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

Irwin Sandler,1 Sharlene A. Wolchik,1 Gracelyn Cruden,2 Nicole E. Mahrer,1 Soyeon Ahn,3 Ahnalee Brincks,2 and C. Hendricks Brown2

1Department of Psychology and Prevention Research Center, Arizona State University, Tempe, Arizona 85287-6005; email: irwin.sandler@asu.edu
2Department of Psychiatry and Behavioral Sciences, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60611
3College of Education, University of Miami, Coral Gables, Florida 33146

Keywords
prevention, meta-analysis, mental health, substance use

Abstract
This review 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 changes that reduce depression, anxiety, antisocial behavior, and substance use. Furthermore, the results were 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 (NRC/IOM) 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 NRC/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 & 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 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 the effects of different prevention programs on the common outcome of economic benefit (Aos et al. 2004).

This article 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, antisocial behavior (including delinquency and violence), and substance use. Two categories of meta-analyses are included in this review. One category focuses on the results of a range of programs targeting one of the specific problems or disorders listed above. The second category focuses on the results of programs that are designed to promote healthy development (e.g., through mechanisms such as teaching social and emotional skills, and 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, antisocial 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 results obtained, whether the effects are heterogeneous or homogeneous, and the 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 the heterogeneity of the results (i.e., moderators) across meta-analyses and use this framework to discuss commonalities and differences in findings across approaches to prevention, as well as 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 search of literature in the following databases: PsycINFO®, PubMed, and Google scholar, searching for relevant meta-analyses published between 2000 and 2013. Key words
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; Center for Evidence-based Policy; and 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 and 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 tables of contents from journals not included in the search results to identify additional meta-analyses. Unpublished meta-analyses were included only if they were on the websites of the organizations described above (e.g., Cochrane, Collaborative for Academic, Social, and Emotional Learning).

The following inclusion criteria were used:

- The meta-analysis was published between 2000 and May 2013.
- The study population was children, adolescents, and/or young adults up to age 26.
- 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 (ORs). Literature reviews and systematic reviews were not included unless they also contained a quantitative analysis of effect sizes across trials.
- 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, the 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 study. 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, all were examined for relevance and nonredundancy by two coders. A total of 48 (71.6%) met our criteria for relevance and nonredundancy and are presented in Table 1. The fact that 50% of these studies 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 PREVENTION OF 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 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 et al. 2009) and exercise programs (Larun et al. 2009). The meta-analyses included studies with youth and young adults, ages 5 through 21, and analyzed effects.
<table>
<thead>
<tr>
<th>Table 1</th>
<th>Meta-analyses included in this overview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
<td>Trials</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Depression</strong></td>
<td></td>
</tr>
<tr>
<td>Brunwasser et al. (2009)</td>
<td>17</td>
</tr>
<tr>
<td>Horowitz &amp; Garber (2006)</td>
<td>30</td>
</tr>
<tr>
<td>Larun et al. (2009)</td>
<td>16</td>
</tr>
<tr>
<td>Merry et al. (2012)</td>
<td>53</td>
</tr>
<tr>
<td>Stice et al. (2009)</td>
<td>47</td>
</tr>
<tr>
<td><strong>Anxiety</strong></td>
<td></td>
</tr>
<tr>
<td>Fisak et al. (2011)</td>
<td>35</td>
</tr>
<tr>
<td>Regehr et al. (2013)</td>
<td>24</td>
</tr>
<tr>
<td>Teubert &amp; Pinquart (2011)</td>
<td>65</td>
</tr>
<tr>
<td><strong>Aggression/antisocial behavior/violence</strong></td>
<td></td>
</tr>
<tr>
<td>Beelmann &amp; Lösel (2006)</td>
<td>84</td>
</tr>
<tr>
<td>Deković et al. (2011)</td>
<td>9</td>
</tr>
<tr>
<td>Derzon et al. (2006)</td>
<td>83</td>
</tr>
<tr>
<td>Farrington &amp; Welsh (2003)</td>
<td>40</td>
</tr>
<tr>
<td>Mytton et al. (2006)</td>
<td>56</td>
</tr>
<tr>
<td>Park-Higgerson et al. (2008)</td>
<td>26</td>
</tr>
<tr>
<td>Piquero et al. (2009)</td>
<td>55</td>
</tr>
<tr>
<td>Piquero et al. (2010)</td>
<td>34</td>
</tr>
<tr>
<td>Wilson &amp; Lipsey (2007)</td>
<td>249</td>
</tr>
<tr>
<td><strong>Drug use</strong></td>
<td></td>
</tr>
<tr>
<td>Cuijpers (2002)</td>
<td>12</td>
</tr>
<tr>
<td>Faggiano et al. (2008)</td>
<td>32</td>
</tr>
<tr>
<td>Gottfredson &amp; Wilson (2003)</td>
<td>94</td>
</tr>
<tr>
<td>Porath-Waller et al. (2010)</td>
<td>15</td>
</tr>
<tr>
<td>Soole et al. (2008)</td>
<td>12</td>
</tr>
</tbody>
</table>

(Continued)
Table 1 (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Trials</th>
<th>Sample size</th>
<th>Age in years</th>
<th>Intervention type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tobler et al. (2000)</td>
<td>207 overall 93 high quality</td>
<td>No record</td>
<td>&lt;18</td>
<td>School-based programs</td>
</tr>
<tr>
<td>Alcohol use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carey et al. (2007)</td>
<td>62</td>
<td>N = 13,750</td>
<td>18–26</td>
<td>Individual-level programs</td>
</tr>
<tr>
<td>Carey et al. (2012)</td>
<td>46</td>
<td>N = 27,460</td>
<td>17–22</td>
<td>Face-to-face versus computer-delivered programs</td>
</tr>
<tr>
<td>Fachini et al. (2012)</td>
<td>18</td>
<td>N = 6,233</td>
<td>17–22</td>
<td>BASICS (brief alcohol screening intervention)</td>
</tr>
<tr>
<td>Scott-Sheldon et al. (2012)</td>
<td>14</td>
<td>N = 1,415</td>
<td>19–21</td>
<td>No record</td>
</tr>
<tr>
<td>Smoking</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thomas &amp; Perera (2006)</td>
<td>94 school-based 23 considered most valid</td>
<td>No record</td>
<td>5–18</td>
<td>School-based programs, some included family and community components. Behavioral, social influence, educational, social competence approaches</td>
</tr>
<tr>
<td>School based</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durlak et al. (2011)</td>
<td>213</td>
<td>N = 270,034</td>
<td>5–18</td>
<td>School-based programs</td>
</tr>
<tr>
<td>After school</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durlak et al. (2010)</td>
<td>75</td>
<td>No record</td>
<td>5–18</td>
<td>Social skills programs Boys &amp; Girls Clubs, 4-H programs, community programs</td>
</tr>
<tr>
<td>Mentoring</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dubois et al. (2011)</td>
<td>73</td>
<td>No record</td>
<td>5–18</td>
<td>Mentoring</td>
</tr>
<tr>
<td>Tolan et al. (2008)</td>
<td>39</td>
<td>N = 9,253</td>
<td>≤18</td>
<td>Mentoring programs</td>
</tr>
<tr>
<td>Parent training</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barlow et al. (2010)</td>
<td>8</td>
<td>N = 410</td>
<td>Birth to 3</td>
<td>Group-based programs, parenting programs</td>
</tr>
<tr>
<td>Burrus et al. (2012)</td>
<td>16</td>
<td>No record</td>
<td>13–18</td>
<td>Interventions targeting parent/caregiver behaviors</td>
</tr>
<tr>
<td>Kaminski et al. (2008)</td>
<td>77</td>
<td>No record</td>
<td>Birth to 7</td>
<td>Parent training</td>
</tr>
<tr>
<td>Nowak &amp; Heinrichs (2008)</td>
<td>55</td>
<td>N = 12,884</td>
<td>2–16</td>
<td>Multiple levels and multiple formats of the Triple P parenting program</td>
</tr>
<tr>
<td>Early developmental</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>programs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manning et al. (2010)</td>
<td>11</td>
<td>N = 3,285</td>
<td>Birth to 5</td>
<td>Home visit, parent training, preschool, family-support, childcare programs</td>
</tr>
<tr>
<td>Nelson et al. (2003)</td>
<td>34</td>
<td>N = 10,805</td>
<td>Birth to 5</td>
<td>Home visitation, parent training, social support, education</td>
</tr>
</tbody>
</table>

(Continued)
Table 1 (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Trials</th>
<th>Sample size</th>
<th>Age in years</th>
<th>Intervention type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweet &amp; Appelbaum (2004)</td>
<td>60</td>
<td>No record</td>
<td>Birth to 8</td>
<td>Parent education, social support, counseling, leadership and advocacy training, adult basic early childhood education, case management services</td>
</tr>
<tr>
<td>Disrupted families</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currier et al. (2007)</td>
<td>13</td>
<td>N = 783</td>
<td>8–13.6</td>
<td>Mostly group interventions, strong psychoeducational component</td>
</tr>
<tr>
<td>Fackrell et al. (2011)</td>
<td>28</td>
<td>No record</td>
<td>No record</td>
<td>Court-affiliated parent-education programs</td>
</tr>
<tr>
<td>Rosner et al. (2010)</td>
<td>15</td>
<td>N = 812</td>
<td>Birth to 18</td>
<td>Support, psychoeducational, CB, family-focused groups</td>
</tr>
<tr>
<td>Siegenthaler et al. (2012)</td>
<td>13</td>
<td>N = 1,490</td>
<td>Birth to adolescence</td>
<td>CB and psychoeducational programs to improve parenting skills Adolescent programs to increase knowledge and resilience</td>
</tr>
<tr>
<td>Stathakos &amp; Roehrle (2003)</td>
<td>24</td>
<td>N = 1,615</td>
<td>3–14</td>
<td>No record</td>
</tr>
</tbody>
</table>

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 postintervention. 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 postintervention 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 subgroup 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 & 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 the diagnosis of depression that lasted 9 months after the programs; the 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 a 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 & 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 it did not report on the maintenance of program effects.

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

Nearly all of the effects at posttest 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 to 120), neither meta-analysis found that length moderated program effects. At posttest, Teubert & 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 at posttest 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 & 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 students.

Aggression, Antisocial 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 (Beelmann & Lösel 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 (Deković 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 posttest for universal programs, selective programs, and programs conducted in special schools or classes; nonsignificant 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 subpopulations. Two meta-analyses (Beelmann & Lösel 2006, Piquero et al. 2010) reported that interventions focused on social skills or a broad range of self-control strategies had small but significant effects
on antisocial behavior at posttest and follow-up. 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 posttest and follow-up for programs that taught social or relationship skills and those that taught nonresponse to provocative situations. Derzon et al. (2006) studied a wider range of antisocial 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, but this may be because they included fewer studies than the other meta-analyses. Although these meta-analyses focused on aggressive and disruptive behavior, four of them (Beelmann & Lösel 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 improved 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 (Beelmann & Lösel 2006; Park-Higgerson et al. 2008; Piquero et al. 2009, 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 (Beelmann & Lösel 2006, Wilson & Lipsey 2007). Wilson & Lipsey (2007) also reported that better implementation was related to larger effects for selective programs and programs delivered in special classes. Beelmann & Lösel (2006) reported that the effects for cognitive-behavioral approaches were moderate and significant at posttest 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 nonresponse to provocative situations.

Two meta-analyses reported on family-based programs to prevent antisocial behavior. Farrington & 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 antisocial behavior and delinquency were found at posttest and follow-up for parent training, home visits, preschool/day care, and multisystemic therapy; the effect for school-based programs that included a parenting component was not significant. However, these subgroup analyses were based on relatively few trials, and the length of follow-up and the targeted population varied greatly across type of approach. Thus, these findings should be viewed with caution. Focusing on family/parent-training programs for children aged five 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 United States versus other countries and those with smaller versus larger samples.

Deković et al. (2011) found small, significant, and heterogeneous effects to reduce criminal behavior of the child when they reached adulthood 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-socioeconomic status (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 programs that focused on social and behavioral skills versus family support or academic skills. Deković et al. (2011) also found small, significant effects to improve positive outcomes, such as the child finishing high school and being employed when they reached adulthood.
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.

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

Tobler et al.’s (2000) review, which included 207 trials, is the most comprehensive. Programs were coded as being either interactive (defined as providing contact among participants; opportunities to exchange ideas; and teaching refusal skills and interpersonal skills; or including system-wide school and community change strategies) or noninteractive 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 1 to 12 months after the intervention were found for programs that involved interactive, skill-building strategies versus noninteractive strategies. The effects were heterogeneous for both interactive and noninteractive 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 the selection of the 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 skills but were nonsignificant and homogeneous for all types of noninteractive 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-Waller 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-Waller et al. (2010) found a moderate reduction in 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, 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 nonequivalent 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 posttest and short-term follow-up. Although the effects diminished over time, the effects on frequency of drinking days and alcohol-related problems remained significant up to four 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, and 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 posttest. 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 that reduced 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 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 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 one year after participation. Although the effects for both outcomes were heterogeneous, moderator analyses were not conducted.

Three meta-analyses focused on the prevention of tobacco use (Hwang et al. 2004, Isensee & Hanewinkel 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 nonequivalent control group designs and included follow-up assessments two 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 had small, significant effects at posttest 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 & 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 postintervention) or long-term assessments. Their results may be attributed to a reduced ability to detect effects because of the low number of trials included in the comparisons, although their findings are consistent with the nonsignificant effects found in one of the largest evaluations of smoking prevention programs, the Hutchison Smoking Prevention Project (Peterson et al. 2000). Isensee & Hanewinkel (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 PROMOTION OF HEALTHY DEVELOPMENT OR RESILIENCE

Programs that are designed to promote the healthy development of youth 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 visit programs. Meta-analyses of resilience promotion programs were identified for youth exposed to each of 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 Social and Emotional Learning

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 preexisting 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 nonrandomized trials with control groups and follow-ups up to three years postintervention.

The results show small but significant effects to reduce conduct problems (e.g., aggression, noncompliance, delinquency) and emotional distress (e.g., depression, anxiety, social withdrawal). Furthermore, the effects were significant at follow-up, a median of one year later. Effect size was not related to the methodological rigor of the studies, children’s ethnicity, or urban versus rural status. Subgroup analyses indicated that the program effects were significant for programs delivered by teachers as well as by 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, multicomponent 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 for 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 nonequivalent control group designs. Program effects were reported at posttest but not at follow-up.

A small, significant but heterogeneous effect was found to reduce conduct problems (e.g., acting out, aggression, delinquency) at posttest. 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 nonsignificant 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.’s (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 nonparental adults (or older youth) who are acting in a nonprofessional helping
capacity” (p. 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
occurring up to four years after the program ended. The meta-analyses differed in the moderators
of program effects that were examined.

Dubois et al. (2011) estimated effect sizes after controlling for the methodological quality of
the study. Small, significant effects were found for psychological/emotional problems and con-
duct problems at posttest, but the effect on drug use was not significant. Six factors were signif-
ically related to average effect size across outcomes; two youth factors (male > versus female
and 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, in-
cluded 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 suc-
cessful 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 RCT’s, 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 pro-
grams with both randomized and quasi-experimental designs and with follow-ups up to 18 months
postintervention. The most comprehensive meta-analysis (Kaminski et al. 2008) included 77 trials
that involved children from birth to age 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 age three and included RCTs, trials that used nonequivalent con-
trol group designs, and studies with follow-up of 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 & Welsh (2003) and
Piquero et al. (2009) focused on parenting and family-focused interventions to prevent antisocial
behavior and delinquency and were reviewed above in the section on antisocial behavior.
& Heinrichs (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 and 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 above-mentioned meta-analyses of parenting programs that targeted antisocial behavior (i.e., Farrington & Welsh 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 for poor behavior, consistent responding to misbehaviors, and practicing program skills with one’s child(ren). Programs with these characteristics had significant and 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 nonsignificant effects on substance use. Nowak & Heinrichs (2008) found a small, significant effect of the Triple P prevention programs (Levels 1–3) to reduce child behavior problems.

Preschool/Home Visitation 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 & Appelbaum’s (2004) meta-analysis included 60 home visitation 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; case management; and child health or developmental screening. Manning et al. (2010) conducted a meta-analysis of the long-term effects of 11 early developmental prevention programs, including preschool, center-based developmental day care, home visitation, family support, and parent-education programs. The meta-analyses included RCTs and quasi-experimental designs and also studies with follow-up into adolescence.

Sweet & Appelbaum (2004) reported a small, significant effect on social emotional outcomes at posttest but did not report effects at follow-up. Although the effect sizes were homogeneous, the lack of specificity of the variables, including social emotional outcomes, wide variability in characteristics of home visitation programs, and lack of tests of program effects at follow-up, make interpretation of these results 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, which were a broad range of variables including behavior problems, social skills, self-esteem, grade retention, employment, and criminal behavior, at medium- (kindergarten to eighth grade) and long-term (high school and beyond) follow-ups. The effects at kindergarten to eighth grade were heterogeneous. Program effects were larger for programs longer than one year versus shorter programs and for programs offered to predominantly African-American families versus other ethnicities. The 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 socioemotional development (e.g., parent and teacher reports 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 & Roehrle (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, nonrandomized 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, psychoeducation, family-focused approaches, and music therapy for children and adolescents from birth to 18 years old. Siegenthaler et al. (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 & Roehrle (2003) found a medium, significant, and heterogeneous effect on anxiety and a nonsignificant effect on depression at posttest. Several factors were related to larger effect sizes across outcome variables, including good training of group leaders, medium duration and number of sessions, a 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 the 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 pretest were related to larger effect sizes. Several other variables were not found to moderate program effects, including youth age, program dosage, time since death, use of confrontation strategies, and publication status of the study. Currier et al. (2007) reported a small, nonsignificant effect on a composite of adjustment outcomes, but they did not report effects on specific outcomes. A shorter time period since the death and a 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.

Siegenthaler 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 that reduced internalizing problems but a nonsignificant effect that reduced 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.
Table 2  Prevention impacts across diverse prevention targets, outcomes based on effect sizes for continuous outcome measures, odds ratios/relative risks, and risk differences for discrete outcome measures

<table>
<thead>
<tr>
<th>Prevention problem or health focus</th>
<th>Effects for meta-analyses with a primary problem or health focus</th>
<th>Effects for all meta-analyses that report on the problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>ES = 0.17 (SD = 0.05, k = 6)</td>
<td>ES = 0.19 (SD = 0.06, k = 10)</td>
</tr>
<tr>
<td></td>
<td>RD = 0.05 (SD = 0.05, k = 5)</td>
<td>RD = 0.052 (SD = 0.05, k = 5)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>ES = 0.37 (SD = 0.34, k = 6)</td>
<td>ES = 0.43 (SD = 0.31, k = 6)</td>
</tr>
<tr>
<td>Aggression/antisocial behavior/violence</td>
<td>ES = 0.19 (SD = 0.17, k = 29)</td>
<td>ES = 0.25 (SD = 0.28, k = 32)</td>
</tr>
<tr>
<td></td>
<td>OR/RR = 1.10 (SD = 0.01, k = 2)</td>
<td>OR/RR = 1.14 (SD = 0.05, k = 2)</td>
</tr>
<tr>
<td>Tobacco</td>
<td>ES = 0.17 (SD = 0.06, k = 6)</td>
<td>ES = 0.19 (SD = 0.08, k = 8)</td>
</tr>
<tr>
<td></td>
<td>OR/RR = 1.07 (k = 1)</td>
<td>OR/RR = 1.07 (k = 1)</td>
</tr>
<tr>
<td>Alcohol</td>
<td>ES = 0.12 (SD = 0.07, k = 16)</td>
<td>ES = 0.13 (SD = 0.08, k = 18)</td>
</tr>
<tr>
<td>Other substance use</td>
<td>ES = 0.15 (SD = 0.16, k = 17)</td>
<td>ES = 0.12 (SD = 0.15, k = 17)</td>
</tr>
<tr>
<td></td>
<td>OR/RR = 1.19 (SD = 0.15, k = 3)</td>
<td></td>
</tr>
<tr>
<td>Any substance use/abuse</td>
<td>ES = 0.14 (SD = 0.12, k = 39)</td>
<td>ES = 0.14 (SD = 0.11, k = 43)</td>
</tr>
<tr>
<td></td>
<td>OR/RR = 1.16 (SD = 0.13, k = 4)</td>
<td>OR/RR = 1.14 (SD = 0.12, k = 5)</td>
</tr>
<tr>
<td>Promotion</td>
<td>ES = 0.30 (SD = 0.17, k = 27)</td>
<td>NA (see text)</td>
</tr>
<tr>
<td></td>
<td>OR/RR = 1.15 (SD = 0.045, k = 3)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ES, effect size; k, number of meta-analyses; RD, risk difference; OR/RR, odds ratio/risk ratio; SD, standard deviation across meta-analyses.

EFFECTS ACROSS OUTCOMES AND APPROACHES TO PREVENTION AND PROMOTION

In this section, we provide two sets of quantitative summaries (see columns 2 and 3 in Table 2) that reflect the preventive impacts on 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 OR, 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 impacts 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 ORs (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 “megaregression” of the effect sizes as this is beyond the scope of this review.

All 124 summary effect size, odds ratio/relative risk (OR/RR), and risk difference (RD) measures that were used to derive the figures in Table 2 are roughly comparable (within effect size, OR/RR, and 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 (i.e., not a subgroup analysis). The first column refers to the broad prevention categories of meta-analyses described above. We present quantitative findings in two ways in the next two columns. The second column refers to all effects reported in meta-analyses organized by the primary focus of the meta-analysis in column one. Thus, for depression there were six separate effect-size summaries in meta-analyses focused on this topic, each with multiple trials. Among these six
meta-analyses for depression was one that examined the impact of these depression-centered interventions on anxiety. Column three presents the effects for each of the problems shown in column one but summarized across all meta-analyses that reported effects on that outcome (e.g., 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 six summary effect sizes, each of which was based on at least four trials, whereas there was a total of nine summary effect sizes across all meta-analyses that included depressive symptoms as an outcome. For the meta-analyses of health promotion programs, the summary effects across all problem outcomes are presented in column two because they targeted multiple problem outcomes, leaving column three blank because these were difficult to separate across all the meta-analyses.

These summary statistics show small to medium overall beneficial effects on symptoms as well as disorders and behavior problems. The largest effect size occurred for the meta-analyses that focused on anxiety (effect size = 0.43, column three). 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, and early childhood programs, as well as programs targeting significant family disruptions, such as bereavement, divorce, and parental mental illness: effect size = 0.30, with the effect sizes for depression, substance use, and crime/antisocial 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 & Appelbaum 2004) were broad composites of indices of internalizing and externalizing problems or socioemotional development rather than measures of specific problem behaviors or indices of severe levels of problems, such as meeting diagnostic criteria, violence, and criminal behavior. The effect sizes for internalizing problems, effect size = 0.30, exceeded those for externalizing problems, effect size = 0.25, and substance use/abuse problems, effect size = 0.14. However, we did not formally test these differences because of the potential overlap in the meta-analyses and lack of reported standard errors (which we discuss as methodological issues at the end of this review). For the dichotomous outcomes, all ORs/relative risks showed small to moderate benefits, with none exceeding 1.4. 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: Seven of the 15 mean effect sizes (47%) 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.
Program Effects

The conceptual model first addresses the study of program effects on theoretical mediators and multiple positive and problem outcomes.

Figure 1
Conceptual framework for meta-analyses of prevention and promotion programs. This shows five sources of heterogeneity of program results on theoretical mediators and positive and problem outcomes.

Figure 1 presents a conceptual model that identifies the key factors that influence the effects of prevention and promotion interventions. The conceptual model addresses issues that are 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 which 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 in Figure 1 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, as well as 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 in detection of program effects and moderators of program effects. 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 the 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

The conceptual model first addresses the study of program effects on theoretical mediators and multiple positive and problem outcomes.
Theoretical mediators. Prevention programs are based on an underlying theory in which the program affects mediators and changes in these mediators lead to changes 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 (Spoth 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 antisocial behavior and aggression assessed program effects on putative mediators of social skills, social adjustment, and social competence (Beelmann & Lösel 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), 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 ability 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.

Multiple positive and 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) and anxiety (Teubert & Pinquart 2011)], or multiple indicators of problems in the same domain [e.g., aggression, crime, violence (Derzon et al. 2006); antisocial behavior, delinquency (Farrington & Welsh 2003); frequency of drinking, and 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), antisocial 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, and social relationships (Beelmann & Lösel 2006; Deković et al. 2011; DuBois et al. 2011; Durlak et al. 2010, 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.

Moderators of Program Effects
The conceptual model identifies five types of moderators of program effects. The effects of each of these moderators are important for understanding the efficacy, effectiveness, and community implementation of prevention programs.
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 program uses 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 nonthreatening 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, 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 the 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, antisocial behavior, and smoking (Beelmann & Lösel 2006, Hwang et al. 2004, Piquero et al. 2010). Motivational interviewing was associated with large effect sizes on alcohol use in college students (Carey et al. 2007, Fachini et al. 2012) 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 the length of program. Several meta-analyses found that shorter programs had larger effects (i.e., Deković et al. 2011, Stathakos & Roehrle 2003, Stice et al. 2009); several found that longer programs had larger effects (e.g., Wilson & Lipsey 2007); and several found nonsignificant 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 on the basis of their target population (NRC/IOM 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 & Pinquart, 2011), aggression/antisocial behavior (Beelmann & Lösel 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 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 & Pinquart 2011), aggression/antisocial problems (Wilson & Lipsey 2007), conduct problems, and emotional distress after participation in SEL programs (Durlak et al. 2011). From a public health perspective, Shamblen & Derzon (2009) have argued that universal programs with a smaller effect size than selective or indicated programs 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 the 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 et al. 2013). 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., Deković et al. 2011, DuBois 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, few community resources, and family dysfunction. They found that the 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 US 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 the 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, Stathakos & Roehrle 2003). These findings may reflect the 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. 2007, 2012; Horowitz & Garber 2006; Scott-Sheldon et al. 2012; Stice et al. 2009; Teubert & Pigaut 2011) and age (Beelmann & Lösel 2006, Horowitz & Garber 2006, Mytton et al. 2006, Scott-Sheldon et al. 2012, Stice et al. 2009, Teubert & Pinquart 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, antisocial behavior, and school-based SEL programs (Durlak et al. 2011, Porath-Waller et al. 2010, Wilson & Lipsey 2007). These findings are consistent with those of Durlak & Dupre (2008), who found that higher levels of implementation were related to better outcomes in 45 of 59 prevention and promotion programs.
Durlak & 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 versus no problems in implementation). Future work needs to study the effects of implementation using more refined measures of the 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 have begun to describe the ways in which the 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 nonprofessionals or teachers (Fisak et al. 2011, Stice et al. 2009, Teubert & Pinquart 2011, Tobler et al. 2000), although no difference was found for the professional level of the provider in a meta-analysis of programs to prevent antisocial behavior (Beelmann & Lösel 2006). Several other meta-analyses found larger effects for peer leaders as compared with teachers or nonpeers 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 findings from several meta-analyses that a higher level of provider training was related to larger effects (Durlak et al. 2010, 2011; Stathakos & Roehrle 2003) have important implications for implementing programs in real-world settings. Theoretically, the effects of providers or of training may be mediated through differences in program implementation (Durlak & Dupre 2008). Similarly, the findings 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 because it is easier for service delivery systems to effectively implement smaller-scale programs. Clearly, research has just begun to identify the impacts of service delivery systems and providers, and future meta-analyses should consider these important moderators.

Contextual factors. Context refers to the historical time and the community in which the program is delivered and includes a broad range of factors, such as the level of problem behaviors occurring in the population at the time the trial is conducted, other services offered in the community (Weiss et al. 2013), and stressors or environmental issues 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 issues. 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). The recent studies on cross-national effects of parenting programs (Knerr et al. 2013) indicate an increasing interest in such contextual issues, 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 the 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 implications for enhancing the public health impact of prevention programs. Illustratively, we focus on the useful 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 are 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 noninteractive 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 $U_3$ index (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 $U_3$ value of 0.75 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 $U_3$ value of 0.55 for participants in parenting programs that did not involve practicing skills with their child(ren) is 0.55, which is only a 5% benefit for the average person in these programs, indicating that there is five 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 Cohen’s $U_3$ index values of the benefit of participating in interactive, skill-building programs as compared to noninteractive programs across meta-analyses in approaches to prevention that focus on specific problems (i.e., anxiety, antisocial behavior, substance use) and those that focus on promoting healthy development (i.e., 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 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, the methodological quality of trials in particular has an important effect on our confidence in the validity of the findings from these studies. The 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., the
Table 3  Cohen’s U₃ index values of the benefits to the average person who participated in interactive, skill-building programs versus noninteractive programs

<table>
<thead>
<tr>
<th>Meta-analysis and referenceᵃ</th>
<th>Interactive, skill-building programs' outcome using Cohen's U₃ index</th>
<th>Noninteractive programs outcome using Cohen's U₃ index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevention of anxiety programs</td>
<td>FRIENDS program: Anxiety Cohen’s U₃ = 0.60</td>
<td>Other programs: Anxiety Cohen’s U₃ = 0.55</td>
</tr>
<tr>
<td>Prevention of antisocial behaviors programs</td>
<td>Cognitive behavioral: Antisocial behavior Cohen’s U₃ = 0.69 at posttest, Cohen’s U₃ = 0.69 at follow-up</td>
<td>Behavioral only: Antiocial behavior Cohen’s U₃ = 0.56 at posttest, Cohen’s U₃ = 0.55 at follow-up Cognitive only: Cohen’s U₃ = 0.56 at posttest, Cohen’s U₃ = 0.48 at follow-up Other approaches: Cohen’s U₃ = 0.65 at posttest, Cohen’s U₃ = 0.56 at follow-up</td>
</tr>
<tr>
<td>Prevention of substance use programs</td>
<td>All interactive programs: Substance use Cohen’s U₃ = 0.55</td>
<td>All noninteractive programs: Substance use Cohen’s U₃ = 0.52</td>
</tr>
<tr>
<td>School-based SEL programs</td>
<td>Used SAFEᵇ practices: Emotional distress Cohen’s U₃ = 0.61 Conduct problems Cohen’s U₃ = 0.59</td>
<td>Did not use SAFE practices: Emotional distress Cohen’s U₃ = 0.57 Conduct problems Cohen’s U₃ = 0.56</td>
</tr>
<tr>
<td>After-school programs</td>
<td>Used SAFEᵇ practices: Conduct problems Cohen’s U₃ = 0.62 Substance use Cohen’s U₃ = 0.56</td>
<td>Did not use SAFE practices: Conduct problems Cohen’s U₃ = 0.53 Substance use Cohen’s U₃ = 0.51</td>
</tr>
<tr>
<td>Parenting programs</td>
<td>Focused on positive interactions with child(ren): Externalizing behavior Cohen’s U₃ = 0.64 Included practice with own child(ren): Externalizing behavior Cohen’s U₃ = 0.75 Focused on time out: Externalizing behavior Cohen’s U₃ = 0.71</td>
<td>Did not focus on positive interactions with child(ren): Externalizing behavior Cohen’s U₃ = 0.57 Did not include practice with own child(ren): Externalizing behavior Cohen’s U₃ = 0.55 Did not focus on time out: Externalizing behavior Cohen’s U₃ = 0.56</td>
</tr>
</tbody>
</table>

ᵃCohen’s U₃ values were derived from effect sizes that were presented in the following tables of these meta-analyses:ᵇtable 2;ᶜtable 3;ᵈtable 9;ᵉtable 3;ᶠtable 3;ᵍtable 7.

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

Abbreviation: SEL, social and emotional learning.
Meta-analyses have accounted for the 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 the 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 & Pinquart 2011), some found that larger effects were associated with indicators of poorer methodological quality (e.g., Merry et al. 2012, Stathakos & Roehrle 2003). The lack of consistency in assessing the methodological quality of trials across meta-analyses is problematic. More systematic approaches to assessing the methodological quality of individual trials need to be considered (e.g., Jadad et al. 1996, Juni et al. 2001). Additional 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 the redundancy of studies included in the analyses. We noted that a sizable portion of the same individual trials was 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 0.06% 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 & Sweeney (1991) posited that the difficulties of correctly specifying the theoretical model might impede researchers when examining the 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 & Schram (1994) claimed that a meta-analysis using an “exploratory model” would help us to build a case for potential causal relations by explaining the 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 the dependent variable is found to be nonsignificant (i.e., the direct effect of the independent variable on the dependent variable) after controlling for the mediating variable, which is suggested by Baron & 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. Additional research is strongly urged to provide statistical justification for examining the existence of mediating variables in meta-analyses.
Examinations of the program effects on multiple outcomes also pose methodological concerns. As discussed by Gleser & 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 (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-analyses, 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, and 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. Furthermore, 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 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 multitrial databases and synthesize the findings (Brown et al. 2012, Perrino et al. 2013), and we believe such syntheses using individual-level data from multiple trials will provide significant advances in understanding who benefits from prevention programs and the underlying mechanisms of these programs.

OVERALL SUMMARY AND IMPLICATIONS

This review used an overview of a 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, antisocial behavior, and substance use. Significant effects were found for both continuous measures and dichotomous measures, such as diagnosed disorder. Furthermore, 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 NRC/IOM (2009) report concerning the benefits of preventive interventions for children and youth.

Are There Common Moderators 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 potentially 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 Are the Methodological Issues in Synthesizing Effects Across Meta-Analyses?

Several methodological issues were discussed in synthesizing findings across meta-analyses, including the inconsistency of meta-analyses use to account for the methodological quality of the studies and the redundancy of 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, the mediation and moderation of the effects of preventive interventions.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We are grateful for support from the National Institute of Mental Health (grant R01MH040859), the National Institute of Drug Abuse (grant R01DA026874), and our colleagues on these grants, Drs. George Howe, Tatiana Perrino, Hilda Pantin, and William Beardslee. We also are grateful for support from the National Center for Advancing Translational Sciences and for the National Institute on Minority Health and Health Disparities grant UL1TR000460-01A1. We thank Ophelia Hernandez for her data management on this project.

LITERATURE CITED


Cooper H, Koenka AC. 2012. The overview of reviews: unique challenges and opportunities when research syntheses are the principal elements of new integrative scholarship. *Am. Psychol.* 67:446–62


Shadish WR, Sweeney RB. 1991. 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. 59:883–93


## Contents

Advances in Cognitive Theory and Therapy: The Generic Cognitive Model  
*Aaron T. Beck and Emily A.P. Haigh* ................................................................. 1

The Cycle of Classification: DSM-I Through DSM-5  
*Roger K. Blashfield, Jared W. Keeley, Elizabeth H. Flanagan, and Shannon R. Miles* ............................................. 25

The Internship Imbalance in Professional Psychology: Current Status and Future Prospects  
*Robert L. Hatcher* ......................................................................................... 53

Exploratory Structural Equation Modeling: An Integration of the Best Features of Exploratory and Confirmatory Factor Analysis  
*Herbert W. Marsh, Alexandre J.S. Morin, Philip D. Parker, and Gurvinder Kaur* ............................................................. 85

The Reliability of Clinical Diagnoses: State of the Art  
*Helena Chmura Kraemer* ........................................................................ 111

Thin-Slice Judgments in the Clinical Context  
*Michael L. Slepian, Kathleen R. Bogart, and Nalini Ambady* .................... 131

Attenuated Psychosis Syndrome: Ready for DSM-5.1?  
*P. Fusar-Poli, W.T. Carpenter, S.W. Woods, and T.H. McGlashan* ............ 155

From Kanner to DSM-5: Autism as an Evolving Diagnostic Concept  
*Fred R. Volkmar and James C. McPartland* ................................................. 193

Development of Clinical Practice Guidelines  
*Steven D. Hollon, Patricia A. Arcán, Michelle G. Craske, Kermit A. Crawford, Daniel R. Kivlahan, Jeffrey J. Magnavita, Thomas H. Ollendick, Thomas L. Sexton, Bonnie Spring, Lynn F. Bufka, Daniel I. Galper, and Howard Kurtzman* .................................................. 213

Overview of Meta-Analyses of the Prevention of Mental Health, Substance Use, and Conduct Problems  
*Irwin Sandler, Sharlene A. Wolchik, Gracelyn Cruden, Nicole E. Mabber, Soyeon Abn, Abnalee Brincks, and C. Hendricks Brown* .................................................. 243
Improving Care for Depression and Suicide Risk in Adolescents: Innovative Strategies for Bringing Treatments to Community Settings

Joan Rosenbaum Asarnow and Jeanne Miranda .......................................................... 275

The Contribution of Cultural Competence to Evidence-Based Care for Ethnically Diverse Populations

Stanley J. Huey Jr., Jacqueline Lee Tilley, Eduardo O. Jones, and Caitlin A. Smith .......................................................... 305

How to Use the New DSM-5 Somatic Symptom Disorder Diagnosis in Research and Practice: A Critical Evaluation and a Proposal for Modifications

Winfried Rief and Alexandra Martin .......................................................... 339

Antidepressant Use in Pregnant and Postpartum Women

Kimberly A. Yonkers, Katherine A. Blackwell, Janis Glover, and Ariadna Forray .......................................................... 369

Depression, Stress, and Anhedonia: Toward a Synthesis and Integrated Model

Diego A. Pizzagalli .......................................................... 393

Excess Early Mortality in Schizophrenia

Thomas Munk Laursen, Merete Nordentoft, and Preben Bo Mortensen .................. 425

Antecedents of Personality Disorder in Childhood and Adolescence: Toward an Integrative Developmental Model

Filip De Fruyt and Barbara De Clercq .......................................................... 449

The Role of the DSM-5 Personality Trait Model in Moving Toward a Quantitative and Empirically Based Approach to Classifying Personality and Psychopathology

Robert F. Krueger and Kristian E. Markon .......................................................... 477

Early-Starting Conduct Problems: Intersection of Conduct Problems and Poverty

Daniel S. Shaw and Elizabeth C. Shelleby .......................................................... 503

How to Understand Divergent Views on Bipolar Disorder in Youth

Gabrielle A. Carlson and Daniel N. Klein .......................................................... 529

Impulsive and Compulsive Behaviors in Parkinson’s Disease

B.B. Averbeck, S.S. O’Sullivan, and A. Djamsbidian .......................................................... 553

Emotional and Behavioral Symptoms in Neurodegenerative Disease: A Model for Studying the Neural Bases of Psychopathology

Robert W. Levenson, Virginia E. Sturm, and Claudia M. Haase .......................................................... 581
Attention-Deficit/Hyperactivity Disorder and Risk of Substance Use Disorder: Developmental Considerations, Potential Pathways, and Opportunities for Research
Brooke S.G. Molina and William E. Pelham Jr. .................................................. 607

The Behavioral Economics of Substance Abuse Disorders: Reinforcement Pathologies and Their Repair
Warren K. Bickel, Matthew W. Johnson, Mikhail N. Koffarnus, James MacKillop, and James G. Murphy ................................................. 641

The Role of Sleep in Emotional Brain Function
Andrea N. Goldstein and Matthew P. Walker ..................................................... 679

Justice Policy Reform for High-Risk Juveniles: Using Science to Achieve Large-Scale Crime Reduction
Jennifer L. Skeem, Elizabeth Scott, and Edward P. Mulvey .............................. 709

Drug Approval and Drug Effectiveness
Glen I. Spielmans and Irving Kirsch ................................................................. 741

Epidemiological, Neurobiological, and Genetic Clues to the Mechanisms Linking Cannabis Use to Risk for Nonaffective Psychosis
Ruud van Winkel and Rebecca Kuepper .......................................................... 767

Indexes

Cumulative Index of Contributing Authors, Volumes 1–10 .............................. 793
Cumulative Index of Articles Titles, Volumes 1–10 .......................................... 797

Errata

An online log of corrections to Annual Review of Clinical Psychology articles may be found at http://www.annualreviews.org/errata/clinpsy
New From Annual Reviews:

**Annual Review of Organizational Psychology and Organizational Behavior**

Volume 1 • March 2014 • Online & In Print • http://orgpsych.annualreviews.org

**Editor:** Frederick P. Morgeson, The Eli Broad College of Business, Michigan State University

The *Annual Review of Organizational Psychology and Organizational Behavior* is devoted to publishing reviews of the industrial and organizational psychology, human resource management, and organizational behavior literature. Topics for review include motivation, selection, teams, training and development, leadership, job performance, strategic HR, cross-cultural issues, work attitudes, entrepreneurship, affect and emotion, organizational change and development, gender and diversity, statistics and research methodologies, and other emerging topics.

Complimentary online access to the first volume will be available until March 2015.

**TABLE OF CONTENTS:**

<table>
<thead>
<tr>
<th>Topic</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>An Ounce of Prevention Is Worth a Pound of Cure: Improving Research Quality Before Data Collection</td>
<td>Herman Aguinis, Robert J. Vandenberg</td>
</tr>
<tr>
<td>Burnout and Work Engagement: The JD-R Approach</td>
<td>Arnold B. Bakker, Evangelia Demerouti, Ana Isabel Sanz-Vergel</td>
</tr>
<tr>
<td>Compassion at Work, Jane E. Dutton, Kristina M. Workman, Ashley E. Hardin</td>
<td></td>
</tr>
<tr>
<td>Constructively Managing Conflict in Organizations, Dean Tjosvold, Alfred S.H. Wong, Nancy Yi Feng Chen</td>
<td></td>
</tr>
<tr>
<td>Coworkers Behaving Badly: The Impact of Coworker Deviant Behavior upon Individual Employees</td>
<td>Sandra L. Robinson, Wei Wang, Christian Kiewitz</td>
</tr>
<tr>
<td>Delineating and Reviewing the Role of Newcomer Capital in Organizational Socialization</td>
<td>Talya N. Bauer, Berrin Erdogan</td>
</tr>
<tr>
<td>Emotional Intelligence in Organizations, Stéphane Côté</td>
<td></td>
</tr>
<tr>
<td>Employee Voice and Silence, Elizabeth W. Morrison</td>
<td></td>
</tr>
<tr>
<td>Intercultural Competence, Kwok Leung, Soon Ang, Mei Ling Tan</td>
<td></td>
</tr>
<tr>
<td>Learning in the Twenty-First-Century Workplace, Raymond A. Noe, Alena D.M. Clarke, Howard J. Klein</td>
<td></td>
</tr>
<tr>
<td>Pay Dispersion, Jason D. Shaw</td>
<td></td>
</tr>
<tr>
<td>Personality and Cognitive Ability as Predictors of Effective Performance at Work</td>
<td>Neal Schmitt</td>
</tr>
<tr>
<td>Perspectives on Power in Organizations, Cameron Anderson, Sebastien Brion</td>
<td></td>
</tr>
<tr>
<td>Psychological Safety: The History, Renaissance, and Future of an Interpersonal Construct</td>
<td>Amy C. Edmondson, Zhike Lei</td>
</tr>
<tr>
<td>Research on Workplace Creativity: A Review and Redirection, Jing Zhou, Inga J. Hoever</td>
<td></td>
</tr>
<tr>
<td>Talent Management: Conceptual Approaches and Practical Challenges, Peter Cappelli, JR Keller</td>
<td></td>
</tr>
<tr>
<td>The Contemporary Career: A Work–Home Perspective, Jeffrey H. Greenhaus, Ellen Ernst Kossek</td>
<td></td>
</tr>
<tr>
<td>The Fascinating Psychological Microfoundations of Strategy and Competitive Advantage, Robert E. Ployhart, Donald Hale, Jr.</td>
<td></td>
</tr>
<tr>
<td>The Psychology of Entrepreneurship, Michael Frese, Michael M. Gielnik</td>
<td></td>
</tr>
<tr>
<td>What Was, What Is, and What May Be in OP/OB, Lyman W. Porter, Benjamin Schneider</td>
<td></td>
</tr>
<tr>
<td>Where Global and Virtual Meet: The Value of Examining the Intersection of These Elements in Twenty-First-Century Teams, Cristina B. Gibson, Laura Huang, Bradley L. Kirkman, Debra L. Shapiro</td>
<td></td>
</tr>
<tr>
<td>Work–Family Boundary Dynamics, Tammy D. Allen, Eunae Cho, Laurenz L. Meier</td>
<td></td>
</tr>
</tbody>
</table>

Access this and all other Annual Reviews journals via your institution at www.annualreviews.org.
New From Annual Reviews:

**Annual Review of Statistics and Its Application**

*Volume 1 • Online January 2014 • http://statistics.annualreviews.org*

**Editor:** Stephen E. Fienberg, Carnegie Mellon University  
**Associate Editors:** Nancy Reid, University of Toronto  
Stephen M. Stigler, University of Chicago

The *Annual Review of Statistics and Its Application* aims to inform statisticians and quantitative methodologists, as well as all scientists and users of statistics about major methodological advances and the computational tools that allow for their implementation. It will include developments in the field of statistics, including theoretical statistical underpinnings of new methodology, as well as developments in specific application domains such as biostatistics and bioinformatics, economics, machine learning, psychology, sociology, and aspects of the physical sciences.

**Complimentary online access to the first volume will be available until January 2015.**

**TABLE OF CONTENTS:**

- A Systematic Statistical Approach to Evaluating Evidence from Observational Studies, David Madigan, Paul E. Stang, Jesse A. Berlin, Martijn Schuemie, J. Marc Overhage, Marc A. Suchard, Bill Dumouchel, Abraham G. Hartzema, Patrick B. Ryan  
- The Role of Statistics in the Discovery of a Higgs Boson, David A. van Dyk  
- Brain Imaging Analysis, F. DuBois Bowman  
- Statistics and Climate, Peter Guttorp  
- Climate Simulators and Climate Projections, Jonathan Rougier, Michael Goldstein  
- Probabilistic Forecasting, Tilmann Gneiting, Matthias Katzfuss  
- Bayesian Computational Tools, Christian P. Robert  
- Bayesian Computation Via Markov Chain Monte Carlo, Radu V. Craiu, Jeffrey S. Rosenthal  
- Build, Compute, Critique, Repeat: Data Analysis with Latent Variable Models, David M. Blei  
- Structured Regularizers for High-Dimensional Problems: Statistical and Computational Issues, Martin J. Wainwright  
- High-Dimensional Statistics with a View Toward Applications in Biology, Peter Bühlmann, Markus Kalisch, Lukas Meier  
- Next-Generation Statistical Genetics: Modeling, Penalization, and Optimization in High-Dimensional Data, Kenneth Lange, Jeannette C. Papp, Janet S. Sinsheimer, Eric M. Sobel  
- Breaking Bad: Two Decades of Life-Course Data Analysis in Criminology, Developmental Psychology, and Beyond, Elena A. Erosheva, Ross L. Matsueda, Donatello Telesca  
- Event History Analysis, Niels Keiding  
- Statistical Evaluation of Forensic DNA Profile Evidence, Christopher D. Steele, David J. Balding  
- Using League Table Rankings in Public Policy Formation: Statistical Issues, Harvey Goldstein  
- Statistical Ecology, Ruth King  
- Estimating the Number of Species in Microbial Diversity Studies, John Bunge, Amy Willis, Fiona Walsh  
- Dynamic Treatment Regimes, Bibhas Chakraborty, Susan A. Murphy  
- Statistics and Related Topics in Single-Molecule Biophysics, Hong Qian, S.C. Kou  
- Statistics and Quantitative Risk Management for Banking and Insurance, Paul Embrechts, Marius Hofert

Access this and all other Annual Reviews journals via your institution at [www.annualreviews.org](http://www.annualreviews.org).