Published in final edited form as:

J Prim Prev. 2013 October ; 34(5): 309–327. doi:10.1007/s10935-013-0316-1.

A Pilot Study of a Primary Prevention Curriculum to Address Preschool Behavior Problems

Carole Upshur, EdD. [Professor],

Department of Family Medicine and Community Health, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655, TEL: 774-443-7267, FAX: 774-441-6212, Carole.upshur@umassmed.edu

Melodie Wenz-Gross, PhD [Research Assistant Professor], and

Department of Psychiatry, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655

George Reed, PhD [Associate Professor]

Division of Behavioral and Preventive Medicine, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655

Early behavioral problems are an important target for prevention and intervention in preschool children, since they have significant potential for long-term effects on both social competence and academic performance. For example, one study (Qi & Kaiser, 2003) found that rates of teacher-reported behavior problems among three- to five- year olds in preschool settings ranged from 14% to 52%. Further, expulsion rates for behavior problems have been found to be 3.2 times higher in preschool than that for K-12 students (Gilliam, 2005). Evidence for persistence of early problems is also striking. One study of hard-to-manage preschoolers found that half had continued adjustment problems in kindergarten, 67% met clinical criteria for externalizing behavior problems at age nine, and 94% of those with problems at both six and nine years continued to meet clinical criteria in early adolescence (Pierce, Ewing, & Campbell, 1999).

In seeking to identify the reasons for early behavioral difficulties, a number of studies have focused on children's social and emotional development. These studies suggest that problem behaviors and resulting social skills deficits are mediated by emotion knowledge (e.g., appropriately identifying emotions in oneself and others) and emotion regulation (e.g., appropriately reacting in emotionally charged situations). For example, Denham et al. (2002) found that lack of emotion knowledge and dysregulated anger among children in preschool predicted poor peer status and teacher ratings of oppositional behavior in kindergarten. These investigators also found that children with the least emotion knowledge at ages 3 and 4 demonstrated higher aggression and poorer social competence at ages 4 and 5. (Denham, Caverly et al., 2002). In addition, Izard's studies of Head Start children (Izard et al., 2001; Schultz, Izard, Ackerman, & Youngstrom, 2001) found that preschool self-regulatory behaviors predicted greater emotion knowledge, which were in turn associated with fewer social problems in first grade. Schultz et al. (2001) also demonstrated that early emotion knowledge predicted behavioral skills at age nine.

Summarizing this literature associating emotion knowledge to child behavior, Trentacosta and Fine (2010) conducted a meta-analysis of studies linking emotion knowledge to measures of social competence and both externalizing and internalizing behavior. Emotion

knowledge was defined as the capacity to understand emotions in facial and behavioral cues and in social contexts. Across a variety of ages, settings, socioeconomic and ethnic backgrounds, the investigators consistently found small to medium effect sizes in the relationships between higher levels of emotion knowledge and children's social competence and lower externalizing behavior problems.

The research evidence thus supports the critical role of lack of emotion knowledge and emotion regulation as key risk factors for problem behavior for young children. There is also accumulating evidence that early emotion regulation and prosocial skills have small but significant associations with school learning success. Grimm et al. (2010), for example, reanalyzed longitudinal data measuring the effects of early behavior problems and social skills on later academic achievement. The patterns they reported were not always consistent, but they did find a small negative effect of externalizing behavior among children in kindergarten on reading achievement and mathematics skills in first grade. Even stronger associations were found between early social skills problems and poor achievement in first grade. Further, Romano et al. (2010), using the same data, reported that kindergarten prosocial skills were among the significant predictors of third grade math and reading achievement.

Two approaches to preventive interventions have been developed in response to the evidence addressing the association of deficits in early social and emotional development with later behavioral and academic problems. One approach has been to provide mental health services to individual children in preschool whose problem behavior is noted by teachers as significant. In recent years, a number of preschool mental health consultation models have been developed, with some success (Alkon, Ramler, MacLennan, 2003; Gilliam, 2007; Perry, Dunne, McFadden & Campbell, 2008; Raver et al., 2009; Upshur, Wenz-Gross, & Reed, 2009; Williford & Shelton, 2008). These models provide a combination of teacher training in classroom management practices and on-site classroom support for management of difficult behavior, often paired with interventions targeting individual children. These programs have yielded evidence of improvement in child behavior and social and emotional development (Gilliam, 2007; Perry et al., 2008; Raver et al., 2009; Upshur et al., 2009). However, most models have found mixed results regarding improvements in teacher skills or classroom climate that might help reduce or prevent some children's poor self-regulation behaviors (Alkon et al., 2003; Gilliam, 2007; Upshur et al., 2009). Finally, some investigators found that more children were identified than could be helped individually by classroom mental health consultants, which suggests a classroom-wide intervention may be beneficial (Upshur et al., 2009).

The second approach to addressing young children's behavioral issues has utilized classroom curricula. These curricula are designed to involve teachers and children in social skill building and in preventing externalizing behavior and school violence (Domitrovich, Cortes, & Greenberg, 2007; Izard et al., 2008; Webster-Stratton, Reid, & Stoolmiller, 2008). Findings related to the effects of such curricula have been inconsistent and the feasibility of implementing them in low-resourced community childcare settings can be problematic. For example, Domitrovich et al. (2007) used the PATHS (Promoting Alternative Thinking Strategies) model adapted from school-age curriculum in Head Start classrooms. This study positively affected children's emotion knowledge and social skills but not their teacher-rated behavior. Izard et al. (2008) implemented a different "emotion-based" curriculum, and found the opposite results: improvements in teacher-rated child behavior, but no consistent findings concerning changes in emotion knowledge or social skills. Finally, Webster-Stratton and colleagues (2008), who evaluated the Dinosaur School curriculum, found improvements in teacher classroom management strategies, classroom climate, and

children's school readiness and conduct problems. However, the curriculum was delivered in part by a member of the research staff, an approach not feasible in most classrooms.

Prior evidence for the curriculum evaluated in the current study, the Second Step Preschool/Kindergarten Social/Emotional Learning curriculum (Committee for Children, 2002), is also weak. For example, McMahon and colleagues (2000) found decreases in observed disruptive classroom behaviors after the intervention, as well as improvements in child emotion knowledge. Teacher ratings, however, did not demonstrate improvements in social skills or problem behavior. This study also used graduate students and mental health staff to help deliver the lessons in classrooms, not just the teachers. Nevertheless, Second Step has the promise to be more widely disseminated than other similar curricula due to its reasonable cost and low implementation burden. In addition, implementing a curriculum like Second Step in preschool classrooms constitutes a current best-practice recommendation to use universal social and emotional learning activities, paired with targeted interventions for individual children with more intense needs, to address challenging classroom behavior. (Baker, Kamphaus, Horne, & Winsor, 2006; Bayat, Mindes, & Covett, 2010; Fox, Dunlap, Hemmeter, Joseph & Strain, 2003; Technical Assistance Center for Social Emotional Intervention (TACSE), http://www.challengingbehavior.org/do/pyramid_model.htm). Thus, developing a stronger evidence base for the efficacy of the preschool Second Step Program is critical.

We hypothesized that classrooms receiving the Second Step curriculum would show improvements in: 1) individual children's teacher-rated behavior problems and teacher-rated prosocial skills; 2) classroom climate defined as classroom interactions and two ratings of classroom disruptive behavior (all rated by independent observers); and 3) teacher interaction skills (also rated by independent observers). In addition, we also assessed teacher satisfaction, ease of implementation, and parent engagement with the curriculum.

Method

We conducted a cluster-randomized design in four community child care centers. The Institutional Review Board of the University of Massachusetts Medical School approved the study.

Intervention

The Second Step Preschool/Kindergarten Kit (Committee for Children, 2002) is recommended for children ages 4 to 6 and includes 25 lesson cards, posters, puppets, a music CD, and small cardboard heart tokens that can be used as reinforcements. The lessons focus on empathy, emotion management, and problem-solving. The cards show a large black and white photo of ethnically diverse children and adults in different situations. The back of the card which faces the teacher includes scripts and bullet points to help the teacher deliver the lesson.

Delivery was adapted from the publisher-recommended weekly schedule to four times a week, in 15-minute sessions. Shorter sessions were recommended by the publisher for younger children, but further guidance on how to structure shorter lessons was not provided. We used teacher feedback to craft a workbook that resulted in a total of 89 individual curriculum lessons. The lessons covered all the specific activities given on the 25 lesson cards, divided over several days. Our adaptation, for example, repeated major curriculum points more than once, added story readings related to the curriculum points, and suggested several ways to use the publisher-provided songs and puppets.

Several Intervention teachers and administrators, along with study staff, attended a two-day train-the-trainer workshop sponsored by the Committee for Children. Subsequently, study staff provided seven monthly two-hour evening training sessions to Intervention teachers during Year 1, and five bi-monthly sessions in Year 2. Follow-up monitoring of continued implementation, but no support, was provided during a third year (see Wenz-Gross & Upshur, 2012, for more information about implementation strategies). In addition, four to six parent-group sessions were provided each year to Intervention Centers, using materials in the Second Step Parent Guide (Committee for Children, 2002). The parent sessions covered topics like the importance of social skills to academic and life success, and reviewed with parents the content the curriculum would cover. Parents were also coached on labeling and recognizing feelings at home and using the problem-solving techniques that were taught to their children in the classroom. Intervention centers were provided funds to pay overtime to teachers for the evening training sessions and to attend the parent sessions. Both intervention and control centers received funds to cover a small portion of administrator time to administer consent forms to families and to support the completion of other data collection activities (i.e., classroom observations and teacher ratings).

Control centers were given \$1000 each year to purchase materials of their choice during the study's first two years, and were provided curriculum kits and training to implement Second Step in the third year. During the first two study years, they were instructed to continue the social and emotional activities they usually implemented. In all cases the control classrooms used the Creative Curriculum (Dodge, Colker, & Heroman, 2002). This curriculum has 38 preschool learning objectives with suggested activities and assessments, of which only three focus on emotion regulation and cooperative play. Thus, we assumed a low level of curricular activities targeting social and emotional skills in the control classrooms, although limitations of study resources prevented specific observations of curricula in control classrooms.

Design

We recruited four community-based childcare centers and then randomly assigned them to either the intervention or control condition (two each). The four centers were located in a mid-size Northeastern city and served a majority of families with state-subsidized childcare vouchers, as well as those from low, middle, and higher income families who paid the centers' fees themselves. The centers were all full-time, full-year programs for preschool children aged two years nine months through five years. All had access to mental health consultation services to which they could refer individual children with significant behavioral issues (see Upshur et al., 2009 for more discussion of mental health consultation models). All Centers were accredited by the National Association for the Education of Young Children (NAEYC). The two intervention centers were A, which had two classrooms and a capacity of 32 children, and B, which had six classrooms and a capacity of 96 children; the two control centers were C, with four classrooms and a capacity of 64 children, and D, with three classrooms and a capacity of 60 children. There were a total of eight intervention and seven control classrooms in Year 1. In Year 2, Center D dropped out due to enrollment and financial problems unrelated to the study. Child enrollment changes also resulted in the reduction of one classroom in intervention Center B in Year 2. As a result, the study included seven intervention and four control classrooms in Year 2. Center administrators completed the informed consent process with families and study enrollment was between 95% and 100% in both years, with the exception of Center D. This program in Year 1 (3 control classrooms) enrolled only 60% of families, and as noted above, Center D dropped out of the study in Year 2.

We collected data from classrooms and teachers twice each year, in the fall and spring. In addition, since many measures required observations, we conducted extensive training each

year with a group of part-time research assistant/observers. These observers were trained to 85% agreement on classroom quality, classroom climate, and teacher interaction skills measures using videotapes and live observational visits to childcare programs not participating in the study. Observers were blind to both the study's hypotheses and the intervention condition of the children and centers.

Measures

Teacher and child demographics—Teachers completed a baseline survey regarding their demographic characteristics and teaching experience. Children's parents provided demographic information about their gender, date of birth, parent and child race/ethnicity, marital status, and family income. For children whose parents consented but did not return the parent questionnaire, information about gender, ethnicity, and whether or not the children were referred for mental health consultation services was collected from center administrators.

Teacher burnout—Teachers completed two subscales of The Maslach Burnout Inventory (Maslach, Jackson & Leiter, 1996) at baseline in Year 1 or when newly hired later in the study, and completed a follow-up in the spring of Years 1 and 2. These two subscales, which were titled "emotional exhaustion" and "personal accomplishment," yielded Cronbach's alphas of .92, and .54 in Year 1, and .93, and .63 in Year 2, respectively.

Classroom quality—Paired observers completed 24 items of the Early Childhood Environmental Rating Scale-Revised (ECERS-R, Harms, Clifford & Cryer, 1998) at one two-hour visit in the fall of Year 1. This measure was used to identify potential confounding differences between the intervention and control classrooms before the intervention started. Each item is rated between 1 (*inadequate*) and 7 (*excellent*). We used the average score across items covering space and furnishings, language-reasoning activities, and program structure. Mean interrater reliability for our study was 94.1% within one-point (range among observers was 80% to 100%). The Cronbach's alpha for these 24 items was .90.

Classroom climate—We used three measures to assess classroom climate: 1) the Interaction scale of the ECERS-R (ECERS-R, Harms, et al., 1998); 2) disruptive behavior counts; and 3) a disruptiveness rating. For the ECERS-R Interaction scale, we calculated the sum of four items each rated on a scale of 1 (*inadequate*) to 7 (*excellent*) (possible range of 4 to 28). The four items were: discipline, general supervision, staff-child interactions, and child-child interactions. The ECERS-R Interaction scale was rated by the trained observers in two-hour visits on four days in the fall and four days in the spring of each year, including two paired observations at each time point to establish interrater reliability. Ratings were averaged over the four observations at each time point. The Cronbach's alphas for these four items were .80, .88, .92 and .94 across our four data collection points, and mean interrater reliability within any given point was 93% (range 86% to 100%).

The second climate measure was disruptive behavior counts. Following the same visit schedule as for the ECERS-R Interaction ratings, the observers completed two 10-minute behavioral counts each day (one during free play, one during a structured activity). We followed the operational definitions and time sampling methods described by McMahon et al. (2000) in their evaluation of the Second Step curriculum. Observers counted problem behavior observed among any of the children in the classroom for 30 seconds, using a stop watch/beeper. They then recorded the number and types of behaviors observed for 30 seconds, before observing again for the next 30-second period. This procedure resulted in five 30-second behavior count sampling periods for each 10-minute observation block, or a total of 10 time samples of problem behavior for each visit day. Categories recorded

included: general disruptive behavior (e.g., disturbs other children during structured or unstructured group activities), verbal aggression (e.g., arguing, teasing, bossing, threatening), and physical aggression (e.g., destroys property, fights, attacks others). We calculated a summary score for disruptive behavior counts by averaging the total number of disruptive time periods observed each of four days in the fall and in the spring for each of the three types of disruptive behaviors (range 0 to 30). Interrater reliability was established by percent agreement of paired observers for four of the observation days (two each in the fall and spring). Mean interrater reliabilities were 92% (range 87%, to 94%) for general disruptive behavior, 95% (range 91% to 97%), for verbal aggression, and 95% (range 91% to 97%) for physical aggression.

The third climate measure was a classroom disruptiveness rating assigned by observers at the end of each observation day. This was rated on a Likert scale of 0 (*not at all*), 1 (*minimally*), 2 (*somewhat*), 3 (*mostly*), and 4 (*very*). Disruptiveness ratings were averaged over the four observation days at each fall and spring time point.

Teacher interaction skills—The Caregiver Interaction Scale (CIS, Arnett, 1989) was used on the same visit days as other observations to assess the quality of teacher interaction skills. The CIS is a widely-used observation tool using a 26-item, four-point Likert scale that assesses four factors or subscales: positive, punitive, permissive, and detached. Scores on the measure have been associated with teacher education and training, child attachment security, and child language development (Arnett, 1989; Howes, Phillips, & Whitebrook, 1992). Most classrooms had two teachers; we observed each teacher twice (in the fall and spring), and then averaged the two sets of scores. Every teacher had at least one paired observation for purposes of establishing interrater reliability: in Year 1 there was 97.3% agreement within one point (range among observers was 92% to 100%), and in Year 2 there was 99.8% agreement (range from 96% to 100%). For analysis purposes, the permissive subscale was dropped due to poor reliability. The other three subscales were combined to provide an overall score (range of possible scores 22 to 88 points). The Cronbach's alpha's for the summary score were .83, .78, .81, and .83 across the four time points.

Teacher-rated behavior problems—Teachers rated all consented children in their classrooms in the fall and spring of each year using the Sutter-Eyberg Student Behavior Inventory-Revised (Eyberg & Pincus, 1999). This inventory is a classroom measure designed to screen for conduct problems in children ages two through 16. We used the Intensity Rating subscale for analysis. The authors specify a score of 151 or greater as at-risk of significant behavior problems. The Cronbach's alphas for the current sample were .97, .97, .98, and .98, respectively, across the four time points.

Teacher-rated prosocial skills—Teachers completed the Adaptive Social Behavior Inventory (ASBI, Hogan, Scott and Bauer, 1992) in the fall and spring of each year for all children whose parents provided consent. This scale uses a rating of: 1 (*rarely*), 2 (*sometimes*), or 3 (*frequently*), and has three factors: express (sociability and empathy); comply (prosocial behavior and cooperation); and disrupt (resistant/agonistic behavior). For the purposes of this study, the 23 items included in the express and comply subscales were summed and used as a measure of prosocial skills, following the developer's instructions. Alpha reliabilities for the prosocial scale in our study across the four time points were: .94, .94, .92, and .91, respectively.

Measures Specific to Intervention Sites Only

Lesson implementation—We constructed checklists to be completed by teachers for each of the 89 adapted lessons. These were used to calculate the number of lessons

completed and the percent of curriculum points covered in both study years, and in the follow-up year.

Lesson fidelity—Each month, study staff observed a Second Step lesson in each Intervention classroom during both study years and in the follow-up year. Lesson fidelity was rated on a five-point scale: 1 (*not at all*) to 5 (*very much*) for 7 items: appropriately gets children's attention before starting; asks children to respond as required; covers all points on the lesson card; properly rewards children during the session; uses Second Step terms throughout the session; reminds children about Second Step language throughout the session; and reminds children to use language and rules for the rest of the day. The Cronbach's alpha for these seven items was .83 in Year 1 and .80 in Year 2. Interrater reliability to a criterion of 85% agreement was established, and item scores were averaged to create a mean fidelity rating.

Teacher satisfaction with Second Step—Teachers completed a questionnaire about the curriculum in the spring of each year. Questions asked on a scale of 1 (*not very helpful or not very successful*) to 4 (*very helpful or very successful*): how helpful the lessons were in improving children's social and emotional skills and positive behavior; how helpful Second Step was to them as a teacher in promoting these skills, and how successful the lessons were in improving the classroom environment. Teachers also indicated how likely they would be to continue implementing Second Step without study support.

Parent engagement with the curriculum—At the end of each year, parents reported their degree of engagement with the curriculum. They were asked: if they received the parent letters sent home; how often they tried any of the at-home activities suggested in the letters; and if they attended a parent group.

Analysis Plan

Analyses were conducted to test baseline differences between intervention and control classrooms in terms of child, teacher, and classroom characteristics and all outcome measures. Any demographic data collected for any children present at the beginning of the school year were used to characterize the children and classrooms, even though some of these children left prior to the completion of baseline classroom climate observations and/or teacher ratings, and other children left prior to completion of end of year follow-up assessments. In addition, teacher baseline information was provided for all teachers who started or joined classrooms during the year. Analyses were also conducted to test for demographic differences between those children who had both baseline and follow-up teacher ratings relative to those who did not. For both sets of analyses, independent samples *t* tests were used for continuous variables and chi-square analyses were used for categorical variables.

Analyses of outcomes were examined in several ways. First, unadjusted analyses were conducted using only intervention status as the independent variable, and examining the effect on each outcome. Our primary approach was to calculate change scores by subtracting fall from spring scores in each year. Independent samples *t* tests were used to analyze group-level differences (intervention vs. control) in these change scores and effect sizes (Cohen's *d*) were calculated. Where there were complete pre- and post-test scores for more than one teacher in the classroom, the lead teacher's scores were used.

Three levels of adjusted analyses were then conducted to examine possible threats to internal validity. First, regression-to-the-mean effects were estimated by using one-way analysis of covariance (ANCOVA), using outcome scores as the dependent measure, group

(intervention or control) as the independent variable, and baseline scores as the covariate. These were conducted for all variables for which there were significant change score differences between intervention and control classrooms. Second, we adjusted for the random effects of classroom on the teacher-rated scores of individual children (behavior problems and prosocial skills) to assess whether the findings would differ after controlling for nesting of children within classrooms. Third, we adjusted for potential baseline modifiers of the intervention effect in two stages. We began by examining correlations between each significant baseline teacher, classroom, and child difference and each of the significant outcome change scores. We then regressed any variables with significant bivariate associations along with intervention status on the relevant change score.

Finally, we calculated partial correlations to assess the relationship between classroom change outcomes and curriculum fidelity ratings, controlling for baseline values of the outcome measures. In these analyses, mean classroom scores were created for teacher-rated behavior problems and prosocial skills at baseline and follow-up, and then change scores were calculated from these means. All analyses were conducted in Stata (Stata Corp, 2011). We established a p level of .05 as a threshold for statistical significance.

Results

Participants

Children—Where families had more than one child in a participating center, one was randomly picked for inclusion in the study. Figure 1 provides a CONSORT flow chart of study enrollment and attrition. Each year of the study was treated as a separate cohort, although a small number of children participated in part or all of both years of the study. As the diagram indicates, some level of detail concerning child demographic variables was available for all children in the study, but due to frequent enrollment changes, there were missing data for baseline teacher-rated behavior problems and teacher-rated prosocial skills. In addition, as can be noted by the numbers used for the follow-up child change analysis, a fairly large number of children did not stay the entire school year.

Table 1 shows the characteristics of enrolled children. Different sample sizes for various variables resulted from different sources of data, including the child care center administrators, teachers, and parents. The most complete data were on child gender and whether or not a child was referred to mental health consultation services. Overall, in both years about half of study children were of Hispanic ethnicity and one-quarter were African American. Over 40% of children also came from families with annual incomes below \$20,000. Further, about one-quarter of study children were referred for mental health consultation services for problem behavior each year during the study period.

Teachers—Table 2 displays teachers' demographic characteristics. We included all primary and assistant teachers in classrooms and any new teachers who replaced those who left during the study. Overall, teachers in both years were predominately female and White. Teachers in both years and in both intervention and control conditions also reported relatively high emotional exhaustion, but also high levels of personal accomplishment in their jobs.

In Year 1, there were 15 intervention teachers in eight classrooms and 15 control teachers in seven classrooms. There were two differences between intervention and control teachers in Year 1. Intervention teachers had significantly lower levels of education than control teachers ($\chi^2 = 11.21$, $df = 5$, $p = .047$), and baseline teacher interaction skills were also significantly higher in Year 1 for intervention teachers [$t(19) = 2.25$, $p = .036$].

In Year 2, there were 17 Intervention teachers in seven classrooms, and nine control teachers in four classrooms. There was relatively high turnover of teachers during the two year study. Only 59% of the intervention and 55% of the control teachers in Year 2 were also present in Year 1. In Year 2, teachers in the intervention classrooms were significantly more ethnically diverse than those in control classrooms ($\chi^2 = 6.12$, $df = 1$, $p = .013$). In addition, intervention teachers again had significantly lower levels of education than control teachers ($\chi^2 = 11.03$, $df = 3$, $p = .011$).

Classrooms—Table 3 describes baseline classroom characteristics. Average class size was 15–17 children for both intervention and control classrooms in both years. Mean classroom quality ratings (ECERS-R space, activities, and program structure items) and the classroom climate rating of interaction (ECERS-R interaction scale) were in the ‘good’ range for both intervention and control classrooms in both years. Only the classroom climate measure of disruptive behavior counts in Year 2 was significantly different: Intervention classrooms had higher disruptive behavior counts than control classrooms [$t(8.71) = 4.50$, $p = .002$].

Baseline Differences between Intervention and Control Children

In terms of significant differences at baseline, in Year 1, intervention children were an average of four months younger than children in the control group, and more often male. Intervention children also had higher teacher-rated prosocial skills [$t(163) = 2.01$, $p = .046$]. In Year 2, intervention children were similarly on average four months younger than control children, and also had higher teacher-rated prosocial skills [$t(146) = 3.12$, $p = .002$]. While there were no significant differences in teacher-rated behavior problems, in Year 2, significantly more intervention children were referred for mental health consultation services than were control children ($\chi^2 = 11.24$, $df = 1$, $p = .001$).

Baseline Differences between Children with and without Two Data Points

In Year 1, there were 125 children who had two teacher-rated behavior problem scores and 78 who had only one (either fall or spring) in both the intervention and control groups. There were no statistically significant differences between those with and without two data points in Year 1 with regard to child age, gender, ethnicity, family income, marital status, or whether they were from an intervention or control classroom. However, children who received child mental health consultation services were more likely to be present at both the fall and spring time points ($\chi^2 = 4.11$, $df = 1$, $p = .043$).

In Year 2, 117 children from both intervention and control groups had both data points, while 60 had only one. There were no significant differences in child age, gender, ethnicity, family income, marital status, or whether they received individual mental health consultation services. However, children in the control classrooms were more likely to be present for both time points (80.3%) than were intervention children (57.7%; $\chi^2 = 9.47$, $df = 1$, $p = .002$).

Unadjusted Group Differences in Outcomes

Table 4 shows the change score and independent sample t tests for each outcome variable. In Year 1, only teacher-rated behavior problems differed significantly between intervention and control classrooms (Cohen's $d = -.29$). Inspection of the means revealed that teacher-rated behavior problems in intervention classrooms declined (e.g., improved) by approximately seven points (one quarter of a standard deviation), whereas teacher-rated behavior problems in control classrooms increased by about the same amount. No other outcome measures were significantly different. However, there were effect sizes favoring intervention classrooms in the medium to high range for one classroom climate measure, the ECERS-R

interaction scale (Cohen's $d = .35$), and three ECERS-R interaction items: discipline (Cohen's $d = .83$), general supervision (Cohen's $d = .32$), and child-child interactions (Cohen's $d = .62$).

In Year 2, we found no differences between intervention and control classrooms on teacher-rated behavior problems or teacher-rated prosocial skills. However, significant differences were found between intervention and control classrooms in one measure of classroom climate, the ECERS-R interaction scale, and for teacher interaction skills. For classroom climate, the intervention classrooms had slightly more positive scores, whereas control classrooms had significant and large decreases in the ECERS-R interaction scale ($p = .043$, Cohen's $d = 1.81$). This represented about a 6% improvement in intervention classrooms and a 30% decline in control classrooms. One ECERS-R interaction item, discipline, was also significant ($p = .005$, Cohen's $d = 2.44$). The change score represented a 23% increase in how well discipline was administered in the intervention classrooms, and a 36% decrease in the control classrooms. Further, there were nonsignificant trends (but large effect sizes) favoring the intervention classrooms in both the general supervision and staff-child interactions items.

One other classroom climate measure, disruptive behavior counts, yielded a nonsignificant trend in Year 2, indicating improvement in intervention classrooms and decreases in control classrooms ($p = .099$, Cohen's $d = -1.22$). In addition, the significant difference in teacher interaction skills showed that teacher interaction skills remained stable for the intervention classrooms, and decreased significantly in control classrooms ($p = .03$, Cohen's $d = 1.74$). This indicated a negligible improvement in teacher interaction skills in the intervention classrooms, but a 20% decline in control classrooms.

Adjusted Outcome Analyses

Adjustment for regression to the mean effects—In Year 1, the significant change in teacher-rated behavior problems for intervention children became nonsignificant after adjusting for baseline scores. In Year 2, significant change between groups was observed in teacher interaction skills, the ECERS-R interaction scale, and the discipline item. All of these significant differences remained after adjusting for baseline scores.

Classroom random effects adjustment—The multi-level model, accounting for classroom random effects for the otherwise unadjusted individual child outcomes, resulted in outcomes similar to the unadjusted results. Nesting within classrooms therefore did not contribute new information to understanding individual child change scores in either study year.

Adjustment for classroom and teacher baseline differences—In Year 1, where there were no significant outcomes after adjustment for regression effects of baseline teacher-rated behavior problems, we did not further investigate baseline group differences as moderators of the outcomes.

In Year 2, we tested for the potential effect of significantly different baseline teacher, classroom, and child characteristics between intervention and control classrooms. Baseline differences in Year 2 included: teacher education, child age, percent of boys in the classroom, and teacher-rated prosocial skills. Because none of these was significantly associated at the bivariate level with any of the Year 2 outcomes, they were not further examined.

The two baseline characteristics in Year 2 that were significantly associated with outcomes were the percent of children receiving mental health consultation services and disruptive

behavior counts. The percent of children receiving mental health consultation services was significantly associated with the teacher interaction skills change score and the ECERS-R interaction scale change score. When our adjusted analyses were performed, in each case the Year 2 intervention effects remained significant.

For disruptive behavior counts in Year 2, we found associations at the bivariate level with both the ECERS-R interaction scale discipline item and teacher interaction skills outcomes. Adjusted analyses showed that controlling for baseline disruptive behavior counts, the change in the ECERS-R interaction scale discipline item remained marginally significant. Thus, the level of disruptive behavior at baseline did partially moderate change in classroom discipline, although there was still a marginally significant improvement in intervention compared to control classrooms. Finally, we examined the moderating effect of disruptive behavior counts on teacher interaction skills. Here, we found that the intervention effect on the teacher interaction skills change score became insignificant when adjusting for baseline disruptive behavior counts. However, when teacher interaction skills outcome scores controlling for *both* baseline teacher interaction skills and baseline disruptive behavior counts were included in the model, the intervention effect on teacher interaction skills remained significant.

Lesson Implementation and Observed Fidelity

In Year 1, teachers completed between 17 and all of the 25 lessons to be taught ($M = 22.06$, $SD = 3.00$) and reported an overall mean of 87% of curriculum points covered (range = 74%–99%). In addition, the mean observed fidelity scores were moderately high: $M = 3.62$ ($SD = .35$; range = 3.11 – 4.25).

In Year 2, all teachers completed at least 24 of the 25 lessons, with the exception of one, who completed only 11 ($M = 22.86$, $SD = 5.24$). This teacher had very young children and a high percentage of children who were diagnosed with special needs. A mean of 86% of curriculum points were covered (range = 49% to 99%). In addition, in Year 2, moderately high observed fidelity scores were also obtained [$M = 3.62$ ($SD = .29$; range = 3.22 – 3.93)].

Finally, during the follow-up year, all but one classroom again completed all 25 lessons ($M = 23.86$, $SD = 3.02$) and 95% of curriculum points were covered. The mean fidelity rating was 4.09 ($SD = .45$, range = 3.12 – 4.37), higher than the two study years.

Relation between Outcomes and Curriculum Fidelity

For these analyses, only intervention classrooms were included. In Year 1, controlling for baseline values of the outcome scores, there were no statistically significant associations between change scores and observer-rated curriculum fidelity ratings (see Table 5). However, while not statistically significant, there were moderately strong associations between fidelity and improvements in two of the ECERS-R interaction scale items: discipline ($r = .61$) and general supervision ($r = .57$) and in teacher interaction skills ($r = .43$).

In Year 2, fidelity ratings were significantly associated with improvements in the disruptiveness rating ($r = -.84$, $p = .04$). In addition, we found a number of other moderately strong, but not statistically significant, associations: teacher-rated behavior problems ($r = -.68$); teacher-rated prosocial skills ($r = .50$); ECERS-R interaction scale and three interaction scale items (range $r = .47$ – $.68$); and disruptive behavior counts ($r = -.68$). Finally, there was a small negative relationship between fidelity and teacher interaction skills. A scatterplot revealed one outlier, and when this was removed the association became moderately positive ($r = .68$).

Teacher Satisfaction with Second Step

In Year 1, teachers ($n = 12$) reported the curriculum was helpful to their children in developing social and emotional skills ($M = 3.25$, $SD = 0.75$, range = 2–4), and was also helpful to them as a teacher ($M = 3.42$, $SD = 0.52$, range = 3–4). In addition, teachers indicated that they believed that Second Step helped to improve the classroom climate ($M = 3.00$, $SD = 0.85$, range = 2–4). However, in response to a question about their willingness to continue to implement Second Step after the study ended, only 25% of teachers in Year 1 reported “Yes, definitely.”

Similar results were found in Year 2; 13 teachers giving a moderately high rating ($M = 3.31$, $SD = 0.63$, range = 2–4) to the helpfulness of the curriculum to their children in developing social and emotional skills, and rating the helpfulness to them as a teacher ($M = 3.31$, $SD = 0.63$, range = 2–4). Year 2 teachers also reported the curriculum helped to improve classroom climate ($M = 3.04$, $SD = 0.72$, range = 2–4). However, a much larger percent of teachers (61.5%) in Year 2 than in Year 1 responded “Yes, definitely,” to the question about their willingness to continue to implement Second Step after the study ended.

Parent Engagement with the Curriculum

Seventy-two Intervention parents returned questionnaires in the spring of Year 1 (72% of the families enrolled at the end of the year). Most parents in Year 1 reported they received the parent letters about Second Step (79%). Most also said they tried the suggested activities with their children at home (88%). However, only 35% reported attending at least one of the Second Step parent sessions.

In Year 2, 39 (45% of enrolled families at the end of the year) returned the spring study questionnaire. Although the sample size was considerably smaller, slightly more respondents (82%) reported receiving the parent letters about Second Step, but fewer (74%) indicated they tried the suggested activities with their children at home. Slightly more, 39%, reported attending at least one of the Second Step parent sessions.

Discussion

We hypothesized that compared to the control classrooms, the intervention classrooms implementing the Second Step social and emotional curriculum activities would show effects on teacher-rated behavior problems and prosocial skills, classroom climate, and teacher interaction skills (hypotheses 1, 2, and 3, respectively). We found moderate to strong evidence to support the second and third hypotheses, but no evidence that children’s teacher-rated behavior problems and prosocial skills improved. A significant difference in Year 1 in teacher-rated behavior problems disappeared after adjusting for baseline scores. These findings are consistent with a prior study of the preschool Second Step curriculum by McMahon et al. (2000). We are aware that we may have had inadequate power to overcome baseline differences in study classrooms. However, it is also possible that the universal, classroom-based focus is not sufficiently targeted to address individual children’s behavior problems. In addition, intervention teachers may have had different expectations of children’s behavior after participating in the curriculum, and therefore did not rate their children the same as did control teachers. These issues were also raised by MacMahon and colleagues’ (2000) evaluation of Second Step conducted in high risk preschool and kindergarten classrooms; we believe that the issues can only be resolved by using multiple assessment methods beyond teacher ratings.

Our findings for the second and third hypotheses, however, are stronger. After adjusting for threats to internal validity (e.g., regression and baseline differences in the classrooms), we found substantial differences in one of the classroom climate measures (hypothesis 2) and in

teacher interaction skills (hypothesis 3), but only in the second year of implementation. The classroom climate measure that showed significant differences in Year 2 for intervention classrooms was the ECERS-R interaction scale, including significant differences in the discipline item. This measure taps how teachers provide general supervision and guidance to the classroom, including handling situations when children's behavior is difficult. Even after adjusting for classroom differences in referral of individual children for mental health consultation, the intervention effect on the ECERS-R interaction scale change score, and each item, remained significant. This suggests an independent effect of the curriculum intervention on classrooms that already had high rates of utilizing mental health consultation services for individual children.

The pattern of additive effects of the curriculum to settings that also have individualized interventions affirms recommended best practices for classroom management of behavior problems by implementing universal classroom programs with additional target services to children with more intense needs. (Baker et al., 2006; Bayat et al., 2010; Fox, Dunlap, Hemmeler, Joseph & Strain, 2003; Technical Assistance Center for Social Emotional Intervention for Young Children (TACSE), http://www.challengingbehavior.org/do/pyramid_model.htm). Fox et al., (2003) and TACSE, for example, describe this as the 'teaching pyramid' for young children. The lower base of the pyramid provides classroom prevention policies and social and emotional teaching strategies to create an environment that reduces most children's challenging behaviors. Then a smaller number of children, representing the next higher level of the pyramid, should be provided individualized services, and an even small number at the peak require intensive services, or services out of the mainstream classroom. An important characteristic of the TACSE model is to incorporate formal teaching about identifying and discriminating different emotions and developing the vocabulary to describe these emotions, and then to directly teach problem-solving skills around strong emotions, all aspects of the Second Step curriculum.

On the other hand, adjusting for baseline differences in disruptive behavior counts did moderate and reduce the statistical significance of the discipline item. This suggests that high rates of disruptive classroom behavior make it more difficult for teachers to implement ideal discipline strategies. Fox et al. (2003) have suggested that some teachers give up or even quit teaching due to being faced with classrooms with high disruptive behavior. Nevertheless, after the adjustment, there was still marginally significant improvement in the way discipline was administered in the intervention relative to the control classrooms, which indicates that even in more difficult classrooms there may be a benefit for teachers in learning and delivering the curriculum.

The findings regarding teacher interaction skills (hypothesis 3) were somewhat more complicated, likely because of sample size and colinearity problems. The significant difference between intervention and control teacher interaction skills change scores in Year 2 remained significant when adjusting for regression effects, when simultaneously adjusting for both baseline disruptive behavior counts and teacher interaction skills, but not when adjusting for baseline disruptive behavior counts alone. This suggests an interaction between teacher baseline skills and level of classroom disruptive behavior that we were not able to tease out. A similar complexity has been found in other classroom studies designed to improve young children's behavior (Raver et al. 2009), and points to the need to pay attention to both baseline teacher skills and classroom characteristics in attempting to improve teacher outcomes.

In sum, we believe the pattern of findings provides moderately strong evidence for intervention effects on classroom climate, and likely effects on teacher interaction skills in Year 2. Effect sizes, even those not statistically significant, suggest consistent associations

favoring intervention classrooms on several dimensions of classroom climate as well as teacher interaction skills. The fact that the effects are all found with directly observed measures, and not just teachers' reports, also strengthened our confidence in the study's findings. Further, support for curriculum effects comes from the strong associations between curriculum fidelity and improvements in teacher-rated behavior problems, teacher-rated prosocial skills, all measures of classroom climate, and teacher interaction skills (once one outlier was removed).

It should be noted that the differences between intervention and control teachers and classrooms in Year 2 were largely due to stability or small improvements from the beginning to the end of the year for the intervention condition. On the other hand, control teachers' interaction skills and classroom climate (as measured by the ECERS-R interaction scale and items) declined substantially, even though they had higher levels of education than intervention teachers. The finding of a decline in classroom climate is similar to results reported by Raver et al. (2008) from the Chicago School Readiness Project. These investigators provided behavioral training and a mental health consultant in a cluster-randomized study of Head Start classrooms. They found intervention and control differences were due to a fall 'honeymoon,' followed by a decrease in the positive climate, teacher sensitivity, and classroom management skills of control compared to intervention classrooms over the school year. Raver et al. (2008) hypothesized that as the challenges of meeting the needs of diverse, high risk children continued over the year, their intervention provided teachers better skills and tools to manage children's behavior and to sustain a positive classroom climate and learning environment. Interventions that provide ongoing assistance to teachers throughout the school year, such as a year-long curricular intervention, thus may have promise to sustain quality in classrooms for high risk children in the same way as found in Raver and colleagues' study.

It is puzzling, however, that the same pattern of control classroom decline was not found in Year 1. We suggest that this is due to considerable excitement and activity related to implementing the study in Year 1, including obtaining parental consent and conducting classroom observations for all classrooms, that created a 'placebo' or 'Hawthorne' effect for control teachers. While the same level of effort was shown by research staff in Year 2, the control centers may have acclimated to the study activities and returned to more typical behavior.

In addition, unlike Raver et al. (2008), gains in teacher interaction skills and classroom climate were evidenced only in the intervention classrooms, which provided *both* Second Step and mental health consultation. Declines were found in the control classrooms that received mental health consultation alone. It is possible that the control classrooms did not receive the same quality or intensity of mental health consultation as in the study by Raver et al. (2008). Another interpretation is that the provision of mental health consultation alone does not provide the requisite foundation suggested by Fox et al. (2003) and TACSE to promote overall positive classroom behavior.

It is also interesting that we found stronger patterns of outcomes in Year 2 than in Year 1. This suggests that it takes time for teachers to learn the curriculum well, and that classroom climate and teacher interaction skills may increase over time because of the increased use and exposure of the teachers to the curricular concepts. In addition, by providing some curricular support and fidelity monitoring for two years, continuing teachers became more 'expert,' and new teachers who joined classrooms where the curriculum had already started appeared to quickly reach the level of competence of continuing teachers. This is particularly important in settings such as the current study, where teacher turnover is high.

Further, our findings on teacher satisfaction with the curriculum were positive and grew over time. In both years, teachers rated the curriculum as helpful to children in developing social and emotional skills, in improving their own ability, and in improving classroom climate. After Year 2, more than double the percent of teachers reported they would ‘definitely’ implement the curriculum in subsequent years than did teachers after Year 1. The positive regard for the curriculum was also reflected in strong rates of implementation and fidelity evidenced in a follow-up year after study support was discontinued. Thoughtful attention to implementation issues is important to making such interventions efficacious and sustainable in community settings in the ‘real world’ (Wenz-Gross & Upshur, 2012). Because we found strong post-study sustainability of the intervention, our approach provides a good model for potential long-term improvements in classrooms to support social and emotional development.

Despite the promising findings, the current study has several limitations. First, sample sizes were small (i.e., numbers of classrooms), resulting in lack of statistical power. Second, there were some significant teacher and child differences at baseline that were difficult to analytically control in a sample that was also relatively small. Finally, due to limited resources, we were not able to accurately document the social and emotional curricula administered in control classrooms, and we also were not able to administer individual assessments of children’s knowledge of emotions or social problem-solving skills.

More evidence is thus needed to document the effectiveness and feasibility of the Second Step preschool curriculum in a larger number of classrooms. Future studies should carefully document control classroom curricula, add direct individual child assessments, and explore mediating and moderating pathways. For instance, is child change primarily mediated by changes in teacher discipline and interactions, or by learning of appropriate social skills, or both? In addition, given the intersection of emotional and attentional control, the latter of which is highly associated with success in elementary school (Grimm et al., 2010), additional measures of children’s self-regulation and attention are needed.

Despite the challenges, there is an urgent need to continue to explore interventions to promote positive social and emotional development of young children. Intervening in the preschool years is a developmentally optimal time, and can prevent adverse long-term consequences for children, families, schools, and the mental health, child welfare, and juvenile justice systems.

Acknowledgments

This study was funded by an Intervention Development Grant, R34 MH073830 from the National Institute of Mental Health. We would like to thank our colleague who oversaw site operations: Lynn Hennigan, LICSW, Director of Services for the Young Child, Community Healthlink, Youth and Family Services; and staff who helped with study management and implementation, classroom observations, and data entry: Gail Sawosik, Judy Davis, Katelyn Allen, Katie Bateman, Courtney Carpenter, Rachel D’Angelo, Alicia Derry, Rachel Friendly, Jenny Hazelton, Karen Nideur, and Laura Seekamp. In addition, we owe special thanks to the staff and families of Worcester Comprehensive Child Care Services, Great Brook Valley; the YWCA of Central Massachusetts, Salem Square; Rainbow Child Development Center; and the YMCA Child Care Center, Worcester Commons.

References

- Alkon A, Ramler M, MacLennan K. Evaluation of mental health consultation in child care centers. *Early Childhood Education Journal*. 2003; 31:91–99.
- Arnett J. Caregiving in day-care centers. Does training matter? *Journal of Applied Developmental Psychology*. 1989; 10:541–552.

- Baker JA, Kamphaus RW, Horne AM, Winsor AP. Evidence for population-based perspectives on children's behavioral adjustment and needs for service delivery in schools. *School Psychology Review*. 2006; 35:31–46.
- Bayat M, Mindes G, Covitt S. What does RTI (Response to Intervention) look like in preschool? *Early Childhood Education Journal*. 2010; 37:493–500.
- Committee for Children. Second Step: Preschool Kindergarten Social/Emotional Learning Kit; Second Step PreK/Teacher Guide; Second Step Parent Guide. Seattle, WA: Committee for Children; 2002.
- Denham SA, Blair K, Schmidt M, DeMulder E. Compromised emotional competence: Seeds of violence sown early? *American Journal of Orthopsychiatry*. 2002; 72:70–82. [PubMed: 14964596]
- Denham SA, Caverly S, Schmidt M, Blair K, DeMulder E, Caal S, et al. Preschool understanding of emotions: Contributions to classroom anger and aggression. *Journal of Child Psychology and Psychiatry*. 2002; 43:901–916. [PubMed: 12405478]
- Dodge, DT.; Colker, LJ.; Heroman, C. *The Creative Curriculum for Preschool*. Washington, DC: Teaching Strategies; 2002.
- Domitrovich CE, Cortes RC, Greenberg MT. Improving young children's social and emotional competence: A randomized trial of the preschool "PATHS" curriculum. *The Journal of Primary Prevention*. 2007; 28:67–91. [PubMed: 17265130]
- Eyberg, S.; Pincus, D. *Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory-Revised: Professional Manual*. Lutz, FL: Psychological Assessments Resources, Inc; 1999.
- Fox L, Dunlap G, Hemmeter JL, Joseph GE, Strain PS. The teaching pyramid: A model for supporting social competence and preventing challenging behavior in young children. *Young Children*. 2003; 58:48–52.
- Gilliam, WS. *Prekindergartners left behind: Expulsion rates in state preK systems*. New Haven, CT: Yale University School of Medicine: Child Study Center; 2005.
- Gilliam, WS. *Early childhood consultation partnership: Results of a random-controlled evaluation. Final Report and Executive Summary*. New Haven, CT: Yale University School of Medicine: Child Study Center; 2007.
- Grimm KJ, Steele JS, Mashburn AJ, Burchinal M, Pianta RC. Early behavioral associations of achievement trajectories. *Developmental Psychology*. 2010; 46:976–983. [PubMed: 20822216]
- Harms, T.; Clifford, RM.; Cryer, D. *Early Childhood Environment Rating Scale-Revised*. New York: Teachers College Press; 1998.
- Hogan AE, Scott KG, Bauer CR. The adaptive social behavior inventory (ASBI): A new assessment of social competence in high-risk three-year-olds. *Journal of Psychoeducational Assessment*. 1992; 10:230–239.
- Howes C, Phillips DA, Whitebrook M. Thresholds of quality: Implications for the social development of children in center-based child care. *Child Development*. 1992; 63:449–460. [PubMed: 1611946]
- Izard C, Fine S, Schultz D, Mostow A, Ackerman B, Youngstrom E. Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychological Science*. 2001; 12:18–23. [PubMed: 11294223]
- Izard CE, King KA, Trentacosta CJ, Morgan JR, Laurenceau J, Kranthamer-Ewing ES, Finlon KJ. Accelerating the development of emotional competence in Head Start children: Effects on adaptive and maladaptive behavior. *Development and Psychopathology*. 2008; 20:369–397. [PubMed: 18211742]
- Maslach, C.; Jackson, SE.; Leiter, MP. *Maslach Burnout Inventory Manual*. 3rd Edition. Palo Alto, CA: Consulting Psychology Press; 1996.
- 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.
- Perry DF, Dunne MC, McFadden L, Campbell D. Reducing the risk for preschool expulsion: Mental health consultation for young children with challenging behaviors. *Journal of Child and Family Studies*. 2008; 17:44–54.
- Pierce EW, Ewing LJ, Campbell SB. Diagnostic status and symptomatic behavior of hard-to-manage preschool children in middle childhood and early adolescence. *Journal of Clinical Child Psychology*. 1999; 28:44–57. [PubMed: 10070606]

- 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 CC, Jones SM, Li-Grining CP, Metzger M, Champion KM, Sardin L. Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly*. 2008; 23:10–26.
- Raver CC, Jones SM, Li-Grining C, Zhai F, Metzger MW, Soloman B. Targeting children's behavior problems in preschool classrooms: A cluster-randomized controlled trial. *Journal of Counseling and Clinical Psychology*. 2009; 77:302–316.
- Romano E, Babchishin LY, Pagani LS, Kohen D. School readiness and later achievement: Replication and extension using a nationwide Canadian survey. *Developmental Psychology*. 2010; 46:995–1007. [PubMed: 20822218]
- Schultz D, Izard CE, Ackerman BP, Youngstrom EA. Emotion knowledge in economically disadvantaged children: Self-regulatory antecedents and relations to social difficulties and withdrawal. *Development and Psychopathology*. 2001; 13:53–67. [PubMed: 11346052]
- Stata Corp.. College Station, TX: 2011. www.stata.com
- Technical Assistance Center on Social Emotional Intervention for Young Children (TACSE). Learn about the pyramid model. http://www.challengingbehavior.org/do/pyramid_model.htm
Downloaded 2/1/10
- Trentacosta CJ, Fine SE. Emotion knowledge, social competence, and behavior problems in childhood and adolescence: A meta-analytic review. *Social Development*. 2010; 19:1–29. [PubMed: 21072259]
- 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.
- Wenz-Gross M, Upshur C. Implementing a primary prevention social skills intervention in urban preschools: Factors associated with quality and fidelity. *Early Education & Development*. 2012; 23:427–450. [PubMed: 22791957]
- Webster-Stratton C, Reid MJ, Stoolmiller M. Preventing conduct problems and improving school readiness: Evaluation of the Incredible Years teacher and child training programs in high-risk schools. *Child Psychology and Psychiatry*. 2008; 49:471–488.
- Williford AP, Shelton TL. Using mental health consultation to decrease disruptive behaviors in preschool children: Adapting an empirically-supported intervention. *Journal of Child Psychology and Psychiatry*. 2008; 49:191–200. [PubMed: 18211278]

Figure 1.
Study Enrollment

Table 1

Baseline Child Characteristics in Years 1 and 2

Child characteristics	Year 1		Year 2	
	Intervention	Control	Intervention	Control
Child age	(<i>n</i> = 113) [*]	(<i>n</i> = 96)	(<i>n</i> = 95) ^{**}	(<i>n</i> = 62)
Mean	46.78	50.72	44.65	48.34
<i>SD</i>	8.26	8.81	7.86	8.27
Child gender	(<i>n</i> = 132) [*]	(<i>n</i> = 100)	(<i>n</i> = 111) ⁺	(<i>n</i> = 65)
% Male	59.1	46.0	64.9	49.2
Child race/ethnicity	(<i>n</i> = 130)	(<i>n</i> = 96)	(<i>n</i> = 104)	(<i>n</i> = 59)
% Hispanic	51.5	42.7	50.0	61.0
% African American	25.4	24.0	24.0	22.0
% White	12.3	18.8	15.4	8.5
% Other/bi-racial	10.8	14.5	10.6	8.5
Annual family income	(<i>n</i> = 125)	(<i>n</i> = 96)	(<i>n</i> = 105)	(<i>n</i> = 57)
% < \$20,000	45.6	43.8	46.7	56.1
Mental health consultation services	(<i>n</i> = 133)	(<i>n</i> = 100)	(<i>n</i> = 111) ^{***}	(<i>n</i> = 66)
% Receiving individual services	31.6	22.0	36.0	12.10
Teacher-rated behavior problems ^a	(<i>n</i> = 96)	(<i>n</i> = 69)	(<i>n</i> = 88)	(<i>n</i> = 60)
Mean	118.85	112.33	107.81	97.60
<i>SD</i>	57.81	45.23	49.76	37.25
Teacher-rated prosocial skills ^b	(<i>n</i> = 94) [*]	(<i>n</i> = 71)	(<i>n</i> = 88) ^{**}	(<i>n</i> = 60)
Mean	55.73	52.51	58.26	53.77
<i>SD</i>	9.97	10.50	8.27	9.07

Note. *SD* = Standard Deviation.

^aNorms for ages 2–6 are *M*=99.2 (*SD*=33.8)

^bNorms for 3 year olds are *M*=55.9 (*SD*=5.97).

⁺*p* < .10,

^{*}*p* < .05,

^{**}*p* < .01,

^{***}*p* < .001.

Table 2

Baseline Teacher Characteristics in Years 1 and 2

Teacher characteristics	Year 1		Year 2	
	Intervention	Control	Intervention	Control
Gender	(n=15)	(n=15)	(n=17)	(n=9)
% Female	93.3	100.0	94.1	100.0
Race/ethnicity	(n=15)	(n=15)	(n=17)	(n=9)*
% White	66.7	73.3	52.9	100.0
% Hispanic	20.0	13.3	23.5	0.0
% African American	6.7	13.3	11.8	0.0
% Other/bi-racial	6.7	0.0	11.8	0.0
Age	(n=15)	(n=15)	(n=17)	(n=9)
Mean	36.93	42.80	35.24	34.11
<i>SD</i>	9.63	11.80	10.90	13.64
Years of teaching experience	(n=15)	(n=15)	(n=16)	(n=9)
Mean	10.20	13.53	8.19	9.44
<i>SD</i>	9.08	6.81	9.17	7.18
Teacher education	(n=15)	(n=15)*	(n=17)	(n=9)*
% No diploma or GED	0.0	6.7	0.0	0.0
% High school grad	26.7	0.0	41.2	0.0
% Some college	26.7	20.0	23.5	33.3
% Associates degree	13.3	46.7	5.9	55.6
% Bachelors degree	33.3	13.3	29.4	11.1
% Masters degree	0.0	13.3	0.0	0.0
Teacher interaction skills ^a	(n=11)	(n=10)*	(n=13)	(n=8)
Mean	73.54	64.12	73.44	68.19
<i>SD</i>	7.01	11.77	15.21	12.14
Teacher burnout	(n=12)	(n=15)	(n=10)	(n=6)
Emotional exhaustion ^b				
Mean	28.33	22.47	29.90	35.33
<i>SD</i>	12.69	13.22	9.28	11.93
Personal accomplishment ^c				
Mean	40.75	40.80	41.00	38.83
<i>SD</i>	4.96	5.76	5.21	5.85

Note. *SD* = Standard Deviation.

^aScores over 70 represent strongly positive interaction skills;

^bScores over 26 reflect high emotional exhaustion;

^cScores over 36 reflect high personal accomplishment.

* $p < .05$.

Table 3**Baseline Classroom Characteristics in Years 1 and 2**

Classroom characteristic	Year 1		Year 2	
	Intervention (n=8)	Control (n=7)	Intervention (n=7)	Control (n=4)
Class size				
Mean	16.00	17.14	16.86	15.00
<i>SD</i>	3.42	2.67	2.61	0.0
% of children meeting at-risk cut-off for teacher-rated behavior problems				
Mean	26.61	22.34	22.31	6.48
<i>SD</i>	20.84	9.56	20.44	9.44
Classroom quality: (ECERS-R space, activities, and structure)				
Mean	5.21	5.01	Not	Not
<i>SD</i>	0.89	0.56	assessed	assessed
Classroom climate: ECERS-R interaction scale				
Mean	22.00	19.62	22.11	18.69
<i>SD</i>	3.02	4.12	6.09	4.97
Disruptive behavior counts				
Mean	1.95	1.59	5.19**	2.38
<i>SD</i>	0.51	0.72	1.43	0.62
Disruptiveness rating				
Mean	0.81	0.69	1.68	0.50
<i>SD</i>	0.24	0.39	1.14	0.47

Note. *SD* = Standard Deviation.

**
p<.01.

Table 4

Classroom Change Scores in Years 1 and 2

Outcome measure	Year 1										Cohen's <i>d</i>
	Intervention				Control		Baseline pooled		<i>t</i>		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>SD</i>				
Teacher-rated behavior problems	72	-7.86	44.88	53	7.72	31.71	53.50	$t(123) = -2.27^*$	<i>ns</i>	-.29	
Teacher-rated prosocial skills	79	4.77	7.29	58	3.10	7.15	9.67			.17	
Classroom climate:											
ECERS-R interaction scale	8	-0.83	4.58	7	-2.08	5.00	3.57		<i>ns</i>	.35	
Discipline	8	0.15	1.33	7	-0.58	1.27	.88		<i>ns</i>	.83	
General supervision	8	-0.32	1.24	7	0.01	1.84	1.06		<i>ns</i>	.32	
Staff-child interactions	8	-0.80	1.16	7	-1.04	1.20	1.33		<i>ns</i>	.18	
Child-child interactions	8	0.15	1.45	7	-0.48	1.15	1.01		<i>ns</i>	.62	
Disruptive behavior counts	8	1.77	1.01	7	1.70	0.90	.96		<i>ns</i>	.07	
Disruptiveness rating	8	0.73	.35	7	0.70	0.77	.32		<i>ns</i>	.09	
Teacher interaction skills	8	-2.29	8.50	7	.95	9.63	11.26		<i>ns</i>	-.29	
Year 2											
Teacher-rated behavior problems	64	-7.02	44.81	53	1.42	33.45	45.46		<i>ns</i>	-.19	
Teacher-rated prosocial skills	64	1.64	7.47	53	1.47	9.32	8.65		<i>ns</i>	.02	
Classroom climate:											
ECERS-R interaction scale	7	1.72	7.28	4	-8.69	6.61	5.74		$t(9) = 2.35^*$	1.81	
Discipline	7	.77	1.62	4	-2.75	1.40	1.44		$t(9) = 3.63^{**}$	2.44	
General supervision	7	.27	2.23	4	-2.38	1.16	1.49		$t(9) = 2.18^+$	1.78	
Staff-child interactions	7	.25	2.29	4	-2.25	1.90	1.68		$t(9) = 1.84^+$	1.49	
Child-child interactions	7	.43	1.30	4	-1.44	2.48	1.44		<i>ns</i>	1.30	
Disruptive behavior counts	7	-1.25	1.90	4	.25	0.76	1.23		$t(8.49) = -1.85^+$	-1.22	
Disruptiveness rating	7	-.41	1.19	4	.12	0.27	.97		<i>ns</i>	-.54	
Teacher interaction skills	7	0.08	11.05	4	-17.63	10.79	10.18		$t(9) = 2.58^*$	1.74	

Note. *SD* = Standard Deviation. *ns* = not significant.

^aECERS-R = Early Childhood Environmental Rating Scale-Revised

^c $p < .10$,

* $p < .05$,

** $p < .01$.

Table 5

Partial Correlations Between Curriculum Fidelity Ratings and Change Scores Intervention Classrooms Only

Outcome change score	<u>Mean observer fidelity rating</u>	
	Year 1 (n=8)	Year 2 (n=7)
Teacher-rated behavior problems	-.05 (df= 5)	-.68 (df= 4)
Teacher-rated prosocial skills	-.26 (df= 5)	.50 (df= 4)
Classroom climate:		
ECERS-R ^a interaction scale	.40 (df= 5)	.52 (df= 4)
Discipline	.61 (df= 5)	.53 (df= 4)
General supervision	.57 (df= 5)	.47 (df= 4)
Staff-child interactions	-.09 (df= 5)	.23 (df= 4)
Child-child interactions	.30 (df= 5)	.68 (df= 4)
Disruptive behavior counts	-.09 (df= 5)	-.68 (df= 4)
Disruptiveness rating	.01 (df= 5)	-.84 (df= 4) *
Teacher interaction skills	.43 (df= 4)	-.11 (df= 4)

Note. SD = Standard Deviation. Partial correlations were run controlling for baseline values of the outcome scores.

^aECERS-R = Early Childhood Environmental Rating Scale-Revised

* $p < .05$.