DOCUMENT RESUME

ED 476 379 EC 309 589

AUTHOR	Neace, William P.; Munoz, Marco A.; Weber, John; Johnson, Knowlton
TITLE	Evaluating a Safety Net for At-Risk Students: Impact of Research-Based Interventions on Non-Cognitive and Cognitive Measures.
INSTITUTION	Jefferson County Public Schools, Louisville, KY.
PUB DATE	2002-04-00
NOTE	26p.; Paper presented at the Annual Meeting of the American Evaluation Association (Washington, DC, November 6-10, 2002).
PUB TYPE	Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Achievement; Attendance; *Behavior Modification; Behavior Problems; Discipline; Early Intervention; Elementary Secondary Education; *Health Promotion; *High Risk Students; Mental Health Programs; *Prevention; *Program Effectiveness; School Safety; Social Development; Substance Abuse
IDENTIFIERS	Social Skills Training

ABSTRACT

An outcome evaluation was conducted for seven Project SHIELD (Supporting Healthy Individuals and Environments for Life Development) research-based and best practices interventions using secondary data from the 26th largest school district in the nation. The SHIELD program is designed to provide students and schools with enhanced infrastructure and comprehensive prevention and early intervention through education, mental health, and social services that promote healthy childhood development and prevent violence, alcohol, and other drug abuse. This study reports the short-term (one-year) and long-term (two-year) results of a three-year program and examines whether there are interim effects of the interventions on K-12 non-cognitive and cognitive outcomes. Non-cognitive outcome measures were the unexcused absences, unexcused tardiness, and disciplinary actions. Cognitive outcomes consisted of grade point average. The evaluation used a scientifically rigorous research design in which students in the various interventions were matched with students in control schools, who were chosen based upon similarities in school-level characteristics (e.g., percentage of students in the free lunch program and school standing on academic achievement tests). A quasi-experimental design was used for the evaluation. Results indicated that the majority of the interventions produced some small positive effects. (Contains 18 references.) (Author/CR)

Evaluating a Safety Net for At-Risk Students: Impact of Research-Based Interventions on Non-Cognitive and Cognitive Measures

William P. Neace, Ph.D.
Pacific Institute for Research and Evaluation, Inc.

Marco A. Muñoz, Ed.D.
Jefferson County Public Schools

John Weber, M.S.
Pacific Institute for Research and Evaluation, Inc.

Knowlton Johnson, Ph.D., MSSW
Pacific Institute for Research and Evaluation, Inc.

Paper presented at the
Annual Meeting of the American Evaluation Association (AEA)
November 6-10, 2002
Washington, D.C.

Jefferson County Public Schools
Shaping the Future

Pacific Institute
for Research and Evaluation
Louisville Center

ABSTRACT

An outcome evaluation was conducted for seven Project SHIELD research-based and best practices interventions using secondary data from the 26th largest school district in the nation. This study reports the short-term (one-year) and long-term (two-year) results of a three-year program, and examines whether there are interim effects of the interventions on K-12 non-cognitive and cognitive outcomes. Non-cognitive outcome measures were the unexcused absences, unexcused tardiness, and disciplinary actions. Cognitive outcomes consisted of grade point averages (GPA). The evaluation used a scientifically rigorous research design in which students in the various interventions were matched with students in control schools who were chosen based upon similarities in school-level characteristics (e.g., percentage of students in the free lunch program, and school standing on academic achievement tests). A quasi-experimental design was used for the evaluation. The data were analyzed using an analysis of covariance in order to statistically control for differences in demographic characteristics and increase confidence that changes in behaviors were due to the interventions rather than some other causes (i.e., extraneous variables). Results indicated that the majority of the interventions produced some small positive effects.

The study of innovations and change in school systems is not new (Hall & Hord, 2001). What is innovative is a rigorous outcome evaluation of a family of school-based interventions that targets the same problem behavior among a school age population. Further, it is innovative for a school system to evaluate the effectiveness of a system of interventions (referred to as research or science-based interventions or best practices) that has varying degrees of empirical evidence of efficacy in a demonstration project or randomized trial. It is also innovative for a school system to reply on evaluation results to make decisions about the sustainability of interventions and best practices.

The opportunity to launch a large scale evaluation of the effectiveness of a family of research-based, school-based interventions and practices that targets a common problem behavior was presented in a large federal grant from a consortium of agencies (Department of Education, Office of Juvenile Justice and Delinquency Prevention, and the Center for Mental Health Services) as part of a Safe Schools/Healthy Students federal initiative. In October 1999, Project SHIELD (Supporting Healthy Individuals and Environments for Life Development) received nearly $3,000,000 per year for three years to provide students and schools with enhanced infrastructure and comprehensive prevention and early intervention through education, mental health, and social services that promote healthy childhood development and prevent violence, alcohol, and other drug abuse. These services target the development of social skills and emotional resilience necessary for youth to avoid violent behavior and drug use, along with establishing safe, disciplined, and drug free areas within school environments. From the school system's perspective the most important purpose for conducting the evaluations of Project SHIELD interventions and best practices was to provide results that could be used in making decisions about which interventions should be sustained using school system funds after the federal funding ended.

EVALUATION FOCUS

The Project SHIELD evaluation of its family of science-based school intervention and best practices that targeted violent behavior and drug use employed a multiple method strategy. One strategy was designed to assess the combined effects of the Project SHIELD interventions and practices on school violence, fear of violence, and alcohol and other drug use that were implemented in a set of middle and high schools, in comparison with a matched set of control schools. Annual school survey data (repeated cross-sectional data) were collected from sixth, eighth, tenth, and twelfth grade students in both the intervention and control school. The second strategy of evaluating Project SHIELD interventions and best practices was to conduct a pre-post, matched control group of individual interventions and best practices using outcome criteria used by the school to evaluate the accountability of all school programs.

This paper presents interim results relating to the second strategy. The outcome measures chosen (unexcused absences, unexcused tardiness, in-school disciplinary referrals, number of suspensions, and suspension length) served as measures of student behavior for two reasons. First, student behaviors and performance are of primary concern to school officials interested in continuing the interventions after Project SHIELD funding has ended, and due to the need to satisfy requirements related to accountability. School systems need quantifiable measures of student performance effectiveness in a high-stakes accountability environment. In this environment, educational policy making is based on objective information, and although no single means of data collection is sufficient, the data generated by well-designed program evaluations are crucial to an understanding of project impact. Policy-makers have to refocus the

educational reform efforts in general, and the educational excellence issues in specific, toward results on school-related indicators (Munoz, 2002). School districts must guarantee that programs have a demonstrably positive effect on students' key non-cognitive (e.g., attendance, tardies, suspensions) and cognitive (e.g., reading, math, science, and social studies scores) indicators. Although many of the interventions implemented through Project SHIELD have produced favorable results in previous efficacy evaluations, these evaluations have focused mainly on outcomes related to indirect measures of behavior, such as changes in attitudes, knowledge, and self-reported delinquency (e.g., Cooper, Lutenbacher, & Faccia, 2000; Mytton, DiGuiseppi, Gough, Taylor, & Logan, 2002). School administrators are far more interested in outcome results directly related to student behavioral problems and academic performance. Under the conceptualization of accountability as performance, output educational indicators are used to track and evaluate program effectiveness based on student results. Accountability is generally conceived as a demand to judge school programs by their outputs. Accountability systems have been designed to track the progress of educational reforms. The function of an accountability system in education is to monitor and evaluate the performance of the educational programs (Wholstetter, 1991). Schools are expected to make wise use of public resources not only by efficient cost accounting procedures but also by increasing attendance, decreasing suspensions, and showing positive impacts on academic performance.

A second reason for using the stated objective measures is that there has been a call for more use of objective outcome measures in violence prevention research literature (Housman, Pierce, & Briggs, 1996; Webster, 1993). Several studies report improvements in knowledge and attitudes related to use of violence to resolve conflicts (e.g., Aber, Jones, Brown, Chaudry, & Samples, 1998; Hausman, Spivak, & Prothrow-Stith, 1995; Oprinas, Parcel, McAlister, & Frankoweski, 1995), or self-reported aggressive behaviors (e.g., Avery-Leaf, Cascardi, O'Leary, & Cano, 1997; DuRant, Barkin, & Krowchuk, 2001; DuRant, Treiber, Getts, McCloud, Linder, & Woods, 1996). Fewer studies have examined changes in relevant student behavioral indicators, such as suspension rates (Farrell, Meyer, & White, 2001; Housman et al., 1996), disciplinary referrals (Ferrell et al, 2001; Twemlow, Fonagy, Sacco, Gies, Evans, & Ewbank, 2001) or academic performance (Twemlow et al., 2001). Only one study (Ferral et al., 2001) has examined the link between changes in knowledge and attitudes and relevant behavioral indicators, with the results indicating no mediating effects. Thus, it is becoming increasingly important to examine the effectiveness of violence prevention interventions on student behavior and performance. This paper will inform school administrators about the impact of these programs on key performance indicators associated with the school environment. The present study was designed to contribute to existing theoretical and practical knowledge about the effectiveness of a family of research-based and best practices interventions on school-related research issues not previously addressed in the literature.

DESCRIPTIONS OF PROJECT SHIELD INTERVENTIONS

The selected research-based and best practices interventions under evaluation fall into two categories: universal interventions (those in which participation is a consequence of class attendance and that target students broadly), and selective interventions (in which participation is a consequence of referral, and that target particular groups of students identified as at-risk for violence and substance abuse). The universal interventions are SMART Moves, and Second Step. The selective interventions are LEEP Mentoring, Functional Family Therapy, Primary Mental Health, Big

Brothers/Big Sisters Mentoring, and Multi-Systemic Therapy. A brief description of each intervention follows.

Universal Interventions

SMART Moves

SMART (Skills Mastery and Resistance Training) Moves is a research-based, universal program (grades 6, 8, 10, and 12) consisting of a series of validated prevention program components for children aged 10-15 years and their parents. The Salvation Army Boys and Girls Clubs are implementing SMART Moves to increase awareness of Alcohol, Tobacco and Other Drug (ATOD) risks, reduce risky behaviors, and improve academic achievement. The program components focus on alcohol, tobacco, other drugs (Stay SMART), pregnancy, peer pressure, decision-making and life planning skills (Start SMART), conflict resolution and dealing with anger (Street SMART), healthy lifestyle choices for girls (SMART Girls), and parental support (SMART Parents). Implementing these program components at the transition period of elementary through middle school ages provides a critical age-appropriate support system that addresses the issues students face as they experiment with their approaching adult rights, roles, and responsibilities.

Second Step

Second Step is a research-based, universal violence prevention program for K-middle school-aged children. Seven Counties Services is implementing Second Step in grades 1 to 3 in order to increase children's ability to identify what others are feeling, take others' perspectives, and respond empathically with others. The program also hopes to decrease impulsive, aggressive, and angry behavior. The 28 lessons focus on precursor behaviors that are incompatible with violence: empathy, impulse control, problem solving, and anger management. Lessons are developmentally appropriate in content and delivery with ample opportunity for students to model, practice, and reinforce their prosocial behavior.

Selective Interventions

LEEP Mentoring

The Louisville Education and Employment Partnership (LEEP) is a selective program that uses a mentoring team concept, and has been implemented at participating middle schools (grades 6-8) in the school district. LEEP, including mentoring, has been expanded in two Project SHIELD high schools (grades 9-12) to include an additional 55 students each. Through LEEP activities (including tours, career workshops, job fairs, and breakfast clubs), group mentoring, and one-on-one mentoring, the program hopes to improve school attendance rates, cognitive outcomes, graduation rates, and employment, post-secondary education, or military transition.

Functional Family Therapy

Functional Family Therapy (FFT) is a nationally recognized research-based, selective program that has been applied successfully in various contexts (e.g., rural, urban, multicultural, international) and treatment systems (e.g., clinics, home-based programs, independent providers). Seven Counties Services is implementing FFT to engage and retain families and targeted youth in prevention/intervention activities, develop long-term behavior change patterns, and enhance families' abilities to impact multiple systems in which the families are embedded. This program's success has been replicated and demonstrated for over twenty-five years. The program is designed for individual youth ages 11-18. The youth present a wide range of acting-out behaviors and are at

risk for, or are presenting behaviors associated with delinquency, violence, substance use, conduct disorder, oppositional-defiant disorder, or disruptive behavior disorder. Therapists administer the program by responding promptly to family expectations, teaching parenting, problem solving and conflict management skills, and initiating clinical and community linkages with the families. The program focuses on enhancing protective factors and reducing risk factors through engagement, motivation, assessment, behavioral change, and generalization of results to apply to the family's needs.

Primary Mental Health

Primary Mental Health Project is a research-based, selective program. This early detection and prevention program for preschool and primary grades is being implemented by JCPS. It is a nationally recognized model out of Rochester, New York that has been replicated in over 200 cities since 1957. Trained child associates (supervised by school psychologists) provide individual interactive play sessions to children (grades 1-2) who begin showing adjustment difficulties. Learning skills and other school-related competencies are enhanced, and social, emotional, and school adjustment problems are reduced.

Big Brothers/Big Sisters Mentoring

Big Brothers/Big Sisters School-Based Mentoring is a research-based, selective program focusing on case managed mentoring by adult volunteers of 100 elementary students (K-5) each year who are considered high risk to engage in risky behaviors which include violence and substance use. Through its mentoring program, Big Brothers/Big Sisters hopes to increase the children's confidence, competence, caring, and cognitive outcomes, and to decrease behavioral problems. The program (1) recruits, screens, and trains adult volunteers, (2) interviews and assesses student applicants and their parents, (3) makes and supervises matches between the adults and students, and (4) closes the matches when timelines or situations dictate.

Multi-Systemic Therapy

Multi-Systemic Therapy (MST) is a nationally recognized, research-based, selective program for chronic, violent, or substance abusing juvenile offenders aged 12-18 (grades 9-12). MST is being implemented by Seven Counties Services, to improve school attendance, and reduce suspensions, antisocial behavior, and youth criminal activity. Replication studies show extensive improvement in family relations, decreased behavioral problems, reduced substance abuse and a reduced three-year recidivism rate. The program considers the complex network of systems that affect individual, family, and other relationships. Parents and students experience strategic and structural family therapy, behavioral parent training, and cognitive behavior therapies to focus attention and direct change.

METHODS

Evaluation Design

The outcome evaluation is necessary in order to determine whether the intervention has the desired effect on relevant criteria by which an intervention is to be judged (Rossi, Freeman, & Lipsey, 1999). The outcome evaluation employed a pre-post matched control group design to assess outcomes for participants in the seven interventions being evaluated. A two-level matching procedure was used to develop the relevant control groups for the intervention groups in order to provide a measure of rigor in terms of the internal validity of the study (Cook & Campbell, 1979).

Due to the inability to randomly assign individuals to intervention and control groups, a two-level matching procedure was employed in order to add rigor to the internal validity (Cook and Campbell, 1979) of the evaluation. The first level involved the selection of a subset of Jefferson County public schools in which Project SHIELD was not being implemented but which were similar in global characteristics (e.g., percentage of students involved in the free and reduced lunch program, percentage of students in special education programs, percentage of students suspended) to Project SHIELD schools. These schools served as the basis for the second level of matching which took place at the individual level. The control group "partners" for each of the intervention group individuals were selected from the subset of control schools by database software authored by Mr. Mark Shuler of MBS Technologies, Inc. Students were matched on three demographic characteristics considered important by key district-level administrators: (1) Exceptional Childhood Education (ECE) status, (2) gender, and (3) age. This procedure resulted in providing a unique, matched control group for each of the interventions being evaluated.

An analysis of the success of the matching procedure is presented in the results section. Overall matching, as well as matching within each intervention, was examined to determine matching success rate, and to verify that the Project SHIELD and control groups do not differ significantly on the three main matching criteria employed. Categorical variables were examined using chi-square analyses, and age was examined using a one-way analysis of variance.

Data Requirements

This evaluation relied on secondary data collected by school personnel and maintained in the school systems' information system referred to as the Teradata system. Detailed student information on unexcused absences and unexcused tardiness is collected on a daily basis by the school system, and maintained in the Teradata electronic database. The Teradata system also tracks a number of more direct indicators of problematic behavior. In-school suspensions and probations (ISAP) are tracked on a daily basis. Generally speaking, the number of days a student spends in this form of punishment could serve as a rough indicator of his or her behavioral problems. Two other measures of behavioral problems are also available. The first measure, number of out-of-school suspensions, serves as an indicator of the extent of problem behaviors exhibited by the student. The more total number of suspensions per semester, the more problematic the behavior of the student. The second measure extracted is the total number of days suspended, which serves as an indicator of the seriousness of the offense which resulted in a suspension. Generally speaking, the more serious the offense, the more days the student is suspended. Infractions for which a suspension could result range in seriousness from failure to sign out of school to terroristic threatening (e.g., bomb threat directed at the school). Finally, the Teradata system also contains some academic information, which can be used as an indicator of academic performance. These data include academic grades in four core content courses (English, Social Studies, Mathematics, and Science), coded on an ordinal scale ranging from F to A+. Grades were recoded into a numeric scale (ranging from 0 to 4.5, representing F to A+, respectively), and used to form GPA (grade point average) by taking the mean of the four core content courses within each semester.

In addition to the outcome measures just described, demographic characteristics of the students were also collected. These included Exceptional Childhood Education (ECE) Status (a general indicator of the student's cognitive, psycho-social and physical functioning), gender, and age (chronological age in years). These measures were collected in order to examine the degree

to which students in the intervention groups were matched to control students. In addition, data on Free-and-Reduced Lunch Status (in which students are classified according to their ability to pay for their own school meals, ranging from those who receive free meals to those who pay full price for meals), and zip code (area of residence within the school district) were also collected to be used as covariates in the analysis.

Analyses

As indicated earlier, the outcome evaluation is an assessment of secondary data maintained in the Teradata system. The outcome measures consisted of a change in the number of unexcused absences, unexcused tardiness, ISAPs, and suspensions both before and after up to three semesters of implementation of the particular Project SHIELD intervention being evaluated.

The data were first examined to test the statistical assumptions (e.g., distributional assumptions of the outcomes, homogeneity of variance, examination of outliers) of the desired analysis procedures (Tabachnik & Fidell, 1996). Since the data were found to be amenable to general linear modeling, the interventions were evaluated using a repeated measures analysis of covariance (ANCOVA) with group (intervention group vs. control group) as a between-subjects factor, and repeated measures on outcome as the within-subject factor. Students were matched on ECE status, gender, and age (see results below); however, pre-existing group differences on eligibility for free or reduced lunch prices and area of residence had to be statistically controlled by entering them as covariates in the models. Covariates were used here principally to rule out the effects of potentially confounding factors, such as differences in participation in the free-and-reduced lunch program (our proxy variable for socio-economic status), on assessing the change in behavior and cognitive outcomes. Short-term outcome change was defined as the difference in outcome measure from the Spring 2000 semester to the Spring 2001 semester, reflecting immediate outcomes. Long-term outcome change was defined as a change in outcome measure from the Spring 2000 semester to the Spring 2002 semester. Separate ANCOVAs were conducted for each of the interventions of interest.

An attrition analysis (see Results) was conducted on two levels. The first analysis examined whether attrition from the study was related to specific demographic characteristics (gender, age, etc.), and to determine whether there were differences in attrition between students in the intervention groups and students in the control groups. The second level of analysis was performed to examine whether pre-intervention levels of the outcome measures affected attrition.

RESULTS

SHIELD–Control Group Equivalence

Prior to addressing the substantive results of this study, it is necessary to demonstrate that the software application designed by Mr. Mark Shuler of MBS Technologies, Inc. successfully extracted a well-matched control group. As mentioned above, the matching was implemented using students from matched control schools along three dimensions: ECE status, gender, and age. The matching results are reported first for the combined sample (i.e., combining the participants from all seven interventions), and then for the seven individual interventions for which an outcome evaluation was conducted. Table 1 below presents the number of participants and control students for each of the interventions evaluated, along with the total number of SHIELD and control students.

Overall, the matching success rate was 96% (3,233 control students were found for 3,372 students participating in Project SHIELD).

Table 1. Number of participants and controls for Project SHIELD interventions.

Intervention	Number of Control Students	Number of Students Served
SMART Moves	1,708	1,814
Second Step	918	922
LEEP Mentoring	102	119
Functional Family Therapy	69	75
Primary Mental Health	233	233
Big Brothers/Big Sisters	151	152
Multi-Systemic Therapy	52	57
TOTAL	**3,233**	**3,372**

Chi-square analysis was used to test for statistically significant differences between the groups on categorical variables (e.g., ECE status, gender), and a one-way analysis of variance was used to determine whether the two groups of students statistically differed in age.

Matching Using the Combined Sample

ECE Status. ECE Status constitutes a general indicator of the student's cognitive, psycho-social, and physical functioning. It is reported in the Teradata system in 16 categories as shown in Table 2.

Table 2. Exceptional Child Education Status Categories.

Code	Description
A	Learning Disability
B	Emotional/Behavioral Disability
C	Mild Mental Disability – Educable
D	Functional Mental Disability – Trainable
F	Functional Mental Disability – Severe/Profound
G	Physical Disability
H	Hearing Impaired – Oral
J	Hearing Impaired – Total Communication
M	Traumatic Brain Injury
N	Visual Disability
O	Other Health Impairments
P	Multiple Disabilities
R/S	Speech/Language Disability
T	Developmentally Delayed Speech Only
U	Autism
Y	Developmentally Delayed Plus Speech

Given the infrequency of some of these categories, it was necessary to collapse this system into a simpler set of three categories, which were developed in consultation with district staff who have expertise in childhood disabilities. The final three categories used were emotional/behavioral disability (comprised of category B above), mental/physical disability (comprised of all other categories shown above), and no disability (comprised of students who do not have an ECE designation). The distribution of students participating in Project SHIELD and the control students are presented in Table 3.

Table 3. ECE Matching Summary.

	Control Group	SHIELD Group
No disability	2,339 (72%)	2,383 (71%)
Mental/Physical disability	819 (25%)	888 (26%)
Emotional behavioral disability	75 (2%)	101 (3%)
TOTAL	**3,233 (100%)**	**3,372 (100%)**

Students in the control group were reasonably well-matched with students in the SHIELD group on ECE status (chi-square $_{(2)} = 4.12$, $p = .128$).

Gender. Students were nearly perfectly matched on gender (chi-square $_{(1)} = .022$, $p = .882$).

Table 4. Gender Matching Summary.

	Control Group	SHIELD Group
Female	1,515 (47%)	1,574 (47%)
Male	1,718 (53%)	1,798 (53%)
TOTAL	**3,233 (100%)**	**3,372 (100%)**

Age. Students in the control group did not significantly differ from students in the SHIELD group on age: $F(1,884) = 1.25$, $p = .264$.

Table 5. Birth Date Matching Summary.

	N	Mean	S.D.
Control Group	3,233	12.34	3.26
SHIELD Group	3,372	12.43	3.25

To sum up, results of the analyses indicate that students in the control group are reasonably matched to students in the program group along the five dimensions used to select control group "partners" for the intervention group.

Matching Results by Intervention

On a per-intervention basis, the same results presented above were obtained. Table 6 (shown on the following page) summarizes the findings. As can be seen, the distribution of students in the control groups was very similar to the distribution of students in the corresponding SHIELD interventions. Chi-square analyses (on categorical items) and analysis of variance (on the continuous measure of age) conducted for each intervention indicated that the control groups did not significantly differ from the corresponding SHIELD groups on ECE status, gender, or age. From these analyses, we conclude that the two groups are well-matched on the selected criteria.

Attrition Analysis

Both panel and differential attrition analyses were conducted. A panel attrition analysis answers one basic question: "Are those who drop out different from those who stay with respect to pretest values of the individual characteristics and outcomes?" A differential attrition analysis answers two basic questions. First, "Are there differences in rates of attrition between intervention and control groups?" Second, "Are pretest scores for background characteristics and baseline outcomes for students that dropped out different across intervention groups?"

Panel Attrition

The overall attrition rate was 4.6% (302 out of 6,605 students did not have data in the Spring 2002 semester). Attrition rates for each of the interventions separately are reported in Table 7 below. As can be seen, attrition was minimal with the exception of FFT and PMH. Attrition

resulted from students who matriculated or otherwise moved out of the school district, resulting in an inability to obtain relevant outcome measures for them in the Spring 2002 semester.

Table 6. Results of the computerized matching procedure for each Project SHIELD intervention.

	ECE Status			Gender		
Intervention	**None**	**Mental / Physical**	**EBD**	**Female**	**Male**	**Age (in years)**
SMART Moves	1,230 (68%)	522 (29%)	62 (3%)	835 (46%)	979 (54%)	14.49
Control	1,199 (70%)	466 (27%)	43 (3%)	793 (46%)	915 (54%)	14.49
Second Step	733 (80%)	177 (19%)	12 (1%)	448 (49%)	474 (51%)	8.29
Control	733 (80%)	175 (19%)	10 (1%)	446 (49%)	472 (51%)	8.30
LEEP	71 (60%)	46 (39%)	2 (2%)	70 (59%)	49 (41%)	18.14
Control	60 (59%)	40 (39%)	2 (2%)	57 (56%)	45 (44%)	18.06
FFT	46 (61%)	19 (25%)	10 (13%)	41 (55%)	34 (45%)	14.60
Control	45 (65%)	16 (23%)	8 (12%)	39 (57%)	30 (44%)	14.62
PMH	169 (73%)	63 (27%)	1 (0.4%)	94 (40%)	139 (60%)	8.82
Control	169 (73%)	63 (27%)	1 (0.4%)	94 (40%)	139 (60%)	8.93
BB/BS	109 (72%)	38 (25%)	5 (3%)	68 (45%)	84 (55%)	10.98
Control	109 (72%)	38 (25%)	4 (3%)	68 (45%)	83 (55%)	10.99
MST	25 (44%)	23 (40%)	9 (16%)	18 (32%)	39 (68%)	17.43
Control	24 (46%)	21 (40%)	7 (14%)	18 (35%)	34 (65%)	17.57

To examine whether age was a factor in attrition, an analysis of variance (ANOVA) was performed using attrition group (stayers vs. droppers) as a between-subjects factor, and age as the dependent variable. Overall, students who left the school system had a mean age of 12.5 years (standard deviation = 3.22), and students who remained in the school system had a mean age of 9.87 years (standard deviation = 3.25). Analysis revealed that this difference was significant ($F_{(1,6603)} = 192.7$, $p = .001$), indicating that older students tended to leave the school system, and younger students remained in the school system through the Spring 2002 semester.

Finally data were examined to determine whether pre-intervention levels on outcome measures (unexcused absences, unexcused tardiness, ISAPs, number of days suspended, number of suspensions, and GPA) were associated with attrition. Multivariate analysis of variance (MANOVA) was performed with attrition group (stayers vs. droppers) as a between-subjects factor and the six pre-intervention outcomes as the dependent variables. A separate MANOVA was performed for each intervention. The results indicated that there were no significant differences in pre-intervention levels of behaviors or of GPA between students who left the school system and those who remained in the school system.

Differential Attrition

In order to determine differential attrition (whether students who remained in the school system at follow-up differed from students who did not), the relationship between demographic factors, group membership (intervention vs. control), scores on pre-intervention outcome

measures and attrition was examined in three separate analyses. Each attrition analysis is discussed in turn.

Dummy coding was used to create an attrition group variable representing students who remained in the school system (coded 0) and students who left the school system (coded 1). This attrition variable was then used as the dependent variable in a logistic regression analysis with the following predictors: Region (dichotomized to reflect students who lived in the city and students who lived outside of the city), ECE status, Gender, SES, and Group (intervention vs. control).

Evident from Table 7 is that differential attrition rates were nearly equal between students in the various interventions and their respective control groups. In confirmation of this observation, the group factor was not significant in the logistic regression performed on the attrition groups. Thus, there was no evidence of differential attrition between students in the interventions and control students. None of the demographic characteristics was significant in any of the logistic regressions, indicating that there were no statistically significant differences on demographic characteristics between students who remained in the school system and those that left the school system.

Table 7. Attrition results for each intervention.

Intervention	Spring 2000	Spring 2001	Spring 2002	Attrition Rate
BBBS	152	152	141	7%
Control	151	151	140	7%
FFT	75	75	48	36%
Control	69	69	44	36%
LEEP	119	119	115	3%
Control	102	102	99	3%
MST	57	57	55	4%
Control	52	52	52	0%
PMH	233	233	129	45%
Control	233	233	128	45%
Second Step	922	922	921	0%
Control	918	918	912	1%
SMART Moves	1,814	1,814	1,813	0%
Control	1,708	1,708	1,706	0%

OUTCOME RESULTS

The results of the main outcome analyses are reported for each of the Project SHIELD interventions separately. The universal interventions are discussed first, followed by the selective interventions. The results tables are organized as follows: column 1 lists the outcomes. Columns 2 and 3 show the means for each outcome in the pre-intervention semester (Spring 2000) and the post intervention semester (Spring 2001 for short-term effects, and Spring 2002 for long-term effects). Columns 4 and 5 show the means for the SHIELD group in the pre-intervention semester (Spring 2000) and the post intervention semester (Spring 2001 for short-term effects, and Spring 2002 for long-term effects). The F-test value for the group by time interaction is shown in column 6, and column 7 presents the effect size. Effect sizes, a measure of the amount of pre-post change attributable to a factor, are discussed in text where results are statistically significant. According to Cohen (1988), a small effect size is a bivariate correlation of less than .10, a medium effect is a correlation between .10 and .25, and a large effect is a

correlation of .50 or greater. Results are presented for the short-term effects followed by the long-term effects for each intervention.

Universal Interventions

The two universal interventions evaluated were SMART Moves and Second Step. The short-term and long-term effects of each of these interventions are discussed in turn.

Short-term Effects of Smart Moves

Table 8 presents the results of the change in behaviors and student performance from Spring 2000 to Spring 2001 for SMART Moves.

Table 8. Change in non-cognitive and cognitive outcomes for students in SMART Moves and matched controls from Spring 2000 to Spring 2001.

	Control (n = 1708)		SMART (n = 1814)			
	Spring 2000	Spring 2001	Spring 2000	Spring 2001	F	ES
Unexcused Absences	3.68	3.55	4.88	5.08	2.39	0.03
Unexcused Tardiness	3.28	3.88	4.65	7.40	57.92***	0.13
ISAPs	.50	1.28	.35	.84	9.36**	0.05
Total Days Suspended	.39	.52	.51	1.05	19.75***	0.08
Total Number of Suspensions	.12	.16	.18	.32	16.80***	0.07
GPA	2.10	2.03	2.03	2.02	2.32	0.03

Note: * = results are significant at $p \leq .05$; ** = results are significant at $p \leq .01$; δ = results indicate a trend in the appropriate direction. GPA analysis is based on N = 862 control students (50%) and N = 1,043 SMART Moves students (57%).

Differential change in behaviors between students in SMART Moves and controls was significant for four of the five non-cognitive outcome measures. For students in SMART Moves, unexcused tardiness, number of days suspended, and total number of suspensions increased more for the intervention group than for the control group. Effect sizes for these differences were small to medium, ranging from .07 for the change in suspensions to .13 for the change in tardiness. ISAPs increased more for control students than for students in SMART Moves, however, indicating that the intervention may have had some positive impact on student behavior while in school. The effect size associated with the differential change in ISAPs was small (.05). Change in cognitive outcomes (GPA) did not differ between intervention and control students; however, this result must be interpreted with caution due to a large amount of missing academic data.

Long-term Effects of SMART Moves

Table 9 on the next page presents the results of the long-term analysis of SMART Moves. The significant increase in unexcused tardiness, number of days suspended, and number of suspensions for the intervention group noted in the short-term analysis was not evident in the long-term analysis. These non-cognitive measures increased for both students in SMART Moves and control students, and the magnitude of the increase was not statistically different between them. As in the short-term analysis, ISAPs increased slightly more for students in the

control group than for students in SMART Moves. The differential change in ISAPs was significant but the effect size was small. While unexcused absences increased for both intervention and control groups, the magnitude of increase was larger for students in SMART Moves, relative to controls. The group by time interaction was significant but the effect size was small (.05).

Table 9. Change in non-cognitive and cognitive outcomes for students in SMART Moves and matched controls from Spring 2000 to Spring 2002.

	Control (n = 1706)		SMART (n = 1813)			
	Spring 2000	Spring 2002	Spring 2000	Spring 2002	F	ES
Unexcused Absences	3.68	6.99	4.88	9.43	10.51***	0.05
Unexcused Tardiness	3.28	8.02	4.65	10.27	3.44	0.03
ISAPs	.50	3.13	.35	2.42	5.47*	0.04
Total Days Suspended	.39	1.45	.51	1.44	.580	0.00
Total Number of Suspensions	.12	.38	.18	.40	1.05	0.00
GPA	2.10	2.03	2.06	1.93	1.49	0.03

Note: * = results are significant at $p \leq .05$; ** = results are significant at $p \leq .01$; *** = results are significant at $p \leq .001$. GPA analysis is based on N = 866 control students (51%) and N = 911 SMART Moves students (50%).

No significant differences in changes in cognitive outcomes was noted between students in SMART Moves and control students; however, since only half of the students had data on GPA, these results must be interpreted with caution.

Short-term Effects of Second Step

The short-term changes in non-cognitive outcomes for students in Second Step and their controls are shown in Table 10. No cognitive outcome measure was available for these students due to the fact that grades are not maintained in Teradata at the elementary school level.

Table 10. Change in non-cognitive outcomes for students in Second Step and matched controls from Spring 2000 to Spring 2001.

	Control (n = 918)		Second Step (n = 922)			
	Spring 2000	Spring 2001	Spring 2000	Spring 2001	F	ES
Unexcused Absences	2.98	2.25	5.21	3.05	33.70***	0.13
Unexcused Tardiness	3.12	3.04	4.78	4.47	.530	0.00
ISAPs	0	0	.001	0	1.07	.03
Total Days Suspended	.01	.03	.02	.02	.771	0.00
Total Number of Suspensions	.004	.01	.008	.007	.832	0.00
GPA+	-	-	-	-	-	-

Note: * = results are significant at $p \leq .05$; ** = results are significant at $p \leq .01$; *** = results are significant oat $p \leq .001$; + = GPA data not maintained for elementary school students.

Evident from Table 10 is that unexcused absences decreased slightly for control students, while students in Second Step exhibited a large decrease. The differences in magnitude of change between intervention and control students was significant, and is associated with a medium effect size. No other group differences were significant.

Long-term Effects of Second Step

Table 11 presents the results of the short-term changes in non-cognitive outcomes for students in Second Step and controls.

Table 11. Change in non-cognitive outcomes for students in Second Step and matched controls from Spring 2000 to Spring 2002.

	Control (n = 912)		Second Step (n = 921)			
	Spring 2000	Spring 2002	Spring 2000	Spring 2002	F	ES
Unexcused Absences	3.00	3.83	5.21	4.20	35.96***	.14
Unexcused Tardiness	3.14	5.75	4.78	6.83	1.21	.03
ISAPs	0	0	0	0	1.07	.03
Total Days Suspended	.01	.10	.02	.07	.315	.00
Total Number of Suspensions	.00	.02	.00	.03	.018	.00
GPA+	-	-	-	-	-	-

Note: * = results are significant at $p \leq .05$; ** = results are significant at $p \leq .01$; *** = results are significant oat $p \leq .001$; n/a = not applicable; + = GPA data not maintained for elementary school students.

In the long-term, unexcused absences decreased for students in Second Step but increased for control students. Differences between the groups over time was significant, with a medium effect size (.14). No other group differences were significant.

Selective Interventions

Short-term Effects of LEEP

Short-term changes in non-cognitive and cognitive outcomes for the LEEP intervention are shown below in Table 12. As can be seen, students in LEEP exhibited a decrease in unexcused tardiness compared to matched controls, who exhibited an increase. The differential group change was significant, and is associated with a medium effect size. No other group differences were significant, indicating that both groups exhibited similar magnitudes of change in the other outcome measures.

Table 12. Change in non-cognitive and cognitive outcomes for students in LEEP and matched controls from Spring 2000 to Spring 2001.

	Control (n = 102)		LEEP (n = 119)			
	Spring 2000	Spring 2001	Spring 2000	Spring 2001	F	ES
Unexcused Absences	5.33	6.13	8.47	9.42	.011	.00

Table continues...

Table 12 continued. Change in non-cognitive and cognitive outcomes for students in LEEP and matched controls from Spring 2000 to Spring 2001.

	Control (n = 102)		LEEP (n = 119)			
	Spring 2000	Spring 2001	Spring 2000	Spring 2001	F	ES
Unexcused Tardiness	3.68	6.12	10.56	8.69	6.32*	.17
ISAPs	2.81	2.05	1.34	1.29	1.03	.07
Total Days Suspended	.21	.32	.57	.65	.005	.00
Total Number of Suspensions	.05	.09	.17	.19	.005	.00
GPA	1.71	1.94	1.65	1.68	1.35	.08

Note: * = results are significant at $p \leq .05$; ** = results are significant at $p \leq .01$; δ = results indicate a trend in the appropriate direction. For GPA, N = 85 control students (83%) and 100 LEEP students (84%).

Long-term Effects of LEEP

Table 13 presents the long-term effects of the LEEP intervention on student behavior and performance outcome measures.

The short-term effect of the intervention on change in unexcused tardiness did not carry-over to the long term. Nevertheless, the results indicate that the intervention did have a positive effect on ISAPs. Students in LEEP exhibit a decrease in ISAPs, relative to the increase found in the control group. The differential change in ISAPs is significant (effect size = .15, a medium effect).

Table 13. Change in non-cognitive and cognitive outcomes for students in LEEP and matched controls from Spring 2000 to Spring 2002.

	Control (n = 99)		LEEP (n = 115)			
	Spring 2000	Spring 2002	Spring 2000	Spring 2002	F	ES
Unexcused Absences	5.14	11.53	8.65	10.25	3.32	0.13
Unexcused Tardiness	3.64	11.01	10.82	14.04	1.36	0.08
ISAPs	2.71	3.99	1.41	.20	4.94*	0.15
Total Days Suspended	.22	.46	.58	.67	.130	0.03
Total Number of Suspensions	.06	.23	.16	.16	1.98	0.09
GPA	1.69	2.11	1.66	1.68	2.55	0.15

Note: * = results are significant at $p \leq .05$; ** = results are significant at $p \leq .01$; *** = results are significant at $p \leq .001$;. For GPA, N = 82 control students (83%) and 36 LEEP students (31%).

Short-term Effects of Functional Family Therapy

Results of the short-term outcome evaluation of Functional Family Therapy (FFT) are presented in Table 14 on the next page. Relative to controls, students in FFT exhibited an increase in unexcused absences (effect size = .21), and unexcused tardiness (effect size = .19). In

addition, significant group differences are noted in the change in total number of days suspended. Relative to the decrease in number of days suspended for control students, the FFT group exhibited an increase in this outcome (effect size = .18).

Table 14. Change in non-cognitive and cognitive outcomes for students in FFT and matched controls from Spring 2000 to Spring 2001.

	Control (n = 69)		FFT (n = 75)			
	Spring 2000	Spring 2001	Spring 2000	Spring 2001	F	ES
Unexcused Absences	4.00	3.47	5.01	7.75	6.59*	.21
Unexcused Tardiness	2.74	3.03	4.43	8.19	5.39*	.19
ISAPs	.82	1.30	1.05	.91	1.08	.09
Total Days Suspended	.64	.31	1.58	2.66	4.51*	.18
Total Number of Suspensions	.19	.09	.50	.68	2.15	.12
GPA	1.86	1.92	1.53	1.82	.735	.11

Note: * = results are significant at $p \leq .05$; ** = results are significant at $p \leq .01$; *** = results are significant at $p \leq .001$. For GPA, N = 33 control students (48%) and N = 31 FFT students (41%).

No other group differences were significant, although there is a positive trend in the change in ISAPs for students in FFT, relative to controls.

Long-term Effects of Functional Family Therapy

Table 15 presents the results of the analysis of long-term changes in non-cognitive and cognitive outcomes for the FFT intervention.

Table 15. Change in non-cognitive and cognitive outcomes for students in FFT and matched controls from Spring 2000 to Spring 2002.

	Control (n = 44)		FFT (n = 48)			
	Spring 2000	Spring 2002	Spring 2000	Spring 2002	F	ES
Unexcused Absences	4.68	10.77	4.82	10.05	.059	0.03
Unexcused Tardiness	3.30	8.78	5.06	8.74	.476	0.07
ISAPs	.78	4.64	.99	1.71	3.75*	0.20
Total Days Suspended	.85	1.69	1.70	2.37	.023	0.00
Total Number of Suspensions	.22	.46	.48	.77	.017	0.00
GPA	1.92	2.27	1.34	1.53	.254	0.06

Note: * = results are significant at $p \leq .05$; ** = results are significant at $p \leq .01$; *** = results are significant at $p \leq .001$ For GPA, N = 32 control students (72%) and N = 32 FFT students (67%).

The positive trend in ISAPs previously noted was also seen in the long-term analysis, and was statistically significant. Students in FFT exhibited a slight increase in ISAPs relative to control students, who exhibited a substantial increase (effect size = .20).

Short-term Effects of Primary Mental Health

The results of the short-term outcome analysis of Primary Mental Health (PMH) are shown in Table 16. Recall that student grades for elementary school students are not maintained in Teradata; thus, no measure of cognitive outcomes were available for students in PMH or their matched controls. There were no statistically significant group differences noted in the short-term analysis of PMH.

Table 16. Change in non-cognitive and cognitive outcomes for students in PMH and matched controls from Spring 2000 to Spring 2001.

	Control (n = 231)		PMH (n = 233)			
	Spring 2000	Spring 2001	Spring 2000	Spring 2001	F	ES
Unexcused Absences	3.00	1.96	4.47	3.33	.047	0.00
Unexcused Tardiness	3.35	3.00	4.81	4.89	.453	0.03
ISAPs	0	.01	0	.03	.011	0.00
Total Days Suspended	0	.03	0	.04	.015	0.00
Total Number of Suspensions	0	.02	0	.01	.184	0.00
GPA+	-	-	-	-	-	-

Note: * = results are significant at $p < .05$; ** = results are significant at $p < .01$; n/a = not applicable; + = GPA data not maintained for elementary school students.

Long-term Effects of Primary Mental Health

Table 17 presents the results of the analysis of long-term effects of PMH. As in the analysis of short-term effects, there were no statistically significant group differences noted in the long-term analysis of the effects of PMH.

Table 17. Change in non-cognitive and cognitive outcomes for students in PMH and matched controls from Spring 2000 to Spring 2002.

	Control (n = 128)		PMH (n = 129)			
	Spring 2000	Spring 2002	Spring 2000	Spring 2002	F	ES
Unexcused Absences	2.82	3.65	4.04	5.82	1.27	0.07
Unexcused Tardiness	2.95	4.87	4.26	6.21	.00	0.00
ISAPs	.00	.00	.00	.00	.00	.00
Total Days Suspended	.00	.01	.00	.08	1.13	0.06

Table continues...

Table 17 continued. Change in non-cognitive and cognitive outcomes for students in PMH and matched controls from Spring 2000 to Spring 2002.

	Control (n = 128)		PMH (n = 129)			
	Spring 2000	Spring 2002	Spring 2000	Spring 2002	F	ES
Total Number of Suspensions	.00	.01	.00	.05	.525	0.04
GPA+						

Note: * = results are significant at $p < .05$; ** = results are significant at $p < .01$; *** = results are significant at $p < .001$;. n/a = not applicable; + = GPA data not maintained for elementary school students.

Short-term Effects of Big Brothers/Big Sisters Mentoring

Table 18 presents the results of the short-term effects of Big Brothers/Big Sisters (BB/BS) mentoring on non-cognitive outcomes. Since the students are at the elementary level, there was no cognitive outcome measure available for them.

Table 18. Change in non-cognitive and cognitive outcomes for students in BB/BS and matched controls from Spring 2000 to Spring 2001.

	Control (n = 151)		BBBS (n = 152)			
	Spring 2000	Spring 2001	Spring 2000	Spring 2001	F	ES
Unexcused Absences	2.84	2.05	3.81	2.72	.419	.03
Unexcused Tardiness	2.33	2.64	5.90	4.12	9.27**	.17
ISAPs	0	.02	0	.05	.443	.03
Total Days Suspended	0	.07	.03	.09	.028	0
Total Number of Suspensions	0	.02	.02	.04	0	0
GPA+	-	-	-	-	-	-

Note: ES = effect size; * = results are significant at $p \leq .05$; ** = results are significant at $p \leq .01$; δ = results indicate a trend in the appropriate direction. n/a = not applicable; + = GPA data not maintained for elementary school students.

The only significant difference between students in BB/BS and control students was for the change in unexcused tardiness. Control students exhibited a slight increase in unexcused tardiness, while students in BB/BS exhibited a decrease (effect size = .17). No other group differences were noteworthy.

Long-term Effects of Big Brothers/Big Sisters Mentoring

Table 19. Change in non-cognitive and cognitive outcomes for students in BB/BS and matched controls from Spring 2000 to Spring 2002.

	Control (n = 140)		BBBS (n = 141)			
	Spring 2000	Spring 2002	Spring 2000	Spring 2002	F	ES
Unexcused Absences	2.94	3.59	3.67	5.66	3.53	0.11

Table continues...

Table 19 continued. Change in non-cognitive and cognitive outcomes for students in BB/BS and matched controls from Spring 2000 to Spring 2002.

	Control (n = 140)		BBBS (n = 141)			
	Spring 2000	Spring 2002	Spring 2000	Spring 2002	F	ES
Unexcused Tardiness	2.43	5.17	6.01	7.52	.779	0.05
ISAPs	.00	.75	.00	.25	3.79	0.11
Total Days Suspended	.00	.46	.04	.84	.853	0.05
Total Number of Suspensions	.00	.15	.02	.23	.301	0.03
GPA+	-	-	-	-	-	

Note: ES = effect size; * = results are significant at $p \leq .05$; ** = results are significant at $p \leq .01$; *** = results are significant at $p \leq .01$; n/a = not applicable; + = GPA data not maintained for elementary school students.

The results of the long-term outcome analysis for BB/BS are presented above in Table 19. None of the group differences was significant in this analysis.

Short-term Effects of Multi-Systemic Therapy

Table 20 presents the short-term effects of Multi-Systemic Therapy (MST) on non-cognitive and cognitive outcomes.

Table 20. Change in non-cognitive and cognitive outcomes for students in MST and matched controls from Spring 2000 to Spring 2001.

	Control (n = 52)		MST (n = 57)			
	Spring 2000	Spring 2001	Spring 2000	Spring 2001	F	ES
Unexcused Absences	6.11	7.74	11.36	9.35	2.25	0.15
Unexcused Tardiness	4.67	8.21	7.98	5.60	5.82*	0.23
ISAPs	2.76	3.26	1.88	1.12	1.26	0.11
Total Days Suspended	.47	.76	1.90	1.83	.111	0.03
Total Number of Suspensions	.17	.26	.53	.44	.442	0.06
GPA	1.58	1.65	.80	1.01	.146	0.04

Note: ES = effect size; * = results are significant at $p \leq .05$; ** = results are significant at $p \leq .01$. *** = results are significant at $p \leq .001$; δ = results indicate a trend in the appropriate direction. GPA data are based on N = 40 control students (77%) and N = 32 students in MST (56%).

A significant group by time interaction was found for unexcused tardiness. Students in MST exhibited a reduction in unexcused tardiness relative to matched controls, who exhibited an increase (effect size = .23). No other group differences were significant.

Long-term Effects of Multi-Systemic Therapy

Table 21 on the next page presents results of the analysis of the long-term effects of MST on non-cognitive and cognitive outcomes. Several long-term results are noteworthy. Significant group differences in the change in unexcused absences (effect size = .33), number of days

suspended (effect size = .21), and total number of suspensions (effect size = .25) were found. Relative to controls, students in MST exhibited decreases in each of these outcomes.

Table 21. Change in non-cognitive and cognitive outcomes for students in MST and matched controls from Spring 2000 to Spring 2002.

	Control (n = 52)		MST (n = 55)			
	Spring 2000	Spring 2002	Spring 2000	Spring 2002	F	ES
Unexcused Absences	5.92	13.37	11.95	7.36	12.41***	0.33
Unexcused Tardiness	4.58	11.24	8.36	11.81	.442	0.06
ISAPs	2.70	2.60	2.01	2.42	.10	0.03
Total Days Suspended	.47	1.45	1.90	.62	4.67*	0.21
Total Number of Suspensions	.17	.52	.53	.13	6.92**	0.25
GPA	1.68	1.61	1.11	1.12	.029	0.03

Note: ES = effect size; * = results are significant at $p \leq .05$; ** = results are significant at $p \leq .01$. *** = results are significant at $p \leq .001$; GPA data are based on N = 38 control students (73%) and N = 9 students in MST (16%).

SUMMARY, LIMITATIONS, AND CONCLUSIONS

In October 1999, Project SHIELD (Supporting Healthy Individuals and Environments for Life Development) received nearly $3,000,000 from a consortium of federal agencies (Department of Education, Office of Juvenile Justice and Delinquency Prevention, and Center for Mental Health Services) as part of a Safe Schools/Healthy Students Federal Initiative. The award will provide three years of funding (nearly $9,000,000) to the school district. Project SHIELD aims to provide students and schools with enhanced infrastructure and comprehensive prevention and early intervention, through education, mental health, and social services that promote healthy childhood development and prevent violence, alcohol and other drug abuse. These services target the development of social skills and emotional resilience necessary for youth to avoid violent behavior and drug use, along with establishing safe, disciplined, and drug free areas within school environments. To that end, Project SHIELD is implementing 17 research-based violence/substance abuse prevention interventions and best practices. An outcome evaluation was conducted for a selected set of research-based and best practices interventions using secondary data available in the school district's Teradata system, a data warehouse maintained centrally.

Outcome data available in Teradata include non-cognitive measures (e.g., absences, tardiness, alternative disciplinary actions [ISAPs], and suspensions) and, where available, cognitive measures (student grades). These outcomes are the most relevant to school district administrators and other school officials who are key to sustaining these interventions, and for purposes of accountability (Munoz, 2002; Wholstetter, 1991). In most outcome evaluation studies of violence prevention programs found in the literature, the criteria used to determine program effectiveness are changes in knowledge, attitudes, self-reported aggressive behaviors, often in conjunction with other self-report scales assessing aggression (see reviews by Cooper et al., 2000 and Mytton et al., 2002). Student behavioral and academic performance indicators are seldom used as outcome criteria (Farrell et al., 2001; Housman, Pierce, & Briggs, 1996;

Twemlow et al., 2001). Thus, one purpose of this evaluation was to examine the effectiveness of the interventions implemented through Project SHIELD on outcomes which are relevant to instilling and maintaining a safe and orderly school environment.

Two classes of interventions were evaluated: universal interventions, in which participation is a consequence of class attendance, and selected interventions, in which participation is based upon referral. The study used a pre-post matched control group design to analyze secondary data obtained from Teradata. The matching procedure used, discussed in greater detail in the method section, is innovative in that specially designed software was used to individually match control students to SHIELD students on important demographic characteristics (ECE status, gender, and age). Such matching lends a level of rigor to the evaluation of intervention effects that is seldom seen in such research, and increases confidence in the internal validity of the results (Cook & Campbell, 1979). The evaluation examined both short-term and long-term intervention effects. Short-term effects were defined as the change in outcomes during the first year of Project SHIELD, that is, Spring 2000 to Spring 2001. Long-term effects were defined as changes in outcomes over two years of Project SHIELD implementation, that is, Spring 2000 to Spring 2002.

The results provide mixed evidence for the effectiveness of Project SHIELD interventions on student behavior and performance. For the universal interventions, SMART Moves and Second Step, there was weak evidence that these interventions were having positive impacts on student behavior. Students in SMART Moves exhibited a smaller magnitude increase in ISAPs (in-school disciplinary referrals), relative to matched controls, in the analysis of long-term effects. Students in Second Step exhibited deceases in unexcused absences for both short-term and long-term analyses, compared to matched controls who exhibited little change in this outcome. There was also evidence that SMART Moves was not having the desired impact on other behaviors, however. The SMART Moves group exhibited increases in unexcused tardiness, number of days suspended, and number of suspensions in the analysis of short-term effects.

There was also evidence that some of the selected interventions are having positive effects on student behavior. LEEP students exhibited decreased unexcused tardiness, relative to matched controls between Spring 2000 and Spring 2001. Moreover, students in LEEP exhibited a decrease in ISAPs from the Spring 2000 to Spring 2002 semesters. Students in FFT exhibited a smaller magnitude of increase in ISAPs, relative to controls, in the analysis of long-term effects. Students in BBBS showed a decrease in unexcused tardiness, relative to the increase exhibited by their controls from Spring 2000 to Spring 2001. Stronger effects were seen in the MST group. Relative to control students, students in MST exhibited a decrease in unexcused tardiness in the short-term, and decreases in unexcused absences, number of days suspended, and total number of suspensions in the long-term analysis.

Little evidence of improvement in outcomes was found in the short-term analysis of FFT, however. Relative to the control group, students in FFT exhibited increases in unexcused absences, unexcused tardiness, and number of days suspended from Spring 2000 to Spring 2001. For PMH, there were no significant differences between students in the intervention group and their matched controls.

Finally, there were no significant effects found for any of the interventions in regard to academic performance; however, many of the students did not have valid academic data. As a result, the evaluation was hindered in its attempts to study changes in academic performance.

Limitations

This study was not conducted as a randomized, controlled trial, which limits the ability to reach firm causal conclusions about intervention effectiveness. Nevertheless, we partially compensated for this limitation by employing a rigorous computerized matching procedure to create individual student controls who were matched to intervention students on such characteristics as learning disability, gender, and age. Thus, the study has strong internal validity, and differential group change may be attributed more to the interventions than to potentially confounding intervening factors (such as differential maturity, history, and the like discussed by Cook and Campbell, 1979). Our ability to randomly assign students to intervention and control groups was limited by the scope and comprehensiveness of Project SHIELD. Future evaluation studies of large comprehensive intervention projects should use similar procedures to control for extraneous factors in the absence of random assignment.

A second limitation of the study is the amount of missing academic data for students. Initially, the investigation planned to use standardized test scores in order to evaluate intervention effects on student academic performance. As the evaluation unfolded, however, it came to our attention that standardized tests are not administered in a manner that would allow for an examination of change in score. For example, standardized tests in mathematics, reading, and language arts are given to students in grades 3, 6, and 9. The following year, a different standardized test is given to students in grades 6, 7, and 10 for reading and science; however, the reading test is not comparable to the reading test taken the prior year. Students in grades 2 and 3 are given standardized mathematics and reading tests for diagnostic rather than achievement purposes. Thus, standardized test scores could not be used to measure change in academic proficiency for students longitudinally. Semester grades in core content areas (mathematics, english, social studies, and science) could serve as measures of change in academic performance over time, and were chosen as the next best indicator. Nevertheless, necessary academic information for many of the students in both intervention and control groups were not available in the central data warehouse. In addition, elementary school students are not given grades in the traditional sense. As a result, student performance measures (grades) for elementary schools are not maintained in the central data warehouse. As a result of the amount of missing academic data, changes in student performance could not be adequately measured.

A third limitation is that the amount of individual exposure to a particular intervention could not be gauged; thus, the average change in outcomes within the intervention groups represents changes for students who have various levels of exposure to the intervention. It is not possible to tease apart students who completed the intervention in question from students who may have participated for only a short period of time. As a result, changes in outcomes may have been attenuated by varying degrees of exposure to the intervention being evaluated. This limitation is more reflective of the evaluation of the selective interventions than of the universal interventions, however. It would be of benefit to future evaluations of large-scale intervention projects to ensure that some mechanism is in place to assist investigators in tracking individuals exposed to interventions in order to examine differential effects of exposure on outcome.

Finally, as noted previously, the present study was designed to examine the effect of violence prevention interventions on indicators of global behavioral and performance measures, rather than to examine outcomes specifically targeted by the interventions (e.g., social skill development, emotional resiliency, conflict resolution skills, knowledge of consequences of violence, and norms related to use of violence). As a result, theoretically relevant intervention

effects have been overlooked in favor of examining school-relevant outcomes. Future studies should include both types of outcome measure in order to more fully examine intervention effects.

Conclusions

The majority of the interventions evaluated in this study appear to have small, positive effects on school-relevant behaviors; however, the results are mixed. With the exception of the two universal interventions, there was little consistency in results between short-term and long-term effects.

Nevertheless, the results indicate that the interventions implemented through Project SHIELD appear to lead to modest changes in school-relevant behaviors, and further study is warranted in order to examine more fully which intervention has the greatest impact on student behaviors at school.

REFERENCES

Aber, J.L., Jones, S. M., Brown, J. L., Chaudry, N., & Samples, F. (1998). Resolving conflict creatively: Evaluating the developmental effects of a school-based violence prevention program in neighborhood and classroom context. *Development and Psychopathology*, 10, 187-213.

Avery-Leaf, S., Cascardi, M., O'Leary, K. D., & Cano, A. (1997). Efficacy of a dating violence prevention program on attitudes justifying aggression. *Journal of Adolescent Health, 21*, 11-17.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally.

Cooper, W. O., Lutenbacher, M., & Faccia, K. (2000). Components of effective youth violence prevention programs for 7- to 14-year olds. *Archives of Pediatric and Adolescent Medicine, 154*, 1134-1139.

DuRant, R. H., Barkin, S., & Crowchuk, D. P. (2001). Evaluation of a peaceful conflict resolution and violence prevention curriculum for sixth-grade students. *Journal of Adolescent Health, 28*, 386-393.

DuRant, R. H., Treiber, F., Getts, A., McCloud, K., Linder, C. W., Woods, E. R. (1996). Comparison of two violence prevention curricula for middle school adolescents. *Journal of Adolescent Health, 19*, 111-117.

Farrell, A. D., Meyer, A. L., & White, K. S. (2001). Evaluation of Responding in Peaceful and Positive Ways (RIPP): A school-based prevention program for reducing violence among urban adolescents. *Journal of Clinical Child Psychology, 30*(4), 451-463.

Hall G. E., & Hord S. M. (2001). *Implementing Change.* Boston: Allyn and Bacon.

Hausman, A., Pierce, G., & Briggs, L. (1996). Evaluation of comprehensive violence prevention education: Effects on student behavior.

Hausman, A. J., Spivak, H., & Prothrow-Stith, D. (1995). Evaluation of a community-based youth violence prevention project. *Journal of Adolescent Health, 17*, 353-359.

Mytton, J. A., DiGuiseppi, C., Gough, D. A., Taylor, R. S., & Logan, S. (2002). School-based violence prevention programs. *Archives of Pediatric and Adolescent Medicine, 156*, 752-762.

Munoz, M. A. (Fall 2002). Accountability in education: Equity as learning for all. Reality Test. American Educational Research Association, Division H School Evaluation and Program Development. Available on-line at www.aera.net/divisions/h

Oprinas, P., Parcel, G. S., McAlister, A., & Frankowski, R. (1995). Violence prevention in middle schools: A pilot evaluation. *Journal of Adolescent Health, 17*, 360-371.

Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation: A systematic approach* (6th ed.). Thousand Oaks, CA: Sage.

Tabachnik, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York, NY: HarperCollins Publishers, Inc.

Twemlow, S. W., Fonagy, P., Sacco, F. C., Gies, M. L., Evans, R., & Ewbank, R. (2001). Creating a peaceful school learning environment: A controlled study of an elementary school intervention to reduce violence. *American Journal of Psychiatry, 158*(5), 808-810.

Webster, D. (1991). The unconvincing case for school-based conflict resolution. *Health Affairs*, 12, 126-141.

Wohlstetter, P. (1991). Accountability mechanisms for state education reform: Some organizational alternatives. *Educational Evaluation and Policy Analysis, 13*, 31-48.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC Educational Resources Information Center

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Evaluating a Safety Net for At-Risk Students

Author(s): William P. Neace, Marco A. Munoz, John Weber, and Knowlton Johnson

Corporate Source: Jefferson County Public Schools; Pacific Institute for Research and Evaluation

Publication Date: November 2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

X

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature: M. Munoz	Printed Name/Position/Title: Marco A. Munoz, Evaluation Specialist
Organization/Address: 3332 Newburg Rd, Louisville, KY 40218	Telephone: / FAX:
	E-Mail Address: / Date: 11/6/02

(Over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
UNIVERSITY OF MARYLAND
1129 SHRIVER LAB
COLLEGE PARK, MD 20742-5701
ATTN: ACQUISITIONS

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfacility.org