



Published in final edited form as:

Early Educ Dev. 2012 January 1; 23(4): 427–450. doi:10.1080/10409289.2011.589043.

Implementing a Primary Prevention Social Skills Intervention in Urban Preschools: Factors Associated with Quality and Fidelity

Melodie Wenz-Gross and Carole Upshur

Family Medicine and Community Health, University of Massachusetts Medical School

Abstract

Research Findings—Preschool behavior problems are of increasing concern to early childhood educators. Preventive interventions are being developed, but implementation in under-resourced child care programs is challenging. This study describes the implementation of an adapted Second Step curriculum to increase children’s social skills and decrease behavior problems in preschool classrooms. Teacher training/coaching, organizational supports, and procedures for sustainability are described. Measures included baseline classroom characteristics, lesson completion and fidelity, and teacher satisfaction. Teachers completed 88% of the lessons across two years. Follow-up monitoring in Year 3, once all training was terminated, found 95% of lessons were delivered with higher fidelity than the first two years. Teachers were satisfied with training, felt the curriculum improved children’s social skills, and stated they would continue to use the curriculum. Lesson completion and fidelity were associated with various teacher characteristics, training attendance, classroom characteristics, teacher and child turnover, and the percent of children with developmental delays in the classroom.

Practice—Results support the implementation strategy and suggest that paying attention to teacher morale, general skill development, and a period of on-going support for training is important and will lead to sustainable and high implementation rates.

Keywords

social skills; behavior problems; preschool; implementation; fidelity; Second Step

Prevalence rates of early childhood behavior problems and preschool expulsion rates are of increasing concern to educators and researchers. Reported prevalence of behavioral problems ranges from 14% to 52% in three to five year olds (Qi & Kaiser, 2003), and rates of children being expelled from preschool are higher than any other grade (Gilliam, 2005). These studies represent both programs that focus on low income children as well as the general preschool population, indicating the wide ranging nature of the problem. Studies show that children’s neurological and cognitive development at this stage are greatly influenced by experience, with parents and teachers playing important roles in the development and behavioral expression of neurological and temperamental attributes (Blair, 2002; O’Neal et al., 2010). Research suggests that broad-based preventive interventions focusing on changing the nature of interactions within the home and/or the school environment, and increasing social skills in the child appear to have lasting benefits (Shure, 1997; Webster-Stratton & Taylor, 2001).

While a number of interventions have been developed to address or prevent behavior problems in children, most have not been studied in randomized effectiveness trials, or

specifically with young preschool children in the classroom (Raver, et al., 2009; Upshur, Wenz-Gross, & Reed, 2009). Further, the few programs with demonstrated effectiveness that do exist are often not widely disseminated (Elliot & Mihalic, 2004). One reason may be feasibility. Transporting research based interventions to real world settings has inherent problems that need to be addressed, including individual and organizational buy-in; organizational capacity and resources to support implementation; adequate training and technical assistance; monitoring of fidelity and adaptations; and ensuring organizational structure and resources necessary for sustainability (Elliot & Mihalic, 2004; Hoagwood & Johnson, 2003; Klein & Sorra, 1996; Rosenheck, 2001; Schoenwald & Hoagwood, 2001). McCall (2009) outlines in detail the problems with replicating evidenced based programs in community settings, stating:

The process of getting communities to want the program, modifying the demonstrated program to fit local circumstances if necessary, and having the program implemented with reasonable enthusiasm and fidelity by agencies and staff who did not create the original program is equally necessary to “bring the program to scale.” (p. 9)

For instance, an ambitious project to implement the Webster-Stratton model in Colorado preschools (Invest in Kids, 2004), resulted in many positive outcomes, yet at the same time revealed a number of challenges to large-scale implementation in child care settings. Among the challenges were adapting a well-developed intervention to the local community and individual developmental needs of children, particularly the youngest preschoolers in the setting, addressing the costs (over \$12,000 per site for full implementation training and materials), and underestimating the need for ongoing technical assistance (Invest in Kids, 2004).

Even with adequate organizational support and training, research suggests that, particularly for school based interventions, the quality and quantity of curriculum implementation varies widely across classrooms (Domitrovich & Greenberg, 2000; Dusenbury, Brannigan, Hansen, Walsh, & Falco, 2005; Pankratz, et al., 2006). Research on implementing interventions at the preschool level is limited, but some studies at the elementary and middle school levels show that teacher characteristics such as teacher attitudes and expectations, years of teaching experience, teacher burnout, and teacher self-efficacy, as well as classroom characteristics (e.g., the number of children with behavior problems, the severity of behavior problems in the classroom) play a role in the degree to which a program is successfully implemented (Dusenbury, et al., 2005; Ghaith & Yaghi, 1997; Han & Weiss, 2005; Rohrbach, Graham, & Hansen, 1993).

For instance, two studies have examined teacher characteristics related to implementation. Dusenbury et al. (2005) found that middle school teachers implemented an average of 65% of teaching objectives and 58% of the main points of observed lessons in a drug prevention program. More experienced teachers had more negative attitudes about the program, but delivered the lessons with greater fidelity. Ghaith and Yaghi (1997) also found that more experienced middle and high school teachers had more negative attitudes toward implementing new practices, with the more experienced teachers rating the implementation of an instructional innovation as more difficult and less important than less experienced teachers.

Similarly, Rohrbach and colleagues (1993) found elementary school teachers delivered between 23% and 100% of the lessons ($M = 75%$) of an alcohol prevention program. They also found that those who delivered at least one lesson compared to those who delivered no lessons had less teaching experience, more experience with the teaching methods used by the program, more enthusiasm for the program, more active participation in training, and

greater self-efficacy for implementation. Further, program integrity was positively related to teacher skill levels, program acceptance, and self-efficacy in implementation, although also related to fewer years of teaching experience. Finally, Han and Weiss (2005) found that professional burnout is negatively related to implementation, while greater numbers of students with behavior problems in the classroom are associated with greater acceptance of interventions. Few studies, however, have focused on factors related to the process and feasibility of implementing universal prevention curriculums in preschool settings where teachers often are undertrained and under-paid, and expulsion or termination of children with behavior problems is a more common solution.

McCall (2009) and others have emphasized the need for greater research on implementation processes and fidelity, even before outcomes are studied. The purpose of this paper is to report on the process and fidelity of implementing an established social skills curriculum, delivered by teachers in eight preschool classrooms located in two different sites over two years. The implementation required all teachers in the intervention sites to be trained to deliver an adapted version of the Second Step curriculum in their classroom (Second Step Preschool/Kindergarten version; Committee for Children, 2002). Second Step is a social-emotional curriculum focusing on emotion regulation, prosocial skills, and problem solving. This implementation study is part of an efficacy pilot study examining classroom behavioral outcomes between four preschool sites randomly assigned to an intervention or control condition. Classroom behavior and teacher observations were conducted in both intervention and control classrooms and will be reported in a separate paper focusing on outcomes (manuscript under review).

For the purposes of the present paper, the focus will be on describing the implementation process in the intervention sites, and investigating: 1) the rates of teacher attendance in training/coaching sessions each year, 2) changes in teacher reported self-efficacy in delivering the lessons over time, and 3) the level of implementation (number of lessons delivered, percent of main lesson points covered), and fidelity (based on independent observer ratings) achieved each year. Further, given previous research on implementation level and fidelity of school intervention curricula, we also investigated the extent to which: a) teacher characteristics (years of experience, observed teacher-child interaction quality, burnout, rate of participation in curricular training, and self-efficacy in implementing the curriculum); and b) classroom characteristics (age of children, number of children with developmental delays, number with behavioral issues, and classroom climate) are associated with implementation and observed fidelity. We also assessed teacher's ratings of satisfaction with the training, ease of implementation, and usefulness of the curriculum. Finally, we collected curriculum implementation and fidelity data over a third year to assess sustainability after training and support had been withdrawn.

Method

Description of Centers

The two participating centers were full time, full year programs enrolling preschool children (2 years 9 months through age 5) of diverse backgrounds (53% Latino, 16% Caucasian, 24% African American, and 7% other ethnicities). In terms of annual household income, 46% had incomes of less than \$20,000, 34% had incomes between \$20,000 and \$40,000, and 20% had incomes greater than \$40,000. The centers were located in a New England city with a population of approximately 780,000. Both were accredited by the National Association for the Education of Young Children (NAEYC). Center A had two classrooms, four teachers, and a total capacity of 32 children, while Center B had six classrooms and eleven teachers, with a total capacity of 96 children.

Baseline Characteristics of Teachers and Teacher Turnover

Table 1 shows the demographics, education, experience, and self-reported burnout of the teachers in the intervention classrooms at baseline for Years 1 and 2. There was some staff turnover in both centers in both years. In Center A, one of the four main teachers left and was replaced in Year 1; in Year 2, two teachers continued and two were new, while one continuing teacher was subsequently replaced midyear. Thus in Center A, teacher turnover was 25% in Year 1, but 75% in Year 2. In Center B, three of the eight main teachers left and were replaced in Year 1; in Year 2, one of the classrooms was combined with another classroom because of lower enrollment (so there were five instead of six classrooms), and seven of the eight teachers continued, but one left mid-year and was replaced. Center B thus had turnover of 37.5% in Year 1, and 25% in Year 2. All of the head teachers in the Year 2 classrooms in both centers continued during the Year 3 sustainability monitoring year.

There were no significant differences between the two centers in Year 1 except in one area: Teachers in Center A had significantly less education than the teachers in Center B ($t = -2.36$, $df = 13$, $p = .034$). They did not differ significantly in number of years of experience, minority status, age, marital status, observed teacher-child interaction quality, or teacher burnout. In Year 2, there were no differences in teacher education, years of experience, age, or teacher burnout. However, Center A had a significantly greater percentage of minority teachers than Center B (75% and 22% respectively; $\chi^2 = 4.74$, $df = 1$, $p = .03$). Further, Center B was significantly higher than Center A in observed teacher-child interaction quality ($t = -3.94$, corrected $df = 4.16$, $p = .016$). (See Measures section below for a description of teacher-child interaction quality.)

Baseline Classroom Characteristics and Child Turnover

Table 2 shows the baseline classroom characteristics of the intervention centers in Years 1 and 2 in terms of class size, student risk factors (percentages of children with behavioral problems, developmental delays, and low income), average child age, and classroom environment ratings. There were two significant differences in classroom characteristics in Year 1. Center A had a greater percentage of children with developmental delays in at least one area ($t = -3.18$, $df = 6$, $p = .02$), and scored significantly lower in terms of quality of classroom interactions ($t = -2.51$, $df = 6$, $p = .05$) than Center B. In Year 2, there were no significant differences in the classroom characteristics. In terms of child turnover, in Year 1 90% of children in Center A and 82% of children in Center B were there for both fall and spring assessments. In Year 2, 75% of Center A children and 72% of Center B children were present at both time points. Thus, there was greater child turnover in both centers in Year 2.

Second Step Curriculum and Adaptations

Second Step is a classroom-based curriculum for children in Pre-K through middle school. It is designed to develop socio-emotional and interpersonal problem solving skills in children. It has been shown to be effective for young children (including preschool age) in increasing both social knowledge and prosocial behavior (McMahon et al, 2000; Moore & Beland, 1992) and decreasing aggressive behavior (Grossman et al., 1997; McMahon et al., 2000), based on behavioral observations. Second Step is rated as a “model” program by the Substance Abuse and Mental Health Services Administration (SAMHSA), is on the National Registry of Effective Prevention Programs for school violence and substance abuse (<http://modelprograms.samhsa.gov>), and is rated as an “exemplary” program by the U.S. Department of Education. In addition to the social skills emphasis, the curriculum activities also align well with early literacy and school readiness requirements now being implemented in preschools.

In the Preschool/Kindergarten version, each classroom uses a kit that includes lesson cards, a teacher manual, puppets, a CD of songs, posters, and small heart-shaped tokens that can be given out as reinforcements. The curriculum is organized into three units: Empathy, Emotion Management, and Problem Solving. Each Lesson Card shows a high quality black and white photo of ethnically diverse children on the front that is designed to demonstrate a key concept being taught in the lesson. On the back of the card (facing the teacher) there are explicit instructions about delivering the lesson, including scripted language and specific activities and materials with which to engage children. Each lesson includes two or three objectives, a warm-up activity, a story and discussion, practice activities, and a wrap-up. Some lessons also have take-home letters to parents with suggestions for reinforcing concepts at home.

The standard procedure is to implement the curriculum once or twice a week in 30 to 45 minute sessions. However, as suggested in the standardized Teacher's Guide, due to the young age of many of the children, lessons for the current study were divided into shorter 15 minute blocks, and teachers were asked to deliver the lessons a minimum of four days per week. Enhancements to the shorter lessons were added to make them flow smoothly and maintain engagement across several lesson days, addressing the concepts in the original lesson cards. Daily lesson checklists were developed outlining the exercises or parts of the lesson to cover in each of the four days. These checklists also served as a measure of what was covered (or not covered), and provided a space for teacher comments. The cardboard hearts that were included in the kits were replaced with wooden hearts that would last longer and that children painted and decorated.

Implementation Process

Acceptance—Klein and Sorra (2004) posit that implementation effectiveness is determined by: 1) the organization's climate for the implementation of the innovation; and 2) staff members' perceptions of its fit with their values. As such, acceptance of the program must include all levels within the organization. Rosenheck (2001) suggests the first step is coalition building that must include not only experts and heads of organizations, but front-line staff, and that new innovations or treatment models must be linked to broader organizational objectives that have "taken for granted legitimacy" (p. 1610).

In this study, both sites were already part of a larger coalition focused on decreasing behavior problems and expulsions in area preschools. As such, the organizational climate in each preschool supported the focus on improving child behavior, and offered the needed "taken for granted legitimacy" for implementation. Further, both centers had previously implemented a mental health consultation model that showed significant improvements for those children who received services in comparison to waitlist controls (Upshur, Wenz-Gross, & Reed, 2009). However, the numbers of children needing services consistently exceeded the capacity of the mental health consultants. In addition, greater teacher training was identified as an ongoing need by center directors. As a result, both sites agreed at the organizational level that implementing a broad-based prevention program had the potential to increase teacher skills and decrease the numbers of children needing the more intensive consultation services. Further, to help to build acceptance and make sure that the intervention fit with their values and organizational structure, preschool administrators and staff were involved in reviewing and choosing a curriculum from several existing programs. Once chosen, feedback was obtained from preschool teachers regarding necessary changes and adaptations needed to make the proposed curriculum workable.

Capacity building—Organizational capacity building is another, often overlooked but crucial step to ensuring implementation effectiveness (Elliott & Mihalic, 2004). Elliott and

Milhaic (2004) found that in replicating and disseminating Blueprint Model Programs for violence prevention, the elements needed for site readiness were strong administrative support, staffing stability, up front commitment of necessary resources, and potential for program routinization. Further problems occur when there is staff turnover and no institutionalized system in place to train new personnel. To address this issue both sites received a small salary offset and tuition support for two administrators or head teachers to attend one of the Committee for Children, three day Second Step Training-of-Trainers workshop. Their role was to provide support and continuity of implementation when teacher and/or administrator turn over occurred.

Training and technical assistance—Adequate training and technical assistance are also crucial for implementation success. Klein and Sorra (1996) suggest that training be available and supported for all targeted employees, additional assistance be available following training, ample time be given so that skills can be practiced and used on an ongoing basis, and that staff concerns and complaints regarding the use of the intervention be responded to at the administrative level.

In this study, teachers and administrators met as a group (both centers together) for training and coaching sessions to foster a sense of cohesiveness and collaboration, and to provide a forum for sharing ideas and troubleshooting. Training schedules were developed collaboratively with the sites to meet staff needs. Teachers attended two hours of initial training provided by study staff and the lead teachers/administrators who attended the “train the trainer” workshops. This was followed by evening group coaching meetings (1.5 hours each) with a catered dinner, each taking place at one of the child care sites at the end of the school day. The group coaching sessions met almost monthly from November to June in Year 1 (seven sessions) and every other month in Year 2 (for a total of five sessions). Teachers received hourly compensation (\$12.00 per hour) for evening group training meetings and were provided continuing education credits for the time spent in training to assist with buy-in.

At each training meeting, teachers shared which lesson they were on, brought in materials they developed to supplement the lessons and demonstrated them for other teachers, discussed what was working well, and what problems they were having. This provided teachers with a basis for comparison (how quickly they were progressing with lessons) and a forum for reinforcing and highlighting progress. During these meetings, study staff and site administrators provided general feedback from their monthly observations of classroom lessons, emphasized key concepts, responded to concerns, and offered suggestions when teachers expressed difficulties they were having with implementation. The group format proved highly useful in providing an opportunity for teachers to offer their own tips and suggestions of what worked for them, to validate each other’s struggles and efforts, and to decrease feelings of isolation. Teachers seemed to value each other’s ideas as much or more than those of the “expert staff.” In addition to the group meetings, each teacher was formally scheduled to watch one lesson delivered by another teacher in a different classroom. The observing teacher was asked to give a brief verbal summary of what worked well and ways the lesson delivery could have been improved.

To estimate sustainability after the initial two years of training and support, we measured lesson implementation and fidelity for one more year (Year 3). During this year, no further training and support were provided, and other than lesson completion and fidelity, no other teacher or child information was collected. Our purpose was to see how well the intensively trained sites would sustain the curriculum.

Measures

Teacher demographics—All teachers were asked to fill out a survey developed for this study that included questions regarding their age, gender, ethnicity, marital status, parenting status, years of education, years of experience in childcare, years working at their present site, current position in the center, and number of children in their classroom. This questionnaire was filled out before curriculum training in Year 1 and as new teachers came into the centers. This measure was used to assess the relation between baseline teacher characteristics and lesson completion and fidelity.

Teacher burnout—The Maslach Burnout Inventory (Maslach, Jackson & Leiter, 1996) was designed to measure teacher burnout. Teachers were administered the measure at baseline in Year 1 (before Second Step trainings) or as new teachers were hired. Teachers also received a follow-up administration in the spring of Years 1 and 2. The Maslach has three subscales that measure Emotional Exhaustion (e.g., “I feel emotionally drained from my work,” “I feel burned out from my work”), Depersonalization (e.g., “I feel I treat some students as if they were impersonal objects,” “I don’t really care what happens to some students”), and Personal Accomplishment (e.g., “I feel I’m positively influencing other people’s lives through my work,” “I feel exhilarated after working closely with my students”). Chronbach’s alphas for the present sample were .92, .34, and .54 in Year 1, and .93, .62, and .63 in Year 2, respectively. Since the internal reliability of the Depersonalization subscale was so low, particularly in Year 1, only the Emotional Exhaustion and Personal Accomplishment subscales were used to assess the relation between baseline teacher burnout and future fidelity and lesson completion.

Teacher satisfaction with training and ratings of self-efficacy to deliver the curriculum—At the end of each teacher training or coaching session, teachers filled out a brief satisfaction questionnaire developed for this study. The following questions were rated on a 5-point scale from 1 (*not at all*) to 5 (*very much*): “My questions were answered;” “The review was helpful in understanding things I was confused about;” “I received some good tips about delivering the lessons;” “I provided good tips about delivering the lessons.” Teachers also rated how skilled they felt in delivering Second Step curriculum as of that day, from 1 (*not at all*) to 5= (*very much*). This last item was used to assess the relation between teacher self-efficacy in implementing the curriculum and lesson completion and fidelity. While not ideal to use a one-item measure, others studying curriculum implementation and self-efficacy have used similar single items to measure attitudes toward implementing a new intervention (Ghaith & Yaghi, 1997).

Number of lessons completed and percent of main lesson points covered—At the end of each lesson delivered, teachers completed a lesson checklist. This form listed the key activities and lesson tasks for that day. For each of these activities or tasks, teachers rated whether they completed it *fully*, *partially*, or *not at all*. The percentage of key tasks covered *fully* and *partially*, were averaged across all lessons to give a mean percent of main lesson points covered in each category. The number of these lesson checklists completed provided a measure of number of lessons completed by each teacher and for each classroom. In Year 3, only the number of lessons completed, but not the individual tasks completed, were collected.

Fidelity ratings by study observers—Each month the co-authors or trained research staff visited each classroom to observe the teacher delivering one of the study-designed daily Second Step lessons. Lesson fidelity was rated on a five point scale from 1 (*not at all*) to 5 (*very much*) for the following seven items: “Appropriately gets children’s attention before starting;” “Asks children to respond as required;” “Covers all points on the card;” “Properly

rewards children during session;” “Uses Second Step terms throughout session;” “Reminds children about Second Step language throughout session;” and “Reminds children to use language and rules for rest of day.” Item scores were averaged to create a mean for each observation. The Cronbach’s alpha for these seven items was .83 in Year 1, .80 in Year 2, and .81 in Year 3 (the range of reliability for individual items was .73 to .87 across years). Twenty percent of the curriculum observations were conducted by two observers to establish interrater reliability. Absolute agreement was reached at 80% or above for all items.

Satisfaction with implementation—In the spring of each year, teachers filled out a questionnaire developed for this study assessing their satisfaction with the Second Step curriculum. The questions about implementation (rated on a 4-point scale) were: “How much time did it take to learn the lessons?” (1 = *very little* to 4 = *too much*); “How hard was it to incorporate Second Step lessons into your daily curriculum?” (1 = *not at all* to 4 = *very*); “If your center was not part of the project next year, would you want to continue to use Second Step next year in your classroom?” (1 = *not at all* to 4 = *definitely*); and, “What were the most difficult things in implementing the curriculum?” (a checklist was provided along with an open ended response). Teachers were also asked questions regarding how helpful the lessons were in promoting children’s socio-emotional skills and positive behavior (1 = *not much* to 4 = *a lot*); how helpful Second Step was to them as a teacher (1 = *not much* to 4 = *a lot*); and how successful the lessons were in improving the classroom environment (1 = *not very* to 4 = *very*).

Teacher interaction quality—The Caregiver Interaction Scale (CIS, Arnett, 1989) was used to assess the quality of teacher-child interactions in the fall and the spring of each year. The CIS is a widely-used observation tool that specifically measures each teacher’s behavior with their students using a 26-item, 4-point Likert scale, yielding four factors: Positive, Punitive, Permissive, and Detached. This measure is associated with teacher education and training, child attachment security, and child language development (Arnett, 1989; Whitebrook, Howes, & Phillips, 1990). Eight observers, recruited from local education and psychology graduate schools, were trained to 85% agreement using centers not part of the study based on three training tapes and two live observations. Observers were blind to the study goals and intervention status of the study centers. Four, 2-hour observations on different days of circle time and free play were made for each classroom at each time point. Most classrooms had two main teachers, and therefore two observations of each teacher were conducted in the 4 days and averaged. Interrater reliability was checked by having two of the observations for each classroom conducted by paired observers; calculating the percent agreement, and resolving differences. Mean interrater reliability was 97.3% agreement within one point (range among observers was 92% to 100%). Cronbach’s alpha reliabilities were .88, .79, .33, and .83 for the Positive, Punitive, Permissive, and Detached subscales respectively. In subsequent analyses, the Permissive subscale was dropped due to poor reliability. The Punitive and Detached subscales were reverse scored and summed with the Positive subscale score. This summed score was used as a measure of baseline teacher-child interaction. The Cronbach’s alpha reliability for this summary score was .83.

Classroom quality—Classroom quality was assessed using the Early Childhood Environmental Rating Scale Revised, (ECERS-R, Harms, Clifford & Cryer, 1998). The ECERSR is a widely used observation measure of classroom environment and interaction quality. It consists of 43 items grouped into seven categories: Space and Furnishings, Personal Care, Language-Reasoning, Activities, Interaction, Program Structure, and Parents and Staff. We used a sum of 24 items covering Space and Furnishings, Language-Reasoning, Activities, and Program Structure as an overall classroom quality measure. This was rated by paired observers and reliability established for each classroom during an extra

(5th) baseline observation visit for each classroom. Each item is rated on a 7-point scale from 1 (*inadequate*) to 7 (*excellent*). Mean interrater reliability for our study was 94.1% within one-point (range among observers was 80% to 100%). The Cronbach's alpha for the present study for these 24 items was .90.

In addition, we used a sum of four items from the Interaction subscale to measure the quality of interactions in the classroom each fall and spring. These items were: Discipline, General Supervision, Staff-Child Interactions, and Child-Child Interactions (the fifth item on this subscale, Gross Motor Supervision, was deleted because observers were not consistently able to see gym or playground time). The Cronbach's alpha for these four items in the present sample was .83 at baseline.

Teacher rated child behavior—Teachers rated all children in their classrooms in the fall and spring of each year using the Sutter-Eyberg Student Behavior Inventory-Revised (Eyberg & Pincus, 1999). The SESBI-R is a widely used classroom screening measure for conduct problems in children ages 2 through 16. The frequency of occurrence of each problem behavior (38 items) is rated from 1 (*Never*) to 7 (*Always*), and summed to create an Intensity Rating. Further, teachers indicate *yes* or *no* to the question “Is this a problem for you?” The number of *yes* responses is summed to create a Problem Rating. For the purposes of this study, only the Intensity Rating subscale was used. The Cronbach's alphas for the current sample were .97 in Year 1 and .98 in Year 2. The manual provides a method for conversion of raw scores to t-scores, and clinical cut offs are given indicating children whose problems are significant. The percentage of children in the classroom above the clinical cutoff levels at baseline was used as a classroom characteristic in analyses.

Child developmental delays—Child developmental delays were assessed using four subscales of the Child Development Inventory (Ireton & Glascoe, 1995) teacher version (the Teacher Observation Guide), which was designed for use by teachers of infants, toddlers and preschoolers. This measure was designed as a screener only, to be followed by referral for full developmental assessment for children of concern to center staff. The measure was normed on a primarily white population and the authors suggest care should be taken with culturally and linguistically diverse children. We did this by using a bilingual research assistant for all Latino children. The study research assistants blind to study hypotheses, individually assessed all children whose parents consented (99% of all children) at baseline in Year 1, using the Fine Motor, Receptive and Expressive Language, and Numeracy subscales. These subscales were chosen because they could be accurately assessed by research staff in an individual session and related to the conceptual and language demands of the curriculum. Staff visited the classrooms, sat with each child one by one away from the ongoing activities for 20-45 minutes, asking them to complete tasks such as counting, color identification, puzzle completion, letter recognition, picture drawing, and answering simple questions (to gauge language skills). Classrooms with a high percentage of Hispanic children were visited by a bilingual research assistant. The percentage of children in each classroom who met criteria for developmental delays on these subscales was calculated and used for analyses. For most of the children, this was their first developmental screening and they did not have prior special education services. However, the information was used by parents and sites to identify and request additional services for children if children's development continued to be of concern..

Analyses

T-tests (for continuous variables) or chi-square (for categorical variables) analyses were used to examine baseline differences between the two sites. Further, descriptive statistics (means and percentages) were used to examine training, implementation and fidelity, as well

as teacher satisfaction. Finally, bivariate correlations were used to analyze the relation between the implementation and fidelity measures and baseline teacher and classroom characteristics. Because of the small sample size and non-normal distribution of the scores, Spearman correlations were used instead of Pearson product moment correlation coefficients which require the assumption of normal distributions (Gibbons, 1993). Further, because of the small sample size and exploratory nature of the study, the statistical significance level for these correlations was set at $p < .10$. It is appropriate to use $p < .10$ to indicate trends when sample size and power may be limited and when the purpose is to establish evidence upon which to conduct more rigorous subsequent studies (Maxwell & Delaney, 2003).

Results

Teacher Participation in Training and Coaching

There were eight lead teachers in Year 1 from the two centers. These teachers attended between 43% and 100% of the seven training/coaching sessions. Three of the eight teachers attended all sessions; three attended all but one session; one teacher attended all but two; and one attended only three sessions. Assistant teachers (who generally taught fewer lessons) attended fewer training/coaching sessions (range from 29% to 71%). At the classroom level (looking at whether at least one teacher came to each session), six out of the eight classrooms had perfect or near perfect attendance (missing only one), while one classroom missed two sessions and one missed three. In Year 2, one of the seven lead teachers attended all of the five training/coaching sessions given, while the other six lead teachers attended four out of five sessions. At the classroom level, two out of the seven classrooms had at least one teacher present at each of the five training sessions and the other classrooms had a teacher represented at four out of five sessions. Therefore, training participation was high in both years but showed some improvement in Year 2.

Teacher Self-Efficacy

In Year 1, most teachers showed increasing self-rated skill levels in delivering the lessons over time (1 = *not at all skilled* to 5 = *very skilled*). Specifically, nine teachers (64%) reported increasing skills over time, four (29%) remained the same, and one (7%) reported decreasing feelings of competence in delivering the lessons (this teacher only attended two training sessions). Mean skill level across teachers and lessons was 3.75 out of 4 points ($SD = .86$). In Year 2, most teachers (66%) remained the same in their self-rated skills in delivering the lessons from the beginning to the end of the year, while a third of the teachers increased in their self-efficacy in delivering the lessons from the beginning to the end of the year (mean across teachers and lessons = 3.87, $SD = .92$).

Implementation and Fidelity

Year 1—Of the 25 lessons, teachers completed between 17 and 25 ($M = 22.06$, $SD = 3.00$) lessons by the end of Year 1. In terms of the percentage of main points covered fully, teachers reported an overall mean of 87% full completion across lessons (range = 74% to 99%) and a mean of 8% partially completed main points (range = 0% to 24%). Across classrooms, the overall mean observed fidelity ratings (on the 1 to 5-point scale) ranged from 3.11 to 4.25 ($M = 3.62$, $SD = .35$). The correlation between observed fidelity and teacher reported percent of main points covered fully was marginally significant ($r_s = .69$, $n = 8$, $p = .058$).

Year 2—Results were similar in Year 2. Of the 25 lessons, teachers completed between 11 and 25 ($M = 22.86$, $SD = 5.24$). Only one classroom with very young children and a high percentage of children with diagnosed special needs completed less than 24 lessons and reported only 49% of main curriculum points covered fully. In terms of the percentage of

main points of lessons that were covered fully, teachers in Year 2 reported a mean of 86%, with all but the classroom previously mentioned reporting between 82% and 99%. In addition, a mean of 8.5% of main lesson points were covered partially (range = 0% to 37%). Across classrooms, mean observed fidelity ratings ranged from 2.90 to 3.93 ($M = 3.48$, $SD = .36$). Observed fidelity was significantly correlated with teacher reported percent of main points covered fully ($r_s = .86$; $n = 7$, $p = .014$).

Year 3—As in Year 2, all but one of the 7 classrooms completed all 25 lessons, $M = 23.86$, $SD = 3.02$, a slightly higher lesson completion rate than Year 2, with a mean of 95% of all lessons delivered. The one classroom that completed only 17 lessons was a small class that had many enrolled children referred for mental health consultation services. The mean observed fidelity ratings in Year 3 improved over both the prior years of implementation to $M = 4.09$ ($SD = .45$, range 3.12 to 4.37).

Correlates of Implementation and Fidelity

Bivariate correlation analyses were conducted to explore the relation between both teacher and classroom characteristics and the number of lessons completed, the percent of main lesson points delivered fully, and the observed fidelity of lesson delivery. Moderate to high associations ($r = .40$) were evident between many of the teacher and classroom characteristics at baseline and the subsequent teacher success in implementing the lessons (e.g., number of lessons delivered, main points covered) and the observed fidelity with which they were delivered over the school year. However, many moderate to high correlations did not reach statistical significance, possibly due to the small sample size.

Teacher characteristics—Table 3 reports the patterns of teacher characteristics related to the three implementation measures (teacher-reported number of lessons delivered, number of main curriculum points covered fully, and independently observed fidelity ratings). The number of lessons completed was significantly and positively related to teacher self-report of baseline Personal Accomplishment, and teacher education level in Year 1; to percent of training sessions attended, and teacher self-efficacy in delivering the lessons in Year 2; and to observer rated teacher-child interaction quality in both years. Percent of full lesson implementation in Year 1 was significantly and positively related to baseline Personal Accomplishment and negatively related to years of teaching experience. In Year 2 percentage of full lesson implementation was positively related to self-efficacy in delivering the lessons. Finally, observed fidelity was significantly and positively related to teacher self-efficacy in both years, and to teacher-child interaction quality in Year 2.

Classroom characteristics—Table 4 shows the pattern of association between baseline classroom characteristics and the three implementation measures. In both years, the number of lessons completed was significantly and positively associated with classroom interaction quality at baseline, and in Year 2 with the overall environmental quality of the classroom (measured at baseline in Year 1). Full implementation of the main lesson points in both years was significantly and positively related to classroom interaction quality, and negatively related to the percent of low income children in the classroom. In addition, less implementation of main lesson points was associated with having greater numbers of children with developmental delay (only measured in Year 1), and higher teacher turnover in Year 2. Finally in Year 2, observed fidelity was significantly and negatively related to both child and teacher turnover.

Teacher Satisfaction with Training and Curriculum Implementation

Satisfaction with training—Overall, teachers ($n = 15$ in both years) reported being satisfied with the training sessions. Mean responses to all four items across the seven

training sessions in Year 1 and the five in Year 2 were all above 4.0 (on the 1 = *not at all* to 5 = *very much* rating scale). Means and standard deviations for Years 1 and 2 respectively for each question were: “My questions were answered” ($M_s = 4.53$ ($SD=.77$) and 4.50 ($SD=.81$); “The review was helpful in understanding things I was confused about” ($M_s = 4.52$ ($SD=.69$) and 4.40 ($SD=.85$); “I received some good tips about delivering the lessons” ($M_s = 4.58$ ($SD=.68$) and 4.45 ($SD=.75$); and “I provided good tips about delivering the lessons” ($M_s = 4.08$ ($SD=1.07$) and 4.11 ($SD=1.13$).

Satisfaction with implementation and curriculum usefulness—Table 5 shows end of year satisfaction data for Years 1 and 2 for all teachers who were still employed at the centers in the spring of each year ($n = 12$ in Year 1 and $n = 13$ in Year 2). Responses show most teachers in both years report that it took *very little* to *some* time to learn the lessons, with fewer than a quarter of teachers reporting that it took *a lot* of time (no teachers said it took *too much* time). Similarly, most said it was *not at all hard* or *a little hard* to incorporate the lessons into their daily curriculum. In the first year, when asked what was difficult about implementing the curriculum, most of the items were seen as problematic for the majority of teachers, whereas in Year 2, only one item was checked by the majority of teachers: 77% indicated “getting children to sit for circle” was difficult. Most classrooms had a free play period before circle time and the transition to a structured teacher-focused activity remained challenging. In addition, while classrooms had a daily circle time prior to Second Step project implementation, it was less formal than when implementing Second Step, and teachers were less concerned with keeping the children’s attention. In terms of sustainability, in Year 1, the majority of teachers said that they would *probably* use the curriculum next year, while in Year 2, most teachers said that they would *definitely* use the curriculum the following year. Finally, across both years, mean teacher scores indicated that teachers felt that the curriculum was helpful to their children, to themselves as teachers, and to improving the classroom environment.

Discussion

This paper describes the implementation process, organizational support, and training procedures used to implement an adapted Second Step curriculum in two urban preschool settings. The implementation process was planned to address issues of acceptance, capacity building, organizational support, and teacher training and support that often impede the adoption and sustainability of interventions within school settings. Using a multi-method approach, implementation quality and fidelity, including number of lessons taught, number of main curriculum points covered, and independently observed adherence to the curriculum content and procedures were documented. In addition, teacher and classroom characteristics that may be associated with successful implementation were identified. Data on teacher satisfaction with the training and implementation process, their intent to continue to use the curriculum, and the extent to which they felt the curriculum was of value to improving child behavior, social skills, and the overall classroom environment were provided. Finally, data on curriculum sustainability and how implementation fared with a more modest level of teacher training at a new site were collected.

Implementation Process, Teacher Training and Coaching

Clearly, curriculum interventions are challenging to successfully implement. In our work we applied theoretical principles from implementation research to overcome these challenges. First, we invested in a strong implementation training model which appears to be effective. Attendance at the training sessions was high in Year 1, and increased to nearly perfect attendance in Year 2. Teachers also rated these training sessions highly, both in terms of the information they obtained and the amount they contributed to them. Teachers seemed

invested in the process and, we believe, came to own the program. While administrative support was not measured independently, the program directors and supervisors participated in training and coaching sessions on a regular basis, completed feedback ratings on those sessions, and otherwise demonstrated ongoing commitment.

Second, the group training format appeared to keep even struggling teachers going. For example, it is likely that several teachers might have given up had they not been asked to continue to attend the group meetings and hear similar struggles of other teachers. Reported self-efficacy in delivering the curriculum also increased over time for most teachers, suggesting that teachers felt more confident and skilled as they attended these training sessions and had the opportunity to use the lessons over time. Similarly, teachers reported fewer areas of difficulty with implementation in the second year, endorsing only that it was “difficult to get children to sit in circle,” a major change from Year 1 where all listed difficulties were endorsed by a majority of teachers. Third, because multiple teachers were trained at each site and sometimes in one classroom, even with staff turnover, new teachers were easily trained and expected to sustain implementation through the culture created at the sites for implementation. This made it relatively painless for teachers and administrators to incorporate ways to introduce new teachers to the curriculum by having them observe other classrooms and assigning a continuing teacher to provide guidance. The self-reinforcing momentum of a large number of well trained teachers seemed to result in continued high lesson completion and fidelity in Year 3.

Finally, the curriculum workbook created by the study team included a daily recipe with the needed materials from the Second Step kit, supplementary activity suggestions, and a limited number of well-scripted curriculum points, along with a checklist where teachers were asked to indicate what they actually delivered. This simple system provided techniques for delivering the lessons and a structured reminder of how well the teacher was progressing through the lessons on a regular basis. Site administrators were also asked to periodically review the checklists. This structure seemed to work well to remind teachers to complete the lessons, and was self-reinforcing. The structured checklists likely also contributed to the continued high implementation rates in Year 3. The teachers also seemed to have a positive view of the curriculum and its benefits for their classroom. The intuitive appeal of the curriculum has been found in other studies, even when outside intervention staff is implementing it in classrooms while the teachers watch (Edwards, Hunt, Meyers, Grogg & Jarrett, 2005).

Implementation Rate, Fidelity, and Sustainability

By focusing on, and addressing up front, common barriers related to acceptance (both on an organizational and staff level), and organizational capacity (for both training and staff turnover), we believe we were able to garner initial enthusiasm for implementation. Further, the training and coaching model used appeared to deal successfully with the need for ongoing support and time to practice and troubleshoot problems as they arose, and served to increase teacher follow through and buy-in. We found relatively high overall implementation rates and fidelity in the first two years of the study, despite significant staff turnover and a planned reduction in site training support in Year 2. In Year 1, a mean of 88%, and in Year 2 a mean of 92% of all lessons were delivered. Fidelity was slightly, but not significantly less in Year 2, most likely due to the high turnover in teachers that year. In Year 3, with complete withdrawal of support, and only monitoring, 95% of all curriculum sessions were delivered. More importantly, independent observations of fidelity found moderately strong fidelity (mean above the midpoint of the rating scale) in all three years, including an increase in fidelity in Year 3. Thus, although there was some variability across classrooms, overall high implementation and fidelity appeared to be attained, and continued

to improve as demonstrated by ratings in Year 3, even though no support or training were provided.

The level of training and financial support for ongoing implementation meetings in Years 1 and 2 may not always be replicable. However, all child care programs in the state have mandatory staff meetings and must deliver 20 hours of continuing education per year. Because of scheduling needs to reach all staff, most sites have evening staff meetings where staff are compensated with overtime hourly pay. Continuing education is most often tied to staff meetings and involves invited speakers on various topics, or re-training on required health and safety issues. The minimal cost is built into operational budgets. Using the opportunity of required staff meetings and continuing education time is one way more widespread dissemination can be achieved without cost or time much beyond what sites already must expend. However, sites should be cautioned not to ignore teacher and classroom factors (discussed below) that we and others have found to be associated with implementation of teaching interventions, even if they must develop somewhat less intense training and support than used in this study.

Correlates of Implementation and Fidelity

Similar to previous studies of curriculum delivery, higher implementation success and fidelity were strongly (though not always significantly) related to both teacher and classroom characteristics. The small sample size of this pilot study limits power for achieving statistical significance, but findings are congruent with previous research using larger samples.

Teacher characteristics—In terms of baseline characteristics of the teachers, we found that teachers who reported a greater sense of personal accomplishment in their work delivered more of the curriculum in the first year. This parallels findings from other school-based mental health interventions (Han and Weiss, 2005). Also, similar to other studies of curriculum implementation in elementary and middle school (Dusenbury et al., 2005; Rohrbach et al., 1993), greater implementation was related to fewer years of teaching experience, again only in Year 1. Ghaith and Yaghi (1997) found in their study of teacher attitudes, more experienced teachers rated the implementation of a new curriculum as more difficult, saw less congruence with their own teaching practices, and felt the innovation to be less important than did less experienced teachers. We found acceptance in Year 2 was greater across teachers of all backgrounds, suggesting that more than one year of implementation may be needed to overcome initial reluctance of more experienced teachers to adopt new practices.

Other teacher characteristics, such as teacher-child interaction quality as rated by outside observers, attendance at training sessions, and teacher rated self-efficacy in delivering the curriculum were also strongly and significantly related to implementation and fidelity, particularly in Year 2. However, education levels were not related to implementation, except in terms of how many lessons were completed in the first year. More than half of the teachers in this study were high school graduates working on 2-year college degrees, so an extensive educational background did not appear to be necessary for successful implementation. Moreover, the types of coaching and daily curriculum logs implemented in this study seem to demonstrate that even paraprofessional-level teachers can reach high quality in adopting evidence-based practices with appropriate support and skill building. However, those teachers who have poorer interactions with children, (regardless of experience or education), and those who have lower rated curriculum self-efficacy may need more ongoing support to successfully implement an intervention. Taken together, these findings suggest that paying attention to issues of morale and general skill development, and

providing adequate support for training over time are important in assisting the broadest group of teaching staff to successfully implement a curriculum intervention.

Classroom characteristics—Congruent with previous studies (Dusenbury, et al., 2005; Ghaith & Yaghi, 1997; Rohbach, et al., 1993), classroom characteristics also played a role in successful implementation. Higher quality classrooms, both in terms of the physical environment and programming, as well as the quality of classroom interactions at baseline were associated with greater implementation success. Similar to having teachers with higher interaction skills, higher quality classroom environments at baseline appear to foster a climate that makes lesson implementation more successful.

Further, two child characteristics were negatively associated with full lesson delivery—having a high percentage of children with developmental screens indicating delays, and having a high percentage of low income children. Having higher percentages of children with these risk factors was associated with fewer lesson points covered fully. Lower socioeconomic status has been linked to cognitive delay (Bradley & Corwyn, 2002), so these factors may be inter-related. The Second Step curriculum was developmentally designed for children aged 4 years old and up and has not been tested specifically on children with developmental delays. Our adaptations were designed to make it accessible to developmentally younger children, but some of the classrooms had large numbers of children 2 years 9 months, and children who exhibited delays in skills of typically developing preschoolers at least at the beginning of preschool (most likely due to environmental risk rather than true developmental delay). The curriculum maybe less well suited to very young and somewhat delayed children without further adaptation, however, we did find that by the end of the school year, almost all children seemed to have learned important aspects of the curriculum regardless of their developmental or behavioral status.

Finally, child and teacher turnover, not surprisingly, was also negatively associated with the percent of curriculum points covered, and observed fidelity in Year 2. Teacher turnover was particularly high in Center A in Year 2, and overall, child turnover was higher in Year 2 in both centers. New teachers coming in needed to be trained, and while structures were in place to support new teachers in learning the curriculum, there was some cost to implementation and fidelity in classrooms with turnover. Teachers also struggled to help newly enrolled children learn early curriculum concepts while continuing to deliver the latter lessons to the continuously enrolled children. However, implementation remained high due to center buy in, training of administrators, as well as continuing on site support into a second year. The benefits seem to be reaped in Year 3 where even higher rates of lesson completion and fidelity were achieved without continued training and support. For broader dissemination, a specific plan for addressing how to train new teachers would be important.

Teacher Satisfaction

Overall, most teachers felt that the curriculum was useful in helping children develop positive socio-emotional skills and behavior, providing them with tools to promote children's skills, and improving the overall classroom environment. Teacher end-of-year ratings in Year 2 showed potential sustainability with almost 62% saying that they would "definitely," and another 23% saying that they would "probably" use the curriculum after the project ended. Thus there was evidence for increased implementation in the second year, with an expressed strong intent to continue implementing the curriculum beyond study funding.

Importantly, we also found some evidence of curriculum effectiveness in improving teacher interactions with children, and child behavior relative to control classrooms. Within intervention classrooms there was a relationship between greater implementation of the

curriculum and improvements in observed child behavior (manuscript under review). While further research is needed, the results of this pilot study suggest that this model of implementation—including preliminary involvement from both administration and staff in choosing an intervention, cross-site group training and coaching that extends to two years, and training of on-site administrators to train and supervise new staff—may be a promising approach for implementing primary prevention interventions in early childhood programs.

Limitations

As an intervention development project, this study is limited by a small sample size (two sites and eight classrooms). This impacts the analyses that can be conducted, the conclusions that can be drawn, and the generalizability of the findings. However, findings were congruent with prior curriculum implementation studies in elementary and middle school, including use of the same curriculum with older children, and paint a consistent picture of some of the teacher and classroom factors that are important in transporting an evidence-based practice into real world settings. While this study also took a multi-method, multi-source approach to assessing key study factors, some of the results could be due to common source variance. For instance, teachers who have a general positive attitude about their work, may have been positive about their implementation, self-efficacy, personal accomplishments, etc., driving the associations between teacher characteristics and implementation. However, implementation fidelity was also observed independently and correlations between teacher reported implementation and observed fidelity were moderate to high. Further, teacher measures of burnout (measured at baseline), curriculum implementation (measured daily), and self-efficacy (measured monthly) were all assessed at different times of the year and over the course of the year. Therefore we believe that there is not a troublesome level of bias in the teacher measures. Another limitation is that the measurement of self-efficacy was based on only one item measured over multiple occasions. One item scales are not ideal, however the results generated were consistent with previous research (Ghaith & Yaghi, 1997). Future studies should use a multi-item measure of this construct. Finally, the Personal Accomplishment subscale had relatively low internal consistency. Therefore, reported results based on this measure should be viewed with caution.

Conclusions

This study demonstrated that evidence-based interventions to prevent behavior problems can be feasibly implemented and sustained in typical urban preschool programs with relatively high turnover of children and teachers; low levels of formal teacher education; and typically much lower staff salaries than in elementary schools. We found strong implementation continuity from Year 1 to Year 2 across most classrooms, even as levels of planned training support lessened over the second year, in addition to sustained high implementation and fidelity in a subsequent monitoring year when all training and support was withdrawn. Two years of some continuous support may be necessary for a project to achieve high teacher buy in. In fact, over twice as many teachers in the second year, as compared to the first, reported that they would “definitely” continue to use the curriculum without project support, and this was substantiated by the rate of lesson completion in Year 3.

It is important for communities, policy makers, and funders to understand that without the commitment to fully support implementation, even well developed interventions that have a research base for effectiveness, are less likely to be adopted, implemented with fidelity, and sustained. Funding an initial training, without follow up and organizational support is likely to be both ineffective and a waste of resources (McCall, 2010). Further, attention must be paid to the types of structures and processes needed. The saying “What gets measured, gets done,” is useful, not only for evaluating implementation, but as a tool for implementation.

The very act of filling out the checklists that were developed to measure implementation, seemed to keep teachers on track, were self-reinforcing, and may at least partially account for continued increases in rates of fidelity and lesson completion in Year 3. Developers of interventions, and those studying implementation processes could make greater use of this strategy. In addition, creating a community of practice (Wenger & Snyder, 2000), may be a more effective training and implementation method than training individual teachers and monitoring their individual practice. Even doing training that is site based lacks the power of teachers seeing that other programs are struggling with similar issues, breaks down isolation, and allows for broader professional support. We believe that the cross-site group training/coaching model used in the current study was effective for building and keeping momentum and overcoming individual difficulties that teachers experienced.

This study focused on one part of the whole puzzle in successful prevention work, that of issues of implementation and fidelity. Feasible and sustainable implementation, as well as adequate intervention outcomes, is required to establish effective models to address early childhood social skills development and behavioral problems. The approach used in this study to structure training and implementation support for the curriculum intervention provides a framework and foundation for a larger scale trial with a greater number of sites to further develop this promising primary prevention model.

Acknowledgments

This study was funded by an Intervention Development Grant, R34 MH073830 from the National Institute of Mental Health. We would like to thank: Lynn Hennigan, LICSW, Director of Services for the Young Child, Community Healthlink, Youth and Family Services; the staff and families of Worcester Comprehensive Child Care Services-Great Brook Valley, the YWCA of Central Massachusetts Salem Square; and our staff of project managers, observers, and data clerks, Gail Sawosik, Judy Davis, Katelyn Allen, Katie Bateman, Courtney Carpenter, Rachel D'Angelo, Alicia Derry, Rachel Friendly, Jenny Hazelton, Karen Nideur, and Laura Seekamp.

References

- Arnett J. Caregivers in day-care centers: Does training matter? *Journal of Applied Developmental Psychology*. 1989; 10:541–552.
- Blair C. School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*. 2002; 57(2):111–127. [PubMed: 11899554]
- Bradley RH, Corwyn RF. Socioeconomic status and child development. *Annual Review of Psychology*. 2002; 53:371–399.
- Committee for Children. *Second Step: A Violence Prevention Curriculum, Preschool/Kindergarten Teacher's Guide*. Seattle, WA: 2002.
- Domitrovich CE, Greenberg MT. The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation*. 2000; 11:193–221.
- Dusenbury L, Brannigan R, Hansen WB, Walsh J, Falco M. Quality of implementation: Developing measures crucial to understanding the diffusion of preventive interventions. *Health Education Research Theory and Practice*. 2005; 20:308–313.
- Edwards D, Hunt MH, Meyers J, Grogg KR, Jarett O. Acceptability and student outcomes of a violence prevention curriculum. *The Journal of Primary Prevention*. 2005; 26(5):401–418. [PubMed: 16228114]
- Elliot DS, Mihalic S. Issues in disseminating and replicating effective prevention programs. *Prevention Science*. 2004; 5:47–53. [PubMed: 15058912]
- Eyberg, S.; Pincus, D. *Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory-Revised: Professional Manual*. Psychological Assessments Resources, Inc; Lutz, FL: 1999.

- Ghaith G, Yaghi H. Relationships among experience, teacher efficacy, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education*. 1997; 13:451–458.
- Gibbons, JD. Nonparametric measures of association. Series: Quantitative Applications in the Social Sciences, #91. Sage Publications; Newbury Park, Ca: 1993.
- Gilliam, WS. Prekindergarteners left behind: Expulsion rates in state prekindergarten systems. Yale University, Child Study Center; New Haven, CT: 2005.
- Grossman DC, Neckerman HJ, Koepsell TD, Liu PY, Asher KN, Beland K, Frey K, Rivara FP. Effectiveness of a violence prevention curriculum among children in elementary school: A randomized controlled trial. *Journal of the American Medical Association*. 1997; 277:1605–1611. [PubMed: 9168290]
- Han SS, Weiss B. Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology*. 2005; 33:665–679. [PubMed: 16328743]
- Harms, T.; Clifford, RM.; Cryer, D. Early Childhood Environment Rating Scale-Revised. Teachers College Press; New York: 1998.
- Hoagwood K, Johnson J. School psychology: A public health framework I. From evidence-based practices to evidence-based policies. *Journal of School Psychology*. 2003; 41:3–21.
- Invest in Kids. Evaluation report of the Incredible Years program in Colorado. Denver: CO; 2004.
- Ireton H, Glascoe FP. Assessing children's development using parents' reports: The Child Development Inventory. *Clinical Pediatrics*. 1995; 34:248–255. [PubMed: 7543037]
- Klein KJ, Sorra JS. The challenge of innovation implementation. *The Academy of Management Review*. 1996; 21:1055–1080.
- Maslach, C.; Jackson, SE.; Leiter, MP. Maslach Burnout Inventory Manual. 3rd ed.. Consulting Psychologists Press; Palo Alto, CA: 1996.
- Maxwell, SE.; Delaney, HD. Designing experiments and analyzing data. Second Edition. Routledge Academic; New York, NY: 2003.
- McCall RB. Evidence-based programming in the context of practice and policy. *Society for Research in Child Development Social Policy Reports*. 2009; 23(3):3–19.
- McMahon SD, Washburn J, Felix ED, Yakin J, Childrey G. Violence prevention: Program effects on urban preschool and kindergarten children. *Applied and Preventive Psychology*. 2000; 9:271–281.
- Moore, B.; Beland, K. Evaluation of Second Step: Preschool-Kindergarten, a violence-prevention curriculum kit: Summary report. Committee for Children; Seattle, WA: 1992.
- O'Neal CR, Brotman LM, Huang KY, Gouley KK, Kamboukos D, Calzada EJ, Pine DS. Understanding relations among early family environment, cortisol response, and child aggression via a prevention experiment. *Child Development*. 2010; 81(1):290–305. [PubMed: 20331668]
- Pankratz MM, Jackson-Newsom J, Giles SM, Ringwalt CL, Bliss K, Bell ML. Implementation fidelity in a teacher-led alcohol use prevention curriculum. *Journal of Drug Education*. 2006; 36:317–333. [PubMed: 17533804]
- Qi CH, Kaiser AP. Behavior problems of preschool children from low income families: Review of the literature. *Topics in Early Childhood Special Education*. 2003; 23:188–216.
- Raver C, Jones SM, Li-Grining C, Zhai F, Metzger MW, Solomon B. Targeting children's behavior problems in preschool classrooms: A cluster-randomized controlled trial. *Journal of Consulting and Clinical Psychology*. 2009; 77:302–316. [PubMed: 19309189]
- Rohrbach LA, Graham JW, Hansen WB. Diffusion of a school-based substance abuse prevention program: Predictors of program implementation. *Preventive Medicine*. 1993; 22:237–260. [PubMed: 8483862]
- Rosenheck RA. Organizational process: A missing link between research and practice. *Psychiatric Services*. 2001; 52:1607–1612. [PubMed: 11726750]
- Schoenwald SK, Hoagwood K. Effectiveness, transportability, and dissemination of interventions: What matters when? *Psychiatric Services*. 2001; 52(9):1190–1197. [PubMed: 11533392]
- Shure, MB. Interpersonal cognitive problem solving: Primary prevention of early high-risk behaviors in the preschool and primary years. In: Albee, G.; Gullotta, T., editors. *Primary Prevention Works*. Sage Publications; Thousand Oaks, CA: 1997. p. 167-188.

- Upshur C, Wenz-Gross M, Reed G. A pilot study of early childhood mental health consultation for children with behavioral problems in preschool. *Early Childhood Research Quarterly*. 2009; 24:29–45.
- Webster-Stratton C, Taylor T. Nipping early risk factors in the bud: Preventing substance abuse, delinquency, and violence in adolescence through interventions targeted at young children (0-8 years). *Prevention Science*. 2001; 2:165–191. [PubMed: 11678292]
- Wenger E, Snyder W. Communities of practice: the organizational frontier. *Harvard Business Review*. 2000; 78:139–145. [PubMed: 11184968]
- Whitebrook, M.; Howes, C.; Phillips, D. Who Cares? Child care teachers and the quality of care in America. Final report of the National Child Care Staffing Study. Child Care Employee Project; Oakland, CA: 1990.

Table 1
Baseline Teacher Characteristics in Years 1 and 2

Teacher Characteristics	Year 1		Year 2	
	Center A	Center B	Center A	Center B
Gender	(n=5)	(n=10)	(n=8)	(n=9)
% Female	80.0	100.0	100.0	88.9
Race/Ethnicity ^a	(n=5)	(n=10)	(n=8)	(n=9)
% African American	0.0	10.0	25.0	0.0
% Caucasian	40.0	80.0	25.0	77.8
% Hispanic	60.0	0.0	50.0	0.0
% Other/Bi-Racial	0.0	10.0	0.0	22.2
Age	(n=5)	(n=10)	(n=8)	(n=9)
Mean	35.80	37.50	34.88	34.67
SD	6.54	11.14	11.27	11.27
Years of Teaching Experience	(n=5)	(n=10)	(n=8)	(n=8)
Mean	13.60	8.50	8.25	8.13
SD	7.67	9.62	9.51	9.46
Teacher Education ^b	(n=5)	(n=10)	(n=8)	(n=9)
% High School Graduate	40.0	20.0	50.0	33.3
% Some College	60.0	10.0	25.0	22.2
% Associates Degree	0.0	20.0	0.0	11.1
% Bachelors Degree	0.0	50.0	25.0	33.3
Teacher Interaction Quality: CIS Sum ^c	(n=4)	(n=8)	(n=5)	(n=8)
Mean	69.63	76.93	58.0	83.08
SD	9.48	5.35	14.08	2.54
Teacher Burnout:	(n=5)	(n=7)	(n=3)	(n=7)
Emotional Exhaustion				
Mean	24.60	31.00	28.67	30.43
SD	14.60	11.53	5.51	10.86
Personal Accomplishment				
Mean	38.40	42.43	38.67	42.0
SD	5.59	4.04	9.07	3.06

^aSignificant difference only in Year 2 ($p = .03$).

^bSignificant difference only in Year 1 ($p = .03$).

^cSignificant difference only in Year 2 ($p = .02$).

Table 2
Baseline Classroom Characteristics in Years 1 and 2

Classroom Characteristic	Year 1		Year 2	
	Center A (n=2)	Center B (n=6)	Center A (n=2)	Center B (n=5)
Class Size				
Mean	16.00	16.00	16.00	17.20
SD	0.0	4.05	0.0	3.11
% of Children with Behavior Problems				
Mean	25.55	26.97	35.70	16.96
SD	20.44	22.89	20.22	19.97
% of Children w/Developmental Delays ^a				
Mean	63.35	22.17	--	--
SD	23.55	13.85		
% Children w/Household Income < \$20K				
Mean	65.95	42.17	50.60	41.46
SD	9.55	20.14	5.52	18.80
Average Child Age				
Mean	48.67	47.49	46.77	44.27
SD	7.45	6.60	1.86	1.87
Environmental Quality:				
ECERS Environment Items				
Mean	4.06	5.59	--	--
SD	0.98	0.44		
Interaction Quality:				
ECERS Interaction Items ^b				
Mean	18.50	23.17	14.75	25.05
SD	0.71	2.47	6.72	2.53

-- Not assessed in Year 2.

^aSignificant difference only in Year 1 ($p = .02$).

^bSignificant difference only in Year 1 ($p = .05$).

Table 3
Spearman Nonparametric Correlations Between Teacher Characteristics and Implementation Success in Years 1 and 2

Teacher Characteristics	Number of lessons completed in this classroom		Mean percent of full implementation across lessons		Mean observer fidelity rating	
	Year 1	Year 2	Year 1	Year 2	Year 1	Year 2
Baseline Emotional Exhaustion ¹	-.110 (n=12)	-.056 (n=10)	.119 (n=11)	-.185 (n=9)	-.207 (n=10)	.042 (n=9)
Baseline Personal Accomplishment ²	.706* (n=12)	.203 (n=10)	.535 ⁺ (n=11)	.489 (n=9)	-.183 (n=10)	.473 (n=9)
Years of Teaching Experience	-.312 (n=15)	-.010 (n=15)	-.526 ⁺ (n=13)	-.136 (n=13)	-.360 (n=12)	.081 (n=12)
Teacher Education	.487 ⁺ (n=15)	.144 (n=16)	-.014 (n=13)	.224 (n=13)	.246 (n=12)	.293 (n=12)
Teacher Interaction Quality: CIS Sum	.649* (n=12)	.806*** (n=13)	.483 (n=9)	.490 (n=12)	.333 (n=8)	.602* (n=12)
% of Training Sessions Attended	.160 (n=16)	.440 ⁺ (n=16)	.428 (n=13)	.461 (n=13)	.436 (n=12)	.197 (n=12)
Self Efficacy in Delivering Lessons	.344 (n=13)	.537* (n=15)	.332 (n=13)	.736*** (n=13)	.882*** (n=12)	.520 ⁺ (n=12)

⁺ $p < .10$

* $p < .05$

** $p < .01$

¹ For this subscale higher scores indicate greater emotional exhaustion

² For this subscale higher scores indicate a higher sense of personal accomplishment

Table 4
Spearman Nonparametric Correlations Between Classroom Characteristics and Implementation Success in Years 1 and 2

Classroom Characteristic	Number of lessons completed in this classroom		Mean percent of full implementation across lessons		Mean observer fidelity rating	
	Year 1 (n=8)	Year 2 (n=7)	Year 1 (n=8)	Year 2 (n=7)	Year 1 (n=8)	Year 2 (n=7)
Class Size	-.298	.159	-.301	-.382	.361	-.346
Percent of Children with Behavior Problems	-.025	-.539	-.167	-.523	.452	-.144
Percent of Children with Developmental Delays	-.513	N/A	-.766*	N/A	-.479	N/A
Percent of Low Income Children	-.272	-.490	-.731*	-.786*	-.575	-.464
Average Child Age	.098	-.401	.333	.036	-.119	-.286
Child Turnover	-.056	-.401	-.301	-.643	.012	-.679 [†]
Teacher Turnover	-.174	-.349	.282	-.791*	.169	-.791*
Environmental Quality: ECERS Environment Items	.491	.802*	.476	.643	.095	.643
Interaction Quality: ECERS Interaction Items	.852**	.809*	.731*	.685 [†]	.611	.649

N/A Not assessed in Year 2

[†] $p < .10$

* $p < .05$

** $p < .01$

Table 5
Satisfaction with Curriculum Implementation in Years 1 and 2

Satisfaction Item	Year 1 (<i>n</i> = 12)	Year 2 (<i>n</i> = 13)
How much time did it take to learn the lessons?		
% Very Little Time	33.3	30.8
% Some Time	50.0	46.2
% A lot of Time	16.7	23.0
% Too Much Time	0.0	0.0
How hard was it to incorporate Second Step into your daily curriculum?		
% Not At All Hard	33.3	38.5
% A Little Bit Hard	50.0	23.0
% Somewhat Hard	16.7	38.5
% Very Hard	0.0	0.0
What were the most difficult things to implementing the curriculum?		
% Learning How to Give the Lessons	60.0	23.1
% Filling Out Daily Reports	71.4	30.8
% Getting Children to Sit for Circle	77.8	76.9
% Thinking Up Different Activities to Make Sure Children Understood	85.7	46.2
% Time for Monthly Meetings	66.7	30.8
If your center were not part of the project, would you want to use Second Step next year?		
% Not At All	0.0	7.7
% Unsure	16.7	7.7
% Probably	58.3	23.1
% Yes, Definitely	25.0	61.5
How helpful have lessons been in helping your children with skills and behavior (1 = <i>not much help</i> to 4 = <i>a lot of help</i>)?		
Mean	3.25	3.31
SD	.754	.630
How helpful have lessons been in helping you as a teacher (1 = <i>not much help</i> to 4 = <i>a lot of help</i>)?		
Mean	3.42	3.31
SD	.515	.630
How successful have lessons been in improving the classroom environment (1 = <i>not very successful</i> to 4 = <i>very successful</i>)?		
Mean	3.00	3.04
SD	.853	.721