Initial Impact of the Fast Track Prevention Trial for Conduct Problems: II. Classroom Effects

Conduct Problems Prevention Research Group

Abstract

This study examined the effectiveness of the universal component of the Fast Track prevention model: the PATHS (Promoting Alternative THinking Strategies) curriculum and teacher consultation. This randomized clinical trial involved 198 intervention and 180 comparison classrooms from neighborhoods with greater than average crime in 4 U.S. locations. In the intervention schools, Grade 1 teachers delivered a 57-lesson social competence intervention focused on self-control, emotional awareness, peer relations, and problem solving. Findings indicated significant effects on peer ratings of aggression and hyperactive–disruptive behavior and observer ratings of classroom atmosphere. Quality of implementation predicted variation in assessments of classroom functioning. The results are discussed in terms of both the efficacy of universal, school-based prevention models and the need to examine comprehensive, multyear programs.

As models of prevention have become better integrated with research on the development of antisocial behavior, the need for addressing the transactional nature of risk factors has emerged. As noted by the Institute of Medicine (1994), “the transactional interaction between the individual child and his or her environment over time is the ecological crucible in which the pathways to the development of a range of positive or negative childhood and adult outcomes are forged” (p. 181). A critical factor in the early development of antisocial behavior for many high-risk children is that they attend schools that have a high density of other high-risk children like themselves and thus present the classroom teacher, who often has fewer resources than do teachers in less high-risk schools, with additional challenges to classroom order. This volatile combination of high-risk child and high-risk classroom has negatively synergistic consequences for both the child and the classroom. The disruptive behavior of high-risk children undermines the social and academic environment for other children. Reciprocally, the high-risk child encounters greater stimulation to disruptive behavior and a less nurturant learning atmosphere, rather than being given the kind of well-structured and nonprovocative classroom that would be needed to compensate for his or her emotional, social, and cognitive deficits.

The primary goal of the Fast Track model is to integrate the provision of universal (all children) and selective (children at some risk) services into a comprehensive model that involves the child, school, family, and community (Institute of Medicine, 1994). Fast Track was designed to provide two levels of child intervention simultaneously during the elementary school years.

Correspondence concerning this article should be addressed to Robert J. McMahon, Seattle Fast Track, University of Washington, 146 North Canal Street, Suite 111, Seattle, Washington 98103. For additional information concerning Fast Track, see http://www.fasttrackproject.org.

Members of the Conduct Problems Prevention Research Group, in alphabetical order, include Karen L. Bierman, Department of Psychology, Pennsylvania State University; John D. Coie, Department of Psychology, Duke University; Kenneth A. Dodge, Sanford Institute of Public Policy, Duke University; Mark T. Greenberg, Department of Human Development and Family Studies, Pennsylvania State University; John E. Lochman, Department of Psychology, University of Alabama; Robert J. McMahon, Department of Psychology, University of Washington; and Ellen E. Pinderhughes, Department of Psychology and Human Development, Vanderbilt University.
Through a multistage screening process involving both teacher and parent reports during kindergarten, the 10% of children demonstrating the greatest degree of early conduct problems were selected for a series of interventions that included weekly parenting support classes, small-group social skills interventions, academic tutoring, and home visiting (Bierman, Greenberg, & Conduct Problems Prevention Research Group [CPPRG], 1996; McMahon, Slough, & CPPRG, 1996). Such interventions are believed to be necessary to both reduce risk factors and promote protective factors in children who are showing the “early starter” model of conduct problems (CPPRG, 1992; Loeber, 1990; Moffitt, 1993).

Simultaneous with the initiation of these interventions for the high-risk children and families, the universal intervention consisting of the Fast Track PATHS (Promoting Alternative THinking Strategies) curriculum and behavioral consultation was started in the classroom. There are two central reasons that integrated delivery of universal and selective interventions should provide an additive effect. First, it is unlikely that effects of the selective interventions with the children and families will generalize to the school and classroom setting without providing support for these new skills in the school (Kazdin, 1990, 1993). By providing similar skills, cues, and a common language in both the selective and universal interventions, teachers and other school staff are able to promote the generalization of skills to the classroom. Second, a universal intervention intended to promote the development of social competence in all children should lead to an improved classroom atmosphere that supports improved interpersonal relations for all students (Battistich, Solomon, Watson, Solomon, & Schaps, 1989; Elias et al., 1998). Reciprocally, more intensive intervention with the highest risk children in these same classrooms may serve to reduce their highly disruptive impact on the classrooms, thereby making it easier for the remaining children to respond to the universal intervention. It was hoped that intensive intervention for higher need students would serve to communicate to teachers that the program staff were committed to helping them with their most difficult problems, thus making teachers more open to participating in a universal intervention. In addition, Fast Track staff provided behavioral consultation to teachers regarding both the high-risk children and the remaining classroom students.

In this report, we consider the effectiveness of the universal intervention using a randomized clinical trial design. We examine its effects at the level of the classroom (Murray & Wolfinger, 1994) in altering the conditions of peer relations and classroom atmosphere through the use of multiple reporters (children, teachers, and observers). Although sets of schools, not classrooms, were the unit of randomization, the classroom is the unit of program delivery and also the level at which measures such as dosage and implementation can be assessed.

Social competence can be broadly conceived as the capacity to integrate cognition, affect, and behavior to achieve specified social tasks and positive developmental outcomes (Waters & Sroufe, 1983; Weissberg & Greenberg, 1998). Meta-analyses of universal prevention programs during the elementary school years indicate that such programs have shown significant and moderate effects on social–cognitive abilities and mild to moderate effects on children's social adjustment (Beelmann, Pfingsten, & Losel, 1994; Denham & Almeida, 1987; Durlak, 1995; Schneider, 1992). A number of other conclusions have been reached. First, traditional prevention models based on single skills (e.g., only social problem solving, self-control, empathy, or behavioral social skills training) have demonstrated less effectiveness than multimodal programs that integrate social problem solving, social skills (peer relations and self-control), and emotional understanding. Second, interventions need to be of significant duration and intensity to show effects (Weissberg & Elias, 1993). Third, there is a need to examine how both dosage and quality of implementation affect outcomes. The PATHS preventive intervention program is based on the ABCD (Affective–Behavioral–Cognitive–Dynamic) model of development (Greenberg & Kusche, 1993; Greenberg, Kusche, & Speltz, 1991), which places primary importance on the developmental integration of affect (and
emotion language), behavior, and cognitive understanding as they relate to social and emotional competence. A basic premise is that a child's coping, as reflected in his or her behavior and internal regulation, is a function of emotional awareness, affective–cognitive control, and social–cognitive understanding.

Between the ages of 5 and 7 years, children undergo a major developmental transformation that generally includes increases in cognitive skills, as well as changes in brain size and function (Kendler, 1963; Pennington & Welsh, 1995; White, 1965). This transition and accompanying developmental changes allow children to undertake major changes in responsibilities, independence, and social roles (Belsky & MacKinnon, 1994; Pianta, Steinberg, & Rollins, 1995). Thus, the relationships between affective understanding, cognition, and behavior are of crucial importance in socially competent action and healthy peer relations (Weissberg & Elias, 1993). Taking this into account, the PATHS curriculum model synthesizes the domains of self-control, emotional awareness and understanding, peer-related social skills, and social problem solving to increase social and emotional competence. In addition to a person-oriented model that focuses primarily on developmental integration, the intervention model incorporates an ecobehavioral systems orientation (Weissberg, Caplan, & Sivo, 1989), which places primacy on the manner in which the teacher uses the curriculum model. That is, program impact may be the greatest when teachers generalize support for curriculum-based skills during the day and build a healthy classroom atmosphere that supports the child's skill use and internalization of skills. It is presumed that improvements in social competence can be a function of changes in the child, changes in the ecology, and their interaction.

Previous field trials with different versions of the PATHS curriculum with both deaf (Greenberg & Kusche, 1993, 1998) and regular- and special-needs children (Greenberg & Kusche, 1997; Greenberg, Kusche, Cook, & Quamma, 1995) have shown that the use of the PATHS curriculum is associated with significantly more mature social cognitions, including better understanding of social problems, higher percentages of effective solutions, lower percentages of aggressive and passive solutions, and increased recognition of emotions. In all three samples studied to date, teachers reported significant improvements in behaviors targeted by PATHS (self-control, emotional understanding, and use of more effective conflict resolution skills). In special-needs children, PATHS also led to significant decreases in self-reported sadness, decreases in teacher reports of internalizing problems, and increases in teacher reports of social competence.

Following from the model, the Fast Track universal prevention strategy operated under the following assumptions. First, the school environment is a fundamental ecology and one that can be a central locus of change (Elias et al., 1998). Second, given the developmental issues facing children in middle childhood, self-control, emotional understanding and awareness, peer-related social skills, and problem solving would all appear to be prime targets for preventive intervention during the elementary school years. Following in the tradition of numerous recent school-based preventive interventions (Elias, Gara, Schuyler, Branden-Muller, & Sayette, 1991; Hawkins & Weiss, 1985; Weissberg & Elias, 1993), the PATHS curriculum was designed to be delivered by teachers with support from project staff, to be taught on a regular basis throughout most of the school year, and to provide daily activities to promote generalization. Using an ecobehavioral systems model, Fast Track staff also consulted with the school principal to bring the philosophy of PATHS to the entire school, resulting in various efforts (on a school-by-school basis), such as placing PATHS posters in school hallways, implementing new school behavior guidelines, and painting problem-solving “stop-lights” on school playgrounds.

The present study advances the knowledge base on school-based universal prevention programs in four ways. First, it examines the efficacy of a universal program in the context of
a comprehensive model that includes targeted intervention. Second, it uses a model in which emotional understanding and regulation are a central focus of intervention. Third, because of its large sample size, it is the first elementary school study in which analyses can be accomplished at the classroom level. This provides two innovations: There is sufficient power to assess measures of the classroom atmosphere, and hierarchical modeling can be used to account for interdependency among scores within classrooms (Bryk & Raudenbush, 1992). Finally, the study can assess the effects of curriculum effectiveness in sites with different populations.

We hypothesized that the introduction of the Fast Track universal prevention program would lead to differences in peer-rated aggression and hyperactive–disruptive behavior, teacher-rated aggression and conduct problems, and observer-rated classroom atmosphere. Further, we hypothesized that both dosage and quality of implementation would predict variation in the above outcomes.

**Method**

**Participants**

The participating schools were selected from four areas of the country, each representing a different cross-section of the U.S. population: (a) Durham, North Carolina, a small city with a large low- to middle-SES (social economic status), primarily African American school population; (b) Nashville, Tennessee, a moderate-sized city with a mix of low- to middle-SES, African American and European American families; (c) Seattle, Washington, a moderate-sized city with a low- to middle-SES, ethnically diverse population; and (d) central Pennsylvania, a mostly rural area with low- to middle-SES, European American families. In the Seattle site, both an urban and a suburban district were chosen; in rural Pennsylvania, three small school districts participated.

Within each site, approximately 12 elementary schools in high-risk neighborhoods (or towns in the case of rural Pennsylvania) were invited to be involved in the Fast Track intervention model. High-risk status was defined from estimated rates of delinquency and juvenile arrest in the neighborhoods. The full Fast Track prevention model was initially described to principals and teachers at each school. After faculty discussion, school-based decisions were made regarding participation. Schools were aware that once they decided to participate, they had a 50% chance of being randomized as a comparison school. After obtaining consensus to participate, we divided schools into matched “sets,” which were equivalent on school sizes, achievement levels, poverty, and ethnic/racial diversity. These sets of schools were then randomly assigned to intervention and control groups. The intervention was conducted in 3 successive years with three cohorts of first graders. There were 198 intervention classrooms and 180 matched comparison classrooms across the three cohorts. The total number of children for whom we obtained consent in these classrooms was 7,560; 845 of these children were either high-risk intervention or high-risk control children. Thus, in analyses that do not include the high-risk children, the sample size was 6,715.

Although there were substantial differences between sites in the degree of risk shown by their respective school locations, there was considerable risk in the typical school selected for this

---

1. Because Nashville schools were substantially larger in size, fewer schools were picked at this site.
2. These classrooms were regular-education classrooms in which PATHS was taught or which were designated as control classrooms. In the accompanying report (CPPRG, 1999a), the number of classrooms differs because that report includes those classrooms (regular and special education) in which high-risk children were present. There were 891 intervention and control children in the high-risk sample described in the accompanying article (CPPRG, 1999a). However, high-risk children who were in a special-education classroom or who had moved to a different school were not included in the high-risk sample in this article.
intervention. Table 1 presents information by site and condition for the variables of poverty, ethnicity, and achievement. The percentage of children receiving free or reduced lunch was 55% (ranging from 39% in rural Pennsylvania to 80% in Durham). The mean percentage of ethnic minority children (primarily African American) attending the schools was 49% (ranging from 1% in rural Pennsylvania to 90% in Durham). The mean reading percentile across three of the sites (excluding Durham\(^4\)) was the 47th percentile (ranging from the 33rd percentile in Nashville to the 57th percentile in rural Pennsylvania). A series of analyses of variance indicated no significant differences between intervention and control schools on the percentage of children who received free or reduced lunch, percentage of ethnic minority children, or academic achievement scores. Approximately 25% of all teachers were of an ethnic minority, and this did not vary by intervention status; 98% of all teachers were female.

The Intervention

The Fast Track PATHS curriculum in Grade 1 contained 57 lessons, approximately 80% of which were drawn from the published version of the curriculum (Kusche & Greenberg, 1994). Previously designed for special-needs populations, this multiyear (Grades 1–5) classroom prevention program was adapted to fit the needs of regular-education students in high-risk schools for the Fast Track program.

In the Fast Track version of PATHS, about 40% of the lessons focus on skills related to understanding and communicating emotions. As a basic step toward self-control, PATHS teaches young children to recognize the internal and external cues of affect and to label them with appropriate terms. In a series of lessons, “feeling” words are identified, along with descriptions of the sorts of situations that may elicit the feeling, descriptions of the external cues to recognize that feeling in others, and the internal cues to identify that feeling in oneself. Additional lessons help children understand the difference between feelings and behaviors. Appropriate and inappropriate behavioral responses are discussed. The teaching of feelings involves a generalization technique (“Feeling Faces”) used to promote the student's use of new knowledge and skills throughout the classroom day. After each emotion concept is introduced, the children personalize their own Feeling Face for that affect; these faces are small cards with idealized line drawings of the affect that are kept on the students’ desks. The faces allow the children to communicate their feelings with minimal difficulty throughout the day, and they facilitate the children's understanding about how feelings change. Teachers have their own set of Feeling Faces and use them as models for their students. Teachers are encouraged to promote generalization at the beginning and at the end of the day, after recesses, and after lunch time by suggesting that the children evaluate how they feel and display the appropriate face(s).

Another 30% of the lessons focus on skills related to the increase of positive social behavior (e.g., social participation, prosocial behavior, and communication skills). Lessons address making and sustaining friendships, using good manners, taking turns and sharing in games, expressing one's viewpoint, and listening to others. In addition, positive behaviors are elicited and reinforced during each lesson. For example, during each lesson, one child serves as the teacher’s helper (the “PATHS Kid of the Day”), and this child receives compliments from classmates, the teacher, and him- or herself.

Finally, about 30% of the lessons focus on self-control and other steps in social problem solving. The development of self-control, affective awareness and communication, and beginning problem-solving skills are integrated with the introduction of the Control Signals Poster (CSP). The CSP is designed like a traffic signal and is a modified version of the Stop

\(^4\)North Carolina schools have initiated their own achievement testing system that presently has no percentile scores and thus cannot be compared with other sites.
Light used in the Yale–New Haven Middle-School Social Problem Solving Program (Weissberg, Caplan, & Bennetto, 1988). The CSP has a red light to signal “Stop—Calm Down,” a yellow light for “Go Slow—Think,” a green light to signal “Go—Try My Plan,” and at the bottom the words “Evaluate—How Did My Plan Work?” Children are taught that when they are in a situation that they find upsetting or frustrating (such as a playground conflict or difficult work situations), the first step toward effective problem solving is to “go to the red light” in order to stop and think before they act. Before they take an action, they should “take a long, deep breath,” calm down, and “say the problem and how they feel.” Once the problem is identified, they can move to the yellow light to “Make a Plan,” considering first the possible solutions and then selecting the best option. The next step is to “Try the Plan” at the green light and evaluate the effectiveness of that plan, recycling through the problem-solving steps if the plan proves ineffective. In addition to scripted lessons teaching children these steps to problem solving, teachers are taught how to hold classroom problem-solving meetings to help children use the problem-solving steps to address current classroom problems.

Typically, skill concepts are presented through direct instruction, discussion, modeling stories, or video presentations. Discussion and role-playing activities follow, giving children a chance to practice the skill and teachers a chance to monitor the level of understanding and skill attained by each class. Although a standard script describes each lesson, teachers are encouraged to adjust the level of presentation and amount of practice as dictated by the responsivity and developmental level of each class.

Although the lessons form an important part of the PATHS program, teachers are also encouraged to generalize their use of PATHS concepts across the school day and to other settings of the school outside the classroom. In particular, teachers are encouraged to help children identify their feelings, communicate clearly with others, use self-control strategies, and apply the three steps of problem solving as frustrations, challenges, and interpersonal problems occur at school. Each classroom has a mailbox in which students can submit problems or concerns that are then discussed in problem-solving meetings.

To generalize concepts to the home situation, the curriculum includes frequent parent updates on curriculum content and suggestions for ways parents can promote their children’s growing competence. Regular homework activities are designed to help children engage their parents in cooperative activities, such as completing drawings or sharing stories related to curriculum components.

**Teacher Training**

The intervention teachers attended a 2.5-day training workshop and received weekly consultation and observation from project staff. The PATHS lessons were taught approximately two to three times per week, with each lesson lasting 20 to 30 min, from mid-September to May. Teachers were either paid for their extra preparation and consultation time or received continuing education credit for their participation. The weekly consultations were intended to enhance the quality of implementation through modeling, coaching, and providing ongoing feedback regarding program delivery. Fast Track staff (termed educational coordinators [ECs]) also provided general feedback on classroom and behavior management. The ECs were experienced teachers hired by the project. They spent an average of 1 to 1.5 hr per week in each classroom observing, demonstrating, or team teaching the PATHS lessons. They also met individually or in groups with teachers on a regular basis.

**Measures of Intervention Dosage and Quality of Implementation**

To assess the amount of dosage, we had teachers report weekly to their assigned ECs on the lessons they had presented. The mean number of lessons taught by Grade 1 teachers was 48.2
Fidelity was assessed through monthly ratings of quality of implementation made by the ECs on the basis of their direct observation of teacher instruction. For all three cohorts, there were four 4-point Likert-scale ratings (ranging from low-skilled to highly skilled performance). The four ratings were (a) quality of teaching of PATHS concepts, (b) modeling of PATHS concepts throughout the day, (c) quality of classroom management (during PATHS lessons), and (d) openness to consultation from the EC. Data aggregation across the year indicated that these measures were highly consistent over time ($\alpha = .88$); thus, a mean score for each rating was computed for each teacher. The first three ratings were moderately correlated (range = .50–.70), and the fourth rating correlated less strongly with the other three (range = .30–.35). The four ratings were kept separate for analyses because they are conceptually distinct domains. Ratings were not available for 14 of the intervention classrooms (7%) because either the teacher changed during the year (e.g., medical/pregnancy leaves) or, in six cases, teachers did not allow ECs to observe their classrooms regularly. It was not feasible to conduct adequate interrater reliability because of both the size of the sample (198 intervention classrooms) and the intensity of data collection; these ratings required knowledge of the curriculum and observations both during and outside of the lesson time and during consultation meetings. In a previous, smaller study of PATHS, adequate reliability ($r = .63–.85$) of these ratings had been demonstrated (Cheney, Greenberg, & Kusche, 1991).

**Measures of Outcome**

Outcome measures were derived from three independent sources. The measures presented here are the only measures available on the entire classroom population. Measures that focused on the targeted population were more extensive and are presented in a companion article (CPPRG, 1999a). First, at the beginning and end of Grade 1, teachers were individually interviewed regarding the behavior of each child in their class using the Teacher Observation of Classroom Adaptation—Revised (TOCA–R; Werthamer-Larsson, Kellam, & Wheeler, 1991) and the Social Health Profile (SHP; CPPRG, 1999b). Second, in the spring of Grade 1, sociometric assessments were collected to assess peer aggression, peer hyperactivity–disruptiveness, and peer social status. Third, also in the spring of Grade 1, observers who were unaware of the status of the schools assessed the quality of the classroom atmosphere using a 10-item scale. Neither the sociometric nor observational measures were used as a pretest assessment. Both measures would have required a month of school adaptation and 2 months to collect; this would have delayed the program, which was designed to begin as early in the school year as possible.

**Teacher report**—Grade 1 teachers completed the TOCA–R and the SHP in the fall and spring during a structured interview. On the TOCA–R, teachers rated the behavior of each child in the class on items using 6-point Likert scales ranging from 0 (*almost never*) to 5 (*almost always*). The interview covering the entire class required about 90 min to complete, and teachers were reimbursed for their time. Two internally consistent factors from the TOCA–R were used in the analyses. The Cognitive Concentration scale (12 items; $\alpha = .97$) assessed concentration, attention, and work completion. The Authority Acceptance scale (10 items; $\alpha = .93$) assessed oppositional and conduct problem behaviors (e.g., takes property, breaks rules, teases, is disobedient). The SHP includes 9 items describing prosocial behaviors and emotion regulation. Items were rated on a 6-point scale and were summed to create a total score for social competence ($\alpha = .87$). Finally, a teacher-rated peer-liking measure asked teachers to rate how liked each child was by his or her peers. All scores were standardized across the sample, and then a mean score was computed for each class. Missing data at either pre- or posttest reduced the sample by 5%, to 357 classrooms. The teacher-rated measure of peer liking was significantly related to peer sociometric ratings of aggression ($r = .30$, $p < .001$) and prosocial behavior ($r = .34$, $p < .001$).

*J Consult Clin Psychol.* Author manuscript; available in PMC 2009 October 14.
**Peer nominations**—Peer reports of aggression, hyperactive–disruptive behavior, prosocial behavior, and likability (“most liked”) were assessed using individual sociometric interviews, which were conducted with each child whose parent had provided consent. The interviewer provided the child with a class roster and read each of the names on the roster aloud in order to ensure that the child was familiar with his or her classmates. Children were then read a series of behavioral descriptors and asked to nominate as many children as they wanted to fit each descriptor. The descriptor for aggression was “Some kids start fights, say mean things and hit other kids. Who are the kids who start fights and say mean things?” The descriptor for hyperactive–disruptive behavior was “Some kids get out of their seat a lot, do strange things, and make a lot of noise. They bother people who are trying to work. Who are the kids that get out of their seats and bother people?” The descriptor for prosocial behavior was “Some kids are really good to have in your class because they cooperate, help others, and share. They let other kids have a turn. Who are the kids who cooperate, help, and share?” The descriptor for “most liked” was “Who are the kids who you like most?” The sociometric scores for each classroom were first corrected for the number of raters. Then the classroom mean scores for all children were standardized across the entire sample within each cohort. Eight classrooms (approximately 2%) did not complete sociometric assessments because fewer than 70% of parents consented to the assessment.

**Observer ratings**—To assess the intervention effects on the high-risk children selected for the additional interventions, observers watched these children and their matched controls twice on different days for 30 min in their classroom (Wehby, Dodge, Valente, & CPPRG, 1993). Using a computer-assisted rating program (ASKER; Tapp & Fiel, 1991), the observers then rated ten 5-point items summarizing the classroom atmosphere during each 30-min observation. These ratings focused on the atmosphere of the entire classroom and ranged from 1 (low) to 5 (high), with behavioral descriptions at the ends and at the midpoint. The classroom atmosphere ratings were derived from the Classroom Rating Form (Solomon, Watson, Delucchi, Schaps, & Battistich, 1988) and assessed the classroom’s (a) level of disruption during academic time, (b) ability to handle classroom transitions, (c) ability to follow rules, (d) level of cooperation, (e) use of problem solving during conflict or need, (f) ability to express feelings appropriately, (g) level of interest and enthusiasm, (h) ability to stay focused and on task, (i) responsiveness to individual student's needs and feelings, and (j) level of criticism versus supportiveness.

Although the number of ratings per class was dependent on the number of “high-risk” children in a particular classroom, there was no difference in the number of times intervention versus control classrooms were rated; the mean number of observations was 4.4 per class. These ratings were completed in 311 classrooms; there were no high-risk intervention or control children in 77 classrooms (19%), and thus observations were not conducted in these classrooms. Six observers were trained at each site each year for approximately 6 weeks using videotapes and in situ practice sessions. To assess reliability, we randomly chose approximately 12% of the sessions to be coded by a second observer. The kappa coefficients ranged from .62 to .81 for the 10 ratings. A mean score was derived across the number of rating occasions in each classroom. The 10 ratings had high internal consistency (α = .92); thus, a mean score for each rating and a total classroom atmosphere score were computed for each classroom. The total classroom atmosphere score (a higher score indicates poorer atmosphere) was significantly related to higher teacher ratings of problem behavior (using the Authority Acceptance scale; \( r = .30, p < .001 \)), poorer Cognitive Concentration (\( r = .26, p < .01 \)), low social competence (\( r = .28, p < .001 \)), and higher peer-rated sociometric aggression (\( r = .21, p < .01 \)).

*J Consult Clin Psychol. Author manuscript; available in PMC 2009 October 14.*
Results

Statistical Models for Outcome Analyses

For the teacher ratings and sociometric data, we used hierarchical linear modeling (HLM) models, with classroom as the second level in a mixed model design. Although random assignment occurred at the level of sets of schools (because the universal intervention was delivered at the classroom level), classroom was included in all analyses as a random factor using HLM (Bryk & Raudenbush, 1992). The rationale here is that classrooms, especially teachers and specific peers, may exert a sizable effect on individual scores; in fact, classrooms within schools might vary as much as classrooms across schools (Kellam, Ling, Merisca, Brown, & Ialongo, 1998). This effect might occur through influence by the teacher, classroom climate and rules, or peer group. In this case, the variation in a dependent variable across children after the effect of intervention is statistically removed, is not random, and thus violates the assumption of independence. If classrooms do indeed exert a strong effect on scores (i.e., the intraclass correlation is high), the individual-level analysis will overestimate the statistical significance of the intervention effect, whereas the classroom-level analysis will take the clustering into account to yield a lower level of statistical confidence (significance).

Furthermore, the use of HLM to compute effect sizes with classroom as the unit will lead to modeling of true score relationships (the constant Level 1 measurement error is estimated and reported as sigma squared) in contrast with observed score relationships for the analysis at the individual child level only. Measurement error acts to attenuate regression coefficients, making effects appear smaller than they really are. With HLM, the coefficients are disattenuated because measurement error has been modeled as a separate parameter. The result is that, with HLM, estimates of the intervention effect will tend to be more accurate than those with individual-level analyses only.

Because the high-risk intervention children also received numerous other interventions that may have affected their scores, we conducted all analyses twice, both with and without the high-risk intervention and control children included in the classroom means. Although no substantive differences were found between these two types of analyses, we present the more conservative data that excluded the high-risk intervention and control children (the highest 10% of each classroom on behavioral risk). In some cases, teachers contributed multiple cases across cohorts; in these analyses, each cohort was treated as a new case.

HLM analyses: Sociometrics—Analyses were conducted using the HLM 4.02 software package. Full maximum likelihood estimation was used for all models. For the HLM analysis of the sociometric variable, an unconditional model was fit as the first step to examine the effects of gender (coded as −1 and 1). The output generated from this model was evaluated to determine the appropriate second step. On the basis of the recommendations of Byrk and Raudenbush (1992), Level 1 predictors were maintained if (a) there was evidence of a significant ($p < .05$) fixed effect or (b) there was evidence of significant ($p < .05$) Level 2 variation. Analyses included an $n$ of 369 classrooms. The full model is below:

$$Y = \beta_{0j} + \beta_{1j}(\text{gender}) + r_{ij}$$
For aggression, at Level 1, both the fixed effect, $t(368) = -4.12, p < .001$, and variance estimates, $\chi^2(1, N = 369) = 515, p < .001$, were significant for gender, with boys showing higher scores for aggression. Thus, gender was treated as a fixed effect, as shown in the full model. The effect of intervention was significant, $t(368) = 2.14, p = .03$; the coefficient for the intervention effect was $-.05$, with a standard error of $.03$. Intervention classrooms had lower aggression scores than did the control classrooms. Cohen's $d$ was $-.22$, as computed from the $T$ ratio and degrees of freedom ($2T / \sqrt{df}$).

For hyperactive–disruptive behavior, at Level 1, both the fixed effect, $t(368) = -5.01, p < .001$, and variance estimates, $\chi^2(1, N = 369) = 522, p < .001$, were significant for gender, with boys showing higher scores for hyperactive–disruptive behavior. Thus, gender was treated as a fixed effect, as shown in the full model. The effect of intervention was significant, $t(368) = 2.23, p = .02$; the coefficient for the intervention effect was $-.05$, with a standard error of $.02$. Cohen's $d$ was $-.22$. For prosocial behavior and “most liked” ratings, the effect of intervention was not significant. There were no Site × Intervention or Cohort × Intervention interaction effects on any sociometric outcome. The intraclass correlation coefficient for sociometric variables ranged from .02 to .07. For descriptive purposes, the aggregated classroom means for sociometric variables for each group are shown in Table 2.

**Teacher ratings of behavior**—The model for both the TOCA–R and SHP analyses was similar to the two-step model shown above for sociometric analyses. However, the pretest covariate (fall of the school year) was entered as an additional Level 1 variable. There were no significant intervention or Intervention × Other Factor effects on any teacher ratings. The intraclass correlation coefficients ranged from .08 to .21 for the teacher ratings. For descriptive purposes, the aggregated classroom means for teacher ratings for each group are shown in Table 3.

**Observer ratings**—Given that the observer ratings already occurred at the classroom level, they were analyzed using the GLM model. We conducted a $4$ (site) $\times$ $2$ (intervention vs. control) ANCOVA with cohort serving as a covariate and classroom assigned as a random effect. We then entered the Intervention $\times$ Site interaction, the Intervention $\times$ Cohort interaction, and the three-way interaction of Intervention $\times$ Site $\times$ Cohort into the model. Other possible interaction effects were not entered into the model because they did not involve effects of the intervention. Analyses of the summary score of classroom atmosphere indicated a significant effect for the intervention, $F(1, 279) = 5.63, p < .01$. Intervention classrooms were rated as having a more positive classroom atmosphere; the intervention classroom mean was $2.68$ ($SD = 0.64$), and

---

5When a general linear modeling (GLM) analysis of covariance (ANCOVA) was used on mean classroom TOCA–R scores (with classroom assigned as a random effect), teachers of intervention classrooms rated their students as having a lower mean conduct problem score (using the Authority Acceptance scale), $F(1, 333) = 3.68, p < .05$, and a higher mean peer-liking score, $F(1, 333) = 3.95, p < .05$, than did teachers in control classrooms. The significant effect on conduct problems indicated that during the school year, intervention classrooms showed less deterioration in behavior than did control classrooms, whereas in the area of peer liking, the intervention classrooms showed improvement relative to the control classrooms. There were no main or interaction effects on teacher ratings of cognitive competence or social competence.

---

*J Consult Clin Psychol. Author manuscript; available in PMC 2009 October 14.*
the control classroom mean was 2.88 ($SD = 0.65$). Exploratory analyses of each of the 10 separate rating scales indicated that the means favored the intervention classrooms in all cases and that 4 of the 10 subratings were statistically significant: students' ability to follow rules, $F(1, 279) = 4.10, p < .05$; students' ability to express feelings appropriately, $F(1, 279) = 4.34, p < .05$; the classroom level of interest and enthusiasm, $F(1, 279) = 8.21, p < .01$; and the classroom's ability to stay focused and on task, $F(1, 279) = 9.52, p < .005$.

**Examination of the role of teacher experience**—Because three cohorts (years of classrooms) of children were used in the project, some teachers taught more than 1 year, and thus it was possible to examine the factor of teacher experience. One hundred thirteen teachers taught PATHS or control classrooms for 1 year, 47 teachers taught PATHS or control classrooms for 2 years, and 54 teachers taught PATHS or control classrooms for all 3 years. Levels of experience were not different for teachers in intervention versus control schools. All GLM and HLM analyses were redone, including years of teacher experience (coded as 1, 2, or 3) as a main effect. In GLM analyses, number of cohorts taught was used as both a main effect and an interaction effect with site and condition. In HLM analyses, number of cohorts taught was added as a Level 2 characteristic. There was only one main effect; For both the intervention and control groups, teachers who taught more cohorts had higher classroom atmosphere ratings (total score). There were no other significant main or interaction effects in these analyses, and the effects of intervention did not change.

**Model for Dosage and Implementation Analyses**

For analyses within the intervention classrooms, GLM analyses examined the relationship between dosage (highest lesson reached), quality of implementation ratings, and outcome. Outcome measures included only those found to have significant intervention versus control differences (in these analyses, we included teacher ratings found to be significant in the GLM analyses; see Footnote 5). First, analyses examined the effect of dosage with variables entered in the following order: Site and cohort were entered first as covariates, followed by the highest number of lessons taught, Highest Number of Lessons Taught × Site, Highest Number of Lessons Taught × Cohort, and the three-way interaction. Second, to examine quality of implementation, we conducted analyses separately for each of the four implementation ratings after controlling for the number of lessons taught. Thus, analyses of quality of implementation used variables entered in the following order: Site and cohort were entered first as covariates, followed by the highest number of lessons taught, implementation rating, Implementation Rating × Site, Implementation Rating × Cohort, and the three-way interaction. Interaction effects are reported only where they alter the interpretation of the main effect of dosage or quality of implementation.

**Dosage**—There was a trend relationship of dosage and sociometric ratings of aggression, $F(1, 169) = 2.61, p < .10$. Higher dosage was related to somewhat lower ratings of aggression. There was also a main effect of dosage on observer ratings of classroom atmosphere, $F(1, 139) = 3.82, p < .05$; higher dosage predicted more positive classroom atmosphere ratings. There were no main effects of dosage on the teacher ratings.

**Quality of implementation**—The teacher's rated skill in teaching PATHS concepts, managing the classroom, and modeling and generalizing PATHS concepts throughout the classroom day were all significantly related to teacher ratings of Authority Acceptance: $F(1, 167) = 9.90, p < .001$; $F(1, 167) = 16.54 p < .001$; and $F(1, 167) = 9.22, p < .001$, respectively. These same three measures were also significantly related to observer ratings of classroom atmosphere: $F(1, 137) = 4.95, p < .01$; $F(1, 137) = 8.87, p < .01$; and $F(1, 137) = 4.94, p < .01$, respectively. There were no main effects of these three implementation ratings on sociometric outcomes. The openness to consultation rating was related only to the classroom
sociometric score of hyperactive–disruptive behavior, $F(1, 168) = 5.48, p < .05$. In all cases, ratings of teacher skill in program implementation and classroom management predicted positive program outcomes.

**Discussion**

The results of this universal intervention model at the end of Grade 1 provide evidence of its effectiveness in the domains of both aggression and peer relations. There were significant effects of the intervention from the viewpoints of both peers and unbiased observers. Effects on aggression and classroom behavior were found from child and peer reports and the observers' ratings (of rule following, better-classroom atmosphere, and more on-task behavior). These findings reflect robust effects of the universal-level prevention activities on classroom behavior with reductions in aggression and increases in self-control and on-task behavior.

This is the first reported study of a universal social competence intervention implemented at the elementary grades that used the classroom, rather than the student, as the unit of analysis. Thus, analysis occurs at the level of implementation. One previous report (Grossman et al., 1997) used the school as the unit of analysis; however, it is at the classroom level and not the school level that intervention is delivered and that intervention delivery and dosage can be reliably assessed.

Given the concordant classroom-level findings for peer and observer ratings, it was surprising that teacher ratings of aggression were not found significant in the HLM analyses. To further understand this puzzling finding, we computed GLM analyses using the classroom aggregate mean score for both teacher ratings and peer sociometric data. For the measures of peer sociometrics, there was convergence between the findings of the two statistical models. This can be attributed to the relatively low intraclass correlation coefficient (Murray & Short, 1997), which demonstrated little dependency between scores of children within an average classroom. Thus, it did not appear that classroom-level effects were exerting a significant influence on individual scores. In contrast, for teacher ratings, there were notable differences between GLM and HLM analyses, with GLM analyses indicating a significant intervention effect on aggression/disruption (using the Authority Acceptance scale) and peer liking, but HLM models indicating no significant intervention effects. In the case of teacher ratings, the intraclass correlations were quite high (mean correlation across measures = .15), indicating high dependence between scores within classrooms. An important difference between these two measures is that peer sociometrics are rated separately by each child (in a manner similar to other measures used in many HLM models, such as academic achievement), whereas the teacher ratings are all completed by one teacher for the entire class. It is likely that this difference in having the same reporter for all class members explains the high intraclass correlation in the teacher ratings and may also account for differences between the HLM and GLM results. That is, the HLM analyses are designed to correct for the effects that children in the classroom have on each other's behavior, removing this source of variation from the estimation of the intervention effects. In the case of teacher ratings, however, HLM corrects not only for the potential dependence between children's actual behavior but also for the fact that teachers are likely to see their classrooms in a certain way and this perception affects the scores they provide for each child. Given the inflation of the intraclass correlation caused by using a single rater, we believe that the HLM analyses are likely to be a conservative estimate of intervention effectiveness.

As two independent sources (peer ratings and observers) showed convergence in the documentation of classroom-level findings, we feel that the findings support the interpretation of a robust impact of the universal prevention activities on classroom processes. In addition, the fact that there were no significant Site × Condition interaction effects indicates no major
differences in effects of intervention as a function of rural versus urban school location, percentage of children below the poverty level, or ethnic composition of the classrooms. It is unfortunate that because of the large size of the sample, we were not able to assess the acquisition of specific skills or test models in which changes in such skills might mediate outcomes.

Pre–post differences in teacher ratings of classroom aggression and disruption indicated an increase in both groups; that is, all teachers saw their classroom as more disruptive in the spring than they did in the previous fall. This is likely due to a combination of greater familiarity with students (and thus observation of more incidences of misbehavior) and the fact that children often show more disruption later in the classroom year than in the first few weeks (at which time they are often “on their best behavior”; Dodge, Coie, & Brakke, 1982).

Although intervention effects were found across raters, these effects were modest in size. This may in part be due to the fact that this was a carefully randomized trial and within schools all teachers, independent of personal interest or ability, were assessed. Thus, teachers who were relatively ineffective, showed little enthusiasm, or completed only a portion of the curriculum were assessed as if they had completed the intervention; no intervention teachers were dropped for poor quality implementation, high resistance, or providing a low dosage. In this sense, this “intent-to-intervene” trial design (Brown, 1993) may provide the highest level of external validity regarding how this universal intervention might affect a typical, entire school community.

The findings on implementation and dosage lend credence to the outcome effects. They indicate that the intervention staffs' ratings of how well teachers understood concepts, generalized skills outside the curriculum time, and managed their classroom were significantly related to decreases in classroom aggression (based on teachers’ mean ratings) as well as observers' ratings of the classroom atmosphere. Further, these implementation effects continued after covarying dosage. As dosage itself had only a mild impact on outcome, it may be less crucial how many lessons are taught (given that a majority are implemented); instead, it may be the teacher’s willingness and ability to accept and use such a classroom model that most affects classroom outcomes (Elias et al., 1998). Because there have been few studies on social competence promotion at any age that have examined how dosage and quality of implementation affect outcomes (Battistich, Schaps, Watson, & Solomon, 1996; Botvin, Baker, Dusenbury, Tortu, & Botvin, 1990; Pentz et al., 1990), these findings indicate that both quantitative and qualitative indexes of curriculum implementation should be assessed in future projects.

Two cautions should be raised, however, in interpreting these findings. First, because of the number of classrooms and the unique relationship of the ECs and teachers, we did not collect interrater reliability data on the implementation ratings. Thus, it is possible that these ratings are in part measuring the perception of the quality of relationship between the EC and the teacher. Second, we cannot rule out the hypotheses that teachers who are better at implementing the universal model are just better teachers in general and that the general quality of the teacher may account for these effects. However, the fact that these ratings, as hypothesized, predicted sociometric scores as well as independent classroom observations supports their validity.

It should be recognized that the Fast Track universal intervention included intensive intervention with high-risk children as an integral part of the overall universal intervention. Although analyses with and without these high-risk children showed similar patterns, it is quite possible that effects of the intervention on the non-high-risk children depend on a simultaneous intervention with the high-risk children (CPPRG, 1999a). Project staff commitment to work with high-risk children may reduce teacher stress and increase teacher interest in implementing
a universal intervention. Likewise, improvements in the high-risk children that are due to the selective intervention may improve classroom peer relations among other children. The present study was not designed to evaluate the effects of a universal intervention that excludes simultaneous intensive intervention with a selected group of high-risk children; thus, it does not assess the use of only the universal intervention alone. Instead, this study provides clear support that an integrated approach that combines universal and selective intervention can have meaningful effects at the universal level of analysis.

These results indicate the effectiveness that a universal intervention, when implemented with fidelity and high dosage, can have in altering the quality of the classroom climate during the 1st year of school. It is the largest study of its kind indicating the efficacy of school-based, universal interventions during the elementary school years for both the promotion of competence (Elias, 1995) and the prevention of maladjustment (Caplan et al., 1992; Dolan et al., 1993; Grossman et al., 1997; O'Donnell, Hawkins, Catalano, Abbott, & Day, 1995). Although 1 year of preventive intervention is of value, the development and evaluation of more comprehensive models that are sustained across multiple years and grade levels are necessary to document the true potential of such models (Weissberg & Greenberg, 1998). In the present project, the universal intervention is provided through Grade 5; in the future, we will examine the effects of multiple years of such exposure.

Acknowledgments

This work was supported by National Institute of Mental Health (NIMH) Grants R18 MH48043, R18 MH50951, R18 MH50952, and R18 MH50953. The Center for Substance Abuse Prevention also has provided support for Fast Track through a memorandum of agreement with the NIMH. This work was also supported in pan by Department of Education Grant S184U30002 and NIMH Grants K05MH00797 and K05MH01027.

We are grateful for the close collaborations with the Durham Public Schools, the Metropolitan Nashville Public Schools, the Bellefonte Area Schools, the Tyrone Area Schools, the Mifflin Country Schools, the Highline Public Schools, and the Seattle Public Schools. We greatly appreciate the hard work and dedication of the many staff members who implemented the project, collected the evaluation data, and assisted with data management and analyses.

References


Weissberg, RP.; Caplan, MZ.; Bennetto, L. The Yale–New Haven Middle School Social Problem Solving (SPS) Program. New Haven, CT: Yale University, Department of Psychology; 1988.

Weissberg RP, Elias MJ. Enhancing young people’s social competence and health behavior: An important
callenge for educators, scientists, policy makers, and funders. Applied & Preventive Psychology:
Weissberg, R.; Greenberg, MT. Community and school prevention. In: Siegel, I.; Renninger, A., editors.
Werthamer-Larsson L, Kellam SG, Wheeler L. Effect of first grade classroom environment on shy
behavior, aggressive behavior, and concentration problems. American Journal of Community
White, SH. Evidence for a hierarchical arrangement of learning processes. In: Lipsett, LP.; Spiker, CC.,
41-89.
Table 1
Means by Site and by Condition of School-Level Variables Indicating Poverty, Ethnicity, and Achievement

<table>
<thead>
<tr>
<th>Site and Condition</th>
<th>% children receiving free or reduced lunch</th>
<th>% minority children</th>
<th>Reading percentile score</th>
<th>Math percentile score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Durham</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>83.8</td>
<td>12.5</td>
<td>90.8</td>
<td>10.4</td>
</tr>
<tr>
<td>Control</td>
<td>75.5</td>
<td>21.2</td>
<td>89.7</td>
<td>17.8</td>
</tr>
<tr>
<td>Nashville</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>78.5</td>
<td>12.4</td>
<td>61.0</td>
<td>22.2</td>
</tr>
<tr>
<td>Control</td>
<td>77.0</td>
<td>10.9</td>
<td>47.3</td>
<td>23.8</td>
</tr>
<tr>
<td>Rural PA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>39.6</td>
<td>16.4</td>
<td>1.0</td>
<td>0.6</td>
</tr>
<tr>
<td>Control</td>
<td>39.1</td>
<td>13.4</td>
<td>1.0</td>
<td>0.9</td>
</tr>
<tr>
<td>Seattle</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>45.4</td>
<td>7.0</td>
<td>50.1</td>
<td>17.3</td>
</tr>
<tr>
<td>Control</td>
<td>46.6</td>
<td>14.2</td>
<td>53.9</td>
<td>22.3</td>
</tr>
</tbody>
</table>

Note. PA = Pennsylvania.

*North Carolina schools have initiated their own achievement testing system that presently has no percentile scores and thus cannot be compared with the other sites.*
Table 2  
Peer Nomination Means and Standard Deviations for Intervention and Control Conditions (Classroom Aggregate Scores)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Intervention</th>
<th></th>
<th>Control</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Aggression</td>
<td>-0.08</td>
<td>0.22</td>
<td>-0.01</td>
<td>0.26</td>
</tr>
<tr>
<td>Hyperactive–disruptive</td>
<td>-0.09</td>
<td>0.22</td>
<td>-0.04</td>
<td>0.24</td>
</tr>
<tr>
<td>Prosocial behavior</td>
<td>0.03</td>
<td>0.33</td>
<td>0.02</td>
<td>0.34</td>
</tr>
<tr>
<td>Most liked</td>
<td>0.04</td>
<td>0.38</td>
<td>0.05</td>
<td>0.43</td>
</tr>
</tbody>
</table>

*Lower scores indicate better functioning.*
### Table 3
Pre- and Posttest Teacher TOCA–R Means and Standard Deviations for Intervention and Control Conditions (Classroom Aggregate Scores)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Intervention</th>
<th></th>
<th>Control</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Authority acceptance&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.81</td>
<td>0.95</td>
<td>0.78</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>0.42</td>
<td>0.44</td>
<td>0.49</td>
<td>0.55</td>
</tr>
<tr>
<td>Cognitive concentration&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.64</td>
<td>1.60</td>
<td>1.64</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td>0.52</td>
<td>0.49</td>
<td>0.61</td>
<td>0.59</td>
</tr>
<tr>
<td>Low social competence&lt;sup&gt;a&lt;/sup&gt;</td>
<td>16.19</td>
<td>15.92</td>
<td>15.58</td>
<td>15.88</td>
</tr>
<tr>
<td></td>
<td>5.33</td>
<td>4.95</td>
<td>6.13</td>
<td>6.08</td>
</tr>
<tr>
<td>Peer liking&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.71</td>
<td>3.77</td>
<td>3.91</td>
<td>3.87</td>
</tr>
<tr>
<td></td>
<td>0.69</td>
<td>0.59</td>
<td>0.69</td>
<td>0.67</td>
</tr>
</tbody>
</table>

*Note.* TOCA–R = Teacher Observation of Classroom Adaptation—Revised.

<sup>a</sup>Higher scores indicate greater problem behavior.

<sup>b</sup>Higher scores indicate greater liking.