 6-month follow-up compared with an assessment-only control group.
Shatte & Seligman, 1997	152 girls and boys	Efficacy trial comparing a selective school-based CBT-based intervention.	Significant effects for depressive symptoms compared with waitlist control group at 8-month, but not at 1-year follow-up or compared with active control group at either 8-month or 1-year follow-ups.
Sheffield et al., 2006	1,226 girls and boys	Effectiveness trial of CBT intervention in both universal and selective subsamples.	No effects for depressive symptoms (CES–D, CDI) at posttest, 3-month or 1-year follow-up compared with assessment-only control group for either universal or selective subsamples.
Shochet et al., 2001	228 girls and boys	Efficacy trial of a universal school-based CBT intervention with a focus on interpersonal and family risk and protective factors compared with a parent version of this intervention.	Significant effects for depressive symptoms (CDI) for both versions of the intervention compared with assessment-only control group at posttest, and significant effects for the child-only version at 1-year follow-up compared with the assessment-only control group.
Spence et al., 2005	751 girls and boys	Follow-ups at 2, 3, and 4 years to Spence et al. (2003) universal effectiveness trial of an intervention focused on cognitive restructuring and problem-solving skills.	No significant effects for depressive symptoms (BDI) at 2- or 3-year follow-ups (Note: This meta-analysis limited analysis of follow-up effects to 3 years) compared with assessment-only control group.
Stice, Burton, et al., 2007	225 young women and men	Efficacy trial comparing selective CBT program to active control groups and waitlist control condition.	Significant effects for depressive symptoms (BDI) for CBT compared with waitlist control group at posttest and 3-month follow-up.

(*table continues*)

Table 2 (*continued*)

Study	Sample	Intervention	Findings
Stice et al., 2008	341 girls and boys	Efficacy trial comparing brief selective CBT program to active control and assessment-only control groups.	Significant effects for depressive symptoms (K–SADS) for CBT compared with assessment-only control group at posttest and 6-month follow-up. Significantly reduced risk for depression onset for CBT versus controls through 6-month follow-up.
Stoppelbein, 2003	59 girls and boys	Universal efficacy trial of school-based CBT intervention.	No significant effects for depressive symptoms (CDI) at posttest, 3-month, and 6-month follow-ups compared with assessment-only control group.
Young, Mufson, & Davies, 2006	41 girls and boys	Efficacy trial of selective interpersonal psychotherapy skills training intervention.	Significant effects for depressive symptoms (CES–D) at posttest, 3-month, and 6-month follow-ups compared with an attention control group. Marginally significant reduced risk for depression onset for CBT versus controls though 6-month follow-up.
Yu & Seligman, 2002	110 girls and boys	Efficacy trial of selective school-based CBT-based intervention targeting Chinese adolescents with elevated depressive symptoms.	Significant effects for depressive symptoms (CDI) at posttest compared with assessment-only control group.

Note. Measures of depression used: CBT = Cognitive–Behavioral Therapy; CDI = Child Depression Inventory; SADS–L = Schedule for Affective Disorders and Schizophrenia—Lifetime Version; BDI = Beck Depression Inventory; RADS = Reynolds Adolescent Depression Scale; RCDS = Reynolds Child Depression Scale; DISC = Diagnostic Interview Schedule for Children; CES–D = Center for Epidemiologic Studies—Depression Scale; SCID = Structured Clinical Interview for *DSM–IV*; K–SADS = Kiddie—Schedule for Affective Disorder and Schizophrenia.

trials with participants below the median age of 13.5 years exhibited negligible effects ($M\ r = .02$, $p = ns$, $n = 26$), whereas those with participants above this median exhibited moderate effects ($M\ r = .23$, $p < .001$, $n = 29$). At follow-up, a quadratic relationship between age and effect size was observed. Tertile splits revealed that effects were similar for the lowest tertile, which was younger than 12.1 years of age ($M\ r = .08$, $p < .01$, $n = 14$) and the middle tertile, which was between 12.1 and 15.1 years of age ($M\ r = .07$, $p < .001$, $n = 16$), but interventions with participants whose average age exceeded 15.1 years exhibited larger effect sizes ($M\ r = .15$, $p < .001$, $n = 15$).

Among moderators reflecting intervention features, only intervention duration and homework were significant predictors of effect size; cognitive change, behavioral activation, problem solving, and social skills content were not. At posttest, interventions below the median duration (12 hr) exhibited larger average effect sizes ($M\ r = .19$, $p < .001$, $n = 23$) than interventions above the median ($M\ r = .07$, $p = ns$, $n = 29$). Use of homework assignments was associated with intervention effects at follow-up, where interventions with homework exhibited larger effects ($M\ r = .13$, $p < .001$, $n = 34$) than those without ($M\ r = .07$, $p < .001$, $n = 15$).

There were no differences in effect sizes for inventions conducted by professional interventionists versus those conducted by endogenous providers for posttest effect sizes, but differences did emerge for follow-up effect sizes. The average effect for trials involving professional interventionists was small and significant ($M\ r = .14$, $p < .001$, $n = 38$); the average effect for trials involving endogenous providers was trivial ($M\ r = .03$, $p < .05$, $n = 11$). Publication status exhibited a main effect, which differed significantly depending on the length of follow-up (i.e., publication status interacted with follow-up duration). Despite the fact that published studies ($M\ r = .09$, $p < .001$, $n = 42$) exhibited smaller average effect sizes than unpublished studies ($M\ r = .19$, $p = ns$, $n = 7$), the effect sizes of the published studies were significantly different than zero, whereas those of the unpublished studies were not, potentially due to an influential outlier. When the one unpublished study with an extremely large effect size (Forsyth, 2000) was excluded, this effect became nonsignificant. The moderators reflecting interview assessment, incorrect unit of analysis, and randomization did not predict effect size.

Table 3
Interrater Agreement for All Moderators Abstracted for the Present Meta-Analytic Review

Moderator	Interclass correlation	κ
Risk status of participants	—	1.00
Participant gender	.95	—
Participant ethnicity	1.00	—
Participant age	1.00	—
Cognitive change content	—	1.00
Behavioral activation content	—	1.00
Problem-solving content	—	.74
Social skills content	—	1.00
Intervention duration	.99	—
Homework	—	1.00
Professional interventionists	—	.90
Interview assessment	—	1.00
Incorrect unit of analysis	—	1.00
Random assignment to condition	—	1.00
Follow-up length	.96	—

Note. Dashes indicate not applicable.

Sensitivity Analyses

We included effect sizes for more than one depression prevention program from 8 of the 47 trials because these 8 trials

Table 4
Moderator Values and Effect Sizes for Depression Prevention Trials

	Participant				Content											Effect size	
Study	Risk status	Gender (% female)	Ethnicity (% White)	Age	Cognitive change	Behavioral activation	Problem solving	Social skills	Intervention duration (hr)	Home-work	Professional interventionist	Interview assessment	Published	Incorrect unit of analysis	Random	Post-test	Follow-up
Barrett et al., 2006	0	54	—	12.0	1	0	1	0	11.7	0	0	0	1	1	1	.05	.06
Beardslee et al., 2006	1	43	94	11.6	1	0	0	1	8.5	1	1	1	1	1	1	—	.00
Bearman et al., 2003	1	100	47	18.9	1	1	0	0	4	1	1	0	1	0	1	.29	.07
Burton et al., 2007	1	100	52	18.6	1	1	0	0	4	1	1	0	1	0	1	.58	.24
Cardemil et al., 2007	0	50	0	11.3	1	0	1	0	18	1	1	0	1	0	1	.27	.19
Chaplin et al., 2006																	
Girls only vs. control	0	100	89	12.2	1	0	1	0	18	1	1	0	1	0	1	−.47	—
Co-ed vs. control	0	44	89	12.2	1	0	1	0	18	1	1	0	1	0	1	−.29	—
Clarke et al., 1993																	
Study 1	0	42	—	15.4	0	0	0	0	2.5	0	0	0	1	0	1	.02	−.03
Study 2	0	46	—	15.1	0	1	0	0	4.2	0	0	0	1	0	1	.02	.05
Clarke et al., 1995	1	70	93	15.3	1	0	0	0	11.2	0	0	1	1	0	1	.18	−.02
Clarke et al., 2001	1	28	90	14.6	1	0	0	0	15	1	1	1	1	0	1	.22	.16
Forsyth, 2000	1	97	93	19.4	1	0	1	1		1	1	0	0	0	1	.68	.76
Garber et al., 2008	1	59	75	14.6	1	0	0	0	15	1	1	1	0	0	1	.14	—
Gillham, 1994, Study 2																	
Child & parent vs. control	0	47	—	14.6	1	0	1	1	16	1	1	0	0	0	1	0.03	−.07
Child only vs. control	0	47	—		1	0	1	1	24	1	1	0	0	0	1	0.25	.21
Gillham & Reivich, 1999	1	46	83		1	0	1	1	18	1	1	0	1	0	0	.18	.24
Gillham, Hamilton, et al., 2006	1	53	73	11.4	1	0	1	1	18	1	1	0	1	0	1	.02	.10
Gillham, Reivich, et al., 2006	1	30	91	11.5	1	1	1	1	12	1	1	0	1	0	1	−.06	−.18
Gillham et al., 2007																	
PRP vs. control	1	46	74	12	1	0	1	1	18	1	1	0	1	0	1	.04	.05
PEP vs. controls	1	46	74	12.1	0	0	0	1	18	1	1	0	1	0	1	.04	.04
Gwynn & Brantley, 1987	1	50	90	12.1	0	0	1	1		0		0	1	0	0	.21	—
Hains & Ellmann, 1994	0	76	90	16	1	0	1	0	11	0	1	0	1	0	1	.57	—
Horowitz et al., 2007																	
CBT vs. control	0	54	79	14.4	1	1	1	0	12	1	1	0	1	1	1	.19	.06
IPT vs. control	0	54	79	14.4	0	0	0	1	12	0	1	0	1	1	1	.16	.02
Johnson, 2000	0	11	—	11.7	1	1	0	1	6.7	1	1	0	0	0	0	.06	−.10
Kellam et al., 1994	0	51	19	6.3	0	0	0	0				0	1	1	1	0	—
Lamb et al., 1998	1	56	95	16.5	1	0	1	0		1	1	0	1	0	1	.28	—
Lowry-Webster et al., 2003	0	54	—	11.5	1	0	0	0	12.5	0	0	0	1	1	1	−.06	.02
Merry et al., 2004	0	52	5	14.2	1	0	0	0	11	0	0	0	1	0	1	.12	.05
Miller, 1999 (Modified PPP; Jaycox et al., 1994)	1	4	5	17.2	1	0	1	1	18	1	1	0	0	0	1	.07	—
Pattison & Lynd-Stevenson, 2001																	
PPP vs. attention control	0	52	—	10.4	1	0	1	1	20	1	1	0	1	0	1	−.06	.15
PPP vs. assessment control	0	52	—	10.4	1	0	1	1	20	1	1	0	1	0	1	.05	.31
Peden et al., 2001	1	100	—	19.3	1	0	0	0	6	1	1	0	1	0	1	.30	.23

(*table continues*)

Table 4 (*continued*)

Study	Participant				Content											Effect size	
	Risk status	Gender (% female)	Ethnicity (% White)	Age	Cognitive change	Behavioral activation	Problem solving	Social skills	Intervention duration (hr)	Home-work	Professional interventionist	Interview assessment	Published	Incorrect unit of analysis	Random	Post-test	Follow-up
Peterson et al., 1997	0	—	—	11.5	1	1	1	1	10.7	0	1	0	1	0	1	−.13	0
Possel et al., 2004	0	51	—	14	1	0	0	1	15	0	1	0	1	0	1	−.30	.18
Quayle et al., 2001	0	100	—	11.5	1	0	1	1	9.3	1	1	0	1	0	1	−.30	.32
Roberts et al., 2003	1	50	84	11.9	1	0	1	1	18	1	0	0	1	0	1	.02	.03
Roosa et al., 1989	1	50	30	10.3	0	0	1	1	36	1	1	0	1	0	1	.14	—
Sandler et al., 1992	1	49	82	12.4	0	1	1	0		0	1	1	1	0	1	−.02	—
Sawyer, 2008	0	53		13.1	1	0	1	1	19	1	0	0	0	1	1	.18	.07
Seligman et al., 1999	1	52	—	18	1	1	0	1	16	1	1	0	1	0	1	.18	.07
Seligman et al., 2007	1	65	—		1	1	0	1	16	1	1	1	1	0	1	.31	.28
Shatte & Seligman, 1997																	
POP vs. control	1	47	—	12.7	1	0	1	1	24	1	1	0	0	0	1	.09	.21
PEP vs. control	1	47	—	12.7	0	0	0	1	24	1	1	0	0	0	1	.21	.17
Sheffield et al., 2006																	
High-risk sample	0	69	—	14.3	1	0	1	1	12	1	0	0	1	0	1	.10	.02
Universal sample	0	54	—	14.3	1	0	1	0	6	1	0	0	1	0	1	.07	.05
Shochet et al., 2001																	
RAP–A vs. control	0	53	—	13.5	1	0	1	1	8.2	0	1	0	1	0	0	.22	.15
RAP–F vs. control	0	53	—	13.5	0	0	1	0	8.2	0	1	0	1	0	0	.21	.10
Spence et al., 2005	0	53	—	12.8	1	0	1	0	6.3	1	0	0	1	0	1	.19	−.01
Stice, Burton, et al., 2007																	
CBT vs. control	1	70	55	18.4	1	1	0	0	4	1	1	0	1	0	1	.48	.22
Sup/exp vs. control	1	70	55	18.4	0	0	0	0	4	0	1	0	1	0	1	.52	.19
Biblio vs. control	1	70	55	18.4	1	0	0	0		0	1	0	1	0	1	.37	.24
Exp writ vs. control	1	70	55	18.4	0	0	0	0	2.25	0	1	0	1	0	1	.40	.06
Journal vs. control	1	70	55	18.4	0	0	0	0		1	1	0	1	0	1	.29	.14
Stice et al., 2008																	
CBT vs. control	1	64	20	16.4	1	1	0	0	6	1	1	1	1	0	1	.25	.16
Sup/exp vs. control	1	64	20	16.4	0	0	0	0	6	0	1	1	1	0	1	.01	.04
Stoppelbein 2003	0	59	88		1	1	0	0	8.3	1	1	0	1	0	1	−.02	−.03
Young et al., 2006	1	85	—	13.4	0	0	1	1	12	1	1	1	1	0	1	.61	.45
Yu & Seligman, 2002	1	45	0	11.8	1	0	1	1	20	1	0	0	1	0	1	.12	.18

Note. Effect size for longest follow-up point is reported. PRP = Penn Resiliency Program; PEP = Penn Enhancement Program; CBT = Cognitive–Behavioral Therapy; IPP = Interpersonal Prevention Program; POP = Penn Optimism Program; RAP–A = Resourceful Adolescent Program—Adolescent; RAP–F = Resourceful Adolescent Program—Family; Sup/exp = Support/expressive group intervention; Biblio = Bibliotherapy; Exp writ = Expressive writing; Journal = Journaling. Dashes indicate not applicable.
$^{*}p = .05$. $^{**}p = .001$. $^{***}p = .001$.

Table 5
Correlations Among the Putative Moderators of Depression Prevention Intervention Effects

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Risk status of participants	—	.19	.13	.37*	−.10	.07	−.18	.07	.10	.28*	.26*	.38*	.08	−.32*	.05
2. Participant gender		—	.16	.41*	.03	.01	−.09	−.27*	−.30*	.00	.10	−.01	.23	−.10	.38*
3. Participant ethnicity			—	−.00	.10	.04	.14	.10	−.10	−.01	.23	.15	−.02	.06	.03
4. Participant age				—	.04	.23	−.43*	−.40*	−.62*	−.08	.15	.08	−.06	−.28*	.15
5. Cognitive change content					—	.14	.23	.10	.08	.41*	−.07	−.05	−.15	−.03	.07
6. Behavioral activation content						—	−.27*	−.14	−.37*	.11	.19	.18	.04	−.08	.06
7. Problem-solving content							—	.37*	.46*	.20	−.05	−.30*	−.06	−.08	−.20
8. Social skills content								—	.57*	.28*	.15	−.16	−.28*	−.04	−.25
9. Intervention duration									—	.38*	.13	−.15	−.29*	−.01	.04
10. Homework										—	.23	.00	−.30*	−.14	.14
11. Professional interventionists											—	.12	−.12	−.25	−.16
12. Interview assessment												—	.08	−.02	.16
13. Publication status													—	.02	.14
14. Incorrect unit of analysis														—	.13
15. Random assignment															—

Note. * $p < .05$.

evaluated more than one program. These effect sizes should be independent in that the effect of one depression prevention program is not dependent on the effect of the other depression prevention program(s) in the trial. However, because the same control group is used as the reference in calculating these effects, the effects may be partially dependent. Dependence across effect sizes may violate the assumption of independent errors and introduce bias in parameters estimates. To examine this possibility, we randomly selected one effect per study and replicated the models presented in Table 6. We compared regression coefficients from these randomly selected models with the confidence intervals presented in Table 6. In each case, the coefficients were within the confidence intervals, indicating that including multiple but orthogonal effects did not result in significantly biased parameter estimates for the relations of the moderators to the effect sizes.

Table 6
Univariate Effects for Moderators

	Posttest				Follow-up			
Moderator	*B*	95% CI *B*	β	Model R^2	*B*	95% CI *B*	β	Model R^2
Risk status of participants[a]	0.19**	0.07–0.31	0.39	.15	0.07*	0.01–0.13	0.26	.12
Participant gender	0.07*	0.01–0.13	0.30	.09	0.08***	0.04–0.12	0.49	.24
Participant ethnicity								
Linear	−0.14*	−0.26 to −0.02	−0.58	.20	−0.01	−0.07 to 0.05	−0.06	.00
Quadratic	−0.08*	−0.14 to −0.02	−0.54					
Participant age								
Linear	0.18***	0.12–0.24	0.64	.40	0.02	−0.02 to 0.06	0.16	.23
Quadratic					0.08***	0.04–0.12	0.41	
Cognitive change content[a]	−0.10	−0.26 to 0.06	−0.18	.03	0.05	−0.03 to 0.13	0.15	.07
Behavioral activation content[a]	0.09	−0.07 to 0.25	0.16	.03	−0.05	−0.13 to 0.03	−0.14	.07
Problem-solving content	−0.12	−0.24 to 0	−0.26	.07	0.03	−0.05 to 0.11	0.10	.01
Social skills content	−0.08	−0.22 to 0.06	−0.15	.02	0.04	−0.02 to 0.10	0.14	.02
Intervention duration	−0.09**	−0.15 to −0.03	−0.35	.12	0.03	−0.01 to 0.07	0.26	.07
Homework[a]	0.04	−0.1 to 0.18	0.07	.00	0.07*	0.01–0.13	0.25	.11
Professional interventionists[a]	0.09	−0.07 to 0.25	0.16	.03	0.10**	0.04–0.16	0.32	.10
Interview assessment	0.16	−0.04 to 0.36	0.22	.05	0.00	−0.10 to 0.010	−0.02	.05
Publication status[†]								
Main effect	−0.04	−0.24 to 0.16	−0.06	.00	−0.37**	−0.65 to −0.09	−0.81	.15
Interaction with follow-up length					0.03*	−0.01 to 0.07	3.19	
Incorrect unit of analysis	−0.07	−0.31 to 0.16	−0.08	.01	−0.08	−0.18 to 0.02	−0.19	.04
Random assignment	−0.004	−0.24 to 0.24	−0.005	.00	0.01	−0.13 to 0.15	0.02	.00
Follow-up length					−0.03	−0.07 to 0.01	−0.22	.05

Note. CI = confidence interval.
[a] Follow-up length included in follow-up model.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Discussion

Summary of Effect Sizes

Among the 32 prevention programs that were evaluated in 60 trials, 13 produced significant reductions in depressive symptoms. Twelve of the trials that produced significant effects found that intervention participants showed greater decreases in symptoms relative to decreases observed in controls, though one found that intervention participants showed a significant decrease in depressive symptoms, whereas controls showed a significant increase (Chaplin et al., 2006). The percentage of programs (41%) that produced effects was larger than the proportion of prevention programs that produced effects for other problems, including HIV (22%; Logan, Cole, & Leukefeld, 2002), eating disorders (29%; Stice, Shaw, & Marti, 2007), and obesity (21%; Stice et al., 2006), though smoking prevention programs have an even higher rate of significant effects (60%; Skara & Sussman, 2003). The average intervention effect size was an $r = .14$ at posttest and $r = .11$ at follow-up, which are small effects. The average posttest effect size for depression prevention programs compares favorably to the average posttest effect size observed for prevention programs for other problems, such as substance abuse ($r = .05$; Tobler et al., 2000), HIV ($r = .05$; Logan et al., 2002), smoking ($r = .07$; Hwang, Yeagley, & Petosa, 2004), eating disorders ($r = .13$; Stice, Shaw, & Marti, 2007), and obesity ($r = .04$; Stice et al., 2006). Of importance, four prevention programs significantly reduced risk for future onset of major depression (Clarke et al., 1995, 2001; Garber et al., 2008; Stice, Rohde, Seeley, & Gau, 2008; Young, Mufson, & Davies, 2006), though other trials found nonsignificant prophylactic effects (Gillham, Hamilton, et al., 2006; Seligman, Schulman, DeRubeis, & Hollon 1999; Seligman, Schulman, & Tryon, 2007; Sheffield et al., 2006).

Moderators of Effect Sizes from Depression Prevention Programs

Overall, 5 of the 15 moderators showed significant relations with effect size at posttest, and 6 showed significant relations with effect size from follow-up assessments. Selective programs offered to high-risk youth produced larger intervention effects than universal programs at both posttest and follow-up, replicating Horowitz and Garber (2006). It was noteworthy that the only programs that produced prophylactic effects were selective or indicated programs. These prophylactic effects are also important because they suggest that the intervention effects are not merely occurring because the programs decrease initial elevations in depressive symptoms, as suggested by Horowitz and Garber (2006). Of interest, several prevention programs were more effective for subgroups of high-risk participants than for the full universal sample (e.g., Clarke et al., 1995; Lowry-Webster et al., 2001). Theoretically, the distress that characterizes high-risk individuals motivates these participants to engage more effectively in the prevention program and the lower levels of depressive symptoms in universal samples attenuate intervention effects. These findings suggest that it may be prudent to focus on selective and indicated prevention programs and to discontinue evaluation of universal prevention programs.

Also as hypothesized, prevention programs were more effective when delivered to samples containing a higher portion of female participants at both posttest and follow-up, replicating the findings of Horowitz and Garber (2006). It is possible that the higher levels of depressive symptoms experienced by females relative to males (Hankin et al., 1998) renders the former more motivated to engage in the intervention, whereas the lower levels of depression for the latter group creates a floor effect. The fact that the impact of participant gender became significantly larger for late versus early adolescence, another novel finding, accords with this interpretation because the gender difference in depression becomes more pronounced during adolescence (Lewinsohn et al., 1994). It is also possible that depression prevention programs are more effective when delivered to groups that are solely composed of females, since some of the largest effect sizes emerged from trials in which this was the case (e.g., Burton, Stice, Bearman, & Rohde, 2007; Forsyth, 2000). Experience suggests that adolescent girls are more likely to discuss sensitive issues that influence their mood (e.g., body image concerns, sexual abuse) in female-only groups. A third interpretation is that current approaches to preventing depression are not well suited to males, potentially because of a limited understanding of the gender-specific risk factors for depression.

There was support for the hypothesis that prevention programs would be more effective for samples with more participants from ethnic minority groups, which is another novel finding. Theoretically, this is because minority youth are at greater risk for depression (Cuffe et al., 1995; Siegel et al., 1998). It is established the preventive effects are typically larger for higher risk samples (Horowitz & Garber, 2006; Stice & Shaw, 2004). These findings might suggest that it may not be necessary to create individually tailored prevention programs for various ethnic groups, yet it is still possible that even more effective prevention programs could be developed for high-risk minority youth.

Support also emerged for the hypothesis that prevention programs would produce larger effects for older adolescents relative to young adolescents and children at both posttest and follow-up, replicating Horowitz and Garber (2006). Theoretically, this effect emerged because the risk for depression increases during adolescence (e.g., Hankin et al., 1998). However, it is possible that older adolescents respond more favorably because they are better able to understand the concepts taught in the prevention programs, due to improved abstract reasoning. These data imply that it will be important to create prevention programs that are more effective for preadolescents and children.

Program content did not show a relation with effect sizes, which has not been tested previously. One interpretation is that these content areas are equally efficacious in preventing depression. Although it might be argued that nonspecific factors (e.g., perceived group support and contact with a caring interventionists) or expectancies account for the majority of the intervention effects, this does not seem to accord with the fact that 59% of the prevention programs evaluated did not reduce depressive symptoms, and 77% did not significantly reduce risk for onset of major depression.

Another novel finding was that relatively shorter prevention programs produced significantly larger intervention effects than did longer prevention programs. Horowitz and Garber (2006) did not observe this effect, possibly because of limited sensitivity due to the lower statistical power or unreliable coding of this moderator. Presumably, extremely long programs may not appeal to youth, which causes greater attrition and attenuated intervention

effects. These data suggest that future studies aimed at preventing depression should use briefer programs.

As hypothesized, prevention programs with homework assignments produced significantly larger effects than those without, which is another novel finding. This finding implies that it may be prudent to include homework exercises regularly in prevention programs, including those that are not primarily cognitive–behavioral. Theoretically, the increased opportunity to acquire intervention skills and apply them in the real world produces larger reductions in current and future depression.

An additional novel contribution is that results supported the hypothesis that prevention programs delivered by professional interventionists produce significantly stronger effect sizes than those delivered by endogenous providers (e.g., teachers), though this was only the case for follow-up effects. A similar finding emerged in a meta-analytic review of eating disorder prevention programs (Stice, Shaw, & Marti, 2007). This effect likely emerged because the professional interventionists have received more training and supervision, accumulated more experience with intervention delivery, and had fewer competing demands for their time. This finding seems to suggest that the importance of providing more detailed training and supervision to endogenous providers who deliver depression prevention programs.

It is noteworthy that none of the design factors were significantly related to the magnitude of the observed effect size, including use of random assignment to condition, use of diagnostic interviews (vs. questionnaires), incorrect unit of analysis, and length of follow-up. The effect sizes in Table 6 indicate that we had sufficient power to detect medium to large effect sizes at posttest but that we did not have sufficient power to detect small effects, particularly effect sizes at follow-up because fewer effect sizes were available. As such, it is conceivable that some null effects may be due to limited power to detect small effects.

Another novel contribution was that we tested whether publication status was correlated with effect sizes. However, publication status did not relate to effect size magnitude once one influential outlier was omitted.

Again, it is reassuring that our results replicated the evidence reported by Horowitz and Garber (2006) that intervention effects were significantly larger for high-risk participants, samples containing more females, and older adolescents. One exception was that although we found that intervention duration was related to effect sizes, Horowitz and Garber (2006) did not observe this effect, perhaps due to limited sensitivity. Our findings also extend the findings from that prior meta-analytic review in several ways. First, our meta-analysis of a larger body of literature revealed that prevention program effects are also moderated by participant ethnicity, intervention duration, use of homework assignments, and program delivery by professional interventionists. The finding that the effect of participant gender was moderated by participant age was also novel. Further, results suggested that program content (e.g., a focus on behavioral activation) and various methodological features of the study (e.g., use of randomization) were not systematically related to intervention effect sizes, which are also unique contributions to the literature as these questions have not been previously addressed.

Limitations

It is important to acknowledge the limitations of the present study. First, we had limited power to detect small effects for moderators because we only had 60 effect sizes. Second, a restriction in range for some of the moderators might have attenuated sensitivity further. These two considerations suggest that the null moderator effects should be interpreted with caution. Third, we were unable to code potentially important moderators, such as extent of training and supervision of facilitators, because insufficient information was provided. Fourth, because we estimated univariate rather than multivariate models, we were unable to investigate which moderators showed unique effects statistically controlling for the effects of the other moderators. Finally, few trials assessed other clinically important outcomes, such as social functioning and days of school missed, limiting our knowledge regarding effects for these outcomes.

Future Directions

The fact that most depression prevention programs produced small effects suggests that it will be important to conduct follow-up trials of enhanced versions of the programs that produced the largest effects and to design new programs that build upon those that worked well. It will also be important to replicate the effects of the most promising programs. Significant intervention effects have replicated across trials for the Coping with Stress Course (Clarke et al., 1995, 2001; Garber et al., 2008) and the Blues Program (Burton et al., 2007; Stice, Burton, Bearman, & Rohde, 2007; Stice et al., 2008). Effects have not replicated across trials of the Penn Prevention Program (Gillham, 1994; Gillham & Revich, 1999; Pattison & Lynd-Stevenson, 2001; Quayle, Dziurawiec, Roberts, Kane, & Ebsworthy, 2001; Roberts, Kane, Thomson, Bishop, & Hart, 2003) or the Penn Resiliency Program (Cardemil et al., 2007; Chaplin et al., 2006; Gillham, Reivich, et al., 2006, 2007).

The modest size of the average intervention effects also implies that it might be advantageous to focus on participant and intervention features that were associated with larger effects. For example, future trials might focus on high-risk youth and use professional interventionists. Nonetheless, future trials should also investigate alternative prevention programs that might be more effective for males, as extant programs appear to be somewhat less effective for this group. Unless efficacious prevention programs are developed for a broad array of individuals, it will be difficult for prevention efforts to reduce the prevalence of depression. Another priority for future research will be to focus on novel approaches to producing larger effects for depression prevention programs, such as monitoring risk status so that selective prevention programs can be delivered when most needed or conducting peer-led prevention programs.

We also believe that it would be useful for future research to experimentally manipulate key moderators of intervention effect sizes, in an effort to confirm the ostensive causal relations. For example, future studies could experimentally manipulate factors such as use of professional interventionists, use of homework, or intervention duration.

Future trials should use more rigorous designs. It would be particularly important to use blinded interviews to test whether

programs reduce the risk for onset of future depressive disorders, which has only been established for four prevention programs. In addition, future studies should use longer follow-up periods, so as to better characterize the persistence of intervention effects. It would also be beneficial to employ active control groups, rather than the assessment-only or waitlist control conditions that are commonly used, to establish the role of nonspecific factors in intervention effects.

It would also be useful to test whether the hypothesized mediators actually account for the effects of depression prevention programs, such as changes in negative cognitions, engagement in pleasant activities, or improved social skills. If the intervention produces change in putative mediators but no depression prevention effects, or if the intervention produces effects for depression but no change in the mediators, this signals that the intervention model may be incorrect or that certain measures are unreliable or invalid. An improved understanding of these processes may aid in the refinement of prevention programs.

Another important direction for future research will be to conduct effectiveness trials that test whether interventions that have produced promising effects within highly controlled efficacy trials continue to do so when endogenous providers are responsible for recruitment, screening, and intervention delivery. There have only been a handful of effectiveness trials (e.g., Gillham, Hamilton, et al., 2006; Yu & Seligman, 2002). It would also be useful to initiate studies on methods for disseminating and implementing effective depression prevention programs that produce effects in efficacy and effectiveness trials. Continued application of rigorous and programmatic research should bring us closer to reducing the incidence of this pernicious mental health problem.

References

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

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist, 55,* 469–480.

Baranowski, T., Cullen, K., Nicklas, T., Thompson, D., & Baranowski, J. (2002). School-based obesity prevention: A blueprint for taming the epidemic. *American Journal of Public Health, 26,* 486–493.

*Barrett. P., Farrell, L., Ollendick, T., & Dadds, M. (2006). Long-term outcomes of an Australian universal prevention trial of anxiety and depression symptoms in children and youth: An evaluation of the friends program. *Journal of Clinical Child and Adolescent Psychology, 35,* 403–411.

*Beardslee, W., Gladstone, T., Wright, E., & Cooper, A. (2003). A family-based approach to the prevention of depressive symptoms in children at risk: Evidence of parental and child change. *Pediatrics, 112,* 119–131.

*Bearman, S. K., Stice, E., & Chase, A. (2003). Evaluation of an intervention targeting both depressive and bulimic pathology: A randomized prevention trial. *Behavior Therapy, 34,* 277–293.

Burns, D. D., & Spangler, D. L. (2000). Does psychotherapy homework lead to improvements in depression in cognitive-behavioral therapy or does improvement lead to increased homework compliance? *Journal of Consulting and Clinical Psychology, 68,* 46–56.

*Burton, E. M., Stice, E., Bearman, S. K., & Rohde, P. (2007). An experimental test of the affect-regulation model of bulimic symptoms and substance use: An affective intervention. *International Journal of Eating Disorders, 40,* 27–36.

*Cardemil, E. V., Reivich, K. J., Beevers, C. G., Seligman. M. E., & James, J. (2007). The prevention of depressive symptoms in low-income, minority children: Two-year follow-up. *Behavior Research and Therapy, 45,* 313–327.

Cardemil, E. V., Reivich, K. J., & Seligman, M. E. (2002). The prevention of depressive symptoms in low-income minority middle school students. *Prevention and Treatment, 5,* Article 8.

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

*Clarke, G. N., Hawkins, W., Murphy, M., & Sheeber, L. (1993). School-based primary prevention of depressive symptomatology in adolescents: Findings from two studies. *Journal of Adolescent Research, 8,* 183–204.

*Clarke, G. N., Hawkins, W., Murphy, M., Sheeber, L., Lewinsohn, P. M., & Seeley, J. R. (1995). 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 and Adolescent Psychiatry, 34,* 312–321.

*Clarke, G. N., Hornbrook, M., Lynch, F., Polen, M., Gale, J., Beardslee, W., et al. (2001). A randomized trial of a group cognitive intervention for preventing depression in adolescent offspring of depressed parents. *Archives of General Psychiatry, 58,* 1127–1134.

Clarke, G. N., Lewinsohn, P. M., Hops, H., Andrews, J. A., Seeley, J. R., & Williams, J. A. (1992). Cognitive–behavioral group treatment of adolescent depression: Prediction of change. *Behavior Therapy, 23,* 341–354.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis*. New York; Russell Sage Foundation.

Cuffe, S. P., Waller, J. L., Cuccaro, M. L., & Pumariega, A. J. (1995). Race and gender differences in the treatment of psychiatric disorders in young adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 1536–1543.

*Forsyth, K. M. (2000). The design and implementation of a depression prevention program. *Dissertation Abstracts International, 61*(12-B), 6704.

*Garber, J., Gladstone, T., Weersing, V., Clarke, G., Brent, D., Beardslee, W., et al. (2008). *The prevention of depression in at-risk adolescents: Rationale, design, and preliminary results*. Presented at the annual meeting of the Society for Prevention Research, San Francisco, California.

*Gillham, J. E. (1994). *Preventing depression symptoms in school children*. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

*Gillham, J. E., Hamilton, J., Freres, D. R., Patton, K., & Gallop, R. (2006). Preventing depression among early adolescents in the primary care setting: A randomized controlled study of the Penn Resiliency Program. *Journal of Abnormal Child Psychology, 34,* 203–219.

*Gillham, J. E., & Reivich, K. J. (1999). Prevention of depressive symptoms in school children: A research update. *American Psychological Society, 10,* 461–462.

*Gillham, J. E., Reivich, K. J., Freres, D., Lascher, M., Litzinger, S., Shatte, A., & Seligman, M. E. (2006). School-based prevention of depression and anxiety symptoms in early adolescence: A pilot of a parent intervention component. *School Psychology Quarterly, 21,* 323–348.

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

Gillham, J. E., Reivich, K. Jaycox, L., & Seligman, M. E. (1995). Preven-

tion of depressive symptoms in schoolchildren: Two-year follow-up. *Psychological Science, 6,* 343–351.

*Gwynn, C. A., & Brantley, H. T. (1987). Effects of a divorce group intervention for elementary school children. *Psychology in the Schools, 24,* 161–164.

*Hains, A. A., & Ellmann, S. W. (1994). Stress inoculation training as a preventive intervention for high school youth. *Journal of Cognitive Psychotherapy, 8, 219*–232.

Hankin, B. L., Abramson, L. Y., Moffitt, T. E., Silva, P. A., McGee, R., & Angell, K. E. (1998). Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study. *Journal of Abnormal Psychology, 107,* 128–140.

Hankin, B. L., Abramson, L. Y., & Siler, M. (2001). A prospective test of the hopelessness theory of depression in adolescence. *Cognitive Therapy and Research, 25,* 607–632.

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

Horowitz, J. L., & Garber, J. (2006). The prevention of depressive symptoms in children and adolescents: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 74,* 401–415.

*Horowitz, J. L., Garber, J., Ciesla, J. A., Young, J., & Mufson, L. (2007). Prevention of depressive symptoms in adolescents: A randomized trial of cognitive–behavioral and interpersonal prevention programs. *Journal of Consulting and Clinical Psychology, 75,* 693–706.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (2nd ed.). New York: Wiley.

Hwang, M. S., Yeagley, K. L., & Petosa, R. (2004). A meta-analysis of adolescent psychosocial smoking prevention programs published between 1978 and 1997 in the United States. *Health Education and Behavior, 31,* 702–719.

Jaycox, L. H., Reivich, K. J., Gillham, J., & Seligman, M. E. (1994). Prevention of depressive symptoms in school children. *Behavior Research and Therapy, 32,* 801–816.

*Johnson, N. C. (2000). *A follow-up study of a primary prevention program targeting childhood depression*. Unpublished doctoral dissertation, Utah State University, Logan.

*Kellam, S., Rebok, G., Mayer, L., Ialongo, N., & Kalodner, C. (1994). Depressive symptoms over first grade and their response to a developmental epidemiological based preventive trial aimed at improving achievement. *Development and Psychopathology, 6,* 463–481.

Klein, D. N., Torpey, D. C., Bufferd, S. J., & Dyson, M. W. (2008). Depressive disorders. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *Child and Adolescent Psychopathology* (pp. 477–509). Hoboken, NJ: Wiley.

*Lamb, J. M., Puskar, K. R., Sereika, S. M., & Corcoran, M. (1998). School-based intervention to promote coping in rural teens. *American Journal of Maternal and Child Nursing, 23,* 187–194.

Lewinsohn, P. M., Roberts, R. E., Seeley, J. R., Rohde, P., Gotlib, I. H., & Hops, H. (1994). Adolescent psychopathology: II. Psychosocial risk factors for depression. *Journal of Abnormal Psychology, 103,* 302–315.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Newbury Park, CA: Sage.

Lock, S., & Barrett, P. M. (2003). A longitudinal study of developmental differences in universal preventive intervention for child anxiety. *Behaviour Change, 20,* 1183–1199.

Logan, T. K., Cole, J., & Leukefeld, C. (2002). Women, sex, and HIV: Social and contextual factors, meta-analysis of published interventions, and implications for practice and research. *Psychological Bulletin, 128,* 851–885.

*Lowry-Webster, H., Barrett, P., & Lock, S. (2003). A universal prevention trial of anxiety symptomatology during childhood: Results at one-year follow-up. *Behaviour Change, 20,* 25–43.

Lowry-Webster, H. M., Barrett, P. M., & Dadds, M. R. (2001). A universal prevention trial of anxiety and depressive symptomatology in childhood: Preliminary data from an Australian study. *Behaviour Change, 18,* 36–50.

McVey, G., Tweed, S., & Blackmore, E. (2007). Healthy schools–healthy kids: A controlled evaluation of a comprehensive eating disorder prevention program. *Body Image, 4,* 115–136.

*Merry, S., McDowell, H., Wild, C., Bir, J., & Cunliffe, R. (2004). A randomized placebo-controlled trial of a school based depression prevention program. *Journal of the American Academy of Child and Adolescent Psychiatry, 43,* 538–547.

Microsoft Corporation (2004). Microsoft Excel for Mac (Version 11). [Computer software]. Redmond, WA: Author.

*Miller, J. B. (1999). The effect of a cognitive-behavioral group intervention on depressive symptoms in an incarcerated adolescent delinquent population (juvenile delinquents). Unpublished doctoral dissertation, Wright Institute Graduate School of Psychology, Berkeley.

Murphy, J. G., Duchnick, J. J., Vuchinich, R. E., Davison, J. W., Karg, R. S., Olson, A. M., et al., (2001). Relative efficacy of a brief motivational intervention for college student drinkers. *Psychology of Addictive Behaviors, 15,* 373–379.

Nolen-Hoeksema, S., Girgus, J. S., & Seligman, M. E. (1992). Predictors and consequences of childhood depressive symptoms: A 5-year longitudinal study. *Journal of Abnormal Psychology, 101,* 405–422.

*Pattison, C., & Lynd-Stevenson, R. M. (2001). The prevention of depressive symptoms in children: The immediate and long-term outcomes of a school based program. *Behaviour Change, 18,* 92–102.

*Peden, A., Rayens, M., Hall, L., & Beebe, L. (2001). Preventing depression in high-risk college women: A report of an 18-month follow-up. *Journal of American College Health, 49,* 299–306.

Perepletchikova, F., Treat, T., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology, 75,* 829–841.

*Petersen, A. C., Leffert, N., Graham, B., Alwin, J., & Ding, S. (1997). Promoting mental healthy during the transition into adolescence. In J. Schulenberg, J. L. Maggs, & A. K. Hierrelmann (Eds.), *Health risks and developmental transitions during adolescence* (pp. 471–497). New York: Cambridge University Press.

*Possel, P., Horn, A., Groen, G., & Hautzinger, M. (2004). School-based prevention of depressive symptoms in adolescents: A six-month follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry, 43,* 1003–1010.

*Quayle, D., Dziurawiec, S., Roberts, C., Kane, R., & Ebsworthy, G. (2001). The effect of an optimism and lifeskills program on depressive symptoms in preadolescence. *Behaviour Change, 18,* 194–203.

Reivich, K. J. (1996). *The prevention of depressive symptoms in adolescents.* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

*Roberts, C. Kane, R., Thomson, H., Bishop, B., & Hart, B. (2003). The prevention of depressive symptoms in rural school children: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 71,* 622–628.

Roberts, R. E., Chen, Y. W., & Solovitz, B. L. (1995). Symptoms of *DSM–III–R* major depression among Anglo, African, and Mexican American adolescents. *Journal of Affective Disorders, 36,* 1–9.

Roberts, R. E., Lewinsohn, P. M., & Seeley, J. R. (1991). Screening for adolescent depression: A comparison of depression scales. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 58–66.

Rooney, B. L., & Murray, D. M. (1996). A meta-analysis of smoking prevention programs after adjustment for errors in the unit of analysis. *Health Education Quarterly, 23,* 48–64.

*Roosa, M., Gensheimer, L., Short, J., Ayers, T., & Shell, R. (1989). A preventive intervention for children in alcoholic families: Results of a pilot study. *Family Relations, 38,* 295–300.

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* Thousand Oaks, CA: Sage.

*Sandler, I. N., West, S. G., Baca, L., Pillow, D. R., Gersten, J. C., Rogosch, F., et al., (1992). Linking empirically based theory and evaluation: The family bereavement program. *American Journal of Community Psychology, 20,* 491–521.

*Sawyer, M., Pfeiffer, S., Spence, S., Bond, L., Graetz, B., Kay, D., et al., (2008). *School-based prevention of depression: A randomized controlled study of the BeyondBlue School Research Initiative*. Manuscript submitted for publication.

*Seligman, M. E., Schulman, P., DeRubeis, R. J., & Hollon, S. D. (1999). The prevention of depression and anxiety. *Prevention and Treatment, 2,* Article 8.

*Seligman, M. E., Schulman, P., & Tryon, A. M. (2007). Group prevention of depression and anxiety symptoms. *Behavior Research and Therapy, 45,* 1111–1126.

*Shatte, A., & Seligman, M. (1997). Prevention of depressive symptoms in adolescents: Issues of dissemination and mechanisms of change. *Dissertation Abstracts International: 57*(11-B), 7236.

*Sheffield, J., Spence, S., Rapee, R., Kowalenko, N., Wignall, A., Davis, A., & McLoone, J. (2006). Evaluation of universal, indicated, and combined cognitive-behavioral approaches to the prevention of depression among adolescents. *Journal of Consulting and Clinical Psychology, 74,* 66–79.

*Shochet, I. M., Dadds, M. R., Holland, D., Whitefield, K., Harnett, P. H., & Osgarby, S. M. (2001). The efficacy of a universal school-based program to prevent adolescent depression. *Journal of Clinical Child Psychology, 30,* 303–315.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–427.

Siegel, J. M., Aneshensel, C. S., Taub, B., Cantwell, D. P., & Driscoll, A. K. (1998). Adolescent depressed mood in a multiethnic sample. *Journal of Youth and Adolescence, 27,* 413–427.

Skara, S., & Sussman, S. (2003). A review of 25 long-term adolescent tobacco and other drug use prevention program evaluations. *Preventive Medicine, 37,* 451–474.

Spence, S. H., Sheffield, J. K., & Donovan, C. L. (2003). Preventing adolescent depression: An evaluation of the problem solving for life program. *Journal of Consulting and Clinical Psychology, 71,* 3–13.

*Spence, S. H., Sheffield, J. K., & Donovan, C. L. (2005). Long-term outcome of a school-based, universal approach to prevention of depression in adolescents. *Journal of Consulting and Clinical Psychology, 73,* 160–167.

*Stice, E., Burton, E., Bearman, S. K., & Rohde, P. (2007). Randomized trial of a brief depression prevention program: An elusive search for a psychosocial placebo control condition. *Behaviour Research and Therapy, 45,* 863–876.

*Stice, E., Rohde, P., Seeley, J., & Gau, J. (2008). Brief cognitive-behavioral depression prevention program for high-risk adolescents outperforms two alternative interventions: A randomized efficacy trial. *Journal of Consulting and Clinical Psychology, 76,* 595–606.

Stice, E., & Shaw, H. (2004). Eating disorder prevention programs: A meta-analytic review. *Psychological Bulletin, 130,* 206–227.

Stice, E., Shaw, H., & Marti, C. N. (2006). A meta-analytic review of obesity prevention programs for children and adolescents: The skinny on interventions that work. *Psychological Bulletin, 132,* 667–691.

Stice, E., Shaw, H., & Marti, C. N. (2007). A meta-analytic review of eating disorder prevention programs: Encouraging Findings. *Annual Review of Clinical Psychology, 3,* 233–257.

Stoolmiller, M., Eddy, J. M., & Reid, J. B. (2000). Detecting and describing preventive intervention effects in a universal school-based randomized trial targeting delinquent and violent behavior. *Journal of Consulting and Clinical Psychology, 68,* 296–306.

*Stoppelbein, L. (2003). Primary prevention: An evaluation of a high-school based cognitive–behavioral program. *Dissertation Abstracts International, 64*(8-B), 4066.

Tobler, N. S., Roona, M. R., Ochshorn, P., Marshall, D. G., Streke, A. V., & Stackpole, K. M. (2000). School-based adolescent drug prevention programs: 1998 meta-analysis. *Journal of Primary Prevention, 20,* 275–336.

Warner, V., Weissman, M. M., Fendrich, M., Wickramaratne, P., & Moreau, D. (1992). The course of major depression in the offspring of depressed parents: Incidence, Recurrence, and recovery. *Archives of General Psychiatry, 49,* 795–801.

Weisz, J. R., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment and outcome studies. *Psychological Bulletin, 117,* 450–468.

*Young, J. F., Mufson, L., & Davies, M. (2006). Efficacy of interpersonal psychotherapy-adolescent skills training: An indicated preventive intervention for depression. *Journal of Child Psychology and Psychiatry, 47,* 1254–1262.

*Yu, D. L., & Seligman, M. E. (2002, May 8.) Preventing depressive symptoms in Chinese children. *Prevention & Treatment, 5,* Article 9.

Received April 21, 2008
Revision received January 5, 2009
Accepted January 8, 2009 ■