

EFFECTIVENESS OF SCHOOL-BASED UNIVERSAL SOCIAL, EMOTIONAL, AND BEHAVIORAL PROGRAMS: DO THEY ENHANCE STUDENTS' DEVELOPMENT IN THE AREA OF SKILL, BEHAVIOR, AND ADJUSTMENT?

MARCIN SKLAD, RENÉ DIEKSTRA, MONIQUE DE RITTER, AND JEHONATHAN BEN

Roosevelt Academy International Honors College of Utrecht University

CAROLIEN GRAVESTEIJN

The Hague University

To answer the question of whether teaching social and emotional skills to foster social–emotional development can help schools extend their role beyond the transfer of knowledge, the authors conducted a meta-analytical review of 75 recently published studies that reported the effects of universal, school-based social, emotional, and/or behavioral (SEB) programs. The analyzed interventions had a variety of intended outcomes, but the increase in social skills and decrease in antisocial behavior were most often reported. Although considerable differences in efficacy exist, the analysis demonstrated that overall beneficial effects on all seven major categories of outcomes occurred: social skills, antisocial behavior, substance abuse, positive self-image, academic achievement, mental health, and prosocial behavior. Generally, immediate effects were stronger than delayed effects, with the exception of substance abuse, which showed a sleeper effect. Limitations of the analysis and moderators of the effectiveness of SEB programs in schools are discussed in the final section of the article. © 2012 Wiley Periodicals, Inc.

Schools are expected to play an important role in promoting the development of children and adolescents by preparing them for their future roles in society. There is increasing consensus that to fulfill that role as best as possible, schools should not limit their focus to the regular academic program per se (see Greenberg et al., 2003; see also article 29 of the CRC, the United Nations Convention on the Rights of the Child, U.N. General Assembly, 1989). To achieve “the development of the child’s personality, talents and mental and physical abilities to their fullest potential” (article 9, 1a, p. 3), the teaching of social and emotional skills can be considered a core task of school systems, in addition to cognitive-academic skills. In the last two decades, a multitude of programs or interventions explicitly addressing social and emotional development have been created and introduced into schools’ curricula. The main goal of the current review is to address the question of whether such school-based social, emotional, and/or behavioral (SEB) programs or interventions are effective in terms of fostering social and emotional development.

There is considerable diversity in the behavioral and attitudinal facets and skills addressed by different programs, their intervention design, and their composition. Whereas some SEB programs mainly consist of classroom curricula, others combine classroom curricula with activities outside of the classroom, involving the entire school, parents, and the community. Programs may also involve children and adolescents in community service, especially if there is an emphasis on education for citizenship or civic engagement in the program.

Similarly, there is variation in the scope and place of social and emotional skills taught in the programs. For example, some programs are mainly focused on teaching students specific skills, such as refusal skills related to drugs, alcohol, premature sexual behavior, or violence (e.g., Agha & Van Rossum, 2004; Botvin, Griffin, & Nichols, 2006). Other programs assume that such skills are mainly

Supporting Information is available in the online issue at wileyonlinelibrary.com.

Correspondence to: Marcin Sklad, Roosevelt Academy International Honors College of Utrecht University at Middelburg–ACC, Lange Noordstraat 1 P.O. Box 94, Middelburg 4330 AB Netherlands. E-mail: m.sklad@roac.nl

derived from more general social and emotional skills or from a positive self-concept or self-esteem (e.g., Battistich, Schaps, Watson, Solomon, & Lewis 2000; Catalano et al., 2003). Therefore, these programs aim to enhance these skills or self-esteem so that children will display more positive and sociable behavior and, consequently, make positive choices more often. Other programs combine both approaches (e.g., Flannery et al., 2003). Although universal school-based SEB programs are founded on a variety of theoretical approaches, most of them cite social learning theory (Bandura, 1977) and cognitive-behavioral approaches as the foundation of their method (Tobler et al., 2000). The core competencies taught in SEB programs, sometimes also referred to as social-emotional learning programs (Collaborative for Academic, Social, and Emotional Learning, 2003) or life skills programs (Botvin & Griffin, 2002), consist of what is often described as emotional intelligence (Goleman, 1995; Mayer & Salovey, 1997). Emotional intelligence includes competencies that allow students to recognize and manage emotions, solve problems effectively, and establish positive relationships with others (Zins & Elias, 2006).

Social and emotional skills can be seen as protective factors that reduce the probability that students exposed to risk factors will engage in problem behavior (Catalano et al., 2002). Authors who take this perspective of preventive science consider programs that target social and emotional skills to be prevention programs. They follow the assumption that the enhancement of universal skills will lead to a reduction in undesirable or (self-) harmful behavior, attitudes, and characteristics.

The main aim of this review of universal school-based SEB programs is to establish whether such programs produce positive effects substantial enough to consider them to be effective tools for extending the school's role beyond traditional cognitively based instruction, namely, toward fostering *both* cognitive as well as social and emotional development. It is also to provide a comprehensive set of intended outcomes, as well as to establish the efficacy of SEB activities in attaining those outcomes both in North America, where most of the programs have been developed and evaluated, and in other parts of the world. An additional goal is to assess the influence of moderating variables, such as the type of program deliverer or program duration, that might be relevant to educational institutions when considering program adoption.

With regard to the efficacy of universal school-based interventions, recent syntheses of literature reviews (Diekstra et al., 2008; Weare & Nind, 2011), such as meta-analyses, indicate that SEB interventions and programs do indeed first and foremost enhance social and emotional competencies. The largest average significant effect sizes are found in these domains (e.g., Beelmann & Losel 2006; Catalano et al., 2002; Durlak et al., 2011; Payton et al., 2008; Wilson & Lipsey, 2007). Significant effects are also observed with regard to enhancing positive self-perception or self-esteem (O'Mara, Marsh, Craven, & Debus, 2006; Payton et al., 2008).

Reviews also point to significant indirect effects: reduction of anxiety and depression or emotional distress (e.g., Durlak et al., 2011; Neil & Christensen, 2007), prevention of conduct problems, such as drug use (Durlak et al., 2011; Faggiano et al., 2008), improved attitudes toward school and enhanced school achievements (e.g., Durlak et al., 2011; Payton et al., 2008; Wilson & Lipsey, 2007), prevention of aggressive and antisocial behavior (Hahn et al., 2007; Wilson & Lipsey, 2007), and promotion of positive or prosocial behavior (Durlak et al., 2011; Payton et al., 2008). On average, the "indirect" effects are, not surprisingly, smaller than the effects on social and emotional competencies.

As the majority of approaches, programs, and interventions aspire to achieve multiple positive outcomes at the same time, single studies differ considerably with regard to the number of reported outcome parameters and labels attached to them. The same is true for meta-analyses of studies, making them not easily comparable to one another (e.g., Durlak et al., 2011, versus Wilson & Lipsey, 2007). In the most recent and comprehensive review of meta-analyses, the authors conclude that the most common outcome parameters can be best categorized into: positive social behaviors

and skills; reduced negative behaviors (e.g., conduct problems and violence); reduction in emotional problems (e.g., depression and anxiety); self-confidence and self-esteem; bonding with the school; and improved academic achievement, test scores, grades, and attendance (Weare & Nind, 2011).

They also concluded that there is considerable variety in efficacy of programs: the main reasons, if not causes, of low or no effects are summarized as follows: "Interventions were not effective if they were only based on loose guidelines and broad principles as, however sound the principles, effective interventions need high quality implementation. High quality implementation included having a sound theoretical base, well defined goals, strong focus and explicit guidelines, thorough training and quality control, feedback on intervention effects, and consistent staffing" (Weare & Nind, 2011, p. 61). Consequently, a number of SEB interventions do not really deserve the label "program," as they are not based on a manual, do not describe a sequence of components, do not train implementers, and do not monitor/evaluate implementation.

Reviews of (meta-analytic) reviews (see Diekstra & Gravesteyn, 2008; Weare & Nind, 2011) also indicate that studies on SEB interventions almost exclusively concern schools in North America. As to the efficacy of SEB elsewhere, such as in Europe or worldwide, the present literature provides hardly any information. In fact, only Weare and Nind (2011), in their review of reviews, suggest effects of SEB programs worldwide, but this suggestion is not based on appropriate meta-analytic procedure.

Due to the dominant preventive science paradigm, existing literature reviews devoted to the assessment of the effectiveness of school-based programs, including those targeted at enhancing social and emotional skills and fostering children's overall development, often concentrate on a specific type of program or intervention. They focus on programs that address children in trouble (e.g., those with emotional and social disorders; Cook et al., 2008; Quinn, Kavale, Mathur, Rutherford, & Forness, 1999) or programs with a single outcome category, such as substance use (Faggiono et al., 2008) nonattendance and delinquency (Wilson, Gottfredson, & Najaka, 2001), violence and aggression (Hahn, Kuzara, 2007; Wilson, Lipsey, & Derzon, 2003), depression (Sutton, 2007), or antisocial behavior (Beelmann & Losel, 2006). In the preventive science paradigm, universal programs are also often analyzed together with programs targeted to high-risk groups. Moreover, programs containing elements of social and emotional learning are commonly analyzed in combination with other preventive programs that are based on different theoretical bases and that may not contain these elements (e.g., Wilson et al., 2001).

The current analysis is aimed at filling in these gaps, examining only school-based, universal programs that concentrate on promoting development rather than merely preventing specific problems. In the case of programs targeting SEB, which are the current focus, there is often very little difference between programs developed from a preventive science perspective and from a positive youth development perspective (Catalano et al., 2002) in terms of contents; the difference between two perspectives lies more in policy and evaluation paradigms, as SEB-based preventive programs attempt to prevent problems through the promotion of positive development by enhancing social and emotional skills. In the present article, we take a positive youth development perspective. An integrative approach is also in line with the main goal of the presented meta-analysis. Answering the central question, whether universal SEB programs may be useful in schools, requires a review of the complete spectrum of intended and achieved outcomes rather than a focus on one specific type of outcome.

Because our analysis is guided by a practical question, we have limited the scope of the analysis to recent universal programs. This is in contrast to recent meta-analyses (such as Durlak et al., 2011), which include reports from recent interventions along with studies conducted in the 1960s. Limiting the scope of our analysis to only recent programs has two advantages. First, the results reflect the effectiveness of contemporary programs or interventions that are currently available for schools.

Second, the meta-analysis includes programs that may have been affected by the shift in attention toward positive youth development and does not include programs that were developed before terms such as *social-emotional skills* (World Health Organization [WHO], 2002) and *emotional intelligence* (Goleman, 1995) were brought to professional public attention.

METHODS

In this study, the major categories of outcomes are analyzed separately to overcome the problem of “mixing apples and oranges,” which is common in meta-analytical studies. The study follows the four basic steps involved in a meta-analysis (Kulik, 1983): (a) locate studies on an issue, using clearly specified procedures; (b) characterize the outcomes of studies in quantitative terms; (c) code as many features of the studies as possible; and (d) use statistical procedures to summarize findings and to relate study features to study outcomes.

Search and Retrieval of Studies

Several approaches were used to identify the relevant literature. Studies were obtained by carrying out searches of large scientific databases, such as ERIC, PsycINFO, EBSCO, and Academic Search Elite, as well as Internet searches on www.google.com, www.scirus.com, and www.altavista.com, using the following key terms: *emotional-skills*, *in-school*, *emotional training*, *school intervention*, *school-based*, *skills-for-life*, *life skills*, *social emotional learning*, *social-skills*, *educational program*, *intervention*, *prevention*, *universal*, and *controlled*. These key terms were used in different combinations to minimize the number of omitted studies.

We also conducted searches of websites of research centers, universities, and private and governmental institutions, including the Substance Abuse and Mental Health Services Administration, WHO, and the American Psychological Association. We obtained additional studies through an online library search using www.picarta.nl, through an examination of bibliographies of earlier meta-analyses and literature reviews, and through direct contact with program coordinators and the programs’ researchers. The final sample of studies was primarily drawn from online scientific databases of peer-reviewed journals. The complete bibliography of reviewed studies is included as a separate supplemental file in the online Supporting Information.

Inclusion Criteria

To be included in this meta-analysis, a study had to meet the following criteria:

1. The study reported a program that taught at least one social-emotional skill (see WHO, 2002).
2. The intervention was school-based (i.e., it was aimed at primary or secondary school students, used school facilities, and took place during regular school hours).
3. The intervention had to be “universal” (i.e., it was aimed at the general school population and not only at “high-risk” or underprivileged children).
4. The study reported the programs’ outcomes in a way that allowed for the calculation of effect sizes, providing statistics that could be mathematically converted to Cohen’s *d* (e.g., standard deviations and means, standard errors and means, or odds ratios).
5. The study reflected current conditions (i.e., it was published in English between 1995 and 2008). The year 1995 was chosen as the starting point because the influential book, *Emotional Intelligence* (Goleman, 1995) was released that year.
6. The study used an experimental or quasi-experimental design with control/comparison group(s).

Coding of Reports

During the summer of 2008, four senior university students familiar with the literature on social and emotional skills coded the studies after having received several training sessions on the use of a coding sheet. The training sessions involved a review of the studies included in this meta-analysis and clarification of the inclusion criteria. To estimate reliability, 10% of the studies were independently scored by two raters. Results demonstrated satisfactory inter-rater reliability across coders (mean kappa alpha = .8 and ICC alpha = .98 for the studies' categorical and scale-level characteristics, respectively). Independent raters' overall agreement was 93%, and all differences were discussed and resolved. A coding sheet was used to code four types of variables from each study included in the analysis: (a) methodological characteristics, (b) intervention/program features, (c) recipients' characteristics, and (d) program outcomes.

Methodological characteristics included: design (randomized vs. nonrandomized); assignment level (schools, classes, or students), implementation integrity (publication did not report treatment or implementation integrity data, publication reported problems, or publication reported integrity as fine), detailed source of outcome report (self-report, teacher, researcher, peers, or other, e.g., criminal record), whether the outcome was self-report, whether the outcome was measured by single or multiple instruments, whether the study used single or multiple settings, publication status, and the number of months that passed between the end of the intervention and the assessment of the outcomes.

The characteristics of the interventions that were coded included the number of sessions; the length of sessions in minutes; whether the intervention was carried out by teachers, psychosocial professionals, or others; whether the intervention was restricted to the school or also involved the community or family; whether the intervention was a part of a "whole school" program of change; and year of implementation.

Coded recipients' characteristics included whether the program was carried out in a primary or secondary school, the country where the intervention took place, the (average) age of participants, the proportion of female/male participants, the proportion of White participants versus other ethnic groups, and socioeconomic status.

Types of Outcomes

After reviewing all outcomes reported in all studies and consulting earlier related meta-analytical studies (e.g., Wilson & Lipsey, 2007), a few major outcome categories were developed, allowing exclusive assignment of each reported outcome to one category: the extracted categories are similar to those used by Payton et al. (2008) and those indicated by Weare and Nind (2011), with some differences. The majority (88%) of reported outcomes were classified into the first seven categories¹:

Social-emotional skills and attitudes (direct outcomes):

1. Social-emotional skills. This category included measurement of general or specific social or emotional skills that were taught or exercised during the intervention. Typical examples of reported outcomes from this category included direct assessment of the skills, for instance, rating of social competence provided by teachers, evaluation of strategy constructiveness in a conflict resolution task, assertiveness skills measure, refusal skills efficacy score, problem orientation assessment, or bully-victim scale score.

¹Other categories of outcomes for which information was assembled but for which there were not enough studies reported to permit analysis included academic attitudes, attitudes toward violence and aggression, physical health, and sexual behavior and attitudes.

2. Positive self-image. This category included all outcomes corresponding to the valence of general self-image, for example, measures of general or global self-esteem and sense of self-efficacy. This category did not include reports of self-assessment on specific skills acquired during the interventions.

Behavioral adjustment (second-order effects):

3. Antisocial behavior. This category included various indicators of antisocial or aggressive behavior, from reported off-task, disruptive, and externalizing behavior to fighting in the past year, hurting someone on purpose, verbal aggression in the past month, active bullying, teachers reporting physical aggression, and committing a felony.
4. Prosocial behavior. This category consisted of behaviors intended to help other people and showing a concern about them. Among other items, it incorporated operational measures of altruistic behavior (teacher rating), empathy (rated by peers), prosocial problem solving, peace building, cooperation, and relations with others.
5. Substance abuse. All outcomes related to the use of alcohol, cigarettes, and drugs were classified into this category.
6. Mental health disorders. Outcomes grouped under the mental health disorders category usually referred to internalizing problems. Particular outcome measures included scores on the Beck Depression Inventory, Reynolds Adolescent Depression Scale, hopelessness scale, degree of stress scale, anxiety level, clinical levels of depressive symptoms, suicide attempts, need for mental health services, and psychological well-being assessment. This category did not include measures of aggressive and antisocial behavior.
7. Academic achievement. Academic achievement consisted of indicators of students' performance in academic areas, for example, grade-point average of core academic subjects, reading achievement score, California Achievement Test score, or academic competence (rated by teacher).

Post-Tests and Follow-Up

Immediate and delayed outcomes of programs were extracted and analyzed separately. The first category consisted of outcomes assessed at post-test, up to and including 6 months after completion of the intervention. Outcomes measured at least 7 months after completion of an intervention were assigned to the category of follow-up outcome. The exact length of time between the intervention and the measurement of the outcome was recorded.

Analysis Plan

The general effectiveness of programs on the different outcome categories was first analyzed separately for post-test and follow-up assessments. This was followed by homogeneity analyses. Then, moderator analyses (characteristics of study methods, interventions/implementation, and participants) were conducted to test whether any of the potential moderators had a significant effect on the programs' effectiveness.

It was not possible to carry out moderator analyses for several outcome-characteristic combinations due to an insufficient number of studies reporting appropriate data. According to Hedges and Pigott (2001), at least five studies are needed for each category to achieve estimated power of .8, even for strong effects.

Authors tend to report only significant outcomes of their programs and to omit outcomes that were not significant from their reports, which can introduce a publication bias that leads to an overestimation of effect size. To address this problem and to substantiate general conclusions about

the effects of programs, a “file drawer analysis” was carried out by means of a failsafe N calculation. The failsafe N can be defined as the number of studies with null effect that would be required to render the overall effect statistically insignificant (Cooper, 1979).

The effect size estimate used here was a standardized difference between means of the intervention and the control or comparison group (Lipsey & Wilson, 2001): Cohen’s d (Cohen, 1988). Multiple studies did not contain information that allowed for the direct calculation of the effect size estimate using the aforementioned formulas. Various statistics (e.g., odds ratios, χ^2 for 2×2 tables, and Pearson’s correlation coefficient) reported in these studies were converted to Cohen’s d (Cohen, 1988) using the Comprehensive Meta-Analysis (CMA) program (Borenstein, & Rothstein, 1999).

For studies reporting more than one program, effects of all programs matching the overall selection criteria were combined and treated as a single effect. Analogically, contrasts separating treatment from booster treatments, females from males, different ethnic groups, and different interventions were averaged, taking into account effect sizes and the relative weight of each group, using the CMA program (Borenstein & Rothstein, 1999). Contrasts comparing different interventions with each other were not included. Several studies reported multiple treatment-comparison contrasts, and a number of studies reported multiple outcomes for various groups within the study. All treatment-comparison contrasts using independent participant groups were included in our analysis.

Effect directions were defined in relation to the meaning of each outcome category measured. Thus, outcomes that were desirable to increase (e.g., social skills or academic achievement) had a positive effect direction when increased; conversely, outcomes that were desirable to decrease (e.g., mental disorders or substance abuse) had a negative effect direction when decreased.

When a study reported more than one outcome from the same category, the outcomes’ effect sizes were averaged to obtain a single estimate for the given study. Different outcome and time categories were examined separately. For example, a study that reported outcomes for prosocial behavior and academic achievement at both time points (post-test and follow-up) would have four outcomes analyzed separately.

The random effects model was used in the analysis because it allows heterogeneity of effect sizes reported across different studies and provides a more conservative estimate of effect size, as it includes study-level sampling error (Lipsey & Wilson, 2001). The random effects model incorporates the assumption that the different studies are estimating different but related treatment effects (Higgins & Green, 2008). All statistical tests were two-tailed and an alpha level of $<.05$ was used to indicate statistical significance. To test for heterogeneity, tests were carried out using the Q test statistic (Lipsey & Wilson, 2001).

RESULTS

The literature search and coding process yielded data from 75 studies of universal school-based programs published from 1995 to 2008 that met all of the inclusion criteria. Of these studies, 72 (96%) were published in peer-reviewed journals, and three were progress reports. The average reported intervention group size was 543 participants ($SD = 1,119.83$); the smallest intervention involved 13 children (one classroom), and the largest involved 8,280 children. Approximately one quarter of the studies were conducted in parts of the world other than North America (16 studies, 21.3%), mostly in continental Europe (11 studies, 14.7%). The remaining five (6.7%) studies originated from other continents. Most non-American studies reported programs implemented in one country, with the exception of one study that included participants from four European countries: Austria, Denmark, Germany, and Luxemburg (Hanewinkel & ABhauer, 2004). One study reported a program that was carried out in both the United States and Canada (Stevahn, Johnson, Johnson, Green, & Laginski, 1997). The inclusion of 16 non-North American studies makes this meta-analysis the only

Table 1
Time of Assessment and Outcomes Reported by Studies

	No. of Programs	Percentage of Programs
Time of Assessment		
Post-Test: 0–6 months	55	73.3
Follow-Up: 7–18 months	27	36.0
Follow-Up: 19+ months	16	21.3
Outcome Reported		
Social–Emotional Skills	35	46.7
Antisocial Behavior	35	46.7
Substance Abuse	21	28.0
Positive Self-Image	14	18.7
Academic Achievement	13	17.3
Mental Health Disorders	13	17.3
Prosocial Behavior	10	13.3
Total	75	

one thus far that allows for comparison between American and non-American socioeconomic status programs, at least with regard to the main outcome examined (social–emotional skills).

Outcome Categories

The most diverse studies reported outcomes belonging to six of the seven outcome categories, whereas the least diverse studies reported outcomes from only one category. The two most often reported outcomes were an increase in social–emotional skills and a reduction in antisocial behavior, which were both assessed in half of the studies. For 29 (38.7%) of the studies, all extracted outcomes belonged to only one category, and 26 (34.7%) of the studies reported outcomes from two categories. Altogether, 93.3% of the studies reported three or fewer outcome categories. The number of studies reporting each outcome category is presented in Table 1.

Post-Test and Follow-Up

Roughly half of the studies (40 studies, 53.3%) reported only immediate effects, defined here as post-tests that occurred no later than half a year after the end of the intervention. Fifteen studies (20%) reported at least one follow-up outcome (assessed at least 7 months after the completion of the intervention), along with the immediate outcome measurement(s). For two studies, it was not clear when the measurement of outcomes took place. The remaining 18 studies (24%) reported only follow-up outcomes.

Randomization

Of the studies included in the meta-analysis, 56% used a randomized experimental design, and 44% used a quasi-experimental design (see Table 2). In only 11 studies (14.7%), individual students were the unit of assignment. Most researchers assigned classes (37.3%) or even schools (17.3%) to intervention or control conditions. In all studies, the analysis was conducted for students. This discrepancy between the levels of the assignment and the analysis has been regarded as common and nonproblematic in the meta-analytic literature (Wilson et al., 2001). Although this discrepancy could result in the overstating of the statistical significance of effects, it does not affect the descriptive statistics extracted.

Table 2
Methodological Features of Studies

	No. of Programs	Percentage of Programs
Experimental Design		
Any Form of a Random Assignment	42	56.0
Nonrandom Assignment	33	44.0
Unit of Assignment		
Matched Pairs	10	13.3
Schools	13	17.3
Classes	28	37.3
Students	11	14.7
Other, e.g., Level of Cohorts	13	17.3
Implementation Integrity		
Not Reported	40	53.3
Reported Problems	3	4
Reported as Fine	32	42.7
Outcome Measurement		
Used Multiple Sources	15	20
Relied Solely on Self-Reports	45	60
Used Multiple Instruments	62	82.6
Intervention Manual		
Unavailable/Availability Not Reported (43%)	55	73.3
Reported as Available	20	26.7
Total	75	100.0

Manuals

The availability of intervention manuals containing the exact description of the content and implementation of an intervention is an important methodological aspect of intervention effect studies. A manual allows for replication of the intervention and permits different instructors to teach in a similar manner, using the same strategies and themes. However, manual availability was explicitly mentioned in only 26.7% of the studies.² In the remaining studies, the existence of a manual was either not mentioned (but likely available, 43%) or not available. Due to the low availability of manuals, it was not possible to reliably discern studies that used methods involving discussion only, discussion plus activities, role-plays with performance feedback, and role-plays with performance feedback and transfer of training instruction.

Characteristics of Programs, Deliverers, and Participants

A quarter of the reported interventions were first and foremost directed at a change in school culture and climate. Among them, 12 programs had a school climate change element, along with class sessions offered to students. However, seven interventions of this type did not have class

²Twenty studies explicitly mentioned existence of intervention manual (Botvin, Griffin, & Nichols, 2006; Cuijpers, Jonkers, Weerd, & de Jong, 2001; Farrell, Meyer, Sullivan, & Kung, 2003; Flay, Acock, Vuchinich, & Beets, 2006; Flay, Allred, & Ordway, 2001; Grossman et al., 1997; Jalongo et al. 1999; Lalongo, Poduska, Werthamer, & Kellam, 2001; Laird, Bradley, & Black, 1998; Linares et al., 2005; Murray & Malmgren, 2005; Quayle, Dziurawiec, Roberts, Kane, & Ebsworthy, 2001; Sanz de Acedo et al., 2003; Shapiro, Burgoon, Welker, & Clough, 2002; Shochet et al., 2001; Smokowski, Fraser, Day, Galinsky, & Bacallao, 2004; Spence, Sheffield, & Donovan, 2003; Twemlow et al., 2001; Van Schoiack-Edstrom, Frey, Beland, 2002).

Table 3
Features of Reported Programs

	No. of Programs	Percentage of Programs
Program Structure Features		
Change of School Culture	19	25.3
Contains Homework Assignments	16	21.3
Program Duration		
Up to 1 Month	8	10.7
Up to 1 Year	49	65.3
More Than 1 Year	18	25.3
Program Delivery		
Only Teachers Involved	41	54.7
Professionals/Researchers Involved	29	38.7
Others	5	6.7
Program Context		
Family Involved	19	25.3
Community Involved	14	18.7
Total	75	100

sessions for students. Four of these interventions encouraged teachers to change their regular classes by introduction of a teaching style and classroom techniques that enhance social skills, for example, cooperative learning and student-centered approach to classroom management. The remaining three offered specific enhancements to be added to the regular curriculum in language arts, mathematics, and physical education.

There was considerable variability in the duration and intensity of reported interventions (see Table 3), ranging from a single 1-day workshop via interventions that consisted of 15 sessions spread over 3 years, to a program of 155 sessions lasting up to 6 years. However, the majority of interventions did not exceed 1 year in length and 18 sessions in number ($M = 27.1$, $SD = 32.4$). The most common length of the intervention was 1 school year, and the average length of the intervention was 396.6 days ($SD = 408.8$). The most common length of a class session was equal to the length of a school lesson; the average session length was equal to 49.8 minutes ($SD = 21.9$).

Most of programs' deliverers were school teachers (see Table 3). In 41 of the studies, which was slightly more than half (54.7%) of the analyzed sample, teachers were the only trainers in direct contact with students. In 29 programs (38.7%), professionals were involved in teaching the program (e.g., psychologists, researchers).

Although all interventions were school based and universal, it should be noted that 14 (18.7%) of the interventions also had community-based elements. Furthermore, 19 studies (25.3%) included programs that also involved families.

More studies were retrieved for secondary schools (62.7%) than for primary schools (see Table 4). Four studies covered both types of schools. The average age of samples of students was 10.5 ($SD = 2.1$). The average age was reported by only five studies.

Of the studies that reported the socioeconomic status of participants, roughly one half addressed students with low socioeconomic status. The other half of the studies included students of mixed socioeconomic status.

Gender and ethnicity were reported by 80% and 48% of the studies, respectively. The average proportion of females was approximately 50% ($M = 51.54\%$, $SD = 9.02\%$). Ethnicity was seldom reported in studies from countries other than the United States. On average, study

Table 4
Participants' Characteristics

Characteristics	No. of Studies	Percentage of All Studies	Percentage of Studies Reporting
School Level			
Primary	31	41.3	41.3
Secondary	48	62.7	62.7
Total Reported	75 ^a		
Socioeconomic Status Reported			
Low(er)	25	33.3	51.0
Mixed	23	30.7	46.9
Total Reported	49 ^b		

^aTwo programs reported interventions including both school levels. ^bOne study reported socioeconomic status in another way.

Table 5
Programs' Efficacy on Major Outcomes

Outcome	N	Effect Size and Its Standard Error			95% Confidence Interval		Test of Null Hypothesis (2-Tail)		Heterogeneity		
		d	SE (d)	Lower Limit	Upper Limit	Z Value	p Value	Q	df (Q)	p Value	I ²
Immediate Outcomes											
Academic Achievement	10	0.46	0.08	0.31	0.61	6.06	<.001	153.62	9.00	<.001	94.14
Antisocial Behavior	39	-0.43	0.06	-0.32	-0.54	-7.59	<.001	985.12	38	<.001	96.14
Mental Disorders	13	-0.19	0.05	-0.09	-0.29	-3.67	<.001	29.11	12	<.001	58.78
Positive Self-Image	8	0.46	0.12	0.22	0.69	3.83	<.001	143.59	7	<.001	95.14
Prosocial Behavior	6	0.39	0.12	0.15	0.63	3.20	<.001	163.37	9	<.001	94.49
Social-Emotional Skills	31	0.70	0.10	0.51	0.89	7.27	<.001	1086.26	31	<.001	97.15
Substance Abuse	10	-0.09	0.02	-0.06	-0.13	-5.06	<.001	25.00	19	.016	24.00
Follow-Up Outcomes											
Academic Achievement	7	.26	0.05	0.16	0.36	5.06	<.001	13.65	6	.03	56.04
Antisocial Behavior	16	-.20	0.05	-0.10	-0.30	-3.90	<.001	132.35	15	<.001	88.67
Mental Disorders	11	-.10	0.03	-0.04	-0.17	-3.21	<.001	8.85	10	.55	0.00
Positive Self-Image	12	.07	0.02	-0.03	0.10	3.42	<.001	5.82	11	.89	0.00
Prosocial Behavior	7	.12	0.03	0.06	0.18	3.94	<.001	8.61	6	.20	30.34
Social-Emotional Skills	15	.07	0.01	0.04	0.09	4.46	<.001	11.97	14	.61	0.00
Substance Abuse	24	-.18	0.04	-0.11	-0.25	-5.06	<.001	200.02	23	<.001	88.50

Note. For outcomes, which SEB programs intend to reduce (antisocial behavior, mental disorders, substance abuse), negative effect sizes represent an improvement; for the remaining, positive outcomes, positive effect sizes indicate improvement.

samples consisted of 44.0% ($SD = 31.12$) White students. Varying definitions of ethnicity and its categories (e.g., White vs. Euro-American) used by the researchers made it difficult to adequately identify ethnic groups. Moreover, it seems quite probable that studies may have included different groups under similar category labels (e.g., White might refer to different ethnicities across studies).

Program Effects

There were a sufficient number of studies for each major category of outcome at both the post-test and the follow-up periods to calculate overall effect sizes. As Table 5 shows, programs show statistically significant effects in the desired direction for all seven outcome categories at the post-test. For social skills measured at the post-test, the effect sizes can be classified as large (Cohen, 1988; Lipsey & Wilson, 2001). On average, social skills of students participating in programs were .7 standard deviation higher than their counterparts, which means that the average program participant had better skills than 76% of regular students.

Programs had moderate immediate effects on positive self-image, prosocial behavior, academic achievement, and antisocial behavior, improving each of these four outcomes by nearly one half a standard deviation. Immediate effects on mental disorders and substance abuse were small (Cohen, 1988; Lipsey & Wilson, 2001), although even effects classified as small can have ample practical relevance (Lipsey & Wilson, 1993). For all seven categories of outcomes, the heterogeneity of effect sizes was significant, with values of the I^2 statistic in the range of 24% to 97%, indicating high heterogeneity (Higgins, Thompson, Deeks & Altman, 2003). This points to the existence of genuine differences in the effectiveness of programs.

Regarding follow-up effects, the largest beneficial effect was found for academic achievement, followed by substance abuse. All follow-up effect sizes were statistically significant, yet their sizes were small. Median improvement at follow-up was equal to .12 standard deviation, which means the average student would outperform an additional 5% of the population as a result of the intervention.

Effect sizes at the follow-up were statistically significantly heterogeneous for academic achievement, antisocial behavior, and substance abuse. The heterogeneity of effect sizes for the antisocial behavior and substance abuse was high: 88% in both cases. At post-test, the failsafe N (number of studies with null effect needed to nullify the general effect) was between 186 for substance abuse and 4,630 for social skills. This means that to attribute the significant effects of the programs to publication bias alone, one would need to assume that between 90% and 99.9% of outcomes were not published because they showed no program effect. For outcomes showing statistically significant effects of programs at follow-up, the failsafe N was in the range of 32 to 111. In summary, we can assume that the observed overall positive effects of interventions cannot be attributed to the publication bias alone for most of the reported outcomes.

The heterogeneity of the effect sizes suggests that there are important factors or moderators that affect the reported effectiveness of programs on analyzed outcome categories. For the moderators related to the studies' methodology, a moderator analysis was carried out for the main outcome—social skills measured at the post-test (see Table 6 for statistically significant methodological moderators).

Table 6
Methodological Level Moderators Analysis, Mixed Effect Model Analysis

Moderator	Effect <i>d</i> (<i>SE</i>)	Heterogeneity Within Group			Effect <i>d</i> (<i>SE</i>)	Heterogeneity Within Group			Heterogeneity Between Groups		
		<i>Q</i> _{within}	<i>df</i>	<i>p</i>		<i>Q</i> _{within}	<i>df</i>	<i>p</i>	<i>Q</i> _{between}	<i>df</i>	<i>p</i>
Outcome Source		One source				Multi-Source					
	.82(.11)**	1074.52	26	<.001	.20(.07)**	6.58	4	.16	21.37	1	<.001
Manual		Not available				Available					
SS	.91(.13)**	1022.57	21	<.001	.21(.07)**	43.09	9	<.001	19.65	1	<.001

Note. SS = social skills.
* $p < .005$. ** $p < .001$.

Table 7
Moderator Analysis, Mixed Effect Model Analysis

Moderators & Outcomes	Heterogeneity Within Group				Heterogeneity Within Group				Heterogeneity Between Groups		
	Effect <i>d</i> (SE)	<i>Q</i> _{within}	<i>df</i>	<i>p</i>	Effect <i>d</i> (SE)	<i>Q</i> _{within}	<i>df</i>	<i>p</i>	<i>Q</i> _{between}	<i>df</i>	<i>p</i>
School Level	Primary school				Secondary school only						
SS	.67(.14)**	457.35	15	<.001	.74(.14)**	590.82	15	<.001	0.16	1	<i>ns</i>
AB	-.59(.07)**	665.09	20	<.001	-.25(.19)**	272.60	17	<.001	7.33	1	.007
Duration	At least 1 year				Less than 1 year						
SS	.20(.05)**	93.08	15	<.001	1.15(.26)**	558.08	13	<.001	13.37	1	.007
AB	-.24(.06)**	262.38	15	<.001	-.50(.11)**	520.46	21	<.001	4.26	1	.039
No. of Sessions	20 sessions or more				<20 sessions						
SS	.24(.07)**	53.54	11	<.001	.80(.19)**	323.57	10	<.001	7.38	1	.007
AB	-.022(.04)**	161.22	19	<.001	-.31(.14)	221.89	11	<.001	0.40	1	<i>ns</i>
Trainers	Only Teachers				Not only Teachers						
SS	.71(.16)**	99.46	10	<.001	.70(.12)**	980.80	20	<.001	0.00	1	<i>ns</i>
AB	-.31(.08)*	81.64	13	<.001	-.48(.07)**	870.67	24	<.001	2.26	1	<i>ns</i>
Professionals Involved	No professional delivering				Professionals delivering						
SS	.82(.11)**	1004.23	21	<.001	.82(.20)**	70.62	9	<.001	0.47	1	<i>ns</i>
AB	-.45(.07)**	896.25	26	<.001	-.39(.11)	76.83	11	<.001	0.24	1	<i>ns</i>
Place	Outside of North America				North America						
SS	.51(.14)**	72.82	6	<.001	0.75(.12)**	1003.82	24	<.001	1.72	1	<i>ns</i>
Focused Study	Focused				Not focused						
SS	0.70 (.15)**	773.21	17	<.001	0.68(.13)**	311.43	13	<.001	0.01	1	<i>ns</i>
AB	-.045(.07)**	896.25	26	<.001	-.039(.11)**	76.83	11	<.001	0.24	1	<i>ns</i>

Note. SS = social skills, AB = antisocial behavior. For outcomes, which SEB programs intend to reduce (antisocial behavior), negative effect sizes represent an improvement; for social skills, positive effect sizes indicate improvement.

* $p < .005$. ** $p < .001$.

For the moderators related to the intervention, two outcomes most often reported at the post-test, social skills and antisocial behavior, were taken into account (see Table 7).

The analysis of methodological features of the studies potentially moderating the reported effects on social skills showed the following results. Studies using multiple settings and multiple instruments to capture the same outcome produced slightly higher effect sizes; nevertheless, the difference was not statistically significant, $Q(N = 32, df = 1) = 1.29, ns$, and $Q(N = 32, df = 1) = 1.74, ns$; similarly, studies providing implementation integrity information reported statistically insignificantly higher effect sizes, $Q(N = 32, df = 1) = 1.74, ns$. However, studies that assessed social skills measured using multiple sources (e.g., self-reports and proxy reports) reported significantly lower effects than those utilizing a single-source outcome measure (see Table 6). Notably, for every subgroup, the effects of the interventions remained statistically significant, at $p < .001$.

Analysis of the interventions' features revealed that programs of short duration (less than 1 year) showed a higher immediate effect on social skills and antisocial behavior than longer programs. Programs carried out only by school teachers did not have significantly lower effects than programs that also involved other types of trainers. Programs involving professionals as deliverers did not have significantly higher effect sizes. There was no significant heterogeneity between programs conducted in primary schools or in secondary schools on social skills, but programs placed at primary schools had significantly larger reported effects than programs at secondary schools on antisocial behavior.

In both cases, social skills and antisocial behavior, programs focused on a specific outcome did not have a statistically significant advantage over more general programs.

It is important to note that there was no significant heterogeneity between the American studies and studies from other parts of the world in effect size for social skills (there was only one non-American study that reported on antisocial behavior; therefore, a comparison was not carried out for this outcome). Studies presenting interventions conducted outside of North America were neither significantly less nor more effective than American studies.

DISCUSSION AND CONCLUSIONS

The main findings suggest that universal school-based SEB programs that have been evaluated in experimental or quasi-experimental studies over the past 13 years generally have positive effects on a number of desirable outcomes. These outcomes include enhancement of social and emotional skills; positive self image; prosocial behavior; reduction or prevention of antisocial behavior, mental problems, and disorders; and promotion of academic achievement. In the short term (up to 6 months), the largest effects were found for social-emotional skills, attitudes toward self, and prosocial behavior, followed by academic achievement and reduction of antisocial behavior. These findings are in line with those found in other meta-analyses that concentrated on a single outcome and synthesized studies from before 1995 (e.g., Wilson et al., 2001) and also with a meta-analysis that covered a period of almost half a century (Durlak et al., 2011). Given that more than half of the studies reported only post-test data collected less than half a year after the end of the intervention, conclusions about lasting effects of these SEB programs need to be made very cautiously.

At follow-up, programs showed positive effects on all outcomes, although some of these effects decreased substantially compared with immediate effects. However, the reduction or prevention of antisocial behavior showed a sleeper effect, increasing at the follow-up measurement. Again, these findings are corroborated by results of earlier studies that evaluated the durability of intervention effects (e.g., Botvin, Baker, Filazzola, & Botvin, 1990; Durlak et al., 2011; O'Mara, Marsh, Craven, & Debus, 2006; Sutton, 2007). In summary, we can conclude that universal SEB programs that are not limited to children at risk have both immediate and longer term positive effects. As to our main question, it can be stated that SEB programs can help schools to extend their role beyond traditional cognitive instruction. However, because there appear to be considerable differences between programs and interventions in composition and in effect, schools should preferably look for positive answers to the following questions when selecting programs (see Durlak et al., 2011): (a) Does the program use a connected and coordinated set of activities to achieve its objectives relative to skill development? (b) Does the program use active forms of learning to help youth learn new skills? (c) Does the program have at least one component devoted to developing personal or social skills? and (d) Does the program target specific skills focused on social and emotional learning rather than targeting skills or positive development in general terms? The answers to these questions should be included in a program manual/training guide. The fact that a majority of studies indicated that a manual/training guide was not available or did not report the availability of a manual is worrisome, in particular, because implementation integrity is to a considerable extent dependent on their availability. Related to this is the issue of cost-effectiveness. As programs differ in effect and in required input in terms of money, facilities, and manpower, an important consideration in selecting a program is cost-effectiveness. In light of the data from the present review, it is difficult to provide any clear-cut guidelines. However, the safest selection seems to be to use programs that have manuals and for which implementation integrity has been reported positively in effect studies.

As this is the first meta-analysis in which effect studies originating in the United States are compared with effect studies from other parts of the world, in particular, the European continent, the finding that overall effect sizes of the two groups of studies are similar is highly relevant (at least for the enhancement of social–emotional skills, the one outcome measure for which comparison was statistically possible). This finding indicates that SEB programs may be beneficial to children from various national and cultural contexts around the globe. Specifically, these children's social and emotional development may be substantially enhanced by these interventions. Such a social and emotional development is crucial to overall youth development, both in terms of personality development as well as academic progress and school career (Zins, Weissbert, Wang, & Walberg, 2004).

In earlier focused reviews, the involvement of teachers in the delivery of a program positively contributed to its effectiveness (Hahn et al., 2007, Wilson et al., 2003). From our perspective, the most relevant question was whether teachers can deliver SEB programs without compromising the programs' effectiveness, in other words, whether the involvement of experts in delivery is a necessary condition for assuring the effectiveness of a program. An earlier meta-analysis, which focused on aggressive and disruptive behavior, did not reveal a significant decrease in effectiveness for programs delivered by teachers (Wilson & Lipsey, 2007). The results of the current review also did not show that students developed substantially fewer social skills during programs in which only teachers fulfilled the role of trainers. The involvement of researchers or psychosocial professionals in the delivery of programs or interventions did not improve their effectiveness.

The results suggest that there are no reasons to believe that programs carried out in secondary schools have a different effectiveness, that a statistically insignificantly larger effect size at the elementary schools might indicate that the age level of students in elementary schools compared with those in secondary school programs had more impact, or that elementary students are more responsive. The apparent lack of a significant effect of school level on the effectiveness of programs is in line with results of earlier meta-analytical studies that focused on problem behavior (Wilson, Gottferson, & Najaka, 2001) and aggression (Hahn et al., 2007; Willson & Lipsey, 2007).

A meta-analysis is limited by the data provided by synthesized studies. There is a wide variety of different interventions addressing social and emotional skills, each intervention with a unique composition. For example, some programs that focused on changing school culture and climate did not even have class sessions. Each study also uses unique measurement of its effects. In this context, any conclusion about causal relations between moderators used in a meta-analysis and programs effectiveness should be made with extreme caution. Although a meta-analysis can show a positive correlation between a moderator and the reported effect size, the correlation does not imply that moderator is the cause of the increase in effects. Therefore, meta-analytical findings need to be substantiated by future experimental research isolating hypothetical factors determining the effectiveness of interventions.

Limitations to the Research

Similar to other meta-analytical studies (e.g., Hahn et al., 2007; Wilson et al., 2003), routine practice programs, that is, programs that are already being regularly conducted in schools without initial scientific research purposes, constituted only a minority of the programs analyzed in this review. Hahn et al. (2007) expressed a concern that programs conducted for research purposes, which constitute a majority of published studies, may be more effective than ongoing, routine practice programs due to stricter monitoring and more intensive implementation. However, this concern was not confirmed by the results of Wilson & Lipsey's (2007) meta-analysis of aggressive and disruptive

behavior, which showed no significant differences in effectiveness between routine/demonstration and research-oriented programs.

The effects of general universal SEB programs on any specific outcome may be smaller than those indicated by this analysis due to the fact that most reported programs have a particular focus and report effects of the program on the intended target outcomes (and sometimes on instrumental outcomes) but rarely on incidental outcomes not directly related to the main goal of the program. Earlier research shows that programs show stronger effects on direct outcomes than on incidental or indirect outcomes (O'Mara et al., 2006).

Future Research Recommendations

To establish an unbiased general estimate of the effectiveness of SEB programs on any particular outcome, more studies should be carried out on multipurpose programs. In addition, authors reporting the results of targeted SEB programs should be encouraged to measure and report a wide spectrum of outcomes rather than focusing on a few target outcomes.

REFERENCES

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

- *Agha, S., & Van Rossem, R. (2004). Impact of a school-based peer sexual health intervention on normative beliefs, risk perceptions, and sexual behavior of Zambian adolescents. *Journal of Adolescent Health, 34*, 441–452.
- Bandura, A. (1977). *Social learning theory*. New York, NY: General Learning Press.
- *Battistich, V., Schaps, E., Watson, M., Solomon, D., & Lewis, C. (2000). Effects of the child development project on students' drug use and other problem behaviours. *The Journal of Primary Prevention, 21*, 75–99.
- Beelmann, A., & Lösel, F. (2006). Child social skills training in developmental crime prevention: Effects on antisocial behavior and social competence. *Psicothema, 18*, 603–610.
- Borenstein, M., & Rothstein, H. (1999). *Comprehensive meta-analysis: A computer program for research synthesis*. Englewood, NJ: BiostatTM.
- Botvin, G. J., Baker, E., Filazzola, A., & Botvin, E. M. (1990). A cognitive-behavioral approach to substance abuse prevention: A 1-year followup. *Addictive Behaviors, 15*, 47–63.
- Botvin, G. J., & Griffin, K. W. (2002). Life skills training as a primary prevention approach for adolescent drug abuse and other problem behaviors. *International Journal of Emergency Mental Health, 4*(1), 41–47.
- *Botvin, G., Griffin, K., & Nichols, T. (2006). Preventing youth violence and delinquency through a universal school-based prevention approach. *Preventive Science, 7*, 403–408.
- Catalano, R. F., Hawkins, J. D., Berglund, M. L., Pollard, J. A., & Arthur, M. W. (2002). Prevention science and positive youth development: Competitive or cooperative frameworks? *Journal of Adolescent Health, 31*, 230–239.
- *Catalano, R. F., Mazza, J. J., Harachi, T. W., Abbott, R. D., Haggerty, K. P., & Fleming, C. B. (2003). Raising healthy children through enhancing social development in elementary school: Results after 1.5 years. *Journal of School Psychology, 41*, 143–164.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Collaborative for Academic, Social, and Emotional Learning. (2003). *Safe and sound: An educational leader's guide to evidence-based social and emotional learning (SEL) programs*. Chicago, IL: Author.
- Cook, C. R., Gresham, F. M., Kern, L., Barreras, R. B., Thornton, S., & Crews, S.D. (2008). Social skills training for secondary students with emotional and/or behavioral disorders: A review and analysis of the meta-analytic literature. *Journal of Emotional and Behavioral Disorders, 16*, 131–144.
- Cooper, H. M. (1979). Statistically combining independent studies: A meta-analysis of sex differences in conformity research. *Journal of Personality and Social Psychology, 37*, 131–146.
- *Cuijpers, P., Jonkers, R., de Weerd, I., & de Jong, A. (2001). The effects of drug abuse prevention at school: The 'Healthy School and Drugs' project. *Addiction, 97*, 67–74.
- Diekstra, R., & Gravesteyn, C. (2008). Effectiveness of school-based social and emotional education programmes worldwide. Part one, a review of meta-analytic literature. In *Social and emotional education: an international analysis* (pp. 255–284). Madrid, Spain: Santander: Fundacion Marcelino Botin.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based Universal interventions. *Child Development, 82*(1), 405–432.

- Faggiano, F., Vigna-Taglianti, F. D., Versino, E., Zambon, A., Borraccino, A., & Lemma, P. (2008) School-based prevention for illicit drugs use: A systematic review. *Preventive Medicine*, 46(5), 385–396.
- *Farrell, A. D., Meyer, A. L., & White, K. S. (2001) Peer and school problems in the lives of urban adolescents: Frequency, difficulty, and relation to adjustment. *Journal of School Psychology*, 44, 169–190.
- *Flannery, D. J., Liao, A. K., Powell, K. E., Vesterdal, W., Vazsonyi, A. T., Guo, S., . . . Embry, D. (2003). Initial behavior outcomes for the Peace Builders universal school-based violence prevention program. *Developmental Psychology*, 39, 292–308.
- *Flay, B. R., Acock, A., Vuchinich, S., & Beets, M. W. (2006). Progress report of the randomized trial of Positive Action in Hawaii: End of third year of intervention. Twin Falls, ID: Positive Action.
- *Flay, B. R., Allred, C. G., & Ordway, N. (2001). Effects of the positive action program on achievement and discipline: Two matched-control comparisons. *Prevention Science*, 2(2), 71–89.
- Goleman, D. (1995). *Emotional intelligence*. New York, NY: Bantam Books.
- Greenberg, T., Weissberg, R. P., Utne O'Brien, M., Zins, J. E., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist*, 58, 466–474.
- *Grossman, D. C., Neckerman, H. J., Koepsell, T. D., Lu, P-Y, Ashter, K. N., Beland, K., . . . Rivara, F. P. (1997). Effectiveness of a violence prevention curriculum among children in elementary school. *Violence Prevention Among Children*, 277, 1605–1611.
- Hahn, R., Fuqua-Whitley, D., Wethington, H., Lowy, J., Liberman, A., Crosby, A., . . . Dahlberg, L. (2007). Effectiveness of universal school-based programs to prevent violent and aggressive behavior. A systematic review. *American Journal of Preventive Medicine*, 33, 11–129.
- *Hanewinkel, R., & Aßhauer, M. (2004). Fifteen-month follow-up results of a school-based life-skills approach to smoking prevention. *Health Education Research*, 19, 125–137.
- Hedges L. V., & Pigott T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods*, 6, 203–217.
- Higgins, J. P. T., & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions* (Version 5.0.0) [updated February 2008]. The Cochrane Collaboration, Retrieved from www.cochrane-handbook.org
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, 327, 557–560.
- *Ialongo, N., Poduska, J., Werthammer, L., & Kellam, S. (2001). The distal impact of two first-grade preventive interventions on conduct problems and disorder in early adolescence. *Journal of Emotional and Behavioural Disorders*, 9, 146–159.
- Kulik, J. A. (1983). How can chemists use educational technology effectively? *Journal of Chemical Education*, 60, 957–959.
- *Laird, M., Bradley, L. R., & Black, S. (1998). *The final evaluation of Lions-Quest's Skills for Action*. Newark, OH: Lions Quest, Lions Clubs International Foundation.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist*, 48, 1181–1209.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.
- Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In J. D. Mayer & P. Salovey (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp. 3–34). New York, NY: Basic Books.
- *Murray, C., & Malmgren, K. (2005). Implementing a teacher-student relationship program in a high-poverty urban school: Effects on social, emotional, and academic adjustment and lessons learned. *Journal of School Psychology*, 43, 137–152.
- Neil, A. L., & Christensen, H. (2007). Australian school-based prevention and early intervention programs for anxiety and depression: A systematic review. *Medical Journal of Australia*, 186(6), 305–308.
- O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist*, 41, 181–206.
- Payton, J., Weissberg, R. P., Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Schellinger, K. B., & Pachan, M. (2008). The positive impact of social and emotional learning for kindergarten to eighth-grade students: Findings from three scientific reviews. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning.
- *Quale, D., Dziurawiec, S., Roberts, C., Kane, R., & Ebsworthy, G. (2001). The effect of an optimism and life skills program on depressive symptoms in preadolescence. *Behaviour Change*, 18, 194–203.
- Quinn, M. M., Kavale, K. A., Mathur, S. R., Rutherford, R. B., & Forness, S. R. (1999). A meta-analysis of social skill interventions for students with emotional or behavioral disorders. *Journal of Emotional and Behavioral Disorders*, 7, 54–64.
- *Sanz de Acedo, M., Ugarte, M., Cardelle-Elawar, M., Iriarte, M., & Sanz de Acedo, M. (2003). Enhancement of self-regulation, assertiveness, and empathy. *Learning and Instruction*, 13, 423–439.
- *Shapiro, J. P., Burgoon, J. D., Welker, C. J., & Clough, J. B. (2002). Evaluation of the Peacemakers program: School-based violence prevention for students in grades four through eight. *Psychology in the Schools*, 39, 87–102.

- *Shochet, I. M., Dadds, M. R., Holland, D., Whitefield, K., Harnett, P. H., & Osgarby, S. M. (2001). The efficacy of a universal school-based program to prevent adolescent depression. *Journal of Clinical Child Psychology*, 30, 303–315.
- *Smokowski, P. R., Fraser, M. W., Day, S. H., Galinsky, M. J., & Bacallao, M. L. (2004). School-based skills training to prevent aggressive behavior and peer rejection in childhood: Evaluation of the Making Choices program. *The Journal of Primary Prevention*, 25, 233–251.
- *Spence, S., Sheffield, J., & Donovan, C. (2003). Preventing adolescent depression: An evaluation of the Problem Solving for Life program. *Journal of Consulting and Clinical Psychology*, 7, 3–13.
- *Stevahn, L., Johnson, D. W., Johnson, R. T., Green, K., & Laginski, A. M. (1997) Effects on high school students of conflict resolution training integrated into English literature. *Journal of Social Psychology*, 137, 302–315.
- Sutton, J. M. (2007). Prevention of depression in youth: A qualitative review and future suggestions. *Clinical Psychology Review*, 27, 552–571.
- Tobler, N. S., Roona, M. R., Ochshorn, P., Marshall D. G, Streke A. V., & Stackpole K. M. (2000). School-based adolescent prevention programs: 1998 meta-analysis. *Journal of Primary Prevention*, 20, 275–336.
- *Twemlow, S. W., Fonagy, P., Sacco, F. C., Gies, M. L., Evans, R., & Ewbank, R. (2001). Creating a peaceful school learning environment: A controlled study of an elementary school intervention to reduce violence. *American Journal of Psychiatry*, 158, 808–810.
- U. N. General Assembly (1989). Convention on the Rights of the Child. Doc A/RES/44/25. Retrieved May 6, 2011, from: <http://www2.ohchr.org/english/law/crc.htm>
- *Van Schoiack-Edstrom, L., Frey, K. S., & Beland, K. (2002). Changing adolescents' attitudes about relational and physical aggression: An early evaluation of a school-based intervention. *School Psychology Review*, 31, 201–217.
- Weare, K., & Nind, M. (2011). Promoting mental health of children and adolescents through schools and school based interventions. School based interventions: Report of Workpackage Three. Manuscript submitted for publication.
- Wilson, D. B., Gottfredson, D. C., & Najaka, S. S. (2001). School-based prevention of problem behaviors: A meta analysis. *Journal of Quantitative Criminology*, 17, 247–272.
- Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior update of a meta-analysis. *American Journal of Preventive Medicine*, 33, 130–143.
- Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 71, 136–149.
- World Health Organization. (2002). Skills for health. Skills-based health education including life skills: An important component of a child-friendly/health-promoting school (Information series on school health. Document 9). Geneva, Switzerland: Author.
- Zins, J. E., & Elias, M. J. (2006). Social and emotional learning. In G. G. Bear & K. M. Minke (Eds.), *Children's needs 3: Development, prevention and intervention* (pp. 1–13). Bethesda, MD: National Association of School Psychologists.
- Zins, J., Weissbert, R., Wang, M., & Walberg, H. (2004). *Building academic success on social and emotional learning: What does the research say?* New York, NY: Teachers College Press.