Published in final edited form as:

Annu Rev Clin Psychol. 2014 March 28; 10: 243–273. doi:10.1146/annurev-clinpsy-050212-185524.

Overview of Meta-Analyses of the Prevention of Mental Health, Substance Use and Conduct Problems

Irwin Sandler¹, Sharlene A. Wolchik¹, Gracelyn Cruden^{2,4}, Nicole E. Mahrer^{1,4}, Soyeon Ahn³, Ahnalee Brincks², and C. Hendricks Brown²

¹Department of Psychology and Prevention Research Center at Arizona State University

²Feinberg School of Medicine, Northwestern University in Chicago, IL

³College of Education, University of Miami, Coral Gables, FL

Abstract

This paper presents findings from an overview of meta-analyses of the effects of prevention and promotion programs to prevent mental health, substance use and conduct problems. The review of 48 meta-analyses found small but significant effects to reduce depression, anxiety, anti-social behavior and substance use. Further, the effects are sustained over time. Meta-analyses often found that the effects were heterogeneous. A conceptual model is proposed to guide the study of moderators of program effects in future meta-analyses and methodological issues in synthesizing findings across preventive interventions are discussed.

Introduction

The past two decades have seen a rapid growth of empirical evidence supporting the effectiveness of preventive interventions for children, youth and young adults. The National Research Council and Institute of Medicine (2009) presented a narrative review of findings from a broad range of trials, including those directed at preventing a single disorder, those directed more broadly at the family, school or community and those designed to promote positive mental health. The report concluded that “substantial progress has been made in demonstrating that evidence-based interventions that target risk and protective factors at various stages of development can prevent many problem behaviors and cases of MEB (mental, emotional and behavioral) disorders” (NRC/IOM, 2009, p. 216). Although the report noted that “the effect sizes for most interventions were small to moderate” (p. 218), it was beyond the scope of that report to quantitatively assess effect sizes across prevention trials or identify factors that were associated with differential effect sizes.

This overview of meta-analyses builds on the IOM report in four ways. First, it summarizes and compares the effects found across a broad array of preventive intervention trials using published meta-analyses. Second, it identifies factors that are related to effect sizes in multiple meta-analyses. Third, it develops a conceptual framework for summarizing factors that are related to effect sizes in prevention and promotion trials. Fourth, it identifies methodological issues in synthesizing effects across prevention trials.

Cooper and Koenka (2012) note that overviews of meta-analyses seek to systematically synthesize findings across reviews, each of which has defined their area of interest much more narrowly. An overview is particularly appropriate to synthesize findings in a broad

⁴The third and fourth authors made equal contributions to the paper. The order of authorship was decided randomly between them.

field such as prevention science where meta-analyses have focused on particular outcomes or specific preventive approaches. This overview summarizes the findings concerning effects across approaches to prevention and identifies issues for future research. Prior overviews of reviews of preventive interventions have focused on a single outcome (e.g. child abuse, Lundahl & Harris, 2006; Mikton & Butchart, 2009), presented a narrative description of characteristics of effective programs across areas of prevention (Nation et al., 2003), or summarized effects of different prevention programs on the common outcome of economic benefit (Aos et al., 2004).

The current paper presents an overview of meta-analytic reviews of preventive interventions in terms of their effects on prevention of the following mental, emotional and behavioral problems of children, youth and young adults (birth to age 26 years): depression, anxiety, anti-social behavior (including delinquency and violence), and substance use. Two categories of meta-analyses are included in this review. One category focuses on the effects of a range of programs targeting one of the specific problems or disorders listed above. The second category focuses on the effects of programs that are designed to promote healthy development (e.g., through mechanisms such as teaching social and emotional skills, parent training) and assessed one or more of the problem outcomes listed above.

The review is presented in three sections. First, we present the methodology used to identify the meta-analyses included in the overview. Second, we present brief reviews of meta-analyses that represent different approaches to prevention, including separate reviews of meta-analyses of prevention of specific problems (i.e., depression, anxiety, anti-social behavior and substance use) and meta-analyses of programs that promote healthy development. In each review, we first describe the characteristics of the meta-analyses. We then discuss the effects obtained, whether the effects are heterogeneous or homogeneous and findings concerning moderators of these effects. In the third section, we present a quantitative summary of the overall magnitude of effects across meta-analyses and discuss factors that account for heterogeneity of effects. We report quantitative assessments to summarize the overall magnitude of the effect sizes on the mental health and substance use problem outcomes across all meta-analyses and for each broad category of meta-analyses. We then present a conceptual framework for understanding factors that account for heterogeneity of effects (i.e., moderators) across meta-analyses and use this framework to discuss commonalities and differences in findings across approaches to prevention and methodological issues relevant for future work. We conclude with a brief summary.

Search methodology

We used five methods to obtain relevant meta-analyses of prevention programs. First, we conducted a computer literature search of the following databases: PsycINFO, PubMed, and Google Scholar, searching for relevant meta-analyses published between 2000 and 2013. Keywords included “prevention” and both specific problem outcomes (e.g., “depression”, “aggression”), general outcomes of interest (e.g., “health promotion”, “emotional learning”) as well as positive outcomes (e.g., “resilience”, “self-efficacy”), and social environmental targets (e.g., “school-based programs”, “parenting”). Second, we examined the resulting lists of studies manually to identify meta-analytic studies of children, adolescents, and young adults. Third, we searched the websites for organizations that have conducted meta-analyses of prevention and promotion programs (i.e., Cochrane; Campbell; Collaborative for Academic, Social, and Emotional Learning [CASEL]; Center for Evidence-based Policy; Annie E. Casey Foundation) to identify additional meta-analyses. Fourth, we examined reference sections of the two most recent meta-analyses on each problem outcome or promotion approach (e.g., depression, social emotional learning) to identify additional meta-analyses. Finally, we conducted a manual search to assess whether all journals in which

meta-analyses on prevention had been published (e.g., *American Psychologist*, *Prevention Science*) were included in the computerized literature search results. We reviewed the table of contents from journals not included in the search results to identify additional meta-analyses. Non-published meta-analyses were included only if they were on the websites of the organizations described above (e.g., Cochrane, CASEL).

The following inclusion criteria were used: (a) The meta-analysis was published between 2000 and May, 2013. (b) The study population was children, adolescents, and/or young adults up to age 26. (c) The meta-analysis involved trials where experimental or quasi-experimental designs were used and included a quantitative measure of program effects, such as effect sizes or odds-ratios. Literature reviews and systematic reviews were not included unless they also contained quantitative analysis of effect sizes across trials. (d) The outcomes examined included symptoms or diagnostic measures of one or more of the following problem outcomes: depression, anxiety, aggression, antisocial behavior, delinquency, criminal behavior, alcohol use, drug use, or cigarette smoking. After all the meta-analyses that met the inclusion criteria were gathered, studies were examined to eliminate redundancy using the following rules. When a meta-analysis involved an update of prior meta-analyses or included the same set of trials, only the more comprehensive meta-analysis was included, typically the most recent or most methodologically rigorous meta-analysis. Two meta-analyses of the same set of studies were included only when they addressed different questions (e.g., examined different moderators). Of the 67 reviews that were reviewed by two coders, all were examined for relevance and non-redundancy by two coders. A total of 48 (71.6%) met our criteria for relevance and non-redundancy and are presented in Table 1. The fact that 50% of these meta-analyses were published in the same year or after publication of the National Research Council and Institute of Medicine report on prevention (NRC/IOM, 2009) indicates the rapid pace of developing statistical summaries of the effects of prevention programs.

Meta-analyses of program effects on a specific problem or disorder

Depression

We identified five meta-analyses of trials on the prevention of depression. Three included a variety of types of programs, most of which were cognitive behavioral or educational in nature (Horowitz & Garber, 2006; Merry et al., 2012; Stice et al., 2009). Two focused on a specific type of program, the Penn Resiliency Program (Brunwasser, Gillham, & Kim, 2009) and exercise programs (Larun et al., 2009). The meta-analyses included studies with youth and young adults ages five through twenty-one, and analyzed effects up to 36 months following the intervention. Three included only randomized controlled trials (RCTs; Horowitz & Garber, 2006; Larun et al., 2009; Merry et al., 2012); the others included RCTs and trials with quasi-experimental designs.

All five meta-analyses found positive program effects of the interventions as compared to no-intervention controls at post-intervention. The most recent and most comprehensive meta-analysis (Merry et al., 2012) included 53 studies involving over 14,000 participants. Merry et al. (2012) found a small but significant effect to prevent diagnosis of depression post-intervention compared with no-intervention controls. Positive effects were observed up to 36 months following program completion. However, when comparisons involved active placebo groups, the effects were not significant for either depressive symptoms or diagnosis of depression. A small but significant effect for physical activity interventions was found but the studies reviewed were reported to be of low quality (Larun et al., 2009).

Four of the five meta-analyses reported heterogeneity of program effects. Moderation analyses and sub-group analyses were used to identify factors associated with larger effect

sizes. One important factor examined in three of the meta-analyses was whether the programs were universal or targeted high risk groups. Horowitz and Garber (2006) and Stice et al. (2009) reported that interventions that targeted high risk groups had larger effect sizes than universal programs, with the effect sizes being moderate for high risk programs and small for universal programs. Merry et al. (2012) reported that both universal and selective/targeted interventions showed significant effects to reduce diagnosis of depression that lasted nine months after the programs; effects on depressive symptoms were significant for universal and targeted interventions 9 and 12 months following the interventions, respectively.

Analyses of other moderators indicated that effect sizes were higher for programs that had higher percentage of minority participants, included older participants, were shorter, involved homework, and were delivered by professionals (Stice et al., 2009). The meta-analyses differed as to whether the programs were more effective for girls than boys (Merry et al., 2012; Stice et al., 2009).

Anxiety

We identified two meta-analyses of programs for preventing anxiety in youth (Fisak et al., 2011; Teubert & Pinquart, 2011), and one meta-analysis of programs to prevent anxiety in college students (Regehr et al., 2013). Most of the studies in the meta-analyses that focused on children and adolescents used cognitive behavioral approaches and worked directly with youth, although some included a component for parents. Both Fisak et al. (2011) and Teubert and Pinquart (2011) included only RCTs and reported on program effects up to 36 months following the interventions. The meta-analysis of programs for college students (Regehr et al., 2013) included RCTs and parallel cohort designs of stress reduction programs that employed cognitive, behavioral and mindfulness approaches, but did not report on maintenance of program effects.

Both meta-analyses for children and adolescents reported small but significant effects on anxiety symptoms at post-test and follow-up, and Teubert and Pinquart (2011) reported small but significant effects on diagnosis of anxiety at post-test and follow-up and significant beneficial effects on depressive symptoms, self-esteem and social competence as well. Regehr et al. (2013) reported moderate to large effects of stress reduction programs to reduce college students' anxiety symptoms at post-test.

Nearly all of the effects at post-test and follow-up were heterogeneous. The meta-analyses on programs for children and adolescents found that larger effects occurred for programs conducted by professional versus nonprofessional leaders. Despite wide variation in the number of sessions (range of 1 – 120), neither meta-analysis found that length moderated program effects. At post-test, Teubert and Pinquart (2011) found significant effects for both universal interventions and selective interventions that targeted at-risk youth. Consistent with findings from meta-analyses on programs for depression, larger effects were found for targeted versus universal programs. However, at follow-up, this difference was not significant. Fisak et al. (2011) found larger effects for programs that specifically targeted anxiety as the primary outcome and used a specific program (i.e., FRIENDS, an interactive cognitive behavioral program) compared to programs that targeted multiple outcomes. Teubert and Pinquart (2011) found that the addition of a parent component did not significantly increase program effects. The two meta-analyses differed on whether effects were moderated by age or gender. Regehr et al. (2013) did not examine factors that were associated with larger effects in their trials with college student.

Aggression, anti-social behavior and violence

We identified nine meta-analyses of prevention of aggressive, violent or antisocial behaviors. Six of these reviewed programs delivered primarily in schools (Beelman & Losel, 2006; Derzon et al. 2006; Mytton et al., 2006; Park-Higgerson et al., 2008; Piquero et al., 2010; Wilson & Lipsey, 2007), two examined family-focused programs (Farrington & Welsh, 2003; Piquero et al., 2009), and one included lengthy programs delivered in childhood (Dekovic et al., 2011). These meta-analyses included studies with youth from birth to age 18; studies that used RCTs and quasi-experimental designs, alternative interventions and placebo controls; and studies that assessed program effects up to 30 years following the interventions.

The six meta-analyses of school-based interventions complement each other by addressing different issues. The most comprehensive review (Wilson & Lipsey, 2007), which included 249 trials, found small but significant effects on aggressive/disruptive behavior at post-test for universal programs, selective programs and programs conducted in special schools or classes; non-significant effects were found for more comprehensive programs that involved multiple components (e.g., social skills groups, parenting skills, programs for school administrators and teachers). The other five trials focused on more specific outcomes, intervention approaches, or sub-populations. Two meta-analyses (Beelman & Losel, 2006; Piquero et al., 2010) reported that interventions focused on social skills (Beelman & Losel, 2006) or a broad range of self-control strategies (Piquero et al., 2010) had small, but significant effects on anti-social behavior at post-test and follow-up (Beelman & Losel, 2006). Mytton et al.'s (2006) meta-analysis, which focused on aggressive youth or youth at risk for becoming aggressive, found small, but significant effects at post-test and follow-up for programs that taught social or relationship skills and those that taught non-response to provocative situations. Derzon et al. (2006) studied a wider range of anti-social outcomes than the other meta-analyses, and reported small, significant effects to reduce criminal behavior and physical violence. Only Park-Higgerson et al. (2008) did not find significant effects on aggression and violence, although this may be due to the fact that they included fewer studies than the other meta-analyses. Although these meta-analyses focused on aggressive and disruptive behavior, four of them (Beelman & Losel, 2006; Derzon et al., 2006; Piquero et al., 2010; Wilson & Lipsey, 2007) also reported small but significant effect sizes to improve other behavior problems, including internalizing problems and school suspension, as well as positive outcomes, such as social skills, social relationships, personal adjustment, school performance and school participation.

Of the five meta-analyses that examined heterogeneity of effects, four found significant heterogeneity (Beelman & Losel, 2006; Park-Higgerson et al., 2008; Piquero et al., 2009; Piquero et al., 2010; Wilson & Lipsey, 2007). However, there was little consistency in the moderators identified across the meta-analyses. Two reported larger effects for more high-risk youth (i.e., lower social class) or youth with higher levels of behavior problems at program entry (Beelman & Losey, 2006; Wilson & Lipsey, 2007). Wilson and Lipsey (2007) also reported that better implementation was related to larger effects for selective programs and programs delivered in special classes. Beelman and Losel (2006) reported that the effects for cognitive-behavioral approaches were moderate and significant at post-test and follow-up whereas the effects for other approaches, such as counseling/psychotherapy, were not significant. Mytton et al. (2006) reported the effects were larger for programs that taught social or relationship skills versus those that taught non-response to provocative situations.

Two meta-analyses reported on family-based programs to prevent antisocial behavior. Farrington and Welsh (2003) included trials that used a variety of approaches to work with families. Although no statistical tests of heterogeneity were conducted, effects on antisocial

behavior were reported separately by approach. Small, significant effects on anti-social behavior and delinquency were found at post-test and follow-up for parent training, home visiting, pre-school/day care and multi-systemic therapy; the effect for school-based programs that included a parenting component was not significant. However, these sub-group analyses were based on relatively few trials and the length of follow-up and targeted population varied greatly across type of approach, thus, these findings should be viewed with caution. Focusing on family/parent training programs for children 5 or younger, Piquero et al. (2009) found a small, significant effect on antisocial/delinquent behavior, which was heterogeneous. Moderator analyses found that larger program effects occurred for studies published in the U.S. versus other countries and those with smaller versus larger samples.

Dekovic et al. (2011) found small, significant and heterogeneous effects to reduce adult criminal behavior for lengthy interventions (e.g., home-based, school-based, preschool) that were delivered to children younger than age 12. The effects were stronger for programs delivered to at-risk versus universal populations, programs for lower SES versus mixed SES groups, programs that included both boys and girls versus only boys, shorter versus longer programs (although all programs lasted at least 1.5 years) and those that focused on social and behavioral skills versus family support or academic skills. Dekovic et al. (2011) also found small, significant effects to improve positive outcomes, such as finishing high school and being employed.

Substance use

We identified 14 meta-analyses of interventions in this area. Six focused on substance use and included alcohol, tobacco and illicit drugs; five focused only on alcohol use; and three focused only on tobacco use. We review the findings on program effects, heterogeneity and moderators separately for the meta-analyses with these separate foci, and then consider cross-cutting findings.

Six meta-analyses focused on preventing substance use, including alcohol, tobacco and drug use (Cuijpers, 2002; Faggiano et al., 2008; Gottfredson & Wilson, 2003; Porath-Waller et al., 2010; Soole et al., 2008; Tobler et al., 2000). These meta-analyses included a range of 12 - 207 trials and involved youth from 11 – 22 years old. All programs were school-based, although some included broader community and family involvement. The programs used both RCTs and non-equivalent control group designs and included follow-ups as long as 3 years after the program.

Tobler et al. (2000)'s review, which included 207 trials, is the most comprehensive. Programs were coded as being either interactive (defined as providing contact among participants and opportunities to exchange ideas and teaching refusal skills, interpersonal skills or including system-wide school and community change strategies) or non-interactive programs (defined as using didactic approaches to educate about drugs, affective awareness and self-esteem or teaching problem solving to make a commitment not to use drugs). Significantly greater effects on drug use one to 12 months after the intervention were found for programs that involved interactive, skill-building strategies versus non-interactive strategies. The effects were heterogeneous for both interactive and non-interactive programs. However, when analyses included only the 93 high quality evaluations (i.e., evaluations that used RCTs and methods that minimized other sources of bias in selection of control group, research design and statistical analysis), the effects were significant and heterogeneous for the interactive programs that used social influence techniques and taught comprehensive life skill but non-significant and homogeneous for all types of non-interactive programs. Analyses that included only high quality evaluations indicated that the effects of the interactive programs were larger for studies with smaller samples, programs led by

clinicians or peers versus teachers and lengthier versus shorter programs. Two other meta-analyses found that peer- versus teacher-led programs had larger effects on drug use (Cuijpers, 2002; Gottfredson & Wilson, 2003).

Two meta-analyses focused on illicit drug use only (Faggiano et al., 2008; Porath et al., 2010). Faggiano et al.'s (2008) meta-analysis, which included 32 trials, reported small, significant effects of skills-based programs to reduce marijuana and hard drug use. Porath et al. (2010) found a moderate effect size to reduce marijuana use. Larger effects occurred for interactive programs versus didactic programs, programs delivered by leaders other than teachers, longer versus shorter programs, programs delivered to youth over age 14 versus younger youth, programs that used multiple methods (e.g., affective, informational, and social-learning methods) versus a single method, and programs that checked the fidelity of implementation versus those that did not. Soole et al.'s (2008) meta-analysis of 12 school-based programs reported small, significant effects on marijuana use and a composite of marijuana, amphetamine, cocaine, and opiate use at both short-term (less than one year after the program) and longer-term follow-up.

Four meta-analyses focused on prevention of alcohol use in college students (Carey et al., 2007; Carey et al., 2012; Fachini et al., 2012; Scott-Sheldon et al., 2012). These meta-analyses included a range of 14 to 62 trials and involved participants between 17 and 26 years old. The programs used both RCTs and non-equivalent control group designs and included follow-ups up to four years after the program.

Carey et al.'s (2007) review of 62 studies examined the effects of individually-focused, short (median of two sessions) programs that included a range of approaches, such as motivational interviewing, blood alcohol content education, normative comparison and feedback on consumption. They found small, significant effects on multiple indicators of alcohol use and alcohol-related problems at post-test and short-term follow-up. Although the effects diminished over time, effects on frequency of drinking days and alcohol-related problems remained significant up to 4 years after the intervention. Most of the effects were homogeneous, although effects on alcohol-related problems at short-term follow-up were heterogeneous. Moderator analysis found that program effects were larger for samples that contained a higher percentage of women; programs delivered in person versus on a computer; and interventions that included motivational interviewing techniques, normative feedback, feedback on expectancies and/or motives for drinking or a decisional balance exercise. Carey et al.'s (2012) meta-analysis of 48 trials examined face-to-face and computer-based interventions and found small, significant effects for both intervention types at post-test. However, only the face-to-face interventions had significant effects at three- and six-month follow-ups. Face-to-face interventions also had larger effects than computer-based interventions in studies that directly compared them. Fachini et al. (2012) conducted a meta-analysis of 12 randomized trials of a single intervention (BASIC) for heavy drinkers that used motivational interviewing and personalized feedback. They reported large significant effects to reduce alcohol consumption and alcohol problems one year after participation. Although the effects for both outcomes were heterogeneous, moderator analyses were not conducted. Scott-Sheldon et al.'s (2012) meta-analysis of 14 trials of programs that challenged alcohol expectancies found small, significant effects at post-test, but the effects were non-significant at follow-ups greater than a month.

Three meta-analyses focused on the prevention of tobacco use (Hwang et al., 2004; Isensee & Hanewinkle, 2012; Thomas & Perera, 2006). These meta-analyses included a range of 5 to 94 trials and involved youth from 5 to 18 years old. The programs used both RCTs and non-equivalent control group designs and included follow-up assessments 2 years after the program.

In the most comprehensive meta-analyses, Hwang et al. (2004) analyzed 65 trials classified according to the theoretical approach of the intervention (i.e., social influence, cognitive behavioral, social skills) and setting (school only, school plus community). They found small, significant effects on smoking up to three years following participation for trials using each theoretical approach. However, subgroup analyses showed that the effects were larger for cognitive behavioral versus the other approaches. Programs delivered in both school only and school plus community settings both had small, significant effects at post-test and follow-up, but those that included both school and community components had slightly larger effects at the three-year follow-up. Notably, the effect for the Life Skills Training program, which included social influence, cognitive behavioral, and affective methods, was larger than the mean weighted effect size across all the studies as well as the effect size of studies that used similar approaches. Thomas and Perera's (2006) meta-analysis included only trials that evaluated social influence or social competence-based approaches using RCTs. Although the effects in half of the trials were significant, the overall effect of programs that used social competence approaches, social influence approaches or a combination of these approaches was not significant at short-term (less than 18 months post-intervention) or long-term assessments. Their results may be attributed to the low power to detect effects due to the low number of trials included in the comparisons, although their findings are consistent with the non-significant effects found in one of the largest evaluations of smoking prevention programs [Hutchinson Smoking Prevention Project, Peterson (2000)]. Isensee and Hanenwinkel (2012) found a small, significant effect in their meta-analysis of a single intervention, the Smoke Free Class Competition. However, the small number of studies and multiple methodological shortcomings of the studies included in this meta-analysis require caution in interpreting these findings.

Meta-analyses of programs that promote healthy development or resilience

Programs that are designed to promote healthy development or resilience of youth who are exposed to stressful situations are theoretically expected to reduce multiple problem outcomes. We identified meta-analyses of five approaches that focus on promoting individual and environmental resources for healthy development: school-based social and emotional learning (SEL), after-school, mentoring, parenting, and preschool/ home visiting programs. Meta-analyses of resilience promotion programs were identified for youth exposed to three family stressors: parental death, parental divorce and parental mental illness. Although the meta-analyses discussed in this section typically assessed program effects on multiple outcomes, some of the studies reviewed in this section included trials of programs that targeted a single problem or disorder.

School-based programs for promoting SEL—We identified one meta-analysis of school-based SEL programs (Durlak et al., 2011). This meta-analysis included 213 universal programs involving 270,304 youth in kindergarten through high school that targeted full classrooms rather than youth with pre-existing problems. All programs focused on SEL skills, such as self-awareness, self-management, emotional awareness, relationship skills and effective decision making. Durlak et al. (2011) reasoned that such programs should lead to reductions in problem outcomes as well as greater academic success and more positive social adjustment. This meta-analysis included randomized and non-randomized trials with control groups and follow-ups up to 3 years post-intervention.

The results show small but significant effects to reduce conduct problems (e.g., aggression, non-compliance, delinquency) and emotional distress (e.g., depression, anxiety, social withdrawal). Further, the effects were significant at follow-up, a median of one year later. Effect size was *not* related to methodological rigor of the studies, children's ethnicity or urban versus rural status. Sub-group analyses indicated that the program effects were

significant for programs delivered by teachers as well as non-school personnel and for programs delivered in the classroom only as well as those that included school-wide components (e.g., school policies or procedures to encourage social and emotional development) or a parent component. However, multi-component programs did not have larger effects than single-component programs. Moderator analysis showed larger effect sizes for programs that used effective teaching practices (i.e., sequenced, active, focused and explicit, SAFE) versus those that did not and programs that did not have implementation problems versus those that had implementation problems

After-school programs—We identified one meta-analysis of after-school programs designed to promote personal and social skills (Durlak et al., 2010). The 75 programs in this meta-analysis included a variety of activities (e.g., academic, social, cultural, and recreational) and primarily targeted elementary and junior high school students who did not have any presenting problems. The meta-analysis included RCTs and trials that used non-equivalent control group designs. Program effects were reported at post-test but not follow-up.

A small, significant but heterogeneous effect was found to reduce conduct problems (e.g., acting out, aggression, delinquency) at post-test. Moderation analyses showed that programs that used SAFE practices had a small significant effect to reduce conduct problems and substance use, whereas those that did not use SAFE practices had non-significant program effects. Although the effects in studies that used SAFE practices were heterogeneous, analyses of eight potential moderators (e.g., randomization, reliable outcome measures, attrition) did not identify factors that accounted for the variability in effect sizes. Small, significant effects occurred for programs that used SAFE practices on positive outcomes, such as school bonding, prosocial behavior, grades, school attendance and achievement test scores.

Mentoring programs—We identified two meta-analyses of mentoring programs. Dubois et al. (2011) meta-analyses included 73 trials. They defined mentoring as “a program or intervention that is intended to promote positive youth outcomes via relationships between young persons (18 years or younger) and specific non-parental adults (or older youth) who are acting in a nonprofessional helping capacity” (pg. 66). Tolan et al. (2008) used a similar definition of mentoring in their meta-analyses of 39 trials. The meta-analyses included RCTs and quasi-experimental studies with follow-ups up to four years after the program. The meta-analyses differed in the moderators of program effects that were examined.

Dubois et al. (2011) estimated effect sizes after controlling for methodological quality of the study. Small, significant effects were found for psychological/emotional problems and conduct problems at post-test, but the effect on drug use was not significant. Six factors were significantly related to average effect size across outcomes; two youth factors (male > female; high risk/problem background > low risk/problem background), two youth-mentor match variables (matched on interests > not matched, not matched on race/ethnicity > matched on race/ethnicity) and two mentor role factors (included advocacy > not included advocacy, included teaching/informational role > not included teaching/information role). Although effects were larger for high risk participants, Dubois et al. (2011) noted that the programs reviewed did not target youth with extremely high levels of problems, so the programs are likely most successful with youth at moderate to severe risk. DuBois et al. (2011) reported significant effects at follow-up an average of 23 months later, but these effects were based on very few studies. Tolan et al. (2008) found small to medium significant effects on delinquency, aggression and substance use. All these effects were heterogeneous. The effects were stronger for RCTs, programs in which emotional support was an important focus, and programs in which professional development was a motivation

for mentoring. Several factors were not significant moderators of program effects, including risk for delinquency, delinquent status at program entry, whether the program included components in addition to mentoring, and whether fidelity was monitored.

Parent training programs—We identified six meta-analyses of parent training programs. These meta-analyses included programs with both randomized and quasi-experimental designs and with follow-ups up to 18 months post intervention. The most comprehensive meta-analysis (Kaminski et al., 2008) included 77 trials that involved children from birth to seven. Two smaller meta-analyses focused on programs for youth in a particular developmental period. Barlow et al. (2010) reported on eight trials for parents of children from birth to three and included RCTs and trials that used non-equivalent control group designs and studies with follow-up up to 18 months. Burrus et al. (2012) reported on 12 trials of programs for parents of adolescents that included teaching a variety of parenting skills (e.g., communication, monitoring). The meta-analyses by Farrington and Welsh (2003) and Piquero et al. (2009) focused on parenting and family-focused interventions to prevent anti-social behavior and delinquency and were reviewed previously in the section on anti-social behavior. Nowak and Heinrich (2008) reported on 55 trials of the Triple P Positive Parenting Program, which included different levels of the program differentiated by intensity and whether they targeted clinical populations or were prevention focused. The studies included RCTs, quasi-experimental and uncontrolled trials.

In the most comprehensive meta-analysis, Kaminski et al. (2008) found small but significant effects to reduce externalizing problems, which is consistent with the findings of the previously discussed meta-analyses of parenting programs that targeted anti-social behavior (i.e., Farrington et al., 2003; Piquero et al., 2009). They also found a medium effect to reduce internalizing problems, as well as small to moderate effects on educational and cognitive outcomes. Greater methodological rigor of the studies was related to larger effect sizes across outcomes. Kaminski et al. (2008) identified four components of programs that were significantly associated with larger effects on externalizing problems; positive interactions with child, time out, consistent responding and practicing program skills with one's child(ren). Programs with these characteristics had significant, medium effect sizes; the effect sizes were small for programs without these components. Barlow et al. (2010) found small, significant effects on both parent and independent observer reports of infant's and toddler's behavior problems. Burrus et al.'s (2012) meta-analysis reported small, significant effects on adolescents' violent behaviors, but non-significant effects on substance use. Nowak and Heinrich (2008) found a small significant effect of the Triple P prevention programs (Levels 1 – 3) to reduce child behavior problems.

Preschool/ home visiting programs—We identified three meta-analyses of programs to promote healthy development between birth and age five. Nelson et al. (2003) reviewed 34 studies of programs that included multiple components (e.g., home visitation, parent training and preschool education). Sweet and Applebaum's (2004) meta-analysis included 60 home visiting programs that targeted at-risk families, defined as having low income, being a teenage parent, being at risk for abuse or neglect, or receiving welfare. The programs provided a range of services, such as parent education, support and counseling; encouraging parent-child activities; and case management and child health or developmental screening. Manning, Homel, and Smith (2010) conducted a meta-analysis of the long-term effects of 11 early developmental prevention programs, including pre-school, center-based developmental day care, home visiting, family support and parent education programs. The meta-analyses included RCTs and quasi-experimental designs and studies with follow-up into adolescence.

Sweet and Applebaum (2004) reported a small, significant effect on social emotional outcomes at post-test but did not report effects at follow-up. Although the effect sizes were homogeneous, the lack of specificity of the variables included as social emotional outcomes, wide variability in characteristics of home visiting programs and lack of tests of program effects at follow-up make interpretation of this effect difficult. Focusing on outcomes in developmental periods following program participation, Nelson et al. (2003) found small but significant effects of preschool programs to improve a construct they labeled as social-emotional outcomes, a broad range of variables including behavior problems, social skills, self-esteem, grade retention, employment and criminal behavior, at medium-(K – 8th grade) and long-term (high school and beyond) follow-ups. The effects at K-8th grade were heterogeneous. Program effects were larger for programs longer than one year versus shorter programs and programs offered to predominantly African-American families versus other ethnicities. Methodological quality of the evaluations was not related to the effects on social-emotional outcomes. Manning et al. (2010) found small, significant but heterogeneous effects in adolescence on socio-emotional development (e.g., parent and teacher report of problem behaviors, obsessive compulsive behavior and self-confidence) and criminal justice involvement (e.g., arrests) and moderate effects on deviance (e.g., delinquency, drug use). However, these effects were based on very few studies and analyses were not conducted to assess moderators of the effects.

Programs targeting family disruptions—Because it was not possible to review research on all factors that could potentially disrupt family life, we selected three such factors: parental divorce, parental bereavement and parental mental illness. We identified five meta-analyses of programs designed to promote resilience of children in families impacted by these factors. Stathakos and Roehrl (2003) reviewed 24 studies of parent- and child-focused programs for divorced families and Fackrell et al. (2011) reviewed 28 studies of court-affiliated parent education programs for divorced parents. The meta-analyses included studies with youth up to age 18. They included RCTs as well as non-randomized trials with control groups and studies with follow-up of up to 40 months following the program. Both meta-analyses of programs for bereaved families (Currier et al., 2007; Rosner et al., 2010) reviewed 13 randomized and quasi-experimental studies that included a wide range of intervention strategies, such as support for children, psycho-education, family-focused approaches and music therapy for children and adolescents from birth to 18 years old. Siegenthaler, Munger and Egger (2012) reviewed 13 studies of RCTs of family-, parent- or couple-, and adolescent-focused interventions for youths (newborns through adolescents) whose parent was experiencing depression, anxiety or a substance abuse disorder.

Reviewing a heterogeneous group of studies, some of which focused on parents and others on children, Stathakos et al. (2003) found a medium, significant and heterogeneous effect on anxiety and a non-significant effect on depression at post-test. Several factors were related to larger effect sizes across outcome variables, including good training of group leaders, medium duration and number of sessions, short period since the divorce (< 30 months), and poorer methodological quality of the studies. Fackrell et al. (2011) reported a small, significant effect to improve child well-being (i.e., child behavior issues, adjustment to divorce) of court affiliated parent-education programs. Although the authors noted that the included studies were not methodologically strong, they did not test whether effects were moderated by quality of the study.

Inconsistent findings were reported in the two meta-analyses of programs for parentally bereaved youths. Rosner et al. (2010) reported small to moderate, significant and heterogeneous effects on depression and anxiety. More severe youth symptoms at pre-test were related to larger effect sizes. The relations between several other variables, including youth age, program dosage, time since death, use of confrontation strategies and publication

status of the study, were not significant moderators of program effects. Currier et al. (2007) reported a small, non-significant effect on a composite of adjustment outcomes, but did not report effects on specific outcomes. Shorter time period since the death and higher level of youth distress at program entry were related to larger effects. Although both meta-analyses noted that the quality of the reviewed studies was uneven, neither examined methodological quality as a moderator.

Seigenthaler et al. (2012) reported a significant effect for interventions to reduce the onset of a diagnosis of mental illness by 40% for youth whose parents had a mental or substance abuse disorder. They also reported a small but significant effect to reduce internalizing problems, but a non-significant effect to reduce externalizing problems. The effects of interventions on externalizing problems were heterogeneous, with smaller studies showing larger effects than larger studies. Interventions involving both parents and youths were not more effective than those involving only parents.

Overview of meta-analyses of programs to prevent mental health and substance use problems and promote healthy development

Effects across outcomes and approaches to prevention and promotion

In this section, we provide quantitative summaries (see Table 2) that reflect preventive impact on the problem outcomes for each broad category of meta-analyses and across all reviewed meta-analyses. For each broad category, we present an overall effect size for symptoms or related measures that used a continuous scale, and where available, an overall odds ratio, relative risk, or risk difference for diagnoses and other dichotomous measures. To make these comparable across broad categories, we used reported effect sizes or summaries for binary outcomes based on the largest population group and assessed impact at one year follow-up or as close to this time as was reported. If these were reported for randomized as well as quasi-experimental studies, we present the summary statistics for the randomized studies. If multiple controls were used we present comparisons against inactive controls. We computed simple averages for effect sizes and odds ratios (first transformed to logs then averaged) as standard errors were not readily available. Although numerous other numerical effect sizes were reported in the original meta-analyses, we did not conduct a comprehensive “mega-regression” of the effect sizes as this is beyond the scope of the current paper.

All 121 summary effect size (ES), odds ratio/ relative risk (OR/RR) or risk difference (RD) measures that were used to derive the figures in Table 2 are roughly comparable (within ES, OR/RR or RD groupings) because they pertain to a similar time frame, up to one year follow-up and compare the preventive intervention against inactive controls in broad populations (as opposed to a subgroup). The first column refers to the broad prevention categories of meta-analyses described in the prior sections. The second column refers to all effects reported in meta-analyses organized by primary focus of the meta-analysis in column one. For the meta-analyses of promotion programs, the summary effects across all problem outcomes are presented in column two because these programs targeted multiple problem outcomes. Column three presents effects for each of the problems shown in column one, but summarized across all meta-analyses that reported on effects on that outcome (e.g., figure includes summary statistics for substance use that were obtained in meta-analyses that primarily focused on other problems, such as conduct disorder). Columns two and three provide different ways to pool the summary statistics. For example, the meta-analyses that focused on depression contributed 6 summary effect sizes (ES), each of which was based on at least four trials, whereas there was a total of 9 summary ES's across all meta-analyses that included depressive symptoms as an outcome.

These summary statistics show small to medium overall beneficial effects on both symptoms as well as disorders and behavior problems. The largest effect size occurred for the meta-analyses that focused on anxiety ($ES=0.42$). The next largest effect size occurred for the broad set of interventions designed to promote healthy development, which included school-based SEL, after-school, mentoring, early childhood programs, and programs targeting significant family disruptions, such as bereavement, divorce and parental mental illness. ($ES=0.30$), with the effect sizes for depression, substance use, and crime/anti-social behavior ranging from 0.12 to 0.19. Note that the outcomes in many of the meta-analyses of programs to promote healthy development (e.g., Durlak et al., 2010; Kaminski et al., 2008; Sweet & Applebaum, 2004) were broad composites of indices of internalizing and externalizing problems or socio-emotional development rather than measures of specific problem behaviors or indices of severe levels of problems, such as meeting diagnostic criteria, violence or criminal behavior. The effect sizes for internalizing problems, $ES=0.30$, exceeded those for externalizing problems, $ES=0.22$, and substance use/abuse problems, $ES=0.14$. However, we did not formally test these differences because of potential overlap in the meta-analyses and lack of reported standard errors. For the dichotomous outcomes, all odds ratios/relative risks showed small benefit, with none exceeding 1.19. All risk differences indicated beneficial outcomes as well, although the overall magnitude of 0.05 is difficult to assess in terms of overall strength. One clear finding regarding heterogeneity is apparent. Nine of the 16 mean effect sizes (57%) had standard deviations that were greater than 0.15, often of the same magnitude as the average effect. This indicates substantial heterogeneity exists even within the summary effect sizes.

Conceptual framework for understanding heterogeneity of program effects

A consistent finding across meta-analyses is that effects are heterogeneous across trials; heterogeneity of effects was reported in 78% of the meta-analyses. Determining what sources of variation account for such heterogeneity is critical for understanding the impact of prevention programs. Because of the large number of reported summary statistics, and the lack of consistency in the sources of variation that were investigated across meta-analyses, we limit our discussion to two types of summaries. We examine commonly assessed factors that contributed to heterogeneity. We also discuss gaps in the reporting in meta-analyses that limit our ability to understand this heterogeneity. Following this, we illustrate how one aspect of interventions – the degree to which interactive teaching strategies are used – affects intervention strength across diverse prevention approaches.

Figure 1 presents a conceptual model that identifies the key factors that influence the effects of prevention and promotion interventions. As shown, the issue of understanding factors associated with program effects is relevant to the full range of prevention and promotion studies, including studies of efficacy (Does the program work under ideal conditions?), effectiveness (Does the program work under natural community service delivery conditions?) and implementation (Does the program work when delivered at scale as a public health intervention?) (Flay et al., 2005). This framework goes beyond the approach used in most meta-analyses in prevention, which are generally organized to assess effects of different approaches on one targeted outcome (e.g., depression) or of a particular approach (e.g., mentoring) on multiple outcomes, and typically give little attention to broader issues that relate to the public health impact of the interventions, such as the conditions of delivery in the community. The model builds on several recently proposed models for understanding sources of variation in the effects of preventive interventions (e.g., Durlak & Dupre, 2008; Weiss et al., 2013) and on the research agenda for the translation of preventive interventions that are effective in laboratory tests into practical programs that can be delivered at scale to the population (e.g., Spoth et al., 2013).

Figure 1 presents five factors as sources of heterogeneity of program effects on theoretical mediators and positive and problem outcomes. The effects of a prevention program are conceptualized to vary depending on characteristics of the program, characteristics of participants who receive the program, variability in program implementation, characteristics of the program delivery system and service providers, and the community and historical context within which the program is delivered. Methodological aspects of the trial (e.g., length of follow-up, nature of the control group) and the computation of meta-analytic summaries (e.g., different inclusion/exclusion criteria, fixed or random effects) also account for variation. We utilize this framework to discuss common moderators of program effects across meta-analyses and to identify gaps that need to be addressed in future trials and meta-analytic syntheses across trials. We then discuss issues in assessment of two sets of outcomes, theoretical mediators and problem outcomes. Finally, we discuss methodological issues that are critical for observing the effects of the trials of programs and the syntheses of findings from these trials

Program effects

Multiple problem outcomes—Meta-analyses report on multiple problem outcomes for several reasons. In some cases, they report on multiple indicators of the same problem, such as continuous measures of symptoms and dichotomous measures of disorder (e.g., depression, Merry et al., 2012; anxiety, Teubert & Pinquart, 2011), or multiple indicators of problems in the same domain (e.g., aggression, crime, violence, Derzon, 2006; antisocial behavior, delinquency, Farrington & Welsh, 2003; frequency of drinking, alcohol-related problems, Carey et al., 2012). Because of the high comorbidity between mental and substance abuse disorders in youth (Kessler et al., 2012) and the finding that many risk and protective factors are significantly related to multiple problems (MacArthur et al., 2012), preventing a single problem outcome or disorder should be associated with preventing other problem outcomes or disorders. Several meta-analyses reported on the effects of prevention programs on comorbid problems or disorders (e.g., anxiety and depression, Larun et al., 2009; Teubert & Pinquart, 2011; use of multiple substances, Soole et al., 2008; anti-social behavior and substance use, Tobler et al., 2000; Tolan et al., 2008). Other meta-analyses reported on indices of life adjustment, such as success in school, work, social relationships (Beelman & Losel, 2006; Dekovic, et al., 2011; DuBois, et al., 2010; Durlak et al., 2010; Durlak et al., 2011), in addition to problem outcomes. Meta-analyses that report on effects on multiple outcomes provide a more accurate representation of the potential public health impact of prevention programs than those that only report on a single outcome. However, as discussed below, there are methodological challenges in including multiple outcomes in meta-analyses.

Theoretical mediators—Prevention programs are based on an underlying theory in which the program affects mediators and change in these mediators lead to change in the development of problem outcomes over time (e.g., Tein et al., 2004). Analysis of mediation informs our understanding of the processes by which prevention programs have their effects and helps identify core components of interventions that are necessary to preserve as programs are disseminated (Spath et al., 2013). Some meta-analyses reported on program effects on theoretical mediators as well as effects on targeted outcomes. For example, several meta-analyses of programs for the prevention of anti-social behavior and aggression assessed program effects on putative mediators of social skills, social adjustment, and social competence (Beelman & Losel, 2006; Piquero et al., 2010; Wilson & Lipsey, 2007). Several meta-analyses of substance use prevention programs also presented evidence of program effects on putative mediators, such as alcohol expectancies (Scott-Sheldon et al., 2012) and refusal skills and attitudes toward smoking (Hwang et al., 2004). Meta-analyses of programs to promote healthy development assessed program effects on factors that were hypothesized

to affect internalizing and externalizing problems, including parenting behavior (Kaminski et al., 2008) and social and emotional skills (Durlak et al., 2011). However, although several of these meta-analyses described theoretical links between the putative mediators and targeted outcomes (e.g., DuBois et al., 2011; Piquero et al., 2010) none conducted summary statistical analyses to test the effects of these variables to mediate program effects on the targeted outcomes. As discussed in the methodology section below, there are serious challenges to conducting meta-analyses on mediational pathways and other methods of synthesizing findings across trials are needed.

Moderators of program effects

Program characteristics—One program characteristic that has been found to explain a significant amount of variation in outcomes across areas of prevention is the degree to which the programs use interactive strategies to teach the skills that are hypothesized to result in reductions in problem outcomes. Programs that involved more active strategies, such as discussion of the program material and practice of program skills, had larger effects than those that did not include these strategies. For example, in the meta-analyses on preventing substance use, interactive programs (i.e., emphasizing exchange of ideas, teaching drug refusal skills and encouraging feedback and constructive criticism in a non-threatening environment) were associated with larger effects compared to programs that were more didactic (Tobler et al., 2000). School-based and after-school SEL programs that involved active learning strategies to teach and practice social skills had larger effects than those that did not use these strategies (Durlak et al., 2010; Durlak et al., 2011). For parenting programs, Kaminski et al. (2008) found that programs that taught specific parenting skills (i.e., positive interactions, time out, consistent responding) and emphasized practice of skills with one's child(ren) were associated with larger effects than those that did not teach these skills or emphasize skills practice. For depression prevention programs, Stice et al. (2009) found that practice of program skills via homework was associated with larger effects. These program characteristics are consistent with cognitive behavioral approaches to behavior change, which are heavily represented across prevention and promotion programs and are associated with larger effect sizes in programs to prevent aggression, anti-social behavior and smoking (Beelman & Losel, 2006; Hwang et al., 2004; Piquero et al., 2010). Motivational interviewing was associated with large effect sizes on alcohol use in college students (Fachini et al., 2012; Carey et al., 2007) but was not assessed in meta-analyses that targeted other problem outcomes. Motivational interviewing is a promising approach that should be investigated in future prevention studies.

A program characteristic that has yielded inconsistent findings across several meta-analyses is length of program. Several meta-analyses found that shorter program had larger effects (i.e., Dekovic et al., 2011; Stahakos & Roehrl, 2003; Stice et al., 2009); several found that longer programs had larger effects (e.g., Wilson & Lipsey, 2007); and several found non-significant associations between length and effect size (e.g., Gottfredson & Wilson, 2003; Teubert & Pinquart, 2011). Although it appears that more is not necessarily better, the current findings provide little guidance as to the conditions under which program length may or may not impact outcomes.

Several limitations of trials included in the meta-analyses present problems for summarizing which program characteristics are associated with outcomes. For example, in meta-analyses of after-school programs and mentoring programs, the nature and content of the programs are not well described, making it difficult to assess the relations between program characteristics and outcomes. Given that program effects represent the difference in the contrast between the program and a control condition, understanding program effects also requires an adequate description of the control condition. With a few exceptions, the meta-

analyses included in this review did not adequately describe the comparison conditions. The importance of this issue is illustrated by Merry et al.'s (2012) finding that programs for preventing depression had a significant effect when compared to no-intervention controls but not when compared to placebo controls. Careful assessment of the experience of participants in comparison conditions becomes particularly important when testing the effects of scaled-up prevention programs in comparison to the programs currently being used in the community, which may include other prevention programs.

Programs can also be differentiated based on their target population (National Research Council/Institute of Medicine, 2009), including the full population (universal), individuals at elevated risk (selective) or individuals with elevated levels of the problem (indicated). Meta-analyses of the effects of programs for prevention of depression (Horowitz & Garber, 2006; Merry et al., 2012; Stice et al., 2009), anxiety (Teubert & Piquart, 2011), aggression/anti-social behavior (Beelman & Losel, 2006) and alcohol use (Carey et al., 2007) reported that selective or indicated programs had larger effects than universal programs. However, these findings should not be over interpreted to indicate that universal programs are not effective. Several meta-analyses found that although the effects of universal programs were smaller than indicated/selective programs, they had significant effects to prevent depression (Merry et al., 2012), anxiety (Teubert & Piquart, 2011), aggression/anti-social problems (Wilson & Lipsey, 2007), and conduct problems and emotional distress after participation in SEL programs (Durlak et al., 2011). From a public health perspective Shamblen and Derzon (2009) have argued that universal programs with a smaller effect size than selective or indicated programs on the population exposed to the intervention might have a larger impact on the public health because they reach a larger proportion of the population.

Participant characteristics—Meta-analyses have studied participant problems as moderators by assessing the relations between level of problems or risk factors at program entry and outcomes. Researchers have used two approaches to assess these effects; one summarizes reported subgroup analyses across trials and the second uses a trial-level proportion of risk, an ecological rather than individual-level factor. Consequently, the interpretation of the findings of these approaches differs (Perrino, Howe, Sperling et al., in press). Several meta-analyses concluded that larger effects were associated with higher levels of initial problems (Carey et al., 2012; DuBois et al., 2011; Nowak & Heinrichs, 2008; Rosner et al., 2010). Only a few of the reviewed meta-analyses examined social class as a moderator (e.g., DuBois et al., 2011; Dekovic et al., 2011). For example, DuBois et al. (2011) assessed social class as part of a broader variable of environmental risk that included social class, low community resources, and family dysfunction. They found that effects were largest for individuals who had a profile of being either high on individual risk and low on environmental risk or high on environmental risk but low on individual risk. Although only a few meta-analyses assessed ethnicity as a moderator, two meta-analyses found larger effects for programs with a larger percentage of ethnic minority participants (Nelson et al., 2003; Stice et al. 2009). Given the relations between social class and problem outcomes (Yoshikawa et al., 2012) and the ethnic heterogeneity of the U.S. population, greater attention is needed to assess effects of preventive interventions across ethnic and racial groups and across social class.

Meta-analyses of prevention programs for youths who experienced family disruptions of divorce and bereavement found that a shorter time between the disruption and program entry was associated with larger effects (Currier et al., 2007; Stahakos & Roehrle, 2003). These findings may reflect trajectories of adaptation to these disruptions over time and indicate the importance of offering prevention programs before maladaptive patterns of behavior become solidified.

Findings on some participant characteristics, including gender (Carey et al., 2012; 2007; Horowitz & Garber, 2006; Scott-Sheldon et al., 2012; Stice et al., 2009; Teubert & Piquart, 2011) and age (Beelman & Losel, 2006; Horowitz & Garber, 2006; Mytton et al., 2006; Scott-Scheldon et al., 2012; Stice et al., 2009; Teubert & Piquart, 2011), were inconsistent across meta-analyses. These differences may reflect differences in age- or gender-related trajectories of the development of the problems targeted by the intervention. Differences in the effects of prevention programs for depression and externalizing problems may reflect differences in rates of these problems across gender (Stice et al., 2009).

Program implementation—The few meta-analyses that have studied implementation as a source of variation of program effects have yielded consistent findings. Better implementation was associated with larger effect sizes in meta-analyses of substance use, anti-social behavior and school-based SEL programs (Durlak et al., 2011; Porath et al., 2010; Wilson & Lipsey, 2007). These findings are consistent with those of Durlak and Dupree (2008) who found that higher levels of implementation were related to better outcomes in 45 of 59 prevention and promotion programs. Durlak and Dupre (2008) point out, however, that understanding the effects of implementation of prevention programs is currently limited because the assessment of implementation has focused on only a few dimensions (i.e., fidelity and dosage) rather than a broader array of dimensions of implementation (e.g., quality, responsiveness, adaptation, Berkel et al., 2011; Durlak & Dupre, 2008). The assessment of implementation in the reviewed meta-analyses was very broad (e.g., problems vs. no problems in implementation). Future work needs to study the effects of implementation using more refined measures of multiple dimensions of implementation.

Service delivery system and service providers—There is considerable evidence that the system in which the intervention is delivered and the providers who deliver the program have important effects on outcomes (for reviews see Durlak & Dupre, 2008; Fixsen et al., 2005; Weiss et al., 2013). Although theory and research has begun to describe the ways in which system (e.g., leadership, capacity, resources, climate) and provider characteristics (e.g., skill proficiency, perceived benefits of the program, perceived efficacy) link to outcomes, consistent patterns between service delivery/provider characteristics and outcomes have not been well-documented by extant meta-analyses. Several meta-analyses report that mental health professionals were more effective providers than non-professionals or teachers (Fisak et al., 2009; Stice et al., 2009; Teubert & Piquart, 2011; Tobler et al., 2000), although no difference was found for professional level of provider in a meta-analysis of programs to prevent antisocial behavior (Beelman & Losel, 2006). Several other meta-analyses found larger effects for peer leaders as compared with teachers or non-peers in substance use prevention programs (Cuijpers, 2002). These findings indicate the importance of identifying the conditions under which different providers may be more effective. The finding from several meta-analyses that a higher level of provider training was related to larger effects (Durlak et al., 2010; Durlak et al., 2011; Stahakos & Roehle, 2003) has important implications for implementing programs in real-world settings. Theoretically, the effects of providers or of training may be mediated through differential program implementation (Durlak & Dupre, 2008). Similarly, the finding from several meta-analyses that smaller size samples were related to larger effect sizes (Piquero et al., 2009; Siegenthaler et al., 2012; Tobler et al., 2000) may be due to it being easier for service delivery systems to effectively implement smaller-scale programs. Clearly, research has just begun to identify the impact of service delivery systems and providers and future meta-analyses should consider this important moderator.

Contextual factors—Contextual factors refer to the historical time and the community context in which the program is delivered and include a broad range of factors, such as the population level of problem behaviors occurring at the time the trial is conducted, other services offered in the community (Weiss et al., 2013), and stressors or environmental factors that make it difficult for participants to benefit from the program (e.g., neighborhood poverty, cultural incompatibility with program). Only a few meta-analyses have analyzed such contextual factors. Some meta-analyses have found that the country in which the program was implemented was related to effect size (Piquero et al., 2009; 2010) and one meta-analysis found no differences in effects between programs delivered in rural and urban settings (Durlak et al., 2011). Recent studies on cross-national effects of parenting programs (Knerr, Gardner & Kluver, 2013) indicate the increasing interest in such contextual factors, although these factors have not yet been studied in meta-analyses of prevention programs. Other contextual factors that influence the development of problems have rarely been studied as moderators of program effects, such as presence of gangs, social network influences on substance use, and school discipline policies. Understanding the effects of such contextual variables will be increasingly important as these programs are applied to address public health problems in heterogeneous community contexts.

Practical implications of findings on factors associated with variability in effects

Understanding the source of variation in effects has important practical implications for enhancing the public health impact of prevention programs. Illustratively, we focus on the practical implications of the program characteristic that was most commonly found to moderate the effects of prevention programs across meta-analyses, the degree to which programs use interactive, skill-building strategies. The findings concerning the significant moderation effects of this program characteristic on multiple outcomes was reviewed above. Table 3 shows the practical implications in terms of the benefit for the average person who participated in an interactive, skill-building program and the benefit for the average person who participated in a non-interactive program. The percentage benefit for the average participant in the program as compared to the median of the comparison group was calculated using Cohen's U_3 (Lipsey et al., 2012). This index provides an interpretable percentile metric comparing how much benefit can be expected from an intervention if the average participant were to score at the median in the control group. The U_3 value estimates the percentile that such a person would be expected to achieve if the intervention were delivered. For example, the Cohen's U_3 of .75 for participation in a parenting program that involved practicing skills with a parent's own child(ren) indicates a 25% benefit for average participant in the program as compared to those who would score at the median in the control condition. In contrast, the Cohen's U_3 of .55 for participants in parenting programs that did not involving practicing skills with their child(ren) is .55, which is only 5% benefit for the average person in these programs, indicating that there is four times the benefit for the average participant in a parenting program that involved skills practice with the parents' child(ren) compared to one that did not involve such skills practice.

Table 3 shows the percent benefit of participating in interactive, skill-building programs as compared to non-interactive programs across meta-analyses in six areas of prevention (i.e., anxiety, antisocial behavior, substance use, school-based SEL programs, after-school programs and parenting programs). Although consistent benefits of interactive, skill-building programs have been found across outcomes and approaches to prevention, in some cases, the differences are particularly dramatic. The benefits of participating in parenting programs that used interactive, skill-building strategies as compared to those that did not use these practices range from 20% for practicing the skills with one's child(ren) to 7% for focusing on positive interactions between the parent and child(ren). The benefits of being in cognitive behavioral programs as compared to other types of programs at follow-up on

antisocial behavior are also substantial, with 21% greater benefit occurring for the average participant. The differential benefit of programs that vary on this moderator illustrates the public health significance of future research that identifies other robust moderators of the effects of prevention programs.

Methodological issues in synthesizing the effects of prevention and promotion programs across meta-analyses

The last part of Figure 1 involves existing methodological practices and limitations in quantifying effects in meta-analyses. As one illustration, methodological quality of trials in particular has an important effect on our confidence in the validity of the findings from these studies. Methodological quality of trials has been coded using standard rating scales that assess risk of bias using indicators such as randomization, allocation concealment, blinding of participants and assessors, methods for analyzing incomplete data and potential selective reporting (e.g., Cochrane Handbook, Higgins, 2008). Meta-analyses have accounted for quality of studies in several ways, such as only including trials that used randomized experimental designs, assessing the relation between quality and effect size, separately analyzing trials with high levels of quality, and statistically accounting for level of quality in the report of effects (Ahn & Becker, 2011).

We found inconsistency in the relation between quality and magnitude of effect. Although a number of meta-analyses reported that methodological quality was not related to effect size (e.g., Gottfredson & Wilson, 2003; Nelson et al., 2003; Teubert & Piquart, 2011), some found that larger effects were associated with indicators of poorer methodological quality (e.g., Merry et al., 2012; Stahakos & Roehrl, 2003). The lack of consistency in assessing methodological quality of trials across meta-analyses is problematic. More systematic approaches to assessing the methodological quality of individual trials (e.g., Jadad et al., 1996; Jüni et al., 2001) need to be considered. Further efforts should also be made on how to account for differences in the quality of meta-analyses when synthesizing meta-analytic results. One critical issue in synthesizing effects across meta-analyses is redundancy of studies included in the analyses. We noted that a sizable portion of the same individual trials were used in multiple meta-analyses. For instance, in the seven depression prevention meta-analyses identified by the study team, we identified 156 unique trials. Of these trials, 69.23% ($n = 108$) appeared in only one of the seven meta-analyses. Of the remaining 48 trials (30.76%), 15.38% appeared in two meta-analyses, 9.62% appeared in three meta-analyses, 5.12% appeared in four meta-analyses, and .06% appeared in five meta-analyses. The implication of this redundancy becomes apparent when we want to synthesize effect sizes across multiple meta-analyses in the same substantive area. Future methodological work is needed to determine the magnitude of bias introduced by this redundancy of trials across meta-analyses.

Methodological challenges also arise in synthesizing findings concerning aspects of our conceptual model, particularly mediation and multiple outcomes. Unlike the literature in primary studies, where the statistical methods for examining the significance of mediating effects have been well-documented (MacKinnon, 2008; Preacher & Hayes, 2004), resulting in their increased utility in practice, tests for mediation have not been widely employed in meta-analytic research (Shadish, 1996; Shadish & Sweeney, 1991). Instead, most meta-analyses focused on a descriptive summary of relations that are quantified by effect size (Shadish, 1996), which in turn limits the capacity to illustrate the intricate mechanisms/processes of the phenomenon. In particular, Shadish and Sweeney (1991) posited that the difficulties of correctly specifying the theoretical model might impede researchers from examining mediating effects in meta-analyses. However, if the underlying theory supports the mediating processes for the given data, the use of a mediational model would be preferable to other potentially misspecified models. In cases where a mediational model is

carefully justified, Becker (1994) claimed that a meta-analysis using an “exploratory model” would help us to build a case for potential causal relations by explaining variation in the observed relation between the independent and dependent variables.

Even if the mediational model can be theoretically and empirically justified, another challenge arises from a lack of research on a test for mediation in a meta-analysis. Although it is rare, meta-analysts exploring mediators have claimed significant mediation, when the effect of the independent on dependent variable is found to be non-significant (i.e., direct effect of independent variable on dependent variable) after controlling for the mediating variable, which is suggested by Baron and Kenny (1986) and also commonly employed in practice (Preacher & Hayes, 2004). Yet, no research currently exists to provide statistical justification for testing mediational effects as is commonly done in primary studies in meta-analysis. Further research is strongly urged to provide statistical justification for examining the existence of mediating variable(s) in meta-analyses.

Examination of program effects on multiple outcomes also pose methodological concerns. As discussed by Gleser and Olkin (1994), although those univariate approaches that most meta-analyses relied upon may be valid in some circumstances, multivariate approaches, including generalized regression (Hedges & Olkin, 1985), generalized-least-squares approach (Raudenbush et al., 1988), and multilevel mixtures (Brown et al., 2008), are recommended to account for possible dependency among the estimated effect sizes and further minimize an increase in Type I error rates.

One of the major limitations of meta-analysis, and hence our overview, occurs because they rely on previously published summaries. Although individual trials report main effects, they often do not report all potentially relevant subgroup, moderator, or mediator analyses, particularly if they are nonsignificant. When subgroup analyses are reported, they are reported differently across meta-analyses. For example different cut points are used to define older versus younger participants, making it difficult to combine or compare such findings across trials. Further, many meta-analyses provide summaries, such as high versus low risk groups, at the ecological level rather than individual level. Using the percent of high-risk participants in the trial as a covariate in a meta-regression is fundamentally different than combining each trial-level moderator analysis involving individual-level risk.

We have been developing methods that integrate findings across trials by pooling together individual-level data from each trial rather than using summary statistics. These methods of integrative data analysis for trial-level data are valuable in overcoming the limitations of individual trials to make inferences about moderators and mediators (Brown, Sloboda, Faggiano et al., 2013). Despite federal policies for data sharing, it is very rare for individual level data from a large number of trials to be included in the same analysis. We have had success in building collaborations between researchers to develop such multi-trial databases and synthesize the findings (Brown, Kellam, Kaupert, 2012; Perrino et al., in press), and believe such syntheses using individual-level data from multiple trials will provides significant advances in understanding who benefits from prevention programs and their underlying mechanisms.

Overall Summary and Implications

This review used an overview of meta-analyses approach to address three questions about research on the prevention of mental health and substance abuse problems. We summarize our findings below.

Do prevention and promotion programs have significant effects?

Findings from the 48 meta-analyses included in the review demonstrated small but significant effects to prevent each problem area that was included in this review, depression, anxiety, anti-social behavior and substance use. Significant effects were found for both continuous measures and dichotomous measures, such as diagnosed disorder. Further, the effects were sustained over time, with benefits on many outcomes lasting one or more years. Significant effects on these problem outcomes were found in meta-analyses of promotion programs as well as meta-analyses of programs that targeted these specific problem outcomes. Also, meta-analyses of programs that targeted a specific problem outcome found effects on other, related-outcomes, indicating that programs often affect multiple outcomes. These findings are consistent with the findings of the National Research Council/Institute of Medicine (2009) report concerning the benefits of preventive interventions for children and youth.

Are there common factors that moderate the effects of prevention and promotion programs?

Heterogeneity of effects is the rule rather than the exception in meta-analyses of prevention programs. There was consistent evidence from multiple meta-analyses for two significant moderators. Programs that used interactive strategies to promote use of program skills were more effective than programs that did not use such approaches. Also, individuals who were at higher risk for problem outcomes generally received more benefits from participation than those at lower risk. Other factors were also found to moderate the effects of prevention and promotion programs, such as differences in implementation and provider characteristics. However, there are fewer studies on the effects of the latter variables. These findings have implications for selecting and implementing programs that are most likely to have public health benefits.

The conceptual framework we used helped to identify gaps in the literature and articulate questions for future research. For example, considerably more work needs to address potential important moderators of program effects, including participant characteristics, quality of implementation, aspects of service delivery systems, provider characteristics, and contextual factors. For example, few meta-analyses examined whether poverty or ethnicity moderated program effects. Also, meta-analyses have not yielded any information on mediators of the effects of prevention and promotion programs. These questions are critical as prevention science moves beyond the question of whether prevention programs work to the question of whether prevention programs actually reduce the public health burden of the problems they are designed to reduce.

What methodological issues need to be addressed to advance research on prevention and promotion?

Several methodological issues were discussed in synthesizing findings across meta-analyses, including inconsistency of meta-analyses to account for the methodological quality of the studies and redundancy in the trials included in meta-analytic reviews. Several methodological challenges were identified and an alternative method of integrating findings across trials, which involves pooling individual-level data from trials rather than summary statistics, was discussed. We believe that this approach will be particularly important in addressing two critical issues in our conceptual model, mediation and moderation of the effects of preventive interventions.

Acknowledgments

We are grateful for support from the National Institute of Mental Health (R01MH040859), the National Institute of Drug Abuse (R01DA026874) and Grant # UL1Tr000460-01A1 to support the Miami Clinical and Translational Science Institute (Szapocznik). We also thank Ophelia Hernandez for her data management on this project and as well as our colleagues on the latter grant, Drs. George Howe, Tatiana Perrino, Hilda Pantin and William Beardslee.

References

- Ahn S, Becker BJ. Incorporating quality scores in meta-analysis. *J Educ Behav Stat.* 2011; 36:555–85.
- Aos, S.; Lieb, R.; Mayfield, J.; Miller, M.; Pennucci, A. Benefits and costs of prevention and Early intervention programs for youth. Olympia, WA: Washington State Institute for Public Policy; 2004. <http://www.wsipp.wa.gov/rptfiles/04-07-3901.pdf>
- Barlow J, Smailagic N, Ferriter M, Bennett C, Jones H. Group-based parent-training programmes for improving emotional and behavioural adjustment in children from birth to three years old. *Cochrane Database Syst Rev.* 2010 Art. No.: CD003680.
- Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J Pers Soc Psychol.* 1986; 51:1173–82. [PubMed: 3806354]
- Becker, BJ.; Schram, CM. Examining explanatory models through research synthesis. In: Cooper, H.; Hedges, LV., editors. *The handbook of research synthesis.* New York, NY, US: Russell Sage Foundation; 1994. p. 357–81.
- Beelmann A, Lösel F. Child social skills training in developmental crime prevention: Effects on antisocial behavior and social competence. *Psicothema.* 2006; 18:603–10. [PubMed: 17296094]
- Berkel C, Mauricio A, Schoenfelder E, Sandler I. Putting the pieces together: An integrated model of program implementation. *Prev Sci.* 2011; 12:23–33. [PubMed: 20890725]
- Brown CH, Kellam SG, Kaupert S, Muthén BO, Wang W, et al. Partnerships for the design, conduct, and analysis of effectiveness, and implementation research: Experiences of the Prevention Science and Methodology Group. *Adm Policy Ment Health.* 2012; 39:301–16. [PubMed: 22160786]
- Brown CH, Sloboda Z, Faggiano F, Teasdale B, Keller F, et al. Methods for synthesizing findings on moderation effects across multiple randomized trials. *Prev Sci.* 2013; 14:144–56. [PubMed: 21360061]
- Brunwasser SM, Gillham JE, Kim ES. A meta-analytic review of the Penn Resiliency Program's effect on depressive symptoms. *J Consult Clin Psychol.* 2009; 77:1042–54. [PubMed: 19968381]
- Burrus B, Leeks KD, Sipe TA, Dolina S, Soler R, et al. Person-to-person interventions targeted to parents and other caregivers to improve adolescent health: A community guide systematic review. *Am J Prev Med.* 2012; 42:316–26. [PubMed: 22341170]
- Carey KB, Scott-Sheldon LAJ, Carey MP, DeMartini KS. Individual-level interventions to reduce college student drinking: A meta-analytic review. *Addict Behav.* 2007; 32:2469–94. [PubMed: 17590277]
- Carey KB, Scott-Sheldon LAJ, Elliott JC, Garey L, Carey MP. Face-to-face versus computer-delivered alcohol interventions for college drinkers: A meta-analytic review, 1998 to 2010. *Clin Psychol Rev.* 2012; 32:690–703. [PubMed: 23022767]
- Cooper H, Koenka AC. The overview of reviews: Unique challenges and opportunities when research syntheses are the principal elements of new integrative scholarship. *Am Psychol.* 2012; 67:446–62. [PubMed: 22352742]
- Cuijpers P. Peer-led and adult-led school drug prevention: a meta-analytic comparison. *J Drug Educ.* 2002; 32:107–19. [PubMed: 12206061]
- Currier JM, Holland JM, Neimeyer RA. The Effectiveness of bereavement interventions with children: A meta-analytic review of controlled outcome research. *J Clin Child Adolesc Psychol.* 2007; 36:253–59. [PubMed: 17484697]
- Dekovi M, Slagt MI, Asscher JJ, Boendermaker L, Eichelsheim VI, Prinzie P. Effects of early prevention programs on adult criminal offending: A meta-analysis. *Clin Psychol Rev.* 2011; 31:532–44. [PubMed: 21239100]

- Derzon, J.; Jimerson, SR.; Furlong, JJ. How effective are school-based violence prevention programs in preventing and reducing violence and other antisocial behaviors? A meta-analysis. In: Jimerson, SR.; Furlong, JJ., editors. *The handbook of school violence and school safety: From research to practice*. Mahwah, NJ: Lawrence Earlbaum; 2006. p. 429-41.
- DuBois DL, Portillo N, Rhodes JE, Silverthorn N, Valentine JC. How effective are mentoring programs for youth? A systematic assessment of the evidence. *Psychol Sci Public Interest*. 2011; 12:57–91.
- Durlak J, DuPre E. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol*. 2008; 41:327–50. [PubMed: 18322790]
- Durlak J, Weissberg R, Pachan M. A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *Am J Community Psychol*. 2010; 45:294–309. [PubMed: 20300825]
- Durlak JA, Weissberg RP, Dymnicki AB, Taylor RD, Schellinger KB. The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Dev*. 2011; 82:405–32. [PubMed: 21291449]
- Fachini A, Aliane PP, Martinez EZ, Furtado EF. Efficacy of brief alcohol screening intervention for college students (BASICS): a meta-analysis of randomized controlled trials. *Subst Abuse Treat Prev Policy*. 2012; 7 <http://www.substanceabusepolicy.com/content/7/1/40>.
- Fackrell TA, Hawkins AJ, Kay NM. How effectiveness of court-affiliated divorcing parents education programs? A meta-analytic study. *Fam Court Rev*. 2011; 49:107–19.
- Faggiano F, Vigna-Taglianti FD, Versino E, Zambon A, Borraccino A, Lemma P. School-based prevention for illicit drugs use: A systematic review. *Prev Med*. 2008; 46:385–96. [PubMed: 18258289]
- Farrington DP, Welsh BC. Family-based prevention of offending: A meta-analysis. *Aust N Z J Criminol*. 2003; 36:127–51.
- Fisak B Jr, Richard D, Mann A. The prevention of child and adolescent anxiety: A meta-analytic review. *Prev Sci*. 2011; 12:255–68. [PubMed: 21437675]
- Fixsen, DL.; Naoom, SF.; Blasé, KA.; Friedman, RM.; Wallace, F. *Implementation research: A synthesis of the literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231); 2005.
- Flay B, Biglan A, Boruch R, Castro F, Gottfredson D, et al. Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prev Sci*. 2005; 6:151–75. [PubMed: 16365954]
- Gleser, LJ.; Olkin, I. Stochastically dependent effect sizes. In: Cooper, H.; Hedges, LV., editors. *The handbook of research synthesis*. New York, NY: Russell Sage Foundation; 1994. p. 339-55.
- Gottfredson D, Wilson D. Characteristics of effective school-based substance abuse prevention. *Prev Sci*. 2003; 4:27–38. [PubMed: 12611417]
- Hedges, LV.; Olkin, I. *Statistical Methods for Meta-analysis*. Boston, MA: Academic Press; 1985.
- Higgins, JP., editor. *Cochrane handbook for systematic reviews of interventions*. Vol. 5. Chichester: Wiley-Blackwell; 2008.
- Horowitz JL, Garber J. The prevention of depressive symptoms in children and adolescents: A meta-analytic review. *J Consult Clin Psychol*. 2006; 74:401–15. [PubMed: 16822098]
- Hwang MS, Yeagley KL, Petosa R. A meta-analysis of adolescent psychosocial smoking prevention programs published between 1978 and 1997 in the United States. *Health Educ Behav*. 2004; 31:702–19. [PubMed: 15539543]
- Isensee B, Hanewinkel R. Meta-analysis on the effects of the smoke-free class competition on smoking prevention in adolescents. *Eur Addict Res*. 2012; 18:110–15. [PubMed: 22285973]
- Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Control Clin Trials*. 1996; 17:1–12. [PubMed: 8721797]
- Juni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ*. 2001; 323:42–46. [PubMed: 11440947]

- Kaminski J, Valle L, Filene J, Boyle C. A meta-analytic review of components associated with parent training program effectiveness. *J Abnorm Child Psychol*. 2008; 36:567–89. [PubMed: 18205039]
- Kessler RC, Avenevoli S, McLaughlin KA, Green JG, Lakoma MD, et al. Lifetime co-morbidity of DSM-IV disorders in the US National Comorbidity Survey Replication Adolescent Supplement (NCS-A). *Psychol Med*. 2012; 42:1997–2010. [PubMed: 22273480]
- Knerr W, Gardner F, Kluver L. Improving positive parenting skills and reducing harsh and abusive parenting in low- and middle-income countries: A systematic review. *Prev Sci*. 2013; 14:352–363. [PubMed: 23315023]
- Larun, L.; Nordheim, L.; Ekeland, E.; Hagen, K.; Heian, F. Exercise for preventing and treating anxiety and depression in children and young people (Review). The Cochrane Library. 2009. <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD004691.pub2/pdf>
- Lipsey, MW.; Puzio, K.; Yun, C.; Hebert, MA.; Steinka-Fry, K., et al. Translating the statistical representation of the effects of education interventions into more readily interpretable forms. Washington, DC: National Center for Special Education Research, Institute of Education Sciences, US Department of Education; 2012. NCSER 2013-3000
- Lundahl B, Harris N. Delivering parent training to families at risk to abuse: Lessons from three meta-analyses. *APSAC Advis*. 2006; 18:7–11.
- MacArthur G, Kipping R, James W, Chittleborough C, Lingam R, et al. Individual-, family-, and school-level interventions for preventing multiple risk behaviours in individuals aged 8 to 25 years. *Cochrane Database Syst Rev*. 2012; 6 Art. No.: CD009927.
- MacKinnon, DP. Introduction to statistical mediation analysis. Mahwah, NJ: Erlbaum; 2008.
- Manning M, Homel R, Smith C. A meta-analysis of the effects of early developmental prevention programs in at-risk populations on non-health outcomes in adolescence. *Child Youth Serv Rev*. 2010; 32:506–19.
- Merry SN, Hetrick SE, Cox GR, Brudevold-Iversen T, Bir JJ, McDowell H. Cochrane Review: Psychological and educational interventions for preventing depression in children and adolescents. *Evid Based Child Health*. 2012; 7:1409–685.
- Mikton C, Butchart A. Child maltreatment prevention: a systematic review of reviews. *Bull World Health Organ*. 2009; 87:353–61. [PubMed: 19551253]
- Mytton J, DiGiuseppi C, Gough D, Taylor R, Logan S. School-based secondary prevention programmes for preventing violence. *Cochrane Database Syst Rev*. 2006 Art. No.: CD004606.
- Nation M, Crusto C, Wandersman A, Kumpfer KL, Seybolt D, et al. What works in prevention: Principles of effective prevention programs. *Am Psychol*. 2003; 58:449–56. [PubMed: 12971191]
- National Research Council and Institute of Medicine. Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth, and Young Adults: Research Advances and Promising Interventions. In: O'Connell, ME.; Boat, T.; Warner, KE., editors. Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities. Board on Children, Youth, and Families, Division of Behavioral and Social Sciences and Education. Washington, D.C.: National Academics Press; 2009.
- Nelson G, Westhues A, MacLeod J. A meta-analysis of longitudinal research on preschool prevention programs for children. *Prev Treat*. 2003; 6:31a.
- Nowak C, Heinrichs N. A comprehensive meta-analysis of Triple P-Positive Parenting Program using hierarchical linear modeling: Effectiveness and moderating variables. *J Clin Child Fam Psychol Rev*. 2008; 11:114–144.
- Park-Higginson H-K, Perumean-Chaney SE, Bartolucci AA, Grimley DM, Singh KP. The evaluation of school-based violence prevention programs: A meta-analysis*. *J Sch Health*. 2008; 78:465–79. [PubMed: 18786039]
- Perrino T, Howe G, Sperling A, Beardslee W, Sander I, et al. Advancing science through collaborative data sharing and synthesis. *Perspect Psychol Sci*. in press.
- Peterson AV, Kealey KA, Mann SL, Marek PM, Sarason IG. Hutchinson Smoking Prevention Project: Long-term randomized trial in school-based tobacco use prevention—results on smoking. *J Natl Cancer Inst*. 2000; 92:1979–91. [PubMed: 11121460]
- Piquero A, Farrington D, Welsh B, Tremblay R, Jennings W. Effects of early family/parent training programs on antisocial behavior and delinquency. *J Exp Criminol*. 2009; 5:83–120.

- Piquero, A.; Jennings, W.; Farrington, DP. Self-control interventions for children under age 10 for improving self-control and delinquency and problem behaviours. Oslo: The Campbell Collaboration; 2010.
- Porath-Waller AJ, Beasley E, Beirness DJ. A meta-analytic review of school-based prevention for cannabis use. *Health Educ Behav.* 2010; 37:709–23. [PubMed: 20522782]
- Preacher K, Hayes A. SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behav Res Methods Instrum Comput.* 2004; 36:717–31. [PubMed: 15641418]
- Raudenbush SW, Becker BJ, Kalaian H. Modeling multivariate effect sizes. *Psychol Bull.* 1988; 103:111–20.
- Regehr C, Glancy D, Pitts A. Interventions to reduce stress in university students: A review and meta-analysis. *J Affect Disord.* 2013; 148:1–11. [PubMed: 23246209]
- Rosner R, Kruse J, Hagl M. A meta-analysis of interventions for bereaved children and adolescents. *Death Stud.* 2010; 34:99–136. [PubMed: 24479177]
- Scott-Sheldon LAJ, Terry DL, Carey KB, Garey L, Carey MP. Efficacy of expectancy challenge interventions to reduce college student drinking: A meta-analytic review. *Psychol Addict Behav.* 2012; 26:393–405. [PubMed: 22428862]
- Seigenthaler E, Munder T, Egger M. Effect of preventive interventions in mentally ill parents on mental health of the offspring: Systematic review and meta-analysis. *J of the Am Acad Child and Adol Psychiatry.* 2012; 51:8–17.
- Shadish WR. Meta-analysis and the exploration of causal mediating processes: A primer of examples, methods, and issues. *Psychol Methods.* 1996; 1:47–65.
- Shadish WR, Sweeney RB. Mediators and moderators in meta-analysis: There's a reason we don't let dodo birds tell us which psychotherapies should have prizes. *J Consult Clin Psychol.* 1991; 59:883–93. [PubMed: 1837794]
- Shamblen SR, Derzon JH. A preliminary study of the population-adjusted effectiveness of substance abuse prevention programming: Towards making IOM types comparable. *J Primary Prev.* 2009; 30:89–107.
- Seigenthaler E, Munder T, Egger M. Effect of preventive interventions in mentally ill parents on mental health of the offspring: Systematic review and meta-analysis. *J of the Am Acad Child and Adol Psychiatry.* 2012; 51:8–17.
- Soole DW, Mazerolle L, Rombouts S. School-based drug prevention programs: A review of what works. *Aust N Z J Criminol.* 2008; 41:259–86.
- Spoth R, Rohrbach L, Greenberg M, Leaf P, Brown CH, et al. Addressing core challenges for the next generation of Type 2 translation research and systems: The translation science to population impact (TSci Impact) framework. *Prev Sci.* 2013; 14:319–51. [PubMed: 23430579]
- Stathakos P, Roehrl B. The effectiveness of intervention programmes for children of divorce - a Meta-Analysis. *Int J Ment Health Promot.* 2003; 5:31–37.
- Stice E, Shaw H, Bohon C, Marti CN, Rohde P. A meta-analytic review of depression prevention programs for children and adolescents: Factors that predict magnitude of intervention effects. *J Consult Clin Psychol.* 2009; 77:486–503. [PubMed: 19485590]
- Sweet MA, Appelbaum MI. Is home visiting an effective strategy? A meta-analytic review of home visiting programs for families with young children. *Child Dev.* 2004; 75:1435–56. [PubMed: 15369524]
- Tein J-Y, Sandler IN, MacKinnon DP, Wolchik SA. How did it work? Who did it work for? Mediation and mediated moderation of a preventive intervention for children of divorce. *J Consult Clin Psychol.* 2004; 72:617–24. [PubMed: 15301646]
- Teubert D, Pinquart M. A meta-analytic review on the prevention of symptoms of anxiety in children and adolescents. *J Anxiety Disord.* 2011; 25:1046–59.
- Thomas R, Perera R. School-based programmes for preventing smoking. *Cochrane Database Syst Rev.* 2006; 3 Art. No.: CD001293.
- Tobler NS, Roona MR, Ochshorn P, Marshall DG, Streke AV, Stackpole KM. School-based adolescent drug prevention programs: 1998 meta-analysis. *J Prim Prev.* 2000; 20:275–336.
- Tolan P, Henry D, Schoeny M, Bass A. Mentoring interventions to affect juvenile delinquency and associated problems. *Campbell Syst Rev.* 2008; 16

- Weiss RP, Bloom HS, Thomas B. A conceptual framework for studying the sources of variation in program effects: MDRC. 2013
- Wilson SJ, Lipsey MW. School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *Am J Prev Med.* 2007; 33:S130–S43. [PubMed: 17675014]
- Yoshikawa H, Aber JL, Beardslee WR. The effects of poverty on the mental, emotional, and behavioral health of children and youth: Implications for prevention. *Am Psychol.* 2012; 67:272–84. [PubMed: 22583341]

Table 1

Meta-analyses included in the overview

Study	Trials	Sample size	Age	Intervention Type
Depression				
Brunwasser et al., 2009	17	N = 2,498	8-18 years	Cognitive behavioral program (CB) (Penn Resiliency Program)
Horowitz & Garber, 2006	30	NR	14-21 years	CB and educational programs
Larun et al., 2009	16	N = 1,191	11-19 years	Exercise programs
Merry et al., 2012	53	N = 14,406	5-19 years	Psychological and educational programs
Stice et al., 2009	47	N = 16872	10 -19 years	CB and psychoeducational programs
Anxiety				
Fisak et al., 2011	35	N=7,735	1-18 years	Mostly CB programs
Regehr et al., 2013	24	N = 1431	17 -21 years College students	Mindfulness, relaxation, CB, art, and computer-based programs
Teubert et al., 2011	65	N=15,713	3-17 years	Mostly CB programs
Aggression/Antisocial behavior/ Violence				
Beelman & Losel, 2006	84	N = 16,723	birth – 18 years	CB programs and counseling/psychotherapy Majority group format/school-based
Dekovic et al., 2011	9	N = 3611	birth – 12 years	Early intervention programs (home visit, parent support groups, preschool, school-based programs)
Derzon et al., 2006	83	Range (n = <51-500)	5 - 18 years	School-based programs
Farrington & Welsh, 2003	40	N = 12,616	1 – 18 years	Family-based programs (Home visiting, preschool/daycare programs, parent training, school-based programs).
Mytton et al., 2006	56	NR	At-risk children grades K –	School-based programs
Park-Higgerson et al., 2008	26	N = 10,773	5 – 18 years	School-based programs
Piquero et al., 2009	55	N = 9958	birth – 8 years	Family/parent training and home visitation programs
Piquero et al., 2010	34	NR	3 – 10 years	Programs to increase self-control. Generally group programs conducted in school settings.
Wilson & Lipsey, 2007	249	NR	5 -18 years	School-based programs Behavioral, cognitive, social skills training, counseling, peer mediation, parent training

Study	Trials	Sample size	Age	Intervention Type
Drug Use				
Cuijpers, 2002	12	N = 12,400	11 - 18 years	Peer-led, teacher-led and expert-led school-based interventions
Faggiano et al., 2008	32	N = 46,539	Mostly 6 th - 7 th grade	School-based programs
Gottfredson & Wilson, 2003	94	NR	NR	School-based programs
Porath-Waller et al., 2010	15	N = 15,571	12-19 years	School-based programs
Soole et al., 2008	12	N = 2346	NR	School-based programs
Tobler et al., 2000	207 overall 93 high quality	NR	<18 years	School-based programs
Alcohol Use				
Carey et al. 2007	62	N = 13,750	18 - 26 years College	Individual-level programs
Carey et al., 2012	46	N = 27,460	17 -22 years College	Face-to-face versus computer-delivered programs
Fachini et al., 2012	18	N = 6233	17 - 22 years College	BASICS (brief alcohol screening intervention)
Scott-Sheldon et al., 2012	14	N = 1415	19 - 21 years	NR
Smoking				
Hwang, Yeagley, & Petosa, 2004 2004	65	NR	10 - 18 years	School and community-based. Social influence, cognitive behavior, life skill modalities.
Isensee & Hanewinkle 2012	5 (1835-4372)	N = 16,302	11 - 17 years	Smoke Free Class Competition
Thomas & Perera, 2006	94 school-based 23 considered most valid	NR	5 - 18 years	School-based programs, some included family and community components. Behavioral, social influence, educational, social competence approaches
School Based				
Durlak et al., 2011	213	N = 270, 034	5 – 18 years	School-based programs
After School				
Durlak et al., 2010	75	NR	5 – 18 years	Social skills programs Boys and Girls, 4-H Clubs, community programs
Mentoring				
Dubois et al., 2011	73	NR	5 – 18 years	Mentoring
Tolan et al., 2008	39	N = 9253	18 years	Mentoring programs
Parent Training				

Study	Trials	Sample size	Age	Intervention Type
Barlow et al., 2010	8	N = 410	birth - 3 years	Group-based programs parenting programs
Burrus et al. 2012	16	NR	13 - 18 years	Interventions targeting parent/caregiver behaviors
Kaminski et al., 2008	77	NR	birth - 7 years	Parent training
Nowak & Heinrichs, 2008	55	N = 12884	2 - 16 years	Multiple levels and multiple formats of the Triple P parenting program
Early Developmental Programs				
Manning et al., 2010	11	N = 3,285	birth - 5 years	Home visit, parent training, preschool, family-support, childcare programs
Nelson et al., 2003	34	N = 10,805	birth - 5 years	Home visiting, parent training, social support, education
Sweet & Applebaum, 2004	60	NR	birth - 8 years	Parent education, social support, counseling, leadership and advocacy training, adult basic early childhood education, case management services
Disrupted Families				
Currier et al., 2007	13	N = 783	8 - 13.6 years	Mostly group interventions, strong psychoeducational component
Fackrell et al., 2011	28	NR	NR	Court-affiliated parent education programs
Rosner et al., 2010	15	N = 812	birth - 18 years	Support, psychoeducational, CB, family-focused groups
Siegenthaler et al., 2012	13	N = 1,490	Birth - adolescence	CB and psycho-educational programs to improve parenting skills, Adolescent programs to increase knowledge and resilience
Stathakos & Roehrle, 2003	24	N = 1,615	3-14 years	NR

Note. CB = cognitive behavioral

Table 2

Prevention Impact Across Diverse Prevention Targets and Outcomes Based on Effect Sizes for Continuous Outcome Measures and Odds Ratios/Relative Risks and Risk Differences for Discrete Outcome Measures

	Effects for meta-analyses with a specific problem focus	Effects for all meta-analyses that report on the problem
Depression	ES= .17 (sd=.05,n=6) RD= .052 (sd=.05,n=5)	ES= .18 (sd=.06,n=9) RD= .052 (sd=.05,n=5)
Anxiety	ES= .43 (sd=.31,n=6)	ES= .43 (sd=.31,n=6)
Crime/Anti-social Behavior/Aggression	ES= .19 (sd=.17,n=29) OR/RR= 1.10 (sd=.01,n=2)	ES= .21 (sd=.18,n=31) OR/RR=1.14 (sd=.05,n=2)
Smoking	ES=.17 (sd=.06,n=6) OR/RR= 1.07 (n=1)	ES= .19 (sd=.08,n=8) OR/RR= 1.07 (n=1)
Alcohol	ES= .12 (sd=.07,n=16)	ES= .12 (sd=.07,n=16)
Other Substance Use	ES = .15 (sd=.16,n=27) OR/RR=1.19 (sd=.15,n=3)	ES=.13 (sd=.15, n=17)
Any Substance use/abuse	ES=.14(sd=.12,n=39) OR/RR=1.16(sd=.13,n=4)	ES=.14 (sd=.11,n=45) OR/RR=1.14 (sd=.12,n=5)
Promotion	ES= .30 (sd=.17,n=27) OR/RR= 1.15 (sd=.045,n=3)	

Note: ES = effect size; RD= risk difference; OR/RR= odds ratio/risk ratio

Table 3

Scores on outcomes of average person who participated in interactive, skill-building programs vs. non-interactive programs

Meta-analysis	Interactive, Skill-Building Programs Outcome Cohen's U_3	Non-interactive Programs Outcome Cohen's U_3
Fisak et al., (2011) ¹ . Prevention of anxiety programs	FRIENDS program Anxiety Cohen's $U_3 = .60$	Other programs Anxiety Cohen's $U_3 = .55$
Beelman & Losel (2006) ² Prevention of antisocial behaviors programs	Cognitive-behavioral Antisocial behavior Cohen's $U_3 = .69$ at post-test, Cohen's $U_3 = .69$ at follow-up	Behavioral only Antisocial behavior Cohen's $U_3 = .56$ at post-test, Cohen's $U_3 = .55$ at follow-up Cognitive only Cohen's $U_3 = .56$ at post-test Cohen's $U_3 = .48$ at follow-up Other approaches Cohen's $U_3 = .65$ at post-test, Cohen's $U_3 = .56$ at follow-up
Prevention of substance use programs (Tobler et al. 2000) ³	All interactive programs Substance use Cohen's $U_3 = .55$	All non-interactive programs Substance use Cohen's $U_3 = .52$
School-based SEL programs (Durlak et al., 2011) ⁴	Used SAFE⁷ practices Emotional distress Cohen's $U_3 = .61$ Conduct problems Cohen's $U_3 = .59$	Did not use SAFE practices Emotional distress Cohen's $U_3 = .57$ Conduct problems Cohen's $U_3 = .56$
After-school programs (Durlak et al., 2010) ⁵	Used SAFE practices Conduct problems Cohen's $U_3 = .62$ Substance use Cohen's $U_3 = .56$	Did not use SAFE practices Conduct problems Cohen's $U_3 = .53$ Substance use Cohen's $U_3 = .51$
Parenting programs (Kaminski, et al. 2008) ⁶	Focused on positive interactions with child(ren) Externalizing behavior Cohen's $U_3 = .64$ Included practice with own child(ren) Externalizing behavior Cohen's $U_3 = .75$	Did not focus on positive interactions with child(ren) Externalizing behavior Cohen's $U_3 = .57$ Did not include practice with own child(ren) Externalizing behavior Cohen's $U_3 = .55$

Meta-analysis	Interactive, Skill-Building Programs Outcome Cohen's U_3	Non-interactive Programs Outcome Cohen's U_3
	Focused on time out Externalizing behavior Cohen's $U_3 = .71$	Did not focus on time out Externalizing behavior Cohen's $U_3 = .56$

Note. Cohen's U_3 were derived from effect sizes that were presented in the following tables of these meta-analyses

¹ Table 2;

² Table 3;

³ Table 9;

⁴ Table 3;

⁵ Table 3;

⁶ Table 7.

⁷ SAFE practices are defined by Durlak et al. (2010) as programs that used pedagogical approaches that are sequenced, active, focused and explicit.