THE JOURNAL OF
CHILD PSYCHOLOGY AND PSYCHIATRY

Journal of Child Psychology and Psychiatry 49:5 (2008), pp 471–488 doi:10.1111/j.1469-7610.2007.01861.x

Preventing conduct problems and improving school readiness: evaluation of the Incredible Years Teacher and Child Training Programs in high-risk schools

Carolyn Webster-Stratton,[1] M. Jamila Reid,[1] and Mike Stoolmiller[2]

[1]University of Washington, School of Nursing, Parenting Clinic, Seattle, USA; [2]Research and Statistical Consulting, Marquette, MI, USA

Background: School readiness, conceptualized as three components including emotional self-regulation, social competence, and family/school involvement, as well as absence of conduct problems play a key role in young children's future interpersonal adjustment and academic success. Unfortunately, exposure to multiple poverty-related risks increases the odds that children will demonstrate increased emotional dysregulation, fewer social skills, less teacher/parent involvement and more conduct problems. Consequently intervention offered to socio-economically disadvantaged populations that includes a social and emotional school curriculum and trains teachers in effective classroom management skills and in promotion of parent–school involvement would seem to be a strategic strategy for improving young children's school readiness, leading to later academic success and prevention of the development of conduct disorders. **Methods:** This randomized trial evaluated the Incredible Years (IY) Teacher Classroom Management and Child Social and Emotion curriculum (Dinosaur School) as a universal prevention program for children enrolled in Head Start, kindergarten, or first grade classrooms in schools selected because of high rates of poverty. Trained teachers offered the Dinosaur School curriculum to all their students in bi-weekly lessons throughout the year. They sent home weekly dinosaur homework to encourage parents' involvement. Part of the curriculum involved promotion of lesson objectives through the teachers' continual use of positive classroom management skills focused on building social competence and emotional self-regulation skills as well as decreasing conduct problems. Matched pairs of schools were randomly assigned to intervention or control conditions. **Results:** Results from multi-level models on a total of 153 teachers and 1,768 students are presented. Children and teachers were observed in the classrooms by blinded observers at the beginning and the end of the school year. Results indicated that intervention teachers used more positive classroom management strategies and their students showed more social competence and emotional self-regulation and fewer conduct problems than control teachers and students. Intervention teachers reported more involvement with parents than control teachers. Satisfaction with the program was very high regardless of grade levels. **Conclusions:** These findings provide support for the efficacy of this universal preventive curriculum for enhancing school protective factors and reducing child and classroom risk factors faced by socio-economically disadvantaged children. **Keywords:** Aggression, behavior problem, prevention, school, teacher, school readiness.

While researchers have long considered intelligence to be a key predictor of school performance, evidence suggests that school readiness (defined here as emotional self-regulatory ability, social competence, the absence of behavior problems, and parent–teacher involvement) are independent and important predictors of future academic achievement even after controlling for variations in cognitive abilities and family resources (Grolnick & Slowiaczek, 1994; Keogh, 1992; Ladd, Birch, & Buhs, 1999; Normandeau & Guay, 1998; Raver & Zigler, 1997). Children with emotional and social problems and 'early onset' conduct problems (defined generically as high rates of aggression, noncompliance, oppositional behaviors) are at high risk for academic failure, school absences, and eventual conduct disorders, school dropout and delinquency (Kellam et al., 1991; Moffitt, 1993; Patterson, Reid, & Dishion, 1992; Tremblay et al., 1996).

Unfortunately, a survey conducted by the National Center for Early Development and Learning indicated that 46% of kindergarten teachers reported that more than half of the children in their classes were not ready for school, that is, they lacked the self-regulatory skills and emotional and social competence to function productively and to learn in kindergarten (West, Denton, & Reaney, 2001).

Conflict of interest statement: The first author of this paper has disclosed a potential financial conflict of interest because she disseminates the Incredible Years interventions and stands to gain from a favorable report. Because of this, she has voluntarily agreed to distance herself from certain critical research activities (i.e., recruiting, consenting, primary data handling, and analysis) and the University of Washington has approved these arrangements.

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

Similarly, several Head Start studies suggested that between 16 and 30% of preschool children in those classes posed ongoing conduct problems for teachers (Kupersmidt, Bryant, & Willoughby, 2000; Lopez, Tarullo, Forness, & Boyce, 2000).

Need for early prevention programs in schools with high-risk populations

The Early Child Longitudinal Survey (ECLS), a nationally representative sample of over 22,000 kindergarten children, suggests that exposure to multiple poverty-related risks increases the odds that children will demonstrate less social competence and emotional self-regulation and more behavior problems than more economically advantaged children (West et al., 2001). While socioeconomic disadvantage does not necessarily lead to social and emotional problems, up to 25% of children living in poverty experience negative social and emotional outcomes (Keenan, Shaw, Walsh, Delliquadri, & Giovannelli, 1997). Low income is also a significant risk factor for the early onset of conduct problems and academic underachievement (Offord, Alder, & Boyle, 1986). Moreover, longitudinal data suggest that these early gaps in social competence for socio-economically disadvantaged children persist and even widen as children progress in school (Huffman, Mehlinger, & Kerivan, 2001).

In addition to poverty-related risks for children's social, emotional and conduct problems, research also shows that teachers with poor classroom management skills have higher overall levels of classroom aggression, peer rejection and exclusion which, in turn, compound the development of individual children's social and conduct problems (Kellam, Ling, Merisca, Brown, & Ialongo, 1998). Moreover, it seems that children who are at highest risk are often taught by teachers who are the least prepared to handle challenging behavior; teachers serving predominantly low-income children use more harsh, detached, and ineffective teaching strategies than those teaching middle-income children (Phillips, Voran, Kisker, Howes, & Whitebrook, 1994; Stage & Quiroz, 1997). Children with conduct problems are also more likely to be disliked by teachers and receive less academic or social instruction, support, and positive feedback from teachers for appropriate behavior (Arnold, Griffith, Ortiz, & Stowe, 1998; Arnold et al., 1999; Campbell & Ewing, 1990; Carr, Taylor, & Robinson, 1991). Consequently, children with conduct problems grow to like school less, have lower school attendance (Birch & Ladd, 1997) and increased risk for underachievement, academic failure, and future adjustment problems.

Conversely, there is substantial evidence indicating that well-trained and supportive teachers, who use high levels of praise, proactive teaching strategies, and non-harsh discipline, can play an extremely important role in fostering the development of social and emotional skills and preventing the development of conduct problems in young children. In fact, longitudinal research demonstrates that low-income children in high quality preschool settings are significantly better off, cognitively, socially, and emotionally, than similar children in low quality settings (Burchinal, Roberts, Hooper, & Zeisel, 2000). Having a supportive relationship with at least one teacher has been shown to be one of the most important protective factors influencing high-risk children's later school success (Pianta & Walsh, 1998; Werner, 1999).

Moreover, there is increasing evidence to suggest that teachers' efforts to involve parents in ways to support their children's learning at home (through newsletters, suggested homework activities, teacher phone calls) and in developing coordinated home/school behavior plans have positive effects on children's academic, social and emotional competence (Henderson & Berla, 1994). However, while most teachers want to be active partners with parents, most have had little training in ways to work collaboratively with families (Burton, 1992; Epstein, 1992).

Thus it is hypothesized that supporting teachers' capacity to manage a classroom with positive behavior management strategies, to deliver a curriculum designed to promote social competence and emotional regulation, and to encourage teacher–parent involvement will lead to fewer conduct problems, increased school readiness and eventual academic success. In particular, supporting teachers who work with young children exposed to poverty-related risks may also help to reduce the academic, social and emotional gap that exists at the starting gate between higher-risk children and their more advantaged peers (Ladd et al., 1999; Skinner, Zimmer-Gembeck, & Connel, 1998).

Emerging research on the effectiveness of early school-based interventions

Mounting evidence from several multifaceted, longitudinal, school-based prevention programs have indicated the promise of prevention programs for reducing risk factors related to academic failure and conduct disorders in adolescence. For example, two large-scale indicated prevention projects, Tremblay et al. (Tremblay, Pagani, Masse, & Vitaro, 1995; Tremblay et al., 1996) and FAST TRACK (Conduct Problems Prevention Research Group, 2002) both selected highly aggressive elementary grade school children and offered comprehensive intervention starting in first grade including social skills training, academic tutoring, teacher classroom management and parent training. Both projects found long-term benefits in school performance and reductions in antisocial behavior such as burglary and theft. Several other prevention trials, LIFT (Reid, Eddy, Fetrow, & Stoolmiller, 1999), Seattle Social

Development Project (Hawkins, Catalano, Kosterman, Abbott, & Hill, 1999) and the Baltimore study (Kellam & Rebok, 1992) addressed similar parent and school risk factors and used universal designs targeting school-age children (grades 1–5) in high-risk neighborhoods or schools. All three projects found benefits for children who had received the intervention, including fewer violent delinquent acts and lower rates of drinking, sexual activity, and pregnancy by 18 years (Hawkins et al., 1999) and lower levels of classroom and playground aggression (Kellam et al., 1998; Reid et al., 1999).

Preschool–kindergarten classroom prevention programs designed specifically to improve young children's social-emotional competence (see review, Webster-Stratton & Taylor, 2001) have also shown promise for improving classroom behavior. For example, programs such as *I can Problem Solve* (Shure, 1997), Second Step (Grossman et al., 1997) and the First Step curriculum (Walker et al., 1998a) combined training for children's emotional and social skills with cognitive strategies to promote school readiness. Results showed improvements in children's school readiness and less aggressive behavior. However, these studies are limited by lack of control groups, small sample sizes, lack of cultural diversity, lack of observational data in the classroom and reliance on teacher reports. Moreover, studies with economically disadvantaged children ages 3 to 6 years are relatively scarce. One exception is an evaluation of the PATHS program (Promoting Alternative Thinking Strategies; Greenberg, Kusche, Cook, & Quamma, 1995), which was originally evaluated in the Fast Track program for older children and was adapted for use in preschool with socio-economically disadvantaged populations (Domitrovich, Cortes, & Greenberg, 2006). This intervention was delivered by teachers in 20 Head Start classrooms. Results showed that intervention children had higher emotion knowledge skills and were rated by teachers and peers as more socially competent compared to peers. The intervention did not produce changes in reports of children's problem-solving abilities or levels of aggressive behavior, and no independent observational data were collected.

The Incredible Years interventions

The Incredible Years (IY) Child Training curriculum (Dinosaur School) was originally developed to treat clinic-referred children (ages 3–7 years) diagnosed with oppositional defiant disorder or early-onset conduct problems. In two randomized trials with clinic populations, results showed improvements in children's conduct problems both at home and at school based on independent observations as well as parent and teacher reports (Webster-Stratton & Hammond, 1997; Webster-Stratton & Reid, 2003; Webster-Stratton, Reid, & Hammond, 2001, 2004). This clinic-based treatment model was revised and adapted so that it could be used by teachers as a preschool and early school-based preventive model. One independent study using an adaptation of this classroom-based curriculum for high-risk students indicated its potential benefits for reducing aggression on the playground (Barrera et al., 2002). The content of the Dinosaur School curriculum is based on cognitive social learning theory (Bandura, 1989; Patterson et al., 1992) and research indicating the kinds of social, emotional and cognitive deficits found in children with conduct problems (e.g., Dodge & Price, 1994; Ladd, Kochenderfer, & Coleman, 1997) as well as on social learning behavior change methods such as videotape modeling, role play and practice of targeted skills, and reinforcement for targeted behaviors. For more details see Webster-Stratton and Reid (2004).

Training for teachers not only involved how to deliver the Dinosaur School curriculum, but also training in utilizing effective classroom management strategies. This included strategies to promote prosocial behaviors and emotional literacy, to prevent or reduce the development of conduct problems and ways to increase parents' involvement in their children's education and behavior planning.

We have grouped risk and protective factors into four categories: (a) teacher classroom management skills and classroom environment; (b) teacher–parent involvement; (c) child school readiness (social competence, emotional self-regulation, and absence of behavior problems); and (d) poverty. If children with social/emotional/behavioral problems are not supported, risk factors can intertwine and cascade to increasingly negative outcomes (Figure 1). The IY child and teacher intervention is designed to target the first three of these more malleable risk factors and it is hypothesized that the increase of protective factors will prevent problematic behavior patterns. The fourth area of risk that we have identified (poverty) is not one that can be easily changed by schools. However, the fact that children living in poverty are at higher-risk points to the need to focus more intervention services in high-need, low-income schools. Thus, this study placed our intervention in schools with high percentages of students who are living in poverty.

We hypothesized that training teachers to deliver the IY Dinosaur School Curriculum utilizing positive classroom behavior management strategies would result in more positive and responsive teaching and less harsh or critical discipline, increased focus on social and emotional teaching, and more focus on parent involvement in children's education than in control classrooms. We hypothesized that children of teachers who received this training would show more school readiness and fewer conduct problems than children in control classrooms. Building on Zigler's (Raver & Zigler, 1997) conceptualization of school readiness, this includes the following domains:

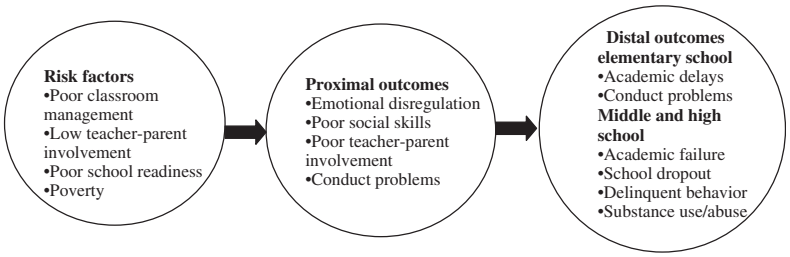

Figure 1 Cascading risk factors

(a) Emotional self-regulation (e.g., attention and persistence, engagement/on-task work, prosocial problem solving, feelings vocabulary, ability to manage anger); (b) Social competence (e.g., sharing, helping, friendship skills, positive peer interactions); (c) Teacher–parent involvement (calls, newsletters, homework activities); and (d) Absence of conduct problems (e.g., compliance to rules and teacher directions, low verbal and physical aggression and negative affect).

Methods

The study design randomly assigned culturally diverse Head Start programs and elementary schools serving low-income populations to intervention or control conditions. Random assignment was conducted separately for matched pairs of Head Start programs and elementary schools. In the intervention schools, all children enrolled in Head Start, kindergarten, or grade one classrooms received the IY social, emotional and problem solving curriculum (Dinosaur School) as a prevention intervention. Control schools followed their usual school curriculum. Informed consent was obtained from all teachers and parents.

School and subject selection

One hundred and twenty classrooms from Seattle area Head Starts and 14 elementary schools were involved in the project. All participating schools served a diverse low-income and multi-ethnic population. Students are typically admitted to Head Start based on low socio-economic status and elementary schools were selected for the project based on higher levels of free and reduced lunch ($M = 59\%$ free lunch). These schools were matched on variables such as size, geographic location, and demographics of the children, and matched pairs were randomly assigned to intervention or control conditions (comparability of intervention and control conditions is reported below). Parents of *all* children in the study classrooms were invited to participate in the research project and 86% of Head Start and 77% of elementary school families who were approached signed consent forms indicating their willingness to participate. Data were collected only on children whose parents had consented, but all children in the intervention classrooms received the classroom intervention. To measure the effectiveness of the prevention program, children were assessed in the fall, the intervention was conducted from November to April, and participants were retested in the spring at the end of the school year. At each assessment period, children, parents, and teachers completed report measures and children and teachers were observed in the classrooms by independent observers (blind to intervention condition) during structured and unstructured times (e.g., playground).

Recruitment of schools and students occurred in each of 4 consecutive years (4 cohorts) in order to ease project burden in each year. By design, schools that served as control participated as intervention in the next year, and also by design, each year a new set of schools were matched and randomly assigned to intervention or control. This procedure was repeated over four consecutive years to fill out the sample.

To keep models as simple as possible, we just used a teacher's first year of participation to evaluate the intervention impact on teacher outcomes (i.e., the TCI, see below). This meant that if a teacher participated for 2 or more years (e.g., year 1 in the control condition and year 2 in the intervention condition), we only used the year 1 TCI data in the models. In some classrooms there were multiple teachers, therefore the nesting structure for the TCI outcomes was repeated measures, pre and post, within teachers, teachers within classrooms and classrooms (i.e., 3 levels). For the TCI outcomes, we had 153 teachers nested within 120 classrooms with 93, 21 and 6 classrooms having 1, 2 and 3 teachers respectively nested within them. The 120 classrooms were 33 Head Start, 42 K and 45 first grade. The numbers of classrooms and teachers in the statistical models, as reported in the results section, were reduced slightly depending on the pattern of missing data for a particular outcome.

For intervention impact on teacher–student (i.e., Teacher MOOSES) and student outcomes, however, the study design lead to a different nesting structure. Teachers who participated for 2 or more years but had different students in year 1 than year 2 contributed 2 classrooms to the analyses. Thus for these models, not only was there an additional level of nesting but there was also a reversal of teachers and classrooms in the nesting hierarchy. The nesting hierarchy was repeated measures nested under students, students nested under classrooms, classrooms nested under teachers (i.e., 4 levels as shown in column 2 of Table 1). Having classrooms nested under teachers allows correlation to arise due to the teacher, the unique composition of the group of students or both. We thought it unwise to a priori assume one or the other would be absent. For teacher–student and student outcomes, we had 1,768

Table 1 Nesting hierarchy

Teacher outcomes	Student outcomes
Classrooms	Teachers
Teachers	Classrooms
Repeated measures	Students
	Repeated measures

students nested within 160 classrooms nested under 119 teachers. Of the 119 teachers, 79 teachers contributed 1 class, 39 contributed 2 classes and 1 contributed 3 classes. Of the 160 classrooms, 42 were Head Start, 59 were K and 59 were 1st grade. Of the 119 teachers, 33 were HS, 37 were K, 44 were 1st and 5 were K/1st. The number of students in the statistical models, as reported in the results section, was reduced slightly depending on the pattern of missing data for a particular outcome.

Using this design some teachers (but no students) crossed over from the control condition to the intervention condition in a subsequent year (with a new group of students). Thus all analyses on student outcomes involved independent groups of students. All analyses on teacher–student outcomes (i.e., teacher behaviors directed to individual students) involved both years of teacher participation and thus the same teacher but new students. Regardless of number of years of participation, all teachers and students were naïve to the intervention at the pre test. For teachers with 2 years of participation, group status (intervention versus control) was included in the models as a time-varying covariate. All models also included potential correlation in outcomes due to the presence of the same teacher across years. Thus, we viewed our design as a simplification of a deliberate cross-over design. Because only control teachers crossed over, we did not include parameters in our models that are usually standard in cross-over designs such as carry over effects, sequence effects or period effects (Milliken & Johnson, 1984).

Elementary intervention and control schools showed no significant differences on key school-level demographic variables. In the intervention schools 56.67% of children received free and reduced lunch compared to 58.75% for control schools. School student enrollment averages were 323 for intervention schools versus 313 for control schools. The percentage of children in study schools who met 4th grade achievement standards were also not significantly different for intervention and control schools; for reading (71% intervention versus 67% control), and math (45% intervention versus control 38%). Head Start uses consistent federal poverty guidelines for enrollment and class size was consistently set at 18 students, one teacher, and one teacher's assistant, thus Head Start intervention and control classrooms in this study were equivalent on classroom variables.

Student-level variables were also comparable across intervention and control conditions. No significant differences were found for any of the following demographic variables. On average students were 63.7 months of age ($SD = 12.7$, range = 35 to 101) and 50% were male. This sample was ethnically diverse (18% Latino, 18% African American, 20% Asian, 27% Caucasian, 8% African, and 9% other minority), and 31% of the children did not speak English as their first language.

Teacher demographic variables were comparable across the intervention and control conditions. No significant differences were found for any of the following teacher variables. Teachers were Caucasian (65%), African American (16%), Asian (12%) and other (8%) and 95% female. Thirty-nine percent of teachers taught in Head Start (HS), 30% kindergarten, and 31% 1st grade. Teachers' level of education was high school (4%), two years of college (13%), bachelor's degree (43%), master's degree (40%).

Selection of moderate- to high-risk group

In addition to the classroom observations and parent and teacher reports, we were interested in directly assessing children's cognitive solutions to social problem-solving situations as well as their feelings vocabulary. Due to budget considerations we were not able to individually test all the children in the study; therefore, we selected an indicated sub-sample of moderate- to high-risk children to participate in these assessments. Since this was a prevention study, a relatively low screening threshold for behavior problems was used to identify the indicated sample. Selection criteria included a method of selecting children who had a higher than average number of behavior problems, but did not limit screening exclusively to a clinical sample. Thus, *either* a parent, teacher, or school counselor report was enough to classify a child as 'indicated.' For the parent rating, children whose parents reported greater than 10 behavior problems on the Eyberg Child Behavior Inventory (ECBI; Robinson, Eyberg, & Ross, 1980) were considered moderate risk. This cut-off has been has been used in our prior prevention studies with low-income families (Webster-Stratton & Hammond, 1998). Teachers' reports were also used to select students who had higher than average levels of problem behaviors in the classroom based on reports on the externalizing scale of the Social Competence Behavior Evaluation (LaFreniere, Dumas, Dubeau, & Capuano, 1992).

Interventions

Teacher training. Intervention teachers participated in 4 days (28 hours) of training spread out in monthly workshops. The training followed the textbook *How to Promote Social and Emotional Competence in Young Children* (Webster-Stratton, 2000). Approximately half of this training focused on classroom management strategies such as ways to develop positive relationships with students and their parents, proactive teaching methods, effective use of praise and encouragement, incentive programs for targeted prosocial skills, setting up discipline hierarchies and individual behavior plans for identified children with problem behaviors. Physical aggression and high levels of oppositional defiant behavior were targeted for close monitoring, incentive and discipline programs in structured and unstructured settings (such as the playground). Teachers learned ways to promote children's self-regulation through persistence, emotion and

problem-solving coaching as well as ways to promote social competence through social peer coaching. Teachers were also encouraged to involve parents in home–school behavior plans as well as classroom learning using regular parent letters about Dinosaur School, weekly Dinosaur homework for children to be completed with parents, and invitations to visit the classroom. More details about this curriculum can be found in Webster-Stratton and Reid (2004).

Dinosaur School. The Dina Dinosaur Social Skills and Problem Solving Curriculum was designed to promote children's social competence, emotional self-regulation (e.g., engagement with classroom activities, persistence, problem solving, anger control), and school behavior (e.g., following teacher directions, cooperation). Our classroom-based version of the Dinosaur Curriculum uses a format of 30 classroom lessons per year and has preschool and primary grade versions. The content is broken into 7 units: (a) Learning school rules; (b) How to be successful in school; (c) Emotional literacy, empathy, and perspective taking; (d) Interpersonal problem solving; (e) Anger management; (f) Social skills; and (g) Communication skills. Teachers followed lesson plans that covered each of these content areas at least 2 times a week, using 15–20-minute large group circle time followed by 20 minutes of small group skill-practice activities. A certified research staff member co-led all the lessons with the teachers to ensure that each classroom received a full dose of intervention. Teachers made the lessons developmentally appropriate for the children in their classrooms by choosing from recommended vignettes and small group activities. There are over 300 small group activities which focus on social emotional skills and cover a wide variety of teaching modalities. The program also consists of over 100 videotaped models of children demonstrating social skills and conflict management strategies. In addition, the program is young child friendly, using life-size puppets, Dinosaur homework activities, picture cue cards for non readers, and games to stimulate group discussion, cooperation, and skill-building. In the classroom, teachers were encouraged to promote the skills taught in circle time lessons throughout the day during less structured settings, such as during choice time, in the lunchroom, or on the playground.

Control classrooms. Families, teachers and children in the control classrooms continued their regular Head Start and elementary school curriculum and services.

Measures

All assessments were conducted on the same time line and frequency in both conditions. The pretests were conducted in the early fall. The intervention ran from November to April, and post-assessments were conducted in late spring. Assessments measured social and emotional competencies, conduct problems, teacher competencies, teacher efforts to involve parents and classroom environment by both teacher reports and independent observations of teachers and all the children in the classroom. A sub-sample of moderate- to high-risk children (216) were selected based on elevated problem scores (selection described above). This sample of children was tested using the Wally Problem-Solving test (Webster-Stratton, 1990; Webster-Stratton et al., 2001). During the last two years of the grant, we developed the Wally Feelings test as a pilot measure to assess the size of children's feelings vocabulary. This test was administered to the moderate- to high-risk children in the last two cohorts, and pilot results are presented for 52 children.

Independent classroom observations

At each assessment period, each child was observed on two separate occasions of 30 minutes each and approximately half of the observation time was in a structured setting and the other half in an unstructured setting. At the same time that coders were conducting observations of child behavior, they also coded teacher behaviors. Since each teacher had multiple children in the classroom, the number of times that teachers were coded varied depending on the number of children in her classroom. Each teacher was coded on at least two and up to eight times at each time point. Coders were blinded to study condition and random reliability checks were completed on approximately 20% of all observations. Coders record discrete behaviors as well as complete rating scales that provide information on teaching style, the quality and duration of children's interactions with teachers and peers. These behavior codes have been used successfully in our studies with over 1,000 Head Start and elementary school students and over 300 teachers (Webster-Stratton & Lindsay, 1999). The specific coding systems are described next.

Observations of teacher classroom management behaviors and teaching style

Multiple Option Observation System for Experimental Studies (MOOSES). The MOOSES classroom observation coding system developed by Tapp, Wehby, and Ellis (1995), revised by our team for use with young children, was used to code teachers' interactions with children as well as children's interactions with teachers and peers. There were three discrete teacher-focused behavior codes: (a) positive reinforcement (praise and encouragement), (b) critical statements, and (c) amount of interaction with students. Coders record frequency of the behavior directly into hand-held computers. Reliability as measured by intraclass correlations were as follows: teacher–child involvement = .94, teacher critical = .83, and teacher praise = .81.

Teacher Coder Impressions Inventory (TCI). The TCI was developed by our group to evaluate teacher's style and classroom management strategies. Coders complete a series of 71 Likert-type questions rating teaching style which were classified into five summary scores based on theoretical considerations and confirmed using factor analyses. The five scales include: (a) Harsh/Critical Style (29 items including threats, criticism, sarcasm, anger, physical aggression and verbal aggression, overly strict, anger, fearful),

(b) Inconsistent/Permissive Style (12 items including no follow-through, failure to monitor, tentative/indecisive, overly permissive), (c) Warm/Affectionate Style (11 items including modeling positive behavior, reinforcing, paying attention, verbal and physical affection, playful, gives rationale), (d) Social/Emotional Teaching (10 items including teaches prosocial behavior, problem-solves, shapes positive peer play, encourages feeling language, promotes social competence), and (e) Effective Discipline (6 items including follow through with threat, warned of consequences, used an incentive program, posted rules and a schedule, used time out, withdrew privileges). Standardized alpha coefficients and intraclass reliability coefficients were Harsh/Critical, alpha = .98, ICC = .83; Inconsistent/Permissive, α = .93, ICC = .73; Warm/Affectionate, α = .90, ICC = .67, Social-Emotional Teaching α = .84, ICC = .62; and Effective Discipline α = .58, ICC = .61. Internal consistency is based on all observations for all teachers on all occasions (N = 1988).

Observations of child conduct problems, emotional self-regulation and social competence

Multiple Option Observation System for Experimental Studies (MOOSES) (Tapp et al., 1995). Coders rated frequency of discrete child behavior codes, which occurred during two separate 30-minute structured and unstructured observations, as well as the duration of a child's involvement with peers and off-task or disengaged behavior. These were summed to form six child variables: (a) total conduct problems, which included physically aggressive behavior (grabbing, hitting, biting, throwing objects), verbal aggressive behavior (yelling, swearing, mocking), and noncompliant or oppositional responses to teacher command or instructions. Two variables measured emotional self-regulation: (b) percent time child disengaged/off-task from classroom activities and (c) percent time in solitary play. Three variables measured social competence: (d) child positive with teacher; (e) child positive with peer; and (f) percent time in peer involvement. Inter-rater reliabilities as measures by intraclass correlations were as follows: child conduct problems = .94, peer involvement = .95, solitary involvement = .94, child positive to teacher = .93, child positive to peer = .88, child disengaged = .88.

School Readiness and Conduct Problems: Coder Observation of Adaptation-Revised (COCA-R). This measure is an observational version of the TOCA-R (Werthamer-Larsson, Kellam, & Oveson-McGregor, 1990). Following the 30-minute observation, coders respond to 36 items to obtain an overall school readiness score. This score includes items on children's emotional self-regulation skills (e.g., concentration, controls temper, expresses feelings appropriately, eagerness to learn, cooperation, task completion, can calm down, and distractibility). It also includes items on social skills (e.g., being friendly, helping others, giving compliments, not bossy with suggestions, liked by classmates, initiating peer interactions), and conduct problems (aggression, noncompliance, teasing, and destructive behavior). Internal consistency (α = .92) and interrater reliability (ICC = .80) were high.

Observations of classroom atmosphere

Classroom Atmosphere Measure (CAS). This 10-item questionnaire (Greenberg et al., 1995) is completed by observers rating the general classroom atmosphere. Observers rate *overall* classroom level of students' cooperation and problem solving, interest in subject matter, focus, responsiveness, on-task behavior, and classroom support. Observers also rate how well the teacher is able to manage overall levels of classroom disruptive behavior, transitions, and follow through on rules. Thus, the measure reflects the interactions between the teacher's style and the children's behaviors as a group rather than focusing on an individual child. In our sample, this scale shows good internal consistency for the 10 items (α = .93) and adequate inter-rater reliability (ICC = .74) based on 281 primary–secondary pairs. Alphas by grade are .92 (Head Start), .94 (kindergarten) and .94 (1st grade).

Child problem solving and feelings testing

Wally Problem Solving and Feelings Tests. The WALLY Problem Solving test (Webster-Stratton, 1990) measures children's problem-solving skills or solutions in response to hypothetical problem situations. Summary scores include the number of different positive and negative strategies that children generate in order to solve the problem. The WALLY was derived from Spivak and Shure's Preschool Problem Solving Test (Spivak & Shure, 1985) and Rubin and Krasnor's Child Social Problem-Solving Test (Rubin & Krasnor, 1986). Inter-rater reliability for number of different positive strategies was ICC .93 and for different negative strategies was ICC .71.

The Wally Feelings test was piloted in the later half of this project to measure whether there was an increase in children's feelings language. Children are shown a series of 8 pictures of children in positive and negative situations and are asked how the characters in the situations feel. The frequencies of different negative and positive feeling words expressed by the children are summed to give a total positive feeling and negative feeling vocabulary score.

Parent involvement

Teacher–Parent Involvement Questionnaire (INVOLVE-T). The INVOLVE-T is a 20-item teacher questionnaire developed by the Oregon Social Learning Center and revised by us for use with young children. The measure asks teachers to report on the extent to which parents participate in school activities, seem comfortable with the teacher and school environment, value education, support the teacher, and assist children with their homework. The questionnaire has 3 subscales: (a) Teacher Bonding With Parent (e.g., teacher called parent, wrote note, invited parent to school, comfortable meeting with parent); (b) Parent Involvement in Education (e.g., parent has same goals as teacher, thinks education important, helps with homework), and (c) Parent Involvement With School/Teacher (e.g., parent calls teacher, parent visits classroom, parent attends conferences, parent volunteers). Alphas are .76, .91, and .84 respectively.

Satisfaction with program

Teacher Satisfaction Questionnaire. Following each day of teacher training, a brief attitude inventory is completed regarding teacher satisfaction in terms of content of program, videotapes shown, methods utilized (e.g., role plays, behavior plans), and group discussion. A more comprehensive teacher satisfaction measure is completed at the end of the year.

Parent Curriculum Involvement and Satisfaction Questionnaire. Parents completed a brief end of the year questionnaire asking about their feelings about the Dinosaur curriculum, how much they talked with their children about the program, the value of the homework and how much their children used the program strategies at home.

Intervention integrity

Fidelity was monitored and measured in the following ways: (a) Teacher training was conducted using a standard protocol and was delivered by certified IY trainers; (b) All trainings were videotaped and reviewed by the program developer; (c) Detailed manuals were provided for all Dinosaur lessons, complete with activities, role plays, and homework assignments; (d) Protocol checklists were completed by the research co-leader after each session, indicating which lessons, small group activities and vignettes were used; (e) Lessons were observed by certified IY supervisors and standardized process and content evaluations were completed after each of these observations; (f) IY Dinosaur research co-leaders met for weekly supervision to review protocols and ensure adherence to the curriculum.

Results

Intervention integrity

Because the classroom intervention was delivered in a partnership between the teacher and the research co-leaders there was a very high degree of intervention integrity. All children in the intervention classrooms participated in the IY Dinosaur intervention and their teachers were offered 4 days of training in the Dinosaur curriculum. Teachers received an average of 3.73 days of training (only 4 teachers attended less than the full four days of training). Checklists completed by the research co-leaders indicated that, on average, 27 of the 30 required lessons were completed in the intervention classrooms and an average of 30 recommended vignettes and 25 small group activities were completed in the intervention classrooms.

Observations of teacher behavior

As explained in the methods section, for the teacher outcome analyses, teachers were nested within classrooms, which predominantly corresponded to a teacher and assistant(s) or in some cases co-teacher(s). Missing data at the pre- and post-test reduced the number of classrooms to 115 and the number of teachers to 139. Approximately 83% (pre-test) and 87% (post-test) of teachers had between 2 and 5 observations (i.e., 13–17% had only one observation or more than 5).

Modeling method. Because teachers within a classroom are observed dealing with the same group of students, some correlation would be expected for teachers within a classroom. Employment selection or shared classroom experience may also produce a shared style of teaching within classrooms. The intervention was aimed at individual teachers, however, so it is desirable to analyze the outcome data at that level. Within teachers, we have up to 8 repeated measures at both the pre and post test. Accordingly, a multi-level random intercept and slope model was used within a pre-post ANCOVA model that allowed for both classroom- and teacher-level variation in intercept and post on pre regression. We did not expect both sets of random effects to be important and we anticipated that teacher-level effects would be more important than classroom-level effects because only 23 classrooms had more than 1 teacher. But we started out by being over-inclusive and then trimmed the models as necessary to eliminate convergence problems and to identify the most important and parsimonious set of random effects. Initial models also included intervention by pre score interactions to test for differential effectiveness of the intervention by initial level.

Background information was available for ethnicity, gender, grade level and number of days of training for each teacher. For simplicity, ethnicity was collapsed to white versus nonwhite, grade level was collapsed to Head Start versus kindergarten and first grade and training was collapsed to 4 versus less than 4 days (only 4 of the intervention teachers missed any training days).

Results of observations of teacher behavior management skills and teaching style

Teacher Coder Impression (TCI) analyses. Coders rated teacher's teaching style on 5 different TCI constructs, Harsh/Critical, Inconsistent/Permissive, Warm/Affectionate, Social/Emotional, and Effective Discipline. Table 2 shows model estimates for fixed and random effects for each construct. Where background covariates (ethnicity, gender, grade and training) were significant, they were included in the final model. Most of the initial models with the full set of classroom- and teacher-level random effects had serious convergence problems or produced non-positive definite information matrices. Using a series of nested chi-square tests, we identified the most parsimonious set of random effects to include and except for one construct this usually

Table 2 Results for coder ratings of teacher behavior (TCI)

Effect	Warmth Value	Inconsistent Value	Harsh/critical Value	Social/emotion Value	Eff. discipline Value
Fixed effects					
Intercept	2.627***	.690***	.806***	1.294***	1.138***
Pre	.364***	.399***	.401***	.296***	.224***
Intervention	.203*	−.112*	−.126**	.231***	.254*
Nonwhite			.129**	−.151*	
1st/K versus HS					.310***
Intervention by 1st/K					−.286*
Random effects					
Teacher intercept	.398***	.178***	.189***	.241[a***]	.204[a***]
Teacher random slope				.258**	.320[a***]
Residual	.554***	.293***	.252***	.396***	.415***

Note: *$p < .05$, **$p < .01$, ***$p < .001$.
[a]Random effect is at classroom not teacher level. Random effect parameters are standard deviations or correlations.

was either classroom-level or teacher-level effects but not a mixture of both. None of the intervention effects shifted from being significant to non-significant or vice versa as a result of trimming the random effects. None of the pre by intervention interaction effects (i.e., differential effectiveness) were significant and none were retained in final models.

Four of 5 constructs had significant main effects such that teachers in the intervention condition became less harsh/critical and inconsistent/permissive, more warm/affectionate, and placed more emphasis on social/emotional teaching. In preliminary models, effective discipline did not show significant main effects of intervention but instead the intervention effect depended on the grade of the teacher (1st/K vs. HS); Accordingly, the model in Table 2 includes a dummy variable indicator of HS vs. 1st/K (HS = 0, 1st/K = 1) and an interaction of the dummy variable with intervention status. As shown in Table 2, there was a significant main effect for grade (fixed effect for 1st/K versus HS = .310, $p < .001$) indicating that 1st/K teachers had higher levels of effective discipline than HS teachers and a significant main effect for intervention (fixed effect for Intervention = .254, $p < .05$) indicating that HS intervention teachers improved more on effective discipline than HS comparison teachers. The interaction effect of grade with intervention (fixed effect for Intervention by 1st/K = −.286, $p < .05$) was negative, indicating no intervention effect for 1st/K teachers. For the other 4 constructs, preliminary analyses indicated there were no significant differences in intervention effects for HS versus K and 1st so these effects were not included in the final models shown in Table 2. Effect sizes, when computed as the adjusted mean level shift in the post score due to intervention divided by the standard deviation of the classroom- or teacher-level random intercept, ranged from medium to large: warmth/affectionate (.51), inconsistent/permissive (.63), harsh/critical (.67), social/emotional teacher (.96), and effective discipline (1.24), for Head Start teachers.

Results of observations of teacher classroom management and children's school readiness and conduct problems

As explained in the methods section, for the teacher–student and student outcome analyses, students were nested within classrooms, which in turn were nested under teachers. The sample size of students who were considered in the study was 1,768. Models were based on 1,746 students who had at least one observational measurement and background covariates of age and gender. These 1,746 students were nested in 160 classrooms, which in turn were nested under 119 teachers (40 teachers had 2 or more classrooms).

Modeling method. We started by fitting a 4-level model (repeated measures within students, students within classrooms, classrooms within teachers, and teachers) for both teacher–student and student outcomes. Our initial model included 3 random effects at the student, classroom, and teacher levels, intercept (initial status), normative slope (time) and intervention slope. The normative slope applied to teachers' or students' growth when they were in the control condition and both the normative and intervention slope applied to teachers' or students' growth when they were in the intervention condition. Thus, the intervention slope was an additional increment in growth over and above the normative slope when the teachers or students were in the intervention condition. By specifying the intervention slope as random, we allowed for differential student response to intervention, that is, we did not expect all teachers or students to derive the same benefit on average from the intervention. Both random slopes were allowed to correlate with initial status but not with each other, for identification purposes. The correlation of the intervention slope with initial status is substantively interesting because it represents differential response to the intervention based on initial level. In universal

preventive interventions aimed at behavior problems, it is not unusual for children with few or no problems to show little benefit while children with initial high levels of behavior problems show considerable benefit. This pattern produces a negative correlation between initial status and the intervention slope. In summary, the key parameters related to intervention effects in Tables 4–6 are the fixed effect for the intervention slope, which represents the main effect of the intervention, and the random effect for the correlation of the intervention slope with initial status, which represents differential effectiveness of the intervention related to initial level. See (Muthen & Curran, 1997; Stoolmiller, Eddy, & Reid, 2000) for more details about models for differential effectiveness in randomized universal prevention trials.

We expected the random effect part of the model to be over-parameterized, and indeed it was, but we started out by being over-inclusive and then trimmed the models as necessary to eliminate convergence problems and to identify the most important and parsimonious set of random effects. For teachers, only teacher-level and in some case classroom-level random effects were significant. None of the student level random effects were significant. This is perhaps not surprising since the intervention targeted teacher behavior and the outcomes were teacher behavior. For students, models with teacher-level initial status, normative and intervention slopes and classroom and student intercepts provided the best combination of fit and parsimony. This is perhaps not surprising given that teachers were the targets of intervention. To probe significant correlations of initial status with intervention slope, we used model results to compute point-wise *t*-tests across the range of observed pre scores and noted where the intervention and control groups first became significantly different at the .05 level. We also computed point-wise effect sizes which we defined as the intervention slope mean (i.e., the mean slope difference between the groups) divided by the standard deviation of the teacher level normative slope.

Observation of Teacher Classroom Management Behaviors (MOOSES). Results for Teacher involvement, teacher critical, and teacher praise from the MOOSES are shown in Table 3. Although initial models included student-level random effects, none were significant so none are shown in Table 3. Teacher critical showed a significant main effect of intervention (fixed effect for Intervention slope = −.181, .001 < p < .01). The correlation of initial status with the intervention slope was negative and significant (Cor (Initial status, Int. slope) = −.434, .01 < p < .05), indicating significant differential effectiveness; the more critical the teacher was initially, the more her score improved at the post-test. The intervention effect first became significant at p < .05 at .24 standard deviations below the pre-score mean (min. significant effect size = −.382 in Table 3), a medium small effect. At 2 standard deviations above the pre-score mean the effect size

Table 3 Multi-level results for teacher MOOSES

Effect	Teacher–child involvement[1] Value	Teacher critical[2] Value	Teacher praise[3] Value
Fixed effects			
Initial status	8.416***	1.514***	2.873***
Normative slope	−.043	.021	−.168
Girl	−.140	−.009	.030
Age at entry	−.148	−.053*	−.051
Observation occasion 1 vs. 2	−1.688***	−.240***	−.588***
Intervention slope	−.465	−.181**	.031
Random effects			
Classroom			
Normative slope			.643***
Initial status	1.011***	.436***	.945***
Intervention slope			.022
Cor(Initial status, norm. slope)			−.458***
Cor(Initial status, int. slope)			
Teacher			
Normative slope	1.564***	.372***	
Initial status	1.699***	.572***	.693***
Intervention slope	1.295***	.302***	
Cor(Initial status, norm. slope)	−.496*	−.188	
Cor(Initial status, Int. slope)	−.302	−.434*	
Within student residual	5.132***	.851***	1.538***
Average effect size		−.486	
Min. significant effect size		−.382	
Max. significant effect size ($z = 2.00$)		−1.370	

Note. Estimated random effects are standard deviations unless otherwise indicated.
Transformations: [1]Root (1.5); [2]Log; [3]Root (2).
*p < .05, **p < .01, ***p < .001.

Table 4 School readiness: Coder Observation of Classroom Adaptation-Revised (COCA-R) (square root transformed)

Effect	Value	Std.Error	t value	p value
Fixed effects				
Girl	−.0800	.0093	−8.5989	.0000
Age at entry	−.0423	.0068	−6.1847	.0000
Observation occasion 1 vs. 2	.0190	.0068	2.7990	.0051
Initial status	.7622	.0119	63.8898	.0000
Normative slope	−.0007	.0117	−.0636	.9493
Intervention slope	−.0349	.0170	−2.0534	.0401
Random effects				
Student intercept	.1257	.0054	23.3938	.0000
Classroom intercept	.0376	.0165	2.2827	.0224
Teacher				
Initial status	.0822	.0105	7.8552	.0000
Normative slope	.0414	.0135	3.0610	.0022
Intervention slope	.0925	.0143	6.4791	.0000
Cor (initial status, int. slope)	−.4551	.2034	−2.2373	.0253
Within student residual	.2749	.0028	97.2359	.0000

Table 5 Multi-level results for child MOOSES

Effect	Child conduct problems[1] Value	Child disengage[1] Value	Peer involve Value	Child solitary[2] Value	Child positive teacher Value	Child positive peer Value
Fixed effects						
Initial status	2.157***	−.848***	33.698***	5.661***	2.167***	3.178***
Normative slope	.035	.118*	3.797***	−.373***	−.014	.265***
Girl	.403***	−.479***	−.554	.131	−.088***	−.122**
Age at entry	.159***	−.141***	2.307***	−.082	−.135***	.186***
Observation occasion 1 vs. 2		.026	8.167***	.046	−.091***	.262***
Intervention slope	.040	−.082	1.552	.090	−.055	.011
Random effects						
Student						
Normative slope	1.240***	.569***			.006	
Initial status	1.066***	.629***	.094	.765***	.244***	.470***
Intervention slope	.606***					
Cor(Initial status, norm. slope)	.596***	−.414*				
Cor(Initial status, int. slope)	.296***					
Classroom						
Normative slope			.044			
Initial Status	.226**	.187**	2.661**	.350***	.071*	.149*
Intervention slope			3.558**			
Cor(initial status, norm. slope)			.013			
Cor(initial status, int. slope)			−.703			
Teacher						
Normative slope		.135			.153***	.296***
Initial status	.347***	.435***	3.690***	.441***	.310***	.283***
Intervention slope	.377***	.401***		.634***	.066	.215*
Cor(initial status, norm. slope)		−.030			−.111	−.104
Cor(Initial status, int. slope)	.522*	−.608**			−.589	
Within student residual	.590[a]	1.353***	22.861***	2.667***	.579***	1.154***
Average effect size	.032	−.141				
Min. significant effect size	.704	−.295				
Max. significant effect size ($z = 2.00$)	1.095	−1.648				

Note. [a]Fixed reliability estimate (20% of the observed variance at pre score). Estimated random effects are standard deviations unless indicated otherwise. *$p < .05$, **$p < .01$, ***$p < .001$. Transformations: [1]Log; [2]Root (1.75).

Table 6 Classroom atmosphere

Effect	Value	Std.error	*t*-value	*p*-value
Fixed effects				
Girl	−.0235	.0140	−1.6829	.0926
Age at entry	−.0079	.0142	−.5523	.5808
Observation occasion 1 vs. 2	.0383	.0130	2.9430	.0033
Initial status	2.3896	.0411	58.2014	.0000
Normative slope	.0176	.0353	.4976	.6188
Intervention slope	−.1500	.0427	−3.5142	.0004
Random effects				
Student intercept	.0769	.0220	3.4974	.0005
Classroom				
Initial status	.2510	.0306	8.2109	.0000
Normative slope	.1947	.0373	5.2257	.0000
Cor (Initial status, norm slope)	−.6181	.2231	−2.7707	.0056
Teacher				
Normative slope	.1463	.0389	3.7623	.0002
Initial status	.3463	.0359	9.6458	.0000
Intervention slope	.1408	.0660	2.1336	.0329
Cor (initial status, int. slope)	−.5813	.4261	−1.3641	.1725
Within student residual	.5074	.0056	91.2291	.0000

Note. Random effect parameters are standard deviations or correlations.

was −1.370, (max. significant effect size in Table 3), a very large effect. No other MOOSES teacher constructs showed significant effects.

School readiness and conduct problems: Coder Observation of Classroom Adaptation (COCA-R). Table 4 shows results for the child school readiness score (lower scores on this measure indicate better school adjustment). Background covariates of age, gender and observation occasion all had strong effects on the outcome but did not interact with intervention status. Girls and older children had lower scores and children had higher scores on the second occasion of observation than the first. The fixed effect for the intervention slope is negative and significant, indicating a greater improvement in school readiness in the intervention than the control condition. Both the teacher-level standard deviation (*SD*) for the intervention slope and the teacher-level correlation of the intervention slope with initial status are significant. The significant *SD* indicates significant individual differences in student outcomes at the teacher level in response to intervention. The significant and negative correlation indicates that teachers whose students

had worse scores on this measure showed more change in the intervention year than those that started with students who had better school readiness scores. Follow-up analyses to probe the dependence of the intervention effect on initial status reveal that the intervention effects were significant at the .05 level starting at about the mean of the pre score with an effect size of −.82 and going down to an effect size of −2.87 at 2 standard deviations above the pre score mean. Thus, the intervention had a large impact on average student scores for teachers with students with average levels of poor school readiness and a very, very large impact on average student scores for teachers with students with very poor initial levels of school readiness.

Observation of Child School Readiness and Conduct Problems (MOOSES). None of the 6 child MOOSES constructs showed significant main effects of intervention (fixed effect for intervention slope in Table 5) but child conduct problems and disengagement scores both showed significant negative correlations of initial status with intervention slope (Cor(init. status, int. slope) in Table 5) signaling differential effectiveness at the teacher level, and child conduct problems showed a significant negative correlation of initial status with intervention slope (Cor(init. status, int. slope) in Table 5), signaling differential effectiveness at the student level as well. For both constructs, findings indicate that the higher the initial average child conduct problems for a teacher, the more improvement in average child scores at the post test. For child conduct problems, this differential effectiveness was also replicated at the student level, indicating that over and above the effect at the teacher level, children with higher baseline conduct problems showed more improvement at post test. For child conduct problems, the intervention effect first became significant at $p < .05$ at 1.42 standard deviations above the pre score mean with an effect size of −.70, (min. significant effect size in Table 5), a medium large effect. At 2 standard deviations above the pre score mean the effect size was −1.10, (max. significant effect size in Table 5), a large effect. For the child disengagement variable, the intervention effect first became significant at $p < .05$ at .20 standard deviations above the pre score mean with an effect size of −.29, (min. significant effect size in Table 5), a small effect. At 2 standard deviations above the pre score mean the effect size was −1.65, (max. significant effect size in Table 5), a very large effect.

Classroom atmosphere. Table 6 shows results for the classroom atmosphere total score (low scores indicate better classroom atmosphere). The intervention slope (fixed effect for Intervention Slope in Table 6) is negative and significant indicating a greater improvement in classroom atmosphere in the intervention than the control condition (effect size was large = 1.03, not shown in Table 6). Of the background covariates, only the observation occasion had a strong fixed effect, similar to the child school readiness total score. The only significant random effect at the student level was the student intercept. The classroom-level correlation of initial status with intervention slope was not significant and eliminated from the model to remedy convergence problems, providing no evidence for differential effectiveness related to initial level. The teacher-level standard deviation of the intervention slope was marginally significant ($z = 2.13$, $p = .0329$), indicating some tendency for differential effectiveness, but the teacher-level correlation of initial status with intervention slope was not significant (Cor(initial status, int. slope) = −.58, $z = -1.36$, $p = .1725$), suggesting that the differential effectiveness was not related to initial status.

Child problem-solving and feelings testing

Wally Problem Solving and Feelings Tests. See Table 7 for means and standard deviations on these measures. Mixed-design ANOVA (time by condition) was used to evaluate the intervention effects on these measures. Children in the intervention condition showed significantly greater improvement than the control children on the number of different positive strategies generated; $F(1,214) = 9.27$, $p < .01$, $\text{Eta}^2 = .041$. On the Wally Feelings test, intervention children showed significantly greater improvement than the control group in the number of positive feelings that they could identify; $F(1,52) = 8.58$, $p < .01$, $\text{Eta}^2 = .14$. These results help to bolster the differential effectiveness findings presented in the multi-level models.

Table 7 Wally problem-solving and feelings at pre and post by condition

	Control					Intervention				
	Pre		Post			Pre		Post		
	M	*SD*	*M*	*SD*	*N*	*M*	*SD*	M	*SD*	*N*
Wally's problem-solving # different positive strategies	5.51	2.09	6.61	1.99	96	5.48	2.47	7.21	2.42	120
Wally's problem-solving # different negative strategies[1]	1.81	1.36	1.39	1.23	96	1.46	1.27	1.27	1.16	120
Wally feelings # positive feelings[1]	1.75	1.07	2.00	.92	20	1.79	1.39	3.71	2.28	34
Wally feelings # negative feelings[1]	2.25	1.21	3.00	1.69	20	2.82	1.78	4.21	2.20	34

Note. [1]Log transformations for analyses.

Parent involvement

Teacher-Parent Involvement Questionnaire (INVOLVE-T). Table 8 shows results for a multi-level pre-post ANCOVA with parents nested within teachers. No intervention effects or interactions of pre test by intervention were detected for parent involvement in education, school or a total score composed of both education and school involvement, although the intervention main effect for parent involvement was marginally significant ($p = .053$). For parent–teacher bonding, however, a significant pre by intervention interaction emerged as well as a significant main effect of intervention favoring the intervention group. The interaction effect was negative, indicating that the post on pre regression was flatter in the intervention group, suggesting differential effectiveness, that is, teacher–parent bonding improved most in the intervention group for teachers who reported low teacher–parent bonding. Follow-up analyses to probe the dependence of the intervention effect on initial status reveal that the intervention effects were significant at the .05 level starting at .39 standard deviations above the mean (2.85) of the pre score with an effect size of .14 and the effects increased for lower pre scores going up to an effect size of .57 at 2 standard deviations below the pre-score mean (1.68). Thus, the intervention had a medium impact on the average parent–teacher bond within teachers at initially low levels of bonding and small impact on the average parent–teacher bond within teachers with average to slightly above average initial levels of bonding.

Satisfaction with program

Family and teacher satisfaction questionnaires. Teachers were very satisfied with both the training they received and the curriculum implementation in their classrooms. Teachers rated 4 aspects of the 4-day training on a 4-point scale (1 = unhelpful and 4 = very helpful): trainer's leadership skills ($M = 3.91$, $SD = .29$), group discussion ($M = 3.72$, $SD = .45$), use of videotape examples ($M = 3.58$, $SD = .59$), and use of role-plays ($M = 3.45$, $SD = .64$). At the end of the year, 83.3% of teachers said that the Dinosaur curriculum was easy to integrate into their regular curricula, 91% said that the program met their social/emotional goals for their students, 73% felt that the content and activities were developmentally appropriate for their students, 75% reported that they would continue the program in the next school year, and 53% reported that they would like ongoing training and technical support regarding the program. Chi square analyses showed no significant differences among grade levels for three of the variables. However, Head Start teachers had significantly more concerns about how to deliver the content and activities in developmentally appropriate ways and requested more ongoing training or technical support than K and 1st grade teachers; Chi square ($1, N = 93$) = 14.45, $p < .01$. Eighty percent of Head Start teachers versus 35.5% of kindergarten teachers and 43.8% of first grade teachers asked for ongoing training.

Parents were also satisfied with the program: 94.1% reported positive overall feelings about the Dinosaur curriculum; 91.4% would recommend the program to other parents; 87.3% found the Dinosaur homework assignments useful; 68.2% often talked at home with their children about the Dinosaur curriculum; and 72.5% said their children used the Dinosaur School strategies at home (e.g., taking deep breaths, talking about feelings, or using problem-solving steps). Chi square analysis showed no differences for parent evaluations across grade levels. These findings indicate that at least 2/3 of the parents were involved in supporting their children's learning in regard to the social and emotional curriculum at home.

Discussion and conclusions

Surveys indicate that kindergarten teachers are very concerned about the number of children who arrive

Table 8 Multi-level ANCOVA results for parent involvement constructs

	Education[a]		School[b]		Total		Teacher bonding	
	Value	Std.error	Value	Std.error	Value	Std.error	Value	Std.error
Fixed effects								
Intercept	1.020	.042	.934	.034	2.837	.051	2.650	.045
Pre score	.594	.043	.831	.074	.589	.046	.439	.053
Intervention	−.020	.040	.061	.031	.025	.045	.099***	.033
Pre by intervention	.098	.052	−.144	.085	.046	.054	−.175***	.060
Random effects								
Teacher level								
Intercept	.275	.029	.224	.024	.349	.035	.342	.032
Post on pre slope	.145	.038	.320	.052	.202	.036	.218	.043
Cor(int, post on pre slope)	.175	.256	−.567	.205	−.426	.210	−.254	.211
Within teacher residual	.433	.010	.336	.008	.472	.011	.324	.007

Notes. [a]Reversed and 2/3rds root transformed. [b]2/3rds root transformed. Random effect parameters are *SD*s or correlations.*$p < .05$, **$p < .01$, ***$p < .001$.

in their classroom without the emotional self-regulatory and social skills to function productively and learn in the classroom (West et al., 2001). They report they are unable to teach academic skills because of difficulty knowing how to manage the increasing numbers of students with conduct problems. Moreover, research has indicated that poor school readiness and increased conduct problems are even more prevalent in classrooms with high percentages of students from poverty situations and where there is a poor connection or involvement between the school and parents. Therefore in this efficacy trial we tested the impact of training teachers in classroom management strategies and in the delivery of a social and emotional curriculum in schools with high percentages of culturally diverse and socio-economically disadvantaged children. The goals of this universal prevention curriculum were to help teachers promote children's social competence and emotional self-regulation, reduce conduct problems, and involve parents in their children's learning. Schools were randomly assigned to the intervention condition or usual school services.

Results of observations of teachers' classroom management style with students indicated significant improvements for teachers in intervention classrooms. Teachers who received intervention were significantly different from control teachers on four of the five TCI variables: harsh/critical, warm/affectionate, inconsistent/permissive, and social/emotional. Intervention teachers used more specific teaching strategies that addressed social and emotional skills than teachers in control classrooms. The effect sizes were moderate to high, indicating that the curriculum and training had robust effects on changing teachers' classroom management approaches. MOOSES ratings confirmed the TCI findings, showing that intervention teachers used significantly fewer critical statements than control teachers.

Results of the observations of students in the classrooms on the COCA-R showed both significant improvement and significant differential improvement in emotional self-regulation, social competence and conduct problems compared with the control students' behaviors. Here the effect size was particularly strong for those students from classrooms with the poorest initial scores. The intervention had a large impact on students from classrooms with average levels of school readiness and conduct problems and a very, very large impact on students from classrooms with very low initial levels of school readiness and high conduct problems. In addition, MOOSES frequency measures of two discrete child behaviors, conduct problems and disengagement (or off-task behavior) in classroom activities, also significantly differentially improved in the intervention classrooms compared to the control condition and showed the same pattern of differential effectiveness. Thus, overall, children from classrooms that were most at initial risk benefited most from the intervention. A global measure of Classroom Atmosphere based on student behaviors of responsiveness, engagement, and cooperativeness and teacher supportive behavior also indicated significant intervention effects. We found no evidence that the student gender, age or grade moderated the effects of the intervention on student outcomes. Indeed, only one of the teacher outcomes (effective discipline) showed HS versus K/1st moderation so it appears that the intervention works equally as well for boys versus girls, and preschool Head Start children versus elementary school children.

All of our student behavioral outcomes showed strong teacher-level effects, meaning that groups of students associated with a particular teacher changed more than groups of students associated with a different particular teacher. In addition, the groups of students that showed the most change due to the intervention were those groups that needed the most improvement to begin with. Because the groups of students are formed about a particular teacher, it is tempting to assume that those groups of students who were most in need of improvement and showed the most change had teachers who also were most in need of improvement and showed the most change as a result of the intervention. That is in fact our hypothesis for future work but it is important to point out that this was never directly demonstrated in any our models and need not be the case.

Results of the individual testing of a subset of high-risk students confirmed the classroom observations of enhanced children's social problem-solving skills and emotional literacy. Students who received intervention had more prosocial solutions to problem situations and an increased positive feeling vocabulary compared with control students. Increasing children's social problem-solving knowledge and emotional language is promising because it increases the likelihood that children exposed to this curriculum will be more successful in solving problems with peers.

Teachers in the intervention group reported feeling more bonded or involved with the parents of children in their classes, with the strongest effects occurring with teachers who reported initial low bonding with parents. This finding indicated that teachers made more efforts to involve parents through newsletters, phone calls and homework. However, intervention teachers did not report a significant improvement in parents' efforts to call them or volunteer in the classroom or attend meetings. Because this intervention was not directly offered to the parents, this might suggest that further studies include an intervention for parents in how to be involved in their children's education and work with the teacher. In fact, in a prior prevention study that offered a 12-week parent training program to parents we did find

that intervention parents were significantly more involved with their children's education and school than control parents on this measure (Reid, Webster-Stratton, & Hammond, in press; Webster-Stratton & Reid, 2008). Since parent involvement and ability to work collaboratively with teachers has been shown to be an important predictor of children's school success (Hawkins et al., 1999), it is important to understand how to promote parental involvement.

Finally, teacher evaluations indicated that teachers were very satisfied with their training and their ability to implement the curriculum in conjunction with their academic curriculum. In addition, parent evaluations indicated that parents were very satisfied with the content of the curriculum and its effects on their children. Interestingly, on these evaluations over 85% found the dinosaur homework useful and over 65% reported using the strategies at home. This would suggest there was a fairly high level of parent involvement at home with the curriculum concepts.

This study contributes to a growing body of literature evaluating the social, emotional and problem-solving classroom curriculum (Domitrovich, Cortes, & Greenberg, in press; Grossman et al., 1997; Walker et al., 1998b) showing promise for improving young children's overall school readiness and reducing conduct problems. Like the PATHS curriculum, the Dinosaur curriculum focused on preschool and kindergarten children who were socioeconomically disadvantaged and showed similar findings in terms of increased emotion knowledge skills as well as enhanced problem-solving strategies for the sub-sample tested with the Wally measures. A strength of the current study is the use of independent classroom observations of teachers' interactions and children's social and emotional behavior. These observations indicated that the intervention resulted in enhanced teacher classroom management skills as well as improvements in children's overall school readiness and reduction of conduct problems. To date, few other studies have used observational methods to measure teacher and child behaviors in the classroom. Instead, most have relied on teacher self-report behavior ratings to measure changes (Domitrovich et al., 2006; Lynch, Geller, & Schmidt, 2004). While teacher report provides important information about teachers' perceptions of children's behavior, these ratings are usually provided by the same teachers who received training and implemented the intervention, and thus may be biased in favor of reporting positive student changes. The addition of independent observations that corroborate teacher report findings strengthens the intervention effects reported in the current study. Further follow-up research is under way to assess whether the changes in the students' social, emotional, and behavioral competence are sustained in subsequent grades, and whether they lead to enhanced academic achievement and reduction of conduct disorders.

Another strength of the current study is very high intervention implementation integrity. Because research staff co-led the Dinosaur curriculum with teachers almost all classrooms received a full dose of intervention delivered using consistent implementation standards. This allowed for an accurate evaluation of the intervention when it is delivered with integrity and with 'full strength,' as intended by the developers. Further research is now needed to conduct an effectiveness trial where the program is evaluated under 'real world' conditions without the research support and careful monitoring that was offered in the current project. It remains to be seen what level of technical support teachers will need to implement the program effectively on their own after receiving the training.

Another limitation of the study is that we cannot determine whether the child behavior improvements occurred outside the classroom environment and whether they generalized to the home environment. Further study should include parent report of behavior change as well.

Children between the ages of 3 and 6 years are developing social and emotional skills at a pace exceeding any other later stage of life. Their behavior is still flexible and their cognitive processes, which vacillate between fantasy and reality, are highly malleable and receptive to adult socialization processes. Teaching and learning that happens in this age range is crucial because it sets either a firm or a fragile foundation for later relationships and socialization, learning, and attitudes toward school. Early childhood learning can be seriously threatened by social, emotional impairments and conduct problems. Intervening early to remediate these difficulties may have lifelong benefits for enhancing children's later success. Research, such as this, that provides empirical information about ways to change these key variables can provide the basis for early intervention plans for schools that will help to benefit children at high risk for later school difficulties. In other words, focusing on promoting social and emotional learning and preventing conduct problems in these early years may put children on a trajectory leading to a cycle of lasting improvements in school achievement and mental health.

Acknowledgements

This research was supported by the NIH/NINR grant # 5 R01 NR001075 and NIH/NIDA grant # 5 R01 DA 12881.

Correspondence to

Carolyn Webster-Stratton, University of Washington, School of Nursing, Parenting Clinic, 1107 NE 45th

St. Suite 305, Seattle, WA 98105, USA; Email: cws@u.washington.edu

References

Arnold, D.H., Griffith, J.R., Ortiz, C., & Stowe, R.M. (1998). Day care interactions and teacher perceptions as a function of teacher and child ethnic group. *Journal of Research in Childhood Education*, *12*, 143–154.

Arnold, D.H., Ortiz, C., Curry, J.C., Stowe, R.M., Goldstein, N.E., Fisher, P.H., et al. (1999). Promoting academic success and preventing disruptive behavior disorders through community partnership. *Journal of Community Psychology*, *27*, 589–598.

Bandura, A. (1989). Regulation of cognitive processes through perceived self-efficacy. *Developmental Psychology*, *25*, 729–735.

Barrera, M., Biglan, A., Taylor, T.K., Gunn, B., Smolkowski, K., Black, C., et al. (2002). Early elementary school intervention to reduce conduct problems: A randomized trial with Hispanic and non-Hispanic children. *Prevention Science*, *3*, 83–94.

Birch, S.H., & Ladd, G.W. (1997). The teacher–child relationship and children's early school adjustment. *Journal of School Psychology*, *35*, 61–79.

Burchinal, M.R., Roberts, J.E., Hooper, S., & Zeisel, S.A. (2000). Cumulative risk and early cognitive development: A comparison of statistical risk models. *Developmental Psychology*, *36*, 793–807.

Burton, C.B. (1992). Defining family-centered early education: Beliefs of public school, child care, and Head Start teachers. *Early education and development*, *3*, 45–59.

Campbell, S.B., & Ewing, L.J. (1990). Follow-up of hard-to-manage preschoolers: Adjustment at age 9 and predictors of continuing symptoms. *Journal of Child Psychology and Psychiatry*, *31*, 871–889.

Carr, E.G., Taylor, J.G., & Robinson, S. (1991). The effects of severe behavior problems in children on the teaching behavior of adults. *Journal of Applied Behavior Analysis*, *24*, 523–535.

Conduct Problems Prevention Research Group. (2002). Evaluation of the first 3 years of the Fast Track prevention trial with children at high risk for adolescent conduct problems. *Journal of Abnormal Child Psychology*, *30*, 3–59.

Dodge, K.A., & Price, J.M. (1994). On the relation between social information processing and socially competent behavior in early school-aged children. *Child Development*, *65*, 1385–1397.

Domitrovich, C., Cortes, R.C., & Greenberg, M. (2006). *Improving young children's social and emotional competence: A randomized trial of the preschool PATHS curriculum*. University Park, PA: Pennsylvania State University.

Domitrovich, C.E., Cortes, R.C., & Greenberg, M.T. (in press). Improving young children's social and emotional competence: A randomized trial of the Preschool PATHS Curriculum. *Journal of Primary Prevention*.

Epstein, A. (1992). School and family partnerships. In M. Alkin (Ed.), *Encyclopedia of educational research* (pp. 1139–1151). New York: MacMillan.

Greenberg, M.T., Kusche, C.A., Cook, E.T., & Quamma, J.P. (1995). Promoting emotional competence in school-aged children: The effects of the PATHS curriculum. Special issue: Emotions in developmental psychopathology. *Development and Psychopathology*, *7*, 117–136.

Grolnick, W.S., & Slowiaczek, M.L. (1994). Parents' involvement in children's schooling: A multidimensional conceptualization and motivational model. *Child Development*, *65*, 237–252.

Grossman, D.C., Neckerman, H.J., Koepsell, T.D., Liu, P., Asher, K.N., Beland, K., et al. (1997). The effectiveness of a violence prevention curriculum among children in elementary school. *Journal of American Medical Association*, *277*, 1605–1611.

Hawkins, J.D., Catalano, R.F., Kosterman, R., Abbott, R., & Hill, K.G. (1999). Preventing adolescent health-risk behaviors by strengthening protection during childhood. *Archives of Pediatrics and Adolescent Medicine*, *153*, 226–234.

Henderson, A., & Berla, N. (1994). *A new generation of evidence: The family is critical to student achievement*. Columbia, MD: National Committee for Citizens in Education.

Huffman, L.C., Mehlinger, S.L., & Kerivan, A.S. (2001). *Risk factors for academic and behavioral problems at the beginning of school. In Off to a good start: Research on the risk factors for early school problems and selected federal policies affecting children's social and emotional development and their readiness for school*. Chapel Hill, NC: Frank Porter Graham Child Development Center.

Keenan, K., Shaw, D.S., Walsh, B., Delliquadri, E., & Giovannelli, J. (1997). DSM-III-R disorders in preschool children from low-income families. *Journal of American Academy of Child and Adolescent Psychiatry*, *36*, 620–627.

Kellam, S., Werthamer-Larsson, L., Dolan, L., Brown, C., Mayer, L., Rebok, G., et al. (1991). Developmental epidemiology based on preventive trials: Baseline modeling of early target behaviors and depressive symptoms. *American Journal of Community Psychology*, *19*, 563–584.

Kellam, S.G., Ling, X., Merisca, R., Brown, C.H., & Ialongo, N. (1998). The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Development and Psychopathology*, *10*, 165–185.

Kellam, S.G., & Rebok, G.W. (1992). Building developmental and etiological theory through epidemiologically based preventive trials. In J. McCord & R.E. Tremblay (Eds.), *Preventing antisocial behavior: Interventions from birth through adolescence* (pp. 139–161). New York: Guilford Press.

Keogh, B.K. (1992). Temperament and teachers' views of teachability. In W. Carey & S. McDevitt (Eds.), *Prevention and early intervention: Individual differences as risk factors for the mental health of children* (pp. 246–254). New York: Bruner/Mazel.

Kupersmidt, J.B., Bryant, D., & Willoughby, M. (2000). Prevalence of aggressive behaviors among preschoolers in Head Start and community child care programs. *Behavioral Disorders*, *26*, 42–52.

Ladd, G.W., Birch, S., & Buhs, E. (1999). Children's social and scholastic lives in kindergarten: Related

spheres of influence? *Child Development*, *70*, 1373–1400.

Ladd, G.W., Kochenderfer, B.J., & Coleman, C. (1997). Classroom peer acceptance, friendship, and victimization: Distinct relational systems that contribute uniquely to children's school adjustment. *Child Development*, *68*, 1181–1197.

LaFreniere, P.J., Dumas, J., Dubeau, D., & Capuano, F. (1992). The development and validation of the preschool socio-affective profile, psychological assessment. *Journal of Consulting and Clinical Psychology*, *4*, 442–450.

Lopez, M.L., Tarullo, L.B., Forness, S.R., & Boyce, C.A. (2000). Early identification and intervention: Head Start's response to mental health challenges. *Early Education and Development*, *11*, 265–282.

Lynch, K.B., Geller, S.R., & Schmidt, M.G. (2004). Multi-year evaluation of the effectiveness of a resilience-based prevention program for young children. *The Journal of Primary Prevention*, *24*, 3353–3353.

Milliken, G.A., & Johnson, D.E. (1984). *Analysis of messy data*. Volume I: Designed Experiments (vol. 1). New York: Van Nostrand Reinhold Company .

Moffitt, T.E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674–701.

Muthen, B., & Curran, P.J. (1997). General growth modeling in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods*, *2*, 371–402.

Normandeau, S., & Guay, F. (1998). Preschool behavior and first-grade school achievement: The mediational role of cognitive self-control. *Journal of Educational Psychology*, *90*, 111–121.

Offord, D.R., Alder, R.J., & Boyle, M.H. (1986). Prevalence and sociodemographic correlates of conduct disorder. *The American Journal of Social Psychiatry*, *6*, 272–278.

Patterson, G., Reid, J., & Dishion, T. (1992). *Antisocial boys: A social interactional approach (vol. 4)*. Eugene, OR: Castalia Publishing.

Phillips, D., Voran, M., Kisker, E., Howes, C., & Whitebrook, M. (1994). Child care for children in poverty: Opportunity or inequity? *Child Development*, *65*, 472–492.

Pianta, R.C., & Walsh, D.J. (1998). Applying the construct of resilience in schools: Cautions from a developmental systems perspective. *School Psychology Review*, *27*, 407–417.

Raver, C.C., & Zigler, E.F. (1997). Social competence: An untapped dimension in evaluating Head Start's success. *Early Childhood Research Quarterly*, *12*, 363–385.

Reid, J.B., Eddy, J.M., Fetrow, R.A., & Stoolmiller, M. (1999). Description and immediate impacts of a preventive intervention for conduct problems. *American Journal of Community Psychology*, *27*, 483–517.

Reid, M.J., Webster-Stratton, C., & Hammond, M. (in press). Enhancing a classroom social competence and problem-solving curriculum by offering parent training to families of moderate-to-high-risk elementary school children. *Journal of Clinical Child and Adolescent Psychology*.

Robinson, E.A., Eyberg, S.M., & Ross, A.W. (1980). The standardization of an inventory of child conduct problem behaviors. *Journal of Clinical Child Psychology*, *9*, 22–28.

Rubin, K.H., & Krasnor, L.R. (1986). Social-cognitive and social behavioral perspectives on problem-solving. In M. Perlmutter (Ed.), *Cognitive perspectives on children's social and behavioral development. The Minnesota Symposia on Child Psychology* (vol. 18, pp. 1–68). Hillsdale, NJ: Lawrence Erlbaum Associates.

Shure, M.B. (1997). Interpersonal cognitive problem solving: Primary prevention of early high-risk behaviors in the preschool and primary years. In G.W. Albee & T.P. Gullotta (Eds.), *Primary prevention works* (pp. 167–188). Thousand Oaks, CA: Sage.

Skinner, E.A., Zimmer-Gembeck, M.J., & Connel, J.P. (1998). Individual differences and the development of perceived control. *Monographs of the Society for Research in Child Development*, *63*, (2–3, Serial No. 254).

Spivak, G., & Shure, M.B. (1985). ICPS and beyond: Centripetal and centrifugal forces. *American Journal of Community Psychology*, *13*, 226–243.

Stage, S.A., & Quiroz, D.R. (1997). A meta-analysis of interventions to decrease disruptive classroom behavior in public education settings. *School Psychology Review*, *26*, 333–368.

Stoolmiller, M., Eddy, J.M., & Reid, J.B. (2000). Detecting and describing preventive intervention effects in a universal school-based randomized trial targeting delinquent and violent behavior. *Journal of Consulting and Clinical Psychology*, *68*, 296–306.

Tapp, J., Wehby, J.H., & Ellis, D. (1995). A multiple option observation system for experimental studies: MOOSES. *Behavior Research Methods, Instruments, and Computers*, *27*, 25–31.

Tremblay, R.E., Pagani, K.L., Masse, L.C., & Vitaro, F. (1995). A biomodal preventive intervention for disruptive kindergarten boys: Its impact through mid-adolescence. Special Section: Prediction and prevention of child and adolescent antisocial behavior. *Journal of Consulting and Clinical Psychology*, *63*, 560–568.

Tremblay, R.E., Vitaro, F., Bertrand, L., LeBlanc, M., Beauchesne, H., Boileau, H., et al. (1996). Parent and child training to prevent early onset of delinquency: The Montreal longitudinal-experimental study. In J. McCord & R.E. Tremblay (Eds.), *Preventing antisocial behavior: Interventions from birth through adolescence* (pp. 117–138). New York: Guilford.

Walker, H., Kavanagh, K., Stiller, B., Golly, A., Severson, H.H., & Feil, E.G. (1998a). First step to success: An early intervention approach for preventing school antisocial behavior. *Journal of Emotional and Behavioral Disorders*, *6*, 66–80.

Walker, H.M., Kavanagh, K., Stiller, B., Golly, A., Severson, H.H., & Feil, E.G. (1998b). First Step to Success: An early intervention approach for preventing school antisocial behavior. *Journal of Emotional and Behavioral Disorders*, *6*, 66–80.

Webster-Stratton, C. (1990). *Wally Game: A problem-solving skills test*. Retrieved July 13, 2007, from http://www.son.washington.edu/centers/parenting-clinic/forms.asp.

Webster-Stratton, C. (2000). *How to promote social and academic competence in young children*. London: Sage Publications.

Webster-Stratton, C., & Hammond, M. (1997). Treating children with early-onset conduct problems: A comparison of child and parent training interventions. *Journal of Consulting and Clinical Psychology*, *65*, 93–109.

Webster-Stratton, C., & Hammond, M. (1998). Conduct problems and level of social competence in Head Start children: Prevalence, pervasiveness and associated risk factors. *Clinical Child Psychology and Family Psychology Review*, *1*, 101–124.

Webster-Stratton, C., & Lindsay, D.W. (1999). Social competence and early-onset conduct problems: Issues in assessment. *Journal of Child Clinical Psychology*, *28*, 25–93.

Webster-Stratton, C., & Reid, M.J. (2003). Treating conduct problems and strengthening social emotional competence in young children (ages 4–8 years): The Dina Dinosaur treatment program. *Journal of Emotional and Behavioral Disorders*, *11*, 130–143.

Webster-Stratton, C., & Reid, M.J. (2004). Strengthening social and emotional competence in young children – the foundation for early school readiness and success: Incredible Years Classroom Social Skills and Problem-Solving Curriculum. *Journal of Infants and Young Children*, *17*, 196–113.

Webster-Stratton, C., & Reid, M.J. (2008). Strengthening social and emotional competence in socio-economically disadvantaged young children: Preschool and kindergarten school-based curricula. In W.H. Brown, S.L. Odom, & S.R. McConnell (Eds.), *Social competence of young children: Risk, disability, and intervention* (pp. 185–203). Baltimore, MD: Paul H. Brookes Publishing Co.

Webster-Stratton, C., Reid, M.J., & Hammond, M. (2001). Social skills and problem solving training for children with early-onset conduct problems: Who benefits? *Journal of Child Psychology and Psychiatry*, *42*, 943–952.

Webster-Stratton, C., Reid, M.J., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology*, *33*, 105–124.

Webster-Stratton, C., & Taylor, T. (2001). Nipping early risk factors in the bud: Preventing substance abuse, delinquency, and violence in adolescence through interventions targeted at young children (ages 0–8 years). *Prevention Science*, *2*, 165–192.

Werner, E.E. (1999). How children become resilient: Observations and cautions. In N. Henderson, B. Benard, & N. Sharp-Light (Eds.), *Resiliency in action: Practical ideas for overcoming risks and building strengths in youth, families, and communities* (pp. 115–134). Gorham, ME: Resiliency in Action.

Werthamer-Larsson, L., Kellam, S.G., & Ovesen-McGregor, K.E. (1990). Teacher interview: Teacher observation of classroom adaptation – Revised (TOCA-R). In S.G. Kellam (Ed.), *Johns Hopkins Prevention Center training manual*. Baltimore, MD: Johns Hopkins University.

West, J., Denton, K., & Reaney, L.M. (2001). *The kindergarten year: Findings from the Early Childhood Longitudinal Study, kindergarten class of 1998–1999 (Publication No. NCES2001-023)*. Washington, DC: Department of Education, National Center for Education Statistics.

Manuscript accepted 25 October 2007