Journal of Research on Educational Effectiveness, 3: 26–55, 2010
Copyright © Taylor & Francis Group, LLC
ISSN: 1934-5747 print / 1934-5739 online
DOI: 10.1080/19345740903353436

Impact of a Social-Emotional and Character Development Program on School-Level Indicators of Academic Achievement, Absenteeism, and Disciplinary Outcomes: A Matched-Pair, Cluster-Randomized, Controlled Trial

Frank Snyder and Brian Flay
Department of Public Health, Oregon State University, Corvallis, Oregon, USA

Samuel Vuchinich, Alan Acock, and Isaac Washburn
Human Development and Family Studies, Oregon State University, Corvallis, Oregon, USA

Michael Beets
Department of Exercise Science, University of South Carolina, Columbia, South Carolina, USA

Kin-Kit Li
Department of Applied Social Studies, City University of Hong Kong, Hong Kong

Abstract: This article reports the effects of a comprehensive elementary school-based social-emotional and character education program on school-level achievement, absenteeism, and disciplinary outcomes utilizing a matched-pair, cluster-randomized, controlled design. The *Positive Action* Hawai‘i trial included 20 racially/ethnically diverse schools (M enrollment = 544) and was conducted from the 2002–03 through the 2005–06 academic years. Using school-level archival data, analyses comparing change from baseline (2002) to 1-year posttrial (2007) revealed that intervention schools scored 9.8% better on the TerraNova (2nd ed.) test for reading and 8.8% on math, that 20.7% better in Hawai‘i Content and Performance Standards scores for reading and 51.4% better in math, and that intervention schools reported 15.2% lower absenteeism and fewer suspensions (72.6%) and retentions (72.7%). Overall, effect sizes were moderate to large (range = 0.5–1.1) for all of the examined outcomes. Sensitivity analyses using permutation models and random-intercept growth curve models substantiated results. The results provide evidence that a comprehensive school-based program, specifically

Address correspondence to Frank Snyder, Department of Public Health, 254 Waldo Hall, Corvallis, OR 97331-6406, USA. E-mail: snyderf@onid.orst.edu

developed to target student behavior and character, can positively influence school-level achievement, attendance, and disciplinary outcomes concurrently.

Keywords: Randomized experiment, matched-pair, academic, achievement, discipline, social and character development

INTRODUCTION

Education has an urgent need to learn more about the role of behavior, social skills, and character in improving academic achievement (Eccles, 2004; Meece, Anderman, & Anderman, 2006). Since the No Child Left Behind Act passed, education has been focused on teaching to core content standards to improve academic achievement scores, particularly in reading and mathematics, for which schools are being held accountable (Hamilton et al., 2007). Teaching to, and support for, the behavioral, social, and character domains have been relegated to no or limited dedicated instructional time (Greenberg et al., 2003). Nevertheless, schools are expected to prevent violence, substance use, and other disruptive behaviors that are clearly linked to academic achievement (Fleming et al., 2005; Malecki & Elliott, 2002; Wentzel, 1993). The prevalence of discipline problems, for example, correlates positively with the prevalence of violent crimes within a school (Heaviside, Rowland, Williams, & Farris, 1999) which, in turn, affects attendance and academic achievement (Eaton, Brener, & Kann, 2008; Walberg, Yeh, & Mooney-Paton, 1974). Further, mental health concerns become more prevalent as students move into adolescence and can contribute to behavioral problems that detract from academic achievement (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003). Disciplinary problems (Dinks, Cataldi, & Lin-Kelly, 2007; Eaton, Kann, et al., 2008; Eisenbraun, 2007) and underachievement abound (Coalition for Evidence-Based Policy, 2002; Perie, Moran, & Lurkus, 2005; Snyder, Dillow, & Hoffman, 2008).

To address these needs, numerous school-based programs have been developed to target problems of academic achievement (Slavin & Fashola, 1998; What Works Clearinghouse, n.d.). In addition, many other types of programs have offered the promise of improving academic performance indirectly through a focus on specific problem behaviors, such as substance use and violence (Battistich, Schaps, Watson, Solomon, & Lewis, 2000; Biglan et al., 2004; DuPaul & Stoner, 2004; Elias, Gara, Schuyler, Branden-Muller, & Sayette, 1991; Flay, 1985, 2009a, 2009b; Horowitz & Garber, 2006; Peters & McMahon, 1996; Sussman, Dent, Burton, Stacy, & Flay, 1995; Tolan & Guerra, 1994). Although some of these programs are promising, most are problem specific and tend to address only the microlevel or proximal predictors (e.g., attitudes toward a behavior) of a single problem (e.g., violent behavior; Catalano, Hawkins, Berglund, Pollard, & Arthur, 2002), not the multifaceted ultimate (e.g., safety of neighborhood) and distal (e.g., bonding to parents) factors

that influence many other important outcomes (Flay, 2002; Flay, Snyder, & Petraitis, in press; Petraitis, Flay, & Miller, 1995; Romer, 2003) Consequently, programs have had limited success (Catalano et al., 2002; Flay, 2002).

As practitioners, policymakers, and researchers have implemented programs and sought to raise academic achievement and address negative behaviors among youth, an increasing amount of evidence indicates a relationship among multiple behaviors (Botvin, Griffin, & Nichols, 2006; Botvin, Schinke, & Orlandi, 1995; Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2004; Flay, 2002). Several mechanisms involving multiple behaviors have been identified in improving student behavior and performance (Greenberg et al., 2003; Zins, Weissberg, Wang, & Walberg, 2004). This suggests that key behaviors do not exist in isolation from each other. Moreover, prevention research offers ample empirical support showing that many youth outcomes, negative and positive, are influenced by similar risk and protective factors (Catalano et al., 2004; Catalano et al., 2002; Flay, 2002). That is, most, if not all, behaviors are linked (Flay, 2002). For example, the early initiation of alcohol and cigarette use and/or abuse is associated with lower academic test scores (Fleming et al., 2005). Further, early initiation of substance use and sexual activity can place youth at a greater risk of mental health disorders and aggressive behaviors (Gustavson et al., 2007; Hallfors, Waller, Bauer, Ford, & Halpern, 2005) and continuation of substance use through adolescence and into adulthood (Merline, O'Malley, Schulenber, Bachman, & Johnston, 2004).

Subsequently, there has been a movement toward more integrative and comprehensive programs that address multiple co-occurring behaviors and that involve families and communities. Such programs generally appear to be more effective (Battistich et al., 2000; Catalano et al., 2004; Derzon, Wilson, & Cunningham, 1999; Elias et al., 1991; Flay, 2000; Flay, Graumlich, Segawa, Burns, & Holliday, 2004; Hawkins, Catalano, Kosterman, Abbott, & Hill, 1999; Hawkins, Catalano, & Miller, 1992; Kellam & Anthony, 1998; Lerner, 1995). One of these programs currently being used nationally is the *Positive Action* (*PA*) program. *PA* is a comprehensive schoolwide social-emotional and character development (SACD) program (Flay & Allred, 2003; Flay, Allred, & Ordway, 2001) developed to specifically target the positive development of student behavior and character.

Based on prior studies, *PA* has been recognized in the character-education report by the U.S. Department of Education's What Works Clearinghouse (2007) as the only "character education" program in the nation to meet the evidentiary requirements for improving both academics and behavior. Preliminary findings indicate that *PA* can positively influence school attendance, behavior, and achievement. Two previous quasi-experimental studies utilizing archival school-level data (Flay & Allred, 2003; Flay et al., 2001) reported beneficial effects on student achievement (e.g., math, reading, and science) and serious problem behaviors (e.g., suspensions and violence rates).

The first study (Flay et al., 2001) used School Report Card (SRC) data from two school districts that had used *PA* within a number of elementary schools for several years in the 1990s. Schools were rank ordered on poverty and mobility, and each *PA* school was matched with the best matched non-*PA* school(s) having similar ethnic distribution. Results indicated that *PA* schools scored significantly better than the non-*PA* schools in their percentile ranking of fourth-grade achievement scores and reported significantly fewer incidences of violence and lower rates of absenteeism. The second study (Flay & Allred, 2003) used a similar methodological approach but expanded the variables on which *PA* and non-*PA* schools were matched to include dependent variables (e.g., reading and math achievement) assessed before the introduction of *PA*. Results confirmed previous findings and demonstrated that involvement in *PA* during elementary school improved academic and disciplinary outcomes at both the elementary and secondary levels.

In sum, the prior quasi-experimental studies provide preliminary evidence regarding the effects of *PA* on academic achievement and disciplinary outcomes. However, these findings are in need of confirmation utilizing a randomized design (Flay, 1986; Flay et al., 2005), a standard considered vital before an intervention is ready for broad dissemination (Flay et al., 2005). Designs that use matching without random assignment leave open the possibility that variables other than those measured were responsible for observed posttest differences, rather than the intervention itself. In addition, the previous quasi-experimental studies lacked data on program implementation, a measurement that is desirable to ensure that implementation occurred and, if so, how well it occurred (Domitrovich & Greenberg, 2000; Durlak & DuPre, 2008; Flay et al., 2005).

Utilizing student self-report data from the current randomized trial, Beets and colleagues (2009) examined the preventive benefits of *PA* on rates of student self-report and teacher reports of student substance use, violence, and voluntary sexual activity. Results indicated lower rates of substance use, violence and sexual activity among students attending *PA* schools. Overall, this randomized trial (a) replicated findings from quasi-experimental studies regarding violence and substance use and (b) found that *PA* can also alter other behaviors, such as sexual activity, that the program does not address directly. Hence, even though *PA* did not teach sexual responsibility, for example, the SACD content produced effects on sexual activity. Previous results suggest a mechanism that leads *PA* to positively affect multiple outcomes, such as sexual responsibility and academic achievement, even though the program does not include explicit discussion of these outcomes.

The purpose of the present study was to apply a matched-pair, cluster-randomized, controlled design to evaluate the effects of *PA* on school-level indicators of academic achievement, absenteeism, and disciplinary outcomes. School-level data are useful for estimating causal effects but are underutilized (Stuart, 2007). The present study builds on extant research and is the first to report the effects of *PA* on school-level outcomes from a randomized, controlled

design; thus, it provides the most rigorous test yet conducted for whether *PA* can improve school-level performance and greatly reduces the possibility that factors other than the *PA* intervention are responsible for observed posttest group differences. *PA* was hypothesized to result in decreased absenteeism, disciplinary referrals and grade retentions and improved academic achievement.

METHODS

Design and Sample

The *PA* Hawai‘i trial was a matched-pair, cluster-randomized, controlled trial, conducted during the 2002–03 through 2005–06 school years, with a 1-year follow-up in 2007, in Hawai‘i elementary schools. The state is one large school district with diverse ethnic groups and a recognized need for improvement (i.e., low standardized test scores and a high percentage of students receiving free or reduced-price lunch). The trial took place in 20 public elementary (K-5 or K-6) schools (10 matched-pairs) on three Hawai‘ian islands. Eligible schools for the study were those elementary schools that (a) were located on O‘ahu, Maui, or Moloka‘i; (b) were K-5 or K-6 community schools (were not academy, charter, or special education); (c) had at least 25% of students receiving free or reduced-price lunch; (d) were in the state’s lower three quartiles of standardized test scores; and (e) had annual student mobility rates under 20%, thereby ensuring that at least 40% of a selected cohort was still in the same school by the end of the trial. To ensure comparability of the intervention and control schools with respect to baseline measures, 2000 SRC data on 111 eligible schools were used to stratify schools into strata ranked on an index based on (a) demographic variables of percentage free or reduced-price lunch, school size, percentage stability, and ethnic distribution; (b) characteristics of the student populations such as percentage special education, and limited English proficiency; and (c) indicators of student behavior and performance outcomes such as standardized test scores, absenteeism, and suspensions (Dent, Sussman, & Flay, 1993; Flay et al., 2004; Graham, Flay, Johnson, Hansen, & Collins, 1984). Schools were matched based on their index score, resulting in 19 utilizable strata. Matched pairs were randomly selected from within strata, with one school of each pair randomly assigned to either the intervention or control condition before recruitment.

Starting with schools only on O‘ahu (to limit travel costs), intervention schools were asked to implement *PA*, whereas the control schools were asked to continue “business as usual” without making any substantial SACD reforms. Once it was evident that no additional schools could be recruited on O‘ahu, recruitment began using strata from Maui and Moloka‘i. The final sample of schools was representative of Hawai‘ian schools, though with higher stability

(as intended) and at higher risk (as intended) as indicated by percent free or reduced-price lunch and standardized test scores, respectively.

Intervention schools were offered the complete *PA* program free of charge and control schools were offered a monetary incentive during the randomized trial and the *PA* program upon completion of the trial. Three of the 10 control schools chose to receive the *PA* program after the formal trial; they were treated as controls at the follow-up to the present study, as anecdotal evidence suggests that they did not fully implement the program, and it is likely that schools need several years to fully implement a comprehensive program to see substantial benefits (Beets et al., 2009; Li et al., 2009).

Program Overview

The *Positive Action* program (http://www.positiveaction.net) is a comprehensive, schoolwide SACD program designed to improve academics, student behaviors, and character (Flay & Allred, in press). The program, developed in 1977 by Carol Gerber Allred, Ph.D., and revised since then as a result of process and outcome evaluations, is grounded in a broad theory of self-concept (Purkey, 1970; Purkey & Novak, 1970); is consistent with integrative, ecological theories of health behavior such as the Theory of Triadic Influence (Flay & Petraitis, 1994; Flay et al., in press), and is described in detail elsewhere (Flay & Allred, 2003; Flay et al., 2001). The full *PA* program consists of K-12 classroom curricula, of which only the elementary curriculum was used in the present randomized trial; a schoolwide climate development component, including teacher/staff training by the developer, a *PA* coordinator's (principal's) manual, school counselor's program, and *PA* coordinator/committee guide; and family- and community-involvement programs.

The sequenced elementary curriculum consists of 140 lessons per grade, per academic year, offered in 15 to 20 min by classroom teachers. When fully implemented, the total time students are exposed to the program during a 35-week academic year is approximately 35 hr. Lessons cover six major units on topics related to self-concept (i.e., the relationship of thoughts, feelings, and actions) physical and intellectual actions (e.g., hygiene, nutrition, physical activity, avoidance of harmful substances, decision-making skills, creative thinking), social/emotional actions for managing oneself responsibly (e.g., self-control, time management), getting along with others (e.g., empathy, altruism, respect, conflict resolution), being honest with yourself and others (e.g., self-honesty, integrity, self-appraisal), and continuous self-improvement (e.g., goal setting, problem solving, courage to try new things, persistence). The classroom curricula utilize an interactive approach, whereby interaction between teacher and student is encouraged through the use of structured discussions and activities, and interaction between students is encouraged through structured or semi-structured small-group activities, including games, role plays,

and practice of skills. For example, students are asked how they like to be treated. Regardless of age, socioeconomic status, gender, or culture, students and adults suggest the same top values of respect, fairness, kindness, honesty, understanding/empathy, and love, consistent with others' findings (Nucci, 1997). These values are then adopted as the code of conduct for the classroom and school (Flay & Allred, in press).

The school-climate kit consists of materials to encourage and reinforce the six units of *PA*, coordinating schoolwide implementation. Included in the kit, the *PA* coordinator's (principal's) manual directs the use of materials such as posters, music, tokens, and certificates. It also includes information on planning and conducting assemblies, creating a *PA* newsletter, and establishing a *PA* committee to create a schoolwide *PA* culture. In addition, a counselor's program, implemented by school counselors, specializes in developing positive actions with students at higher risk and their classrooms, families, and the school as a whole. The family-involvement program is available in various levels of involvement and promotes the core elements of the classroom curriculum and reinforces schoolwide positive actions. The parent manual is designed for parents to use at home and includes materials that parallel the classroom curriculum. The present study did not include the more intensive family kit. The community-development component of *PA* was not used in this trial.

Prior to the beginning of each academic year, teachers, administrators, and support staff (e.g., counselors) attended *PA* training sessions conducted by the program developer. The training sessions lasted approximately 3 to 4 hr for the initial year, and 1 to 2 hr for each successive year. Booster sessions, conducted by the Hawai'i-based project coordinator and lasting approximately 30 to 50 min, were provided an average of once per academic year for each school. In addition, mini-conferences were held in February of each year to bring together five or six leaders and staff (e.g., principals, counselors, teachers) from each of the 10 participating schools to share ideas and experiences as well as to get answers to any concerns regarding implementing the program.

Data and Measures

Archival School-Level Indicators. Archival school-level data were obtained from the Hawai'i Department of Education (HDE) as part of the state's SRC data accountability system (HDE, n.d.-b), with different indicators available at different time points as shown in Table 1. The SRC data were included in schools' School Status and Improvement Report, designed to provide information on schools' performance and progress. Absenteeism, suspensions, retention in grade, and four academic achievement indicators served as the dependent variables for the present study; these were chosen because they were the publically available indicators of school performance. Corresponding classroom- and student-level data were not available due to privacy considerations.

Table 1. Timeline of the *Positive Action* Hawai'i trial and compilation of school-level indicators

	Year										
	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Positive Action trial						Baseline	x	x	x	x	†
School-level indicator											
Standardized Test Math[a] (% ≥ average)						x	x	x	x	x	x
Math HCPS II (% proficient)						x	x	x	x	x	x
Standardized Test Reading[a] (% ≥ average)						x	x	x	x	x	x
Reading HCPS II (% proficient)						x	x	x	x	x	x
Absenteeism (average days absent)	x	x	x	x	x	x	x	x	x	x	x
Suspensions (% of students)			x	x	x	x	x	x	x	x	x
Retentions (% of students)						x	x	x	x	x	x

Note. HCPS II = . Hawai'i Content and Performance Standards.

†*Positive Action* schools continued to implement the program after completion of the randomized trial; three control schools chose to receive *Positive Action* after the trial. [a]Standardized test scores included SAT (Stanford 9) for 2002–06 and TerraNova (2nd ed.) for 2007.

School-level performance is an appropriate measure of program effectiveness because the *PA* Hawai'i trial tested a schoolwide implementation of the program and whole schools were randomized to condition (Stuart, 2007).

The four school-level academic achievement variables included the Grade 5 math and reading standardized test (percentage scoring average or above; the HDE switched from the Stanford Achievement Test [SAT] to the TerraNova [2nd ed.] test at 1-year follow-up during the current study), and the Grade 4 math and reading Hawai'i Content and Performance Standards (HCPS II; percentage proficient). The math and reading SAT and TerraNova (2nd ed.) are national norm-referenced tests that are utilized by school districts in the United States to assess achievement of students from kindergarten through high school. The math and reading HCPS II were developed by the HDE through a collaborative process involving teachers and HDE curriculum specialists and represent the HDE performance standards to meet No Child Left Behind mandates (Hawai'i Department of Education, n.d.-a). The archival school-level academic achievement data were available continuously, from 2002 to 1-year post-trial, as intervention schools continued to implement the *PA* program. Achievement scores were not reported for one of the 10 pairs of schools because they had too few students at each grade level, so these schools were not included in the primary analysis. There were no missing data for the other dependent variables.

The other three school-level indicators used in this study included (a) absenteeism (average number of days absent per year, (b) suspensions (percentage suspended), and (c) retentions (percentage retained in grade, i.e., kept back a grade). Student suspensions may have occurred due to, for example, disorderly conduct, burglary, truancy, and contraband (e.g., possession of tobacco). Suspension data represent all grade levels at each school, and the retention variable included students who were retained in all grades except kindergarten. The archival school-level absenteeism data were available annually from 1997 to 2007, the suspension data from 1999 to 2007, and the retention data from 2002 to 2007.

Thus, the archival data utilized in the present analysis were collected from schools with a different student body each academic year, and intervention schools, over time, had increasing exposure to *PA*. For example, archival school-level data collected for *PA* schools during the 2005–2006 academic year represented schools with students who were exposed to the intervention for up to 4 years compared to the 2002–2003 academic year.

Implementation. As part of the *PA* Hawai'i trial, sufficient data from year-end process evaluation surveys were collected from teachers at the end of the second (2004), third (2005), and final year (2006) of program implementation and are described in detail elsewhere (Beets et al., 2008). We used three school-level implementation indicators related to program exposure and adherence: (a) exposure, measured by seven items (i.e., six items referred to the six units in the *PA* curriculum and asked about how often the teachers taught the concept

throughout the school day, and an additional item assessed the amount of *PA* workbooks and activity sheets used during a typical day); (b) classroom material usage, measured by three items (i.e., how often teachers used *PA* materials/activities); and (c) schoolwide material usage, measured by three items (i.e., how often *PA* materials/activities were used throughout the school). All item responses ranged from 1 (*never*) to 5 (*always*). Alpha reliabilities were adequate (Beets et al., 2008).

The three school-level implementation indicators and an overall school-level implementation indicator were calculated at the 2nd (2004), 3rd (2005), and final year (2006) of program implementation using several steps. First, based on teachers' responses to the items that comprised each of the different implementation indicators, we calculated mean teacher-level indicator scores. Second, using the teacher-level indicator scores, a mean school-level implementation indicator was calculated for every school each year. Last, an overall school-level implementation indicator was calculated by computing the mean across all schools for each year of program implementation.

During the spring of the final year of the 4-year randomized trial, data were collect from one school leader (i.e., principal, vice principal, counselor) from each treatment and control school regarding the SACD programs and/or activities that were conducted in their school during the prior 3 academic years. Respondents were asked to list up to 16 SACD programs. For each program, respondents indicated the number of weeks the program was offered, the amount of time (minutes) devoted to the program per week, and whether teachers attended/received training to deliver the program (yes/no).

Analyses

For our primary analysis, we used matched-paired *t* tests, Hedges' adjusted *g* as a measure of effect size (Grissom & Kim, 2005; Hedges & Olkin, 1985), and percentage relative improvement (RI). To assess the robustness of results, permutation tests and random-intercept growth curve models were used for sensitivity analyses. The random-effects growth curve models provide some statistical control beyond randomization for potentially confounding unmeasured variables in case randomization was not totally successful with 10 schools per condition. This battery of statistical approaches was used separately for each of the outcomes and was applied to end-of-study (2006) and 1-year post trial (2007) outcomes.

Primary Analysis. First, matched-paired *t* tests of difference scores were used to examine change in school-level outcomes by condition. For each outcome, two difference scores [posttest (2006) − baseline (2002) and 1-year post trial (2007) − baseline (2002)] were calculated for each pair of intervention and control schools and a paired *t* test was performed. In a randomized design,

the difference in means provides an unbiased estimate of the true average intervention effect (Stuart, 2007).

Second, effect sizes for absenteeism, suspensions, retentions, and each of the four achievement outcomes were calculated by subtracting the mean difference of control schools from the mean difference of *PA* schools and dividing by the pooled posttest standard deviation. Hedges' *g* (as well as other measures of effect size such as Cohen's *d* and Glass' *d*) has some positive bias; therefore, Hedges' approximately unbiased adjusted *g* was calculated. Moreover, the adjusted *g* is an appropriate effect size calculation when the sample size is small (Grissom & Kim, 2005). Effect sizes were examined at posttest and at one-year post trial and were interpreted as small (0.2), moderate (0.5) or large (0.8; Cohen, 1977).

In addition, we calculated RI as an indicator of effect size that may be more understandable to practitioners. RI is the posttest difference between groups minus the baseline difference between groups, divided by the control group posttest level, that is, $[(PA_{\text{mean}} - \text{C}_{\text{mean}})_{\text{posttest}} - (PA_{\text{mean}} - \text{C}_{\text{mean}})_{\text{baseline}}]/\text{C}_{\text{mean posttest}}$, expressed as a percentage.

Sensitivity Analysis. Subsequently, to avoid reliance on *t*-test assumptions alone and as a sensitivity analysis, permutation tests were conducted with Stata v10 *permute*, which estimates *p* values based on Monte Carlo simulations (Stata Corp., College Station, TX). Both paired *t* tests of differences and permutation models have demonstrated good performance in randomized trials when the number of pairs is small (Brookmeyer & Chen, 1998).

Last, random-intercept growth curve models (see the appendix) were conducted with Stata v10 *xtmixed* (Rabe-Hesketh & Skrondal, 2008) to account for all observations and to model school differences. That is, this allows a more complete analysis of the multiple waves of available data (five waves of data at posttest; six waves of data at 1-year post trial) and takes into account the pattern of change over time. The random-intercept model allows the intercept to vary between schools, which indicates that some schools tend to have, on average, better outcomes and other schools have worse outcomes. The random coefficient is fixed, which reflects that intervention effects are similar for all schools. To estimate effects with missing values present, full information maximum likelihood estimation was used which utilizes all available data to provide maximum likelihood estimation (Acock, 2005). For the present analyses, each growth curve involved approximately 100 observations (5 waves $\times$ 20 schools at posttest; 6 waves $\times$ 20 schools at 1-year post trial). Although this sample size is at the lower end of some suggested guidelines for this estimator, it is adequate as a supplementary sensitivity analysis, as different views exist regarding appropriate sample size (Singer & Willett, 2003).

For each outcome, from baseline through both posttest and 1-year post trial, we tested whether a quadratic term for time was significant using the likelihood-ratio (LR) test (Rabe-Hesketh & Skrondal, 2008). Through posttest,

results indicated that a quadratic model provided a significantly better fit for the data on reading HCPS II, LR $\chi^2(1) = 14.92$, $p < .001$, and absenteeism, LR $\chi^2(1) = 6.25$, $p < .05$. Through 1-year post trial, results showed that a quadratic model fit significantly better for math TerraNova, LR $\chi^2(1) = 4.04$, $p < .05$; reading TerraNova, LR $\chi^2(1) = 4.56$, $p < .05$; math HCPS II, LR $\chi^2(1) = 17.04$, $p < .001$; and absenteeism, LR $\chi^2(1) = 19.39$, $p < .001$.

For the remaining outcomes (school suspensions and retentions), from baseline through both posttest and 1-year post trial, we conducted random-intercept Poisson models with Stata v10 *xtpoisson* (Rabe-Hesketh & Skrondal, 2008). As is common with elementary school-level data, frequency distributions for school suspensions and retentions were skewed at both posttest and 1-year post trial. Hence, a random-intercept Poisson model was used to account for this skewed distribution. The mean and variance of the suspension and retention variables were similar through posttest (suspensions [$M = 0.95$; variance = 1.09], retentions [$M = 0.99$; variance = 0.92]) and 1-year post trial (suspensions [$M = 1.07$; variance = 1.72], and retentions [$M = 0.94$; variance = 0.88]), an assumption of the Poisson model (Snijders & Bosker, 1999); therefore, we did not adjust for overdispersion. Similarly, as previously discussed, an LR test was used to compare random-intercept Poisson models with the inclusion of a quadratic term. Only the result for suspensions, LR $\chi^2(1) = 4.85$, $p < .05$, at 1-year post trial demonstrated a quadratic model provided a better fit for the data.

In addition, to test whether the pattern of curvilinear change was different in *PA* and control schools, a Year Squared $\times$ Condition interaction term was included in the quadratic models, and an LR test was performed. Results indicated that the inclusion of an interaction term did not significantly improve any of the quadratic models and, hence, was not included in the final models.

RESULTS

Baseline Equivalency

At the 2002 baseline no significant differences ($p \geq .05$) existed between intervention and control schools on any of the SRC variables (Table 2; Table 4 displays outcome variables). Thus, the methods of developing strata and random selection and assignment were effective for these variables. Schools were racially/ethnically diverse with a mean enrollment of 544 ($SD = 276.41$).

Implementation

There was some variability in school-level implementation between schools, with small improvements across years (Table 3). Regarding the three

Table 2. Characteristics of study schools at baseline and posttest

	2002 (Baseline)					2006 (Posttest)				
	Control		Positive Action			Control		Positive Action		
Characteristics	*M*	*SD*	*M*	*SD*	p^a	*M*	*SD*	*M*	*SD*	p^a
Enrollment	478.80	207.06	609.40	330.07	0.303	427.70	184.99	575.40	341.42	0.245
Racial/Ethnic Distribution (%)										
African American	1.79	3.20	1.66	2.03	0.915	2.28	3.35	1.76	2.75	0.709
Chinese	2.05	3.66	1.88	2.75	0.908	2.06	3.12	2.01	2.68	0.970
Filipino	11.61	14.20	15.83	9.75	0.449	8.46	12.39	16.80	10.35	0.120
Hawai‘ian	5.61	5.98	5.74	4.16	0.956	7.23	8.55	6.29	5.53	0.774
Hispanic	2.45	2.35	3.28	3.11	0.510	3.50	3.12	3.62	4.40	0.945
Indochinese	2.02	5.62	0.34	0.69	0.361	2.01	3.75	1.56	2.59	0.759
Japanese	4.26	3.57	6.50	6.16	0.333	3.77	2.69	5.57	6.09	0.404
Korean	1.19	2.12	1.71	3.50	0.692	1.07	1.78	1.55	3.10	0.676
Native American	0.44	0.37	0.47	0.47	0.876	0.73	0.64	0.55	0.53	0.503
Part Hawai‘ian	31.86	28.37	28.81	21.61	0.790	30.99	28.92	29.53	24.42	0.904
Portuguese	1.41	1.94	1.99	1.77	0.494	0.99	1.48	0.81	0.85	0.742
Samoan	3.11	4.83	5.23	8.78	0.512	3.1	4.03	4.83	7.96	0.547
White	17.52	18.05	13.05	10.81	0.510	15.51	18.89	12.27	11.15	0.646
Other	14.69	14.01	13.48	8.61	0.879	18.29	22.49	12.76	8.82	0.479
Stability (%)	90.82	2.36	91.71	3.18	0.487	89.92	4.44	88.35	4.31	0.433
Free/reduced lunch (%)	54.32	26.40	59.78	22.95	0.628	54.80	25.08	56.38	18.88	0.875
Limited English proficiency (%)	11.83	15.30	15.58	14.10	0.576	11.00	12.59	14.19	12.32	0.574
Special education (%)	10.56	5.41	9.76	2.99	0.687	8.16	2.29	9.88	4.15	0.267

[a]Two-tailed *t* test; 18 *df*.

Table 3. School-level implementation

	2004		2005		2006	
Indicator	M^a	*SD*	M^b	*SD*	M^b	*SD*
Exposure/amount	3.08	0.65	3.29	0.52	3.14	0.50
Classroom material usage	2.58	0.71	2.65	0.76	2.66	0.46
Schoolwide material usage	3.41	1.02	3.62	0.88	3.63	0.84
Overall	3.02	0.74	3.18	0.64	3.14	0.54

Note. Means correspond to item scale: 1 (*never*) to 5 (*always*).
$^aN = 8$. $^bN = 10$.

school-level indicators examined, schoolwide material usage demonstrated the highest school-level implementation. Implementation was adequate for each indicator; however, results indicated that schools could have implemented *PA* with greater fidelity.

We found that control schools reported implementing an average of 10.2 SACD programs compared with 4.2—in addition to *PA*—in the intervention schools. Teachers in control schools spent an average of 108 min per week on SACD-related activities. *PA*-school teachers spent the expected amount of time on *PA* (55.1 min/week), yet overall they still spent only 35 min/week more on SACD-related activities than teachers in control schools. Control schools reported that teachers were involved in SACD-related activities for an average of 24 weeks per school year. In contrast, teachers in intervention schools reported delivering *PA* almost every week of the school year as well as being involved in other SACD-related activities for 25 weeks/year. Both *PA* and control school teachers reported receiving training to implement approximately half of the SACD-related programs (52.3% and 53.3%, respectively) that they reported implementing other than *PA* (100% trained).

School-Level Raw Means

Raw means for school-level academic achievement, absenteeism, suspensions, and retentions are presented in Figures 1 and 2, respectively. Overall, for the academic achievement outcomes, raw means for *PA* and control schools were statistically similar at baseline and demonstrated a clearly discernable divergence over time. State averages for academic achievement are shown for comparison. Although the *PA* schools were well below state averages at baseline (as planned), they nearly met or exceeded the state averages for academic achievement at posttest and 1-year post trial.

Likewise, for the other school-level outcomes, *PA* and control schools diverged between baseline and posttest. For absenteeism and suspensions,

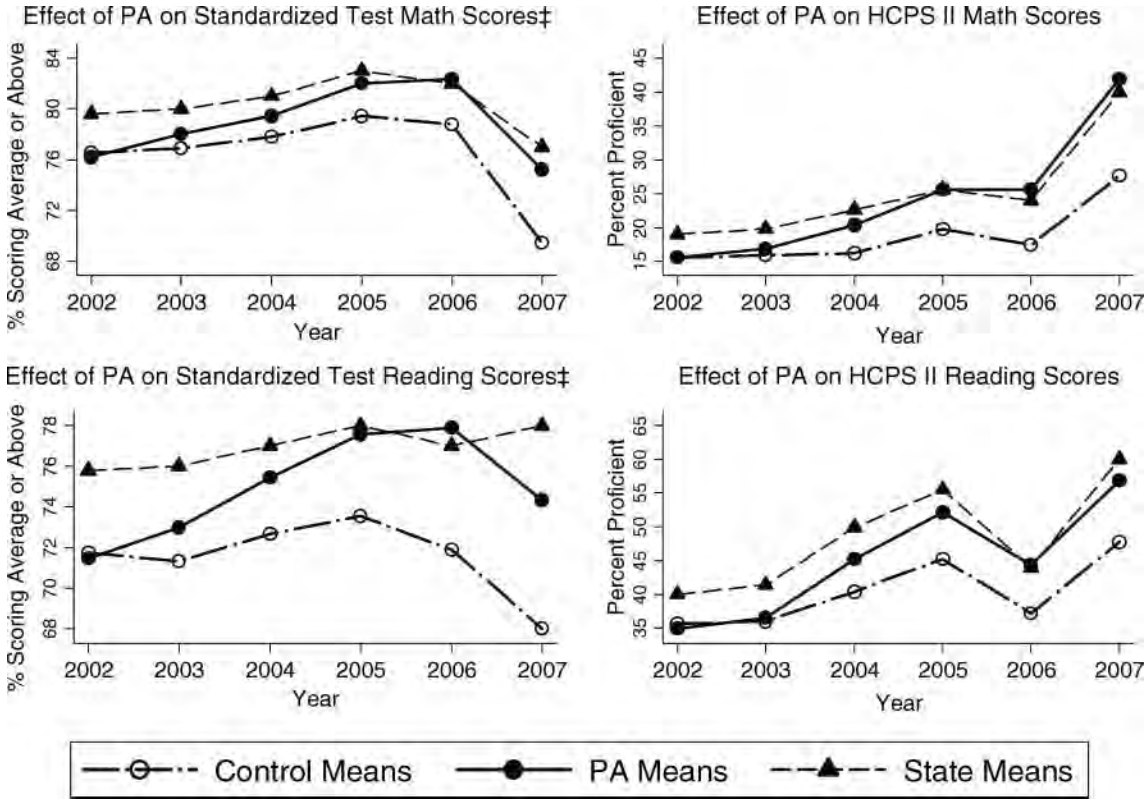


Figure 1. School-level means for math and reading achievement. *Note*. Hawai'i Randomized Trial occurred 2002–03 to 2005–06. ‡Standardized test scores included SAT (Stanford 9) for 2002–06 and TerraNova (2nd ed.) for 2007. *PA* = *Positive Action*; HCPS II = Hawai'i Content and Performance Standards.

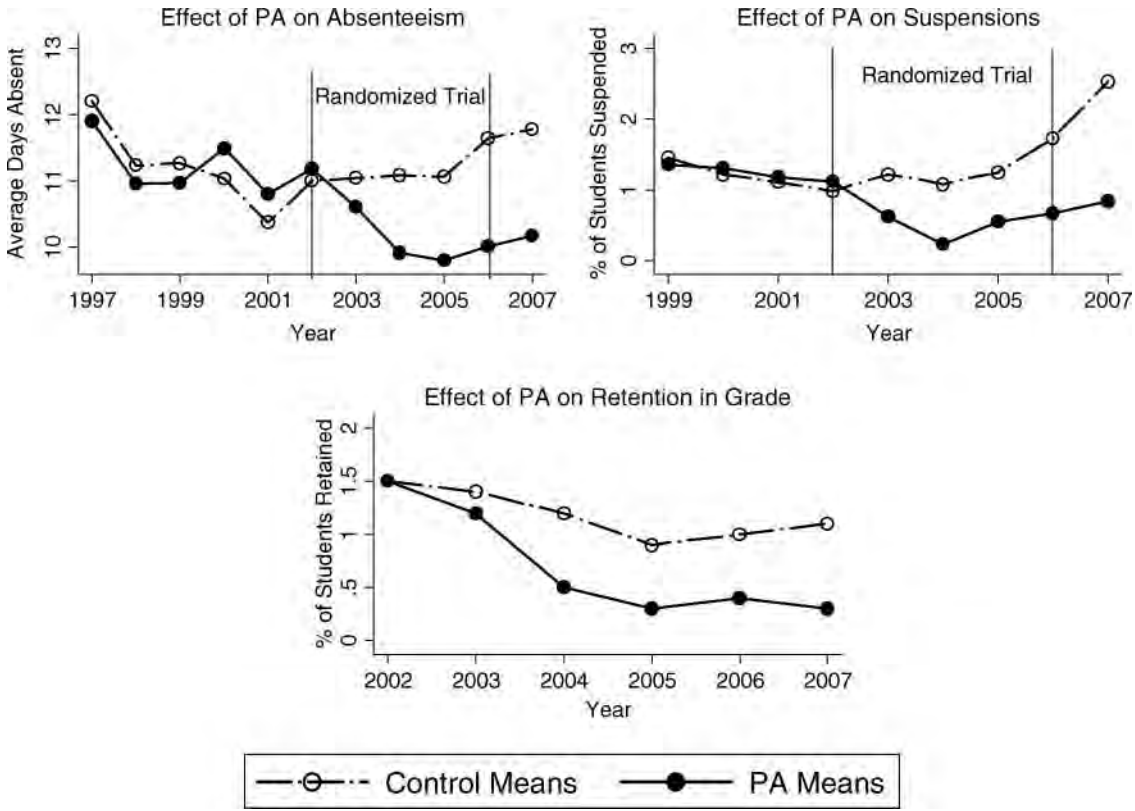


Figure 2. School-level means for absenteeism, suspensions, and retention. *Note*. Hawai‘i Randomized Trial occurred 2002–03 to 2005–06. *PA = Positive Action*.

prebaseline years of archival school-level data were available and provide an interrupted time series presentation. As expected, these outcomes were stable for several preprogram years with divergence occurring after the intervention.

Matched Paired *t* Tests and Effect Sizes

The results of the matched paired *t* tests of difference scores and effect size calculations at posttest and 1-year post trial are presented in Table 4. At posttest, results indicated that *PA* schools had significantly higher math ($p < .05$) and reading ($p < .05$) HCPS II scores; and significantly lower absenteeism ($p < .001$), with marginally fewer suspensions ($p = .056$). After completion of the randomized trial, at 1-year post trial as *PA* schools continued to implement the *PA* program, reading TerraNova ($p < .05$) and math ($p < .01$) and reading ($p < .05$) HCPS II were significantly higher among *PA* schools, and absenteeism ($p < .001$) and suspensions ($p < .05$) were significantly lower for *PA* schools. Overall, results indicated higher achievement and lower absenteeism and suspension outcomes for the *PA* schools. The permutation models provided similar statistically significant results as the matched paired *t* tests at both posttests. That is, permutation tests at posttest indicated statistically significant results for math (marginal $p = .054$) and reading ($p < .01$) HCPS II and absenteeism ($p < .01$), and at 1-year post trial reading ($p < .05$) TerraNova, math ($p < .001$), and reading ($p < .05$) HCPS II, absenteeism ($p < .001$), and suspensions ($p < .05$) were significantly different for *PA* schools as compared to control schools.

To provide a basis for comparing the magnitude of the intervention effects we found with effects found in other trials, effect sizes were calculated. As shown in Table 4, all of the effect sizes were moderate to large, regardless of the level of significance. Corresponding effect size calculations demonstrated moderate to large treatment effects for the academic achievement, absenteeism, and disciplinary outcomes at posttest, with larger effects at 1-year post trial. Similarly, RIs were larger at 1-year post trial.

Random-Intercept Growth Curve Models

The estimates for the intervention effect on academic achievement scores (random-intercept models) from baseline through posttest and 1-year post trial are presented in Table 5. At posttest, the intraclass correlation coefficient (ICC; expressed as the proportion of the total outcome variation that is attributable to differences among schools) for the unconditional means models (Singer & Willett, 2003) were. 72, .67, .87, and .72 for math SAT and HCPS II and reading SAT and HCPS II, respectively. At 1-year post trial, the ICC for the unconditional means models were .68, .46, .87, and .66 for math TerraNova and HCPS II and reading TerraNova and HCPS II, respectively,

Table 4. Baseline measures, school-level matched paired *t* tests of difference scores, and effect sizes for math achievement, reading achievement, absenteeism, suspensions, and retentions

		2002 (Baseline)			2006 (Posttest)						2007 (1-Year Post Trial)					
Outcome	Condition	*M*	*SD*	*p*[a]	*M*	*SD*	M_{diff}[b]	*p*[c]	ES[d]	RI[e]	*M*	*SD*	M_{diff}[b]	*p*[c]	ES[d]	RI[e]
Stand. Test Math[f]	Control	76.56	13.73	0.957	78.77	7.26	2.22	0.495	0.50	4.9%	69.44	11.57	–7.11	0.291	0.52	8.8%
(% ≥ average)	*PA*	76.22	11.72		82.33	7.48	6.11				75.22	10.85	–1.00			
Math HCPS II	Control	15.56	10.01	1.000	17.44	9.80	1.89	0.040	0.69	46.5%	26.67	7.79	12.11	0.006	1.10	51.4%
(% proficient)	*PA*	15.56	7.81		26.56	12.48	10.00				41.89	15.59	26.33			
Stand. Test Reading[f]	Control	71.74	13.36	0.962	71.89	12.29	0.14	0.108	0.58	8.8%	68.00	13.61	–3.74	0.028	0.54	9.8%
(% ≥ average)	*PA*	71.44	13.12		77.89	7.75	6.44				74.33	9.33	2.89			
Reading HCPS II	Control	35.67	16.67	0.904	37.22	10.81	1.56	0.029	0.72	21.2%	47.78	14.38	12.11	0.043	0.65	20.7%
(% proficient)	*PA*	34.89	9.37		44.33	10.00	9.44				56.89	14.55	22.00			
Absenteeism (Average	Control	11.00	2.27	0.872	11.64	3.17	0.64	0.001	−0.63	−15.5%	11.78	3.07	0.78	0.001	−0.65	−15.2%
days absent)	*PA*	11.18	2.66		10.01	2.21	−1.17				10.17	2.13	−1.01			
Suspensions	Control	0.98	1.11	0.777	1.72	1.55	0.74	0.056	−0.96	−69.4%	2.53	2.80	1.55	0.028	−0.87	−72.6%
(% of students)	*PA*	1.12	1.10		0.67	0.64	−0.45				0.84	0.61	−0.29			
Retentions (% of	Control	1.50	0.972	1.000	1.00	0.82	−0.50	0.313	−0.84	−60.0%	1.10	0.88	−0.40	0.210	−1.08	−72.7%
students)	*PA*	1.50	1.08		0.40	0.52	−1.10				0.30	0.48	−1.20			

Note. ES = effect size; RI = Relative Improvement; *: PA = Positive Action*; HCPS II = Hawai'i Content and Performance Standards.

[a]Two-tailed *t* test; 18 *df* for all variables except achievement-related variables (16 *df*). [b]Mean difference = posttest − baseline; 1-year post trial − baseline. [c]Two-tailed paired *t*-test difference score; 8 *df* for achievement-related variables and 9 *df* for other variables.

[d]Hedges' *g* effect size (unbiased adjusted *g*) of mean difference. [e]$RI = [(PA_{mean} - C_{mean})_{posttest} - (PA_{mean} - C_{mean})_{baseline}]/C_{mean\ posttest}$.

[f]Standardized test scores included Stanford Achievement Test (Stanford 9) for 2002–06 and TerraNova (2nd ed.) for 2007.

Table 5. School-level random-intercept growth model estimates for math and reading achievement

	Stand. Test Math[a] (% Average or Better)		Math HCPS II (% Proficient)		Stand. Test Reading[a] (% Average or Better)		Reading HCPS II (% Proficient)	
	B[b]	SE	*B*[b]	SE	*B*[b]	SE	*B*[b]	SE
2006 (Posttest)								
Fixed effects								
Intercept	75.79***	3.26	14.68***	2.97	71.48***	3.92	25.96***	4.68
Year	0.70	0.52	0.77	0.47	0.25	0.41	9.16***	2.02
Year^2							–1.32****	0.32
Condition (0 = C; 1 = *PA*)	–2.39	4.60	–2.50	4.21	–1.65	5.54	–2.92	5.79
Year × Condition	1.19†	0.73	2.10**	0.67	1.49**	0.58	2.21**	0.77
Random effects								
School-level variance	69.13	24.65	57.44	20.50	121.79	41.60	121.55	42.30
Residual variance	23.89	33.98	20.17	3.36	14.95	2.49	26.55	4.43
2007 (1-year post trial)								
Fixed effects								
Intercept	70.91***	3.76	20.21***	3.66	67.87***	4.08	33.18***	4.00
Year	5.49***	1.48	–4.28**	1.50	3.50**	1.09	1.96***	0.47
Year^2	–0.90***	0.20	0.89***	0.21	–0.57***	0.15		
Condition (0 = C; 1 = *PA*)	–2.73	4.41	–3.87	4.41	–1.33	5.43	–1.14	5.66
Year × Condition	1.34*	0.59	2.69***	0.60	1.35**	0.44	2.10**	0.66
Random effects								
School-level variance	71.63	25.41	62.67	22.49	119.62	40.70	121.72	40.85
Residual variance	27.48	4.10	28.55	4.26	14.91	2.22	35.09	5.22

Note. C = control; *PA* = *Positive Action*.

[a]Standardized test scores included SAT (Stanford 9) for 2002–06 and TerraNova (2nd ed.) for 2007. [b]B estimate based on random-intercept model.

†$p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$; all two-tailed.

indicating that most of the variation in academic achievement lies between schools, rather than within schools over time. Overall, through both posttest and 1-year post trial, the random-intercept models' Year × Condition interactions substantiated results of the matched-paired *t* tests and permutation models, indicating higher achievement increases in *PA* schools. For change from baseline through 1-year post trial, the time by condition interactions for math TerraNova ($B = 1.34$, $p < .05$) and HCSPII ($B = 2.69$, $p < .001$) and reading TerraNova ($B = 1.35$, $p < .01$) and HCPS II ($B = 2.10$, $p < .05$) were all statistically significant. These effects indicate about a 2 percentage point advantage per year for the *PA* group compared to the control group due to the intervention, or about a 12 percentage point advantage across the 6-year period.

The estimates for the intervention effect on the absenteeism, suspension, and retention outcomes (random-intercept and random-intercept Poisson models) from baseline through both posttest and 1-year post trial are presented in Table 6. Parameter estimates and incidence rate ratios (IRR) are each presented for the random-intercept Poisson models, as an intercept parameter is not calculated for IRR estimates and, additionally, a residual variance estimate is not part of such models (Rabe-Hesketh & Skrondal, 2008). At posttest, the ICCs for the unconditional means models were .88, .52, and .47 for absenteeism, suspensions, and retentions, respectively. The ICC values for the Poisson models are approximations and were calculated utilizing a similar approach as used for the random-intercept models (Goldstein, Browne, & Rasbash, 2002). At 1-year post trial, the ICCs for the unconditional means models were .88, .52, and .41 for absenteeism, suspensions, and retentions, respectively. Thus, much of the variation in absenteeism, nearly half of the variation in suspensions, and less than half the total variation in retentions can be attributable to differences between schools.

Regarding absenteeism, from baseline through both posttest (Year × Condition $B = -0.45$, $p < .001$) and 1-year post trial (Year × Condition $B = -0.36$, $p < .001$), the random-intercept growth models substantiated results of the matched-paired *t* tests, demonstrating a significant reduction in absenteeism among *PA* schools relative to control schools. However, as compared to the matched-paired *t* tests, inconsistent results emerged for the suspension and retention outcomes. The random-intercept growth curves indicated a marginally significant ($B = -0.20$, $p = .06$; IRR [95% CI] = 0.82 [0.67, 1.01]) Year × Condition interaction for the suspension outcome from baseline to 1-year post trial, where the *t* tests did not. Further, inconsistent with the nonsignificant matched-paired *t* test, the Retention Year × Condition interactions through posttest ($B = -0.30$, $p < .05$; IRR = 0.74 [0.54–1.00]) and 1-year post trial ($B = -0.30$, $p < .05$; IRR = 0.74 [0.58–0.95]) were statistically significant. Therefore, overall, the random-intercept and random-intercept Poisson models demonstrate decreased absenteeism, disciplinary, and retention outcomes among *PA* schools relative to control schools.

Table 6. School-level random-intercept growth model estimates for absenteeism, suspensions, retentions

	Absenteeism (Average Days Absent/Year)		Suspensions (% Suspended)				Retentions (% Retained)			
	B^a	SE	B^b	SE	IRR[c]	(CI)	B^b	SE	IRR[c]	(CI)
2006 (Posttest)										
Fixed effects										
Intercept	11.56***	0.80	–0.15	0.38			0.54	0.34		
Year	–0.54*	0.27	0.12	0.09	1.13	(0.37, 3.62)	–0.13	0.09	0.88	(0.74, 1.06)
Year^2	0.11**	0.04								
Condition (0 = C; 1 = *PA*)	0.47	1.04	0.15	0.58	1.16	(0.95, 1.35)	0.32	0.50	1.37	(0.52, 3.64)
Year × Condition	–0.45***	0.10	–0.28†	0.16	0.76†	(0.56, 1.03)	–0.30*	0.16	0.74*	(0.54, 1.00)
Random effects										
School-level variance	4.85	1.57	0.44	0.22	0.44	(0.16, 1.18)	0.32	0.18	0.32	(0.10, 0.97)
Residual variance	0.54	0.09								
2007 (1-year post trial)										
Fixed effects										
Intercept	11.52***	0.78	0.49	0.50			0.45	0.31		
Year	–0.45*	0.20	–0.39	0.26	0.68	(0.40, 1.14)	–0.08	0.07	0.92	(0.80, 1.05)
Year^2	0.09***	0.03	0.08*	0.04	1.08*	(1.01, 1.16)				
Condition (0 = C; 1 = *PA*)	0.28	1.05	–0.05	0.54	0.95	(0.33, 2.71)	0.32	0.47	1.38	(0.55, 3.43)
Year × Condition	–0.36***	0.08	–0.20†	0.10	0.82†	(0.67, 1.01)	–0.30*	0.12	0.74*	(0.58, 0.95)
Random effects										
School-level variance	5.01	1.61	0.54	0.23	0.54	(0.24, 1.23)	0.32	0.18	0.32	(0.11, 0.94)
Residual variance	0.54	0.08								

Note. C = control; *PA* = *Positive Action*; IRR = incidence rate ratio; CI = 95% confidence interval.

[a]*B* estimate based on random-intercept model. [b]*B* estimate based on random-intercept Poisson model. [c]IRR estimate based on random-intercept Poisson model.

†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$; all two-tailed.

DISCUSSION

The present study extends previous research on the capabilities of school-based interventions targeting social-emotional and character development to improve academic performance and attendance and reduce disciplinary problems and grade retention in schools. This study also confirms earlier preliminary findings of beneficial results of the *PA* program from quasi-experimental studies (Flay & Allred, 2003; Flay et al., 2001) using a matched-pair, cluster-randomized, controlled trial. Specifically, as indicated by matched-paired *t* tests and permutation models, *PA* schools scored significantly better than control schools in reading TerraNova and math and reading HCPS II and significantly lower absenteeism and suspensions at 1-year post trial. Moreover, random-intercept growth models demonstrated that *PA* schools showed significantly greater growth in math and reading TerraNova, math and reading HCPS II, and significantly lower absenteeism and retentions through 1-year post trial, with suspensions showing marginal significance. Indeed, school-level means for math and reading achievement demonstrated that *PA* schools, which were below state averages at baseline, nearly met or exceeded state averages by posttest and 1-year post trial. These findings were especially noteworthy because many of the schools were in low income areas and had a high level of racial/ethnic diversity.

The present results demonstrated moderate to large effect sizes on all of the observed outcomes and were likely the result of several notable attributes of the *PA* program. First, *PA* addresses distal influences on behavior in a multifaceted way; *PA* is a comprehensive approach that involves providing the curriculum to all grades in the school at once, involving all teachers and staff in the school, and involving parents and the community. The *PA* program assists students and adults to gain not only the knowledge, attitudes, norms, and skills that they might gain from other programs but also improved values, self-concept, family bonding, peer selection, communication, and appreciation of school, with the expected result of improvement in academic performance and a broad range of behaviors. These improved outcomes may occur because positive behaviors tend to correlate negatively with negative behaviors (Flay, 2002). More specifically, with regards to academic achievement, for example, *PA* increases positive behaviors and decreases disruptive behaviors, which in turn lead to more time on task for teaching and, in turn, more opportunity for student learning (Flay & Allred, in press). Also, improvements in students' positive behaviors, such as attention and inhibitory control, can lead to increased academic achievement throughout formal schooling (McClelland, Acock, & Morrison, 2006).

Second, *PA* is "interactive" in delivery, using methods that integrate teacher/student contact and communication opportunities for the exchange of ideas, and utilize feedback and constructive criticism in a nonthreatening atmosphere (Tobler et al., 2000). Third, the results observed may also have been a consequence of the intensive nature of the program, with students receiving approximately 1 hr of exposure during a typical week over multiple school years.

Last, in the present study, we believe that the beneficial effects of the *PA* program could have been even greater if the fidelity of implementation was excellent.

This analysis has some limitations. First, data regarding academic achievement, absenteeism, suspensions, and retention outcomes were not available at the student or classroom level. Because of this, variation in scores within students across years, or variation between students within schools could not be examined. As a result, individual student or classroom characteristics could not be included as predictors in the models to reduce unexplained variation. However, with random assignment, student and classroom characteristics should be about the same in the intervention and control groups. In addition, random-intercept models provide some statistical control for unmeasured differences between schools. Because every student's score contributes to a school's mean score, the design and analysis in this study provides a good test for intervention effects (Stuart, 2007). Future work that utilizes multilevel analysis of student-level indicators of academic achievement, absenteeism, and disciplinary outcomes would be beneficial.

Second, although school-level data are useful for estimating causal effects (Stuart, 2007), there may be inconsistencies among schools regarding how data, such as disciplinary-related referrals, are reported. Furthermore, it is possible that an intervention could influence how these data are reported. For example, a negative behavior that resulted in a disciplinary referral before an intervention may be handled in a different way after an intervention like *PA*.

A third limitation of our analyses is that only 20 schools participated in the study, with five waves of data resulting in 100 observations per random-effects growth curve model. Under conditions of small effect size and high ICC, this could result in relatively low statistical power to detect differences between treatment and control schools. This study found moderate to large effect sizes, but also large ICCs, so power was a concern. However, a successful matched-pair design can improve statistical power (Raudenbush, Martinez, & Spybrook, 2007), and our findings demonstrate a successful matched-pair design as well as its ability to detect statistical significance.

Fourth, there were a limited number of observations available for the random-effects growth curve models. With full information maximum likelihood estimation used in those models, a large sample is desirable (Hayes, 2006) to guarantee the accuracy of the estimates, although there are various viewpoints on what constitutes a large sample size (Singer & Willett, 2003). Our sample was large enough to use these models to compare the sensitivity of the matched-paired *t* tests and permutation tests to an alternative statistical model, with different assumptions. The random-intercept models substantiated our findings from the more basic tests.

Fifth, although we demonstrated adequate implementation of *PA* and realize the importance of implementation fidelity (Flay et al., 2005), we had insufficient data (i.e., insufficient variation given a sample of only 10 *PA* schools) to examine implementation as a covariate. Also, we did not have data to observe the change in SACD-related activities in control schools. As indicated

by the data procured during the last year of the 4-year trail, the widespread self-initiation of SACD-related activities, especially in control schools, can reduce the possible effect size that can be detected when evaluating school-based interventions (Hulleman & Cordray, 2009). In addition, because implementation data were not collected after completion of the randomized trial, we could not examine implementation at 1-year post trial. Future studies with larger samples of schools would be valuable to examine the effects of implementation fidelity on school-level outcomes.

Last, as with all other similar studies, results can be generalized only to schools that are willing to conduct such a program. Though our sample was adequate for this study, a larger representative sample of schools, or randomized trials at different locations, would allow generalization of results to a broader population.

These limitations notwithstanding, this study is the first to examine the effects of *PA* on school-level achievement, absenteeism, and disciplinary outcomes using a matched-pair, cluster-randomized, controlled design. The study extends research on the ways that changing a child's developmental status in nonacademic areas can significantly enhance academic achievement (Catalano et al., 2004; Catalano et al., 2002; Flay, 2002) and actually may be essential for it. Future research should examine the specific mechanisms, moderators and mediators of social and character development intervention effects. Such knowledge would allow adjustments to *PA* that might increase the beneficial effect.

Unfortunately, elementary schools, with many demands for accountability, may concentrate solely on math, reading, and science achievement, and due to resource and time constraints, instruction regarding social and character development may be abandoned. The findings of this study provide evidence that the *PA* program, which has demonstrated effects on improving student behavior and character (Beets et al., 2009; Li et al., 2009), can also reduce school-level absenteeism and disciplinary outcomes and, concurrently, positively influence school-level achievement. Indeed, this study makes clear that a comprehensive school-based program that addresses multiple co-occurring behaviors can positively affect both behavior and academics.

ACKNOWLEDGMENTS

This project was funded by the National Institute on Drug Abuse (R01-DA13474). Additionally, The National Institute on Drug Abuse (DA018760) provided financial support for the completion of the work on this manuscript. The authors would like to thankfully recognize the support and involvement of the Hawaii school district and the principals, administrators, teachers, staff, students and their families at the participating schools. We also thank Howard Humphreys and Jonathan Wang for help with data collection and management.

Notice of potential conflict of interest: The research described herein was done using the program and the training and technical support of Positive Action, Inc. Dr. Flay's spouse holds a significant financial interest in Positive Action, Inc.

REFERENCES

Acock, A. (2005). Working with missing values. *Journal of Marriage & Family, 67*, 1012–1028.

Battistich, V., Schaps, E., Watson, D., Solomon, D., & Lewis, C. (2000). Effects of the child development project on students' drug use and other problem behaviors. *Journal of Primary Prevention, 21*(1), 75–99.

Beets, M. W., Flay, B. R., Vuchinich, S., Acock, A. C., Li, K.-K., & Allred, C. G. (2008). School climate and teachers' beliefs and attitudes associated with implementation of the positive action program: A diffusion of innovations model. *Prevention Science*. doi:10.1007/s11121-11008-10100-11122

Beets, M. W., Flay, B. R., Vuchinich, S., Snyder, F. J., Acock, A., Li, K.-K., et al. (2009).Use of a social and character development program to prevent substance use, violent behaviors, and sexual activity among elementary-school students in Hawaii. *American Journal of Public Health, 99*, 1438–1445.

Biglan, A., Brennan, P. A., Foster, S. L., Holder, H. D., Miller, T. L., Cunningham, P. B., et al. (2004). *Multi-problem youth: Prevention, intervention, and treatment*. New York: Guilford.

Botvin, G. J., Griffin, K. W., & Nichols, T. D. (2006). Preventing youth violence and delinquency through a universal school-based prevention approach. *Prevention Science, 7*, 403–408.

Botvin, G. J., Schinke, S., & Orlandi, M. A. (1995). School-based health promotion: Substance abuse and sexual behavior.. *Applied & Preventive Psychology, 4*, 167–184.

Brookmeyer, R., & Chen, Y. Q. (1998). Person-time analysis of paired community intervention trials when the number of communities is small. *Statistics In Medicine, 17*, 2121–2132.

Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (2004). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *Annals of the American Academy of Political and Social Science, 591*, 98–124.

Catalano, R. F., Hawkins, J. D., Berglund, M. L., Pollard, J. A., & Arthur, M. W. (2002). Prevention science and positive youth development: Competitive or cooperative frameworks? *Journal of Adolescent Health, 31*(6 Suppl), 230–239.

Coalition for Evidence-Based Policy. (2002). *Bringing evidence-driven progress to education: A recommended strategy for the U.S. Department of Education*. Washington, DC: Coalition for Evidence-Based Policy.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Costello, E. J., Mustillo, S., Erkanli, A., Keeler, G., & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry, 60*, 837–844.

Dent, C. W., Sussman, S., & Flay, B. R. (1993). The use of archival data to select and assign schools in a drug prevention trial. *Evaluation Review*, *17*, 159–181.

Derzon, J. H., Wilson, S. J., & Cunningham, C. A. (1999). *The effectiveness of school-based interventions for preventing and reducing violence*. Nashville, TN: Center for Evaluation Research and Methodology, Vanderbilt University.

Dinks, R., Cataldi, E. F., & Lin-Kelly, W. (2007). *Indicators of school crime and safety: 2007 (*NCES *2008–021/NCJ 219553)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice.

Domitrovich, C. E., & Greenberg, M. T. (2000). The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation*, *11*, 193–221.

DuPaul, G. J., & Stoner, G. (2004). *ADHD in the schools: Assessment and intervention strategies*. New York: Guilford.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*, 327–350.

Eaton, D. K., Brener, N., & Kann, L. K. (2008). Associations of health risk behaviors with school absenteeism. Does having permission for the absence make a difference? *Journal of School Health*, *78*, 223–229.

Eaton, D. K., Kann, L., Kinchen, S., Shanklin, S., Ross, J., Hawkins, J., et al. (2008). Youth risk behavior surveillance—United States, 2007. *Morbidity & Mortality Weekly Report*, *57*, 1–131.

Eccles, J. S. (2004). Schools, academic motivation, and stage-environment fit. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 125–153). Hoboken, NJ: Wiley.

Eisenbraun, K. D. (2007). Violence in schools: Prevalence, prediction, and prevention. *Aggression and Violent Behavior*, *12*, 459–469.

Elias, M. J., Gara, M., Schuyler, T., Branden-Muller, L., & Sayette, M. (1991). The promotion of social competence: Longitudinal study of preventive school-based program. *American Journal of Orthopsychiatry*, *61*, 409–417.

Flay, B. R. (1985). Psychosocial approaches to smoking prevention: A review of findings. *Health Psychology*, *4*, 449–488.

Flay, B. R. (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine*, *15*, 451–474.

Flay, B. R. (2000). Approaches to substance use prevention utilizing school curriculum plus social environment change. *Addictive Behaviors*, *25*, 861–885.

Flay, B. R. (2002). Positive youth development requires comprehensive health promotion programs. *American Journal of Health Behavior*, *26*, 407–424.

Flay, B. R. (2009a). The promise of long-term effectiveness of school-based smoking prevention programs: A critical review of reviews. *Tobacco Induced Diseases*, *5*(7). doi:10.1186/1617-9625-1185-1187

Flay, B. R. (2009b). School-based smoking prevention programs with the promise of long-term effects. *Tobacco Induced Diseases*, *5*(6). doi:1110.1186/1617-9625-1185-1186

Flay, B. R., & Allred, C. G. (2003). Long-term effects of the Positive Action program. *American Journal of Health Behavior*, *27*, S6.

Flay, B. R., & Allred, C. G. (in press). The *Positive Action* program: Improving academics, behavior and character by teaching comprehensive skills for successful learning and living. In T. Lovat & R. Toomey (Eds.), *International handbook on values education and student well-being*. Dirtrecht, Germany: Springer.

Flay, B. R., Allred, C. G., & Ordway, N. (2001). Effects of the Positive Action program on achievement and discipline: Two matched-control comparisons. *Prevention Science*, *2*(2), 71–89.

Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., et al. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science*, *6*(3), 151–175.

Flay, B. R., Graumlich, S., Segawa, E., Burns, J. L., & Holliday, M. Y. (2004). Effects of 2 prevention programs on high-risk behaviors among African American youth: A randomized trial. *Archives of Pediatrics & Adolescent Medicine*, *158*, 377–384.

Flay, B. R., & Petraitis, J. (1994). The theory of triadic influence: A new theory of health behavior with implications for preventive interventions. *Advances in Medical Sociology*, *4*, 19–44.

Flay, B. R., Snyder, F. J., & Petraitis, J. (2009). The theory of triadic influence. In R. J. DiClemente, R. A. Crosby & M. C. Kegler (Eds.), *Emerging theories in health promotion practice and research* (2nd ed., pp. 451–510). San Francisco, CA: Jossey-Bass.

Fleming, C. B., Haggerty, K. P., Catalano, R. F., Harachi, T. W., Mazza, J. J., & Gruman, D. H. (2005). Do social and behavioral characteristics targeted by preventive interventions predict standardized test scores and grades? *Journal of School Health*, *75*, 342–349.

Goldstein, H., Browne, W., & Rasbash, J. (2002). Partitioning variation in multilevel models. *Understanding Statistics*, *1*, 223–231.

Graham, J. W., Flay, B. R., Johnson, C. A., Hansen, W. B., & Collins, L. M. (1984). Group comparability: A multiattribute utility measurement approach to the use of random assignment with small numbers of aggregate units. *Evaluation Review*, *8*, 247–260.

Greenberg, M. T., Weissberg, R. P., O Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., et al. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist*, *58*, 466–474.

Grissom, R. J., & Kim, J. J. (2005). *Effect sizes for research: A broad practical approach*. Mahwah, NJ: Erlbaum.

Gustavson, C., Stahlber, O., Sjodin, A. K., Forsman, A., Nilsson, T., & Anckarsater, H. (2007). Age at onset of substance abuse: a crucial covariate of psychopathic traits and aggression in adult offenders. *Psychiatry Research*, *153*, 195–198.

Hallfors, D. D., Waller, M. W., Bauer, D., Ford, C. A., & Halpern, C. T. (2005). Which Comes First in Adolescence–Sex and Drugs or Depression? *American Journal of Preventive Medicine*, *29*, 163–170.

Hamilton, L. S., Stecher, B. M., Marsh, J. A., Sloan McCombs, J., Robyn, A., Russell, J., et al. (2007). *Standards-based accountability under no child left behind: Experiences of teachers and administrators in three states*. Santa Monica, CA: RAND Corporation.

Hawai‘i Department of Education. (n.d.-a). *Hawai‘i content & performance standards*. Honolulu: Author. Retrieved December 22, 2008, from http://doe.k12.hi.us/curriculum/hcps.htm

Hawai‘i Department of Education. (n.d.-b). *School accountability: School status & improvement report*. Honolulu: Author. Retrieved March 2008, from http://arch.k12.hi.us/school/ssir/ssir.html

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

Hawkins, J. D., Catalano, R. F., & Miller, J. Y. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychological Bulletin*, *112*, 64–105.

Hayes, A. F. (2006). A primer on multilevel modeling. *Human Communication Research*, *32*, 385–410.

Heaviside, S., Rowland, C., Williams, C., & Farris, R. (1999). *Violence and discipline problems in U.S. public schools: 1996–1997* (NCES 98-030). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Horowitz, J. L., & Garber, J. (2006). The prevention of depressive symptoms in children and adolescents: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, *74*, 401–415.

Hulleman, C., & Cordray, D. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness*, *2*, 88–110.

Kellam, S. G., & Anthony, J. C. (1998). Targeting early antecedents to prevent tobacco smoking: Findings from an epidemiologically based randomized field trial. *American Journal of Public Health*, *88*, 1490–1494.

Lerner, R. M. (1995). Developing individuals within changing contexts: Implications of developmental contextualism for human development research, policy, and programs. In T. A. Kinderman & J. Valsinar (Eds.), *Development of person-context relations* (pp. 227–240). Hillsdale, NJ: Erlbaum.

Li, K.-K., Washburn, I. J., DuBois, D. L., Vuchinich, S., Ji, P., Brechling, V., et al. (2009). *Effects of the Positive Action program on problem behaviors in elementary school students: A matched-pair randomized control trial in Chicago*. Manuscript submitted for publication.

Malecki, C. K., & Elliott, S. N. (2002). Children’s social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly*, *17*(1), 1–23.

McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly*, *21*, 471–490.

Meece, J. L., Anderman, E. M., & Anderman, L. H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology*, *57*, 487–503.

Merline, A. C., O’Malley, P. M., Schulenber, J. E., Bachman, J. G., & Johnston, L. D. (2004). Substance use among adults 35 years of age: Prevalence, adulthood predictors, and impact of adolescent substance use. *American Journal of Public Health*, *94*(1), 96–102.

Nucci, L. P. (1997). Moral development and character formation. In H. J. Walberg & G. D. Haertel (Eds.), *Psychology and educational practice* (pp. 127–157). Berkeley, CA: MacCarchan.

Perie, M., Moran, R., & Lurkus, A. D. (2005). *NAEP 2004 Trends in Academic Progress: Three Decades of Student Performance in Reading and Mathematics* (NCES 2005-464). Washington, DC: U.S. Government Printing Office, U.S. Department of Education, Institute of Education Science, National Center for Education Statistics.

Peters, R. D., & McMahon, R. J. (1996). *Preventing childhood disorders, substance abuse, and delinquency*. Thousand Oaks, CA: Sage.

Petraitis, J., Flay, B. R., & Miller, T. Q. (1995). Reviewing theories of adolescent substance use: Organizing pieces in the puzzle. *Psychological Bulletin*, *117*(1), 67–86.

Purkey, W. W. (1970). *Self-concept and school achievement.* Englewood Cliffs, NJ: Prentice-Hall.

Purkey, W. W., & Novak, J. (1970). *Inviting school success: A self-concept approach to teaching and learning*. Belmont, CA: Wadsworth.

Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using Stata* (2nd ed.). College Station, TX: Stata Press.

Raudenbush, S. W., Martinez, A., & Spybrook, J. (2007). Strategies for improving precision in group-randomized experiments. *Educational Evaluation and Policy Analysis*, *29*, 5–29.

Romer, D. (2003). *Reducing adolescent risk: Toward an integrated approach*. Thousand Oaks, CA: Sage.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis*. New York: Oxford University Press.

Slavin, R. E., & Fashola, O. S. (1998). *Show me the evidence! Proven and promising programs for America's schools*. Thousand Oaks, CA: Sage.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Thousand Oaks, CA: Sage.

Snyder, T. D., Dillow, S. A., & Hoffman, C. M. (2008). *Digest of education statistics 2007* (NCES 2008-022). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Stuart, E. A. (2007). Estimating causal effects using school-level data sets. *Educational Researcher*, *36*(4), 187.

Sussman, S., Dent, C. W., Burton, D., Stacy, A. W., & Flay, B. R. (1995). *Developing school-based tobacco use prevention and cessation programs*. Thousand Oaks, CA: Sage.

Tobler, N. S., Roona, M. R., Ochshorn, P., Marshall, D. G., Streke, A. V., & Stackpole, K. M. (2000). School-based adolescent drug prevention programs: 1998 meta-analysis. *Journal of Primary Prevention*, 20, 275–336.

Tolan, P. H., & Guerra, N. (1994). *What works in reducing adolescent violence: An empirical review of the field*. Boulder: Center for the Study and Prevention of Violence, University of Colorado.

Walberg, H. J., Yeh, E. G., & Mooney-Paton, S. (1974). Family background, ethnicity, and urban delinquency. *Journal of Research in Crime and Delinquency*, *56*, 80–87.

Wentzel, K. R. (1993). Does being good make the grade? Social behavior and academic competence in middle school. *Journal of Educational Psychology*, *85*, 357–364.

What Works Clearinghouse. (2007, June). *WWC topic report: Character education*. Retrieved January 20, 2009, from http://ies.ed.gov/ncee/wwc/reports/character_education/topic/

What Works Clearinghouse. (n.d.). *Topic reports*. Retrieved March 2009, from http://ies.ed.gov/ncee/wwc/publications/topicreports/

Zins, J., Weissberg, R., Wang, M., & Walberg, H. J. (Eds.). (2004). *Building academic success on social and emotional learning*. New York: Teachers College Press.

APPENDIX

1. Random intercept mixed linear models:
   a. Random-intercept model

$$Y_{ij} = \beta_{oj} + \beta_1(\text{condition}_j) + \beta_2(\text{year}_{ij}) + \beta_3(\text{year}_{ij} \times \text{condition}_j) + \zeta_j + \epsilon_{ij}$$

   b. Random-intercept quadratic model

$$Y_{ij} = \beta_{oj} + \beta_1(\text{condition}_j) + \beta_2(\text{year}_{ij}) + \beta_3(\text{year}_{ij}^2) + \beta_4(\text{year}_{ij} \times \text{condition}_j) + \zeta_j + \epsilon_{ij}$$

   Y_{ij} = estimated outcome
   β_{oj} = mean intercept
   ζ_j = random intercept
   ϵ_{ij} = level-1 residual

2. Random-intercept Poisson models:
   The estimated outcome, Y_{ij}, is assumed to have a Poisson distribution with expectation μ_{ij}.
   a. Random-intercept Poisson model

$$\mu_{ij} = \exp\{\beta_{oj} + \beta_1(\text{condition}_j) + \beta_2(\text{year}_{ij}) + \beta_3(\text{year}_{ij} \times \text{condition}_j) + \zeta_j\}$$

   b. Random-intercept Poisson quadratic model

$$\mu_{ij} = \exp\{\beta_{oj} + \beta_1(\text{condition}_j) + \beta_2(\text{year}_{ij}) + \beta_3(\text{year}_{ij}^2) + \beta_4(\text{year}_{ij} \times \text{condition}_j) + \zeta_j\}$$

   μ_{ij} = mean rate at which outcome occurs.

Long-term Effects of the *Positive Action*® Program

Brian R. Flay, DPhil, FAAHB, FSBM, FSCRA; Carol G. Allred, PhD

***Objective*: To report long-term effectiveness of the *Positive Action* program. *Methods*: Used matched-schools design and school-level achievement and disciplinary data to evaluate program effects on student performance and behavior in elementary schools. *Results*: Participation in the *Positive Action* program improved student behavior, school involvement, and academic achievement at all 3 levels of schools, with the results showing a clear dose-response relationship. *Conclusion*: Results provide clear evidence that a coherent, comprehensive, and integrated program can have enduring effects in multiple domains.**

***Key words*: positive youth development, comprehensive programs, problem behaviors, prevention, health promotion, academic achievement**

***Am J Health Behav* 2003;27(Supplement 1):S6-S21**

We and others have argued that a wide range of youth behaviors are related and have common causes[1-4] and that effective positive youth development needs to combine the principles of effective character development, health promotion, disease prevention, and academics.[1,5,6] However, schools cannot afford to use different programs to address each of these areas. Administrators, teachers, legislators, and the public are calling for a comprehensive approach.

A number of different kinds of programs have been developed to address problems of academic achievement,[7] smoking,[8,9] substance use,[10] violence,[11] and many other areas. Although many of these programs are initially promising, most are problem specific and unable to provide sustainable effects. Most programs address the micro-level predictors of problem behavior and do not attempt to affect the multifaceted, distal factors. A comprehensive approach that includes self-concept development, schoolwide environmental change, and parental and community involvement may successfully affect all outcomes together. Recent changes in Title 1 legislation have acknowledged and facilitated the development/funding of comprehensive school reform programs; however, there are few that have been fully evaluated.

The behaviors of children and adolescents are highly correlated and have many of the same risk and protective factors; and behavior, school involvement, and academic achievement are related (see Flay[1] for a review). We need to address student character development, behavior, school involvement, and learning in a comprehensive and integrated way. The present paper reports on the long-term effectiveness of one program that pro-

Brian R. Flay, UIC Distinguished Professor, Public Health (Community Health Sciences) and Psychology, Health Research and Policy Centers, University of Illinois at Chicago, Chicago, IL. Carol G. Allred, President/Developer, Positive Action, Inc., Twin Falls, ID.

Address correspondence to Dr. Flay, Health Research and Policy Centers, University of Illinois at Chicago, 850 W. Jackson Blvd., Suite 400, M/C 275, Chicago, IL 60607. E-mail bflay@uic.edu.

vides schools with the means to achieve this.

The *Positive Action* Program

A detailed description of the theoretical basis, program structure, and prior evaluations of the *Positive Action* program (PA) can be found elsewhere.[6] Here we summarize aspects of PA, with an emphasis on its comprehensiveness.

Theoretical basis. The PA program is grounded in a broad theory of self-concept[12-14] that posits that people determine their self-concepts by what they do; that actions, more than thoughts or feelings, determine self-concept; and that making positive and healthy behavioral choices results in feelings of self-worth. Recent studies in positive psychology[15] support this notion; eg, Fredrickson[16] found that when children feel positive, they have more positive thoughts and engage in more positive behavior.

PA is also consistent with educational theories of brain development,[17] higher-level thinking skills,[18] multiple intelligences,[19,20] and social and emotional learning.[21] PA teaches children what actions are positive, that they feel good when they do positive actions, and that they then have more positive thoughts and future actions. By explicitly linking thoughts, feelings, and actions, the program is believed to enhance the development and integration of affective and cognitive brain functions.[22]

Consistent with multiple social learning theories[23-26] and a wide array of theories of behavior change integrated into Flay's theory of triadic influence,[27,28] PA also trains teachers, other school staff, and parents to identify and reinforce positive feelings, thoughts, and actions by students, leading to continual reinforcement of positive behavior and enhanced student bonding with parents and school. PA is also consistent with other current approaches to social development, health promotion, and prevention of unhealthy behaviors.[10,29,30]

The PA model is very comprehensive, integrated, and holistic. Current mental-health problem, drug abuse or violence-prevention programs rely on providing knowledge, correcting normative beliefs, and teaching self-management and social skills.[10] Recent approaches to improving academic achievement, even many of those classified as whole school reform, focus on enhancing particular curricular content and instruction methods[18] or particular skills such as reading,[7] but not many other needs of students. Current approaches to school ecology focus on parent involvement in school governance and reorganization, although not addressing the students' needs very effectively.[31-33] Each of these approaches attempts to identify and correct particular risk or protective factors.

PA is designed to affect more distal (and ultimately more important) influences on behavior and performance than most other programs affect. This is consistent with Flay's[1] suggestion that broad and long-term effectiveness in reducing problem behaviors and increasing school performance will require addressing more distal factors in a more comprehensive and integrated way. PA attempts this with a holistic approach to school reorganization, teacher-student relations, parent involvement, instructional practices, and development of the self-concept of students, teachers and parents.

Program structure. The PA program includes a detailed curriculum with almost daily lessons, a schoolwide climate program, and family- and community-involvement components, each of which uses research-proven educational strategies and methods such as active learning and positive classroom management. The program has goals and components for each of the individual, family, school, and community levels. Central to all components of the program are 6 program units (Table 1): (1) self-concept; (2) positive actions for one's mind and body; and 4 units that teach social/emotional positive actions for (3) managing oneself responsibly; (4) getting along with others; (5) being honest with oneself and others; and (6) improving oneself continuously.

Schools integrate the program units in a scoped-and-sequenced classroom curriculum and a school-climate program. The K–6 classroom curriculum consists of over 140 lessons per grade. Using teacher's kits (that include teacher's manuals and all materials needed for all activities for a whole class), classroom teachers present 15- to 20-minute lessons almost every day. Scripted lessons are completely prepared and teacher-friendly, employing a variety of methodologies and addressing different learning styles. Activities include stories, role-

Table 1
Content of All Components (Classroom Curriculum, School-Climate Materials, Family Kit, and Community Kit) of the Positive Action Program

Unit # and Topic	Content
Unit 1: Self-concept: What It Is, How It's Formed, and Why It's Important	The relationship of thoughts, feelings and actions (behavior). Units 2-6 teach children what actions are positive in various domains of life, that they feel good when they do positive actions, and that they then have more positive thoughts and future actions.
Unit 2: Positive Actions for Body (Physical) and Mind (Intellectual)	Physical: exercise, hygiene, nutrition, avoiding harmful substances, sleeping and resting enough, safety. Intellectual: creative thinking, learning/studying, decision making, problem solving.
Unit 3: Social/Emotional Positive Actions for Managing Yourself Responsibly	Manage human resources of time, energy, thoughts, actions, feelings (anger, fear, loneliness, others), talents, money, possessions. Includes self-control.
Unit 4: Social/Emotional Positive Actions for Getting Along With Others	Treat others the way you like to be treated, code of conduct (respect, fairness, kindness, honesty, courtesy, empathy, caring, responsible, reliable), conflict resolution, communicating positively (communication skills), forming relationships, working cooperatively, community service. [These are the essence of character education.]
Unit 5: Social/Emotional Positive Actions for Being Honest with Yourself & Others	Self-honesty, doing what you will say you will do (integrity), not blaming others, not making excuses, not rationalizing; self-appraisal (look at strengths and weaknesses); and being in touch with reality. [These are the essence of mental health.]
Unit 6: Social/Emotional Positive Actions for Improving Yourself Continually	Goal setting (physical, intellectual and social/emotional), problem solving, decision making, believe in potential, have courage to try, turn problems into opportunities, persistence.
Unit 7: Review	Review of all of above.

playing, modeling, games, music, questions/answers, activity booklets and sheets, posters, and manipulatives. The program content teaches students how to use positive actions, to recognize feeling good about themselves, to manage themselves (including thoughts, actions, and feelings), and to treat others the way they want to be treated.

The school-climate program encourages and reinforces the practice of positive actions schoolwide and extends the program to families and the community. For each school, a principal's kit[34] provides directions for a school-climate program to promote the practice and reinforcement of positive actions in the entire school. It also includes parent- and community-involvement activities.

The parent program (family kit,[35] see Gorsky[36] for a review) includes coordinated weekly lessons and links the family to the school activities. The family kit contains a manual with 42 multi-age, weekly lessons based on the 6 units and 6 review lessons with enough materials for 6 individuals. This kit coordinates family activities with the PA school curriculum and school-climate activities. It contains all the materials required in the lessons: colorful posters and visuals, hands-on materials, activity worksheets, and music. It contains Words of the Week and the "ICU Doing Something Positive

Box" like those used in the school.

The community program includes a community kit and combines with the school and parent programs to align all the environments (schools, families, and community) involved in the program. The community kit includes a guide, the *Positive Actions for Living*[35] text, music CDs and books, family kits, and other materials. It provides community leaders, public servants, social service workers, and business executives with the tools to plan and cultivate positive actions in every aspect of the community while encouraging development in every aspect of the individual citizen.

Prior evaluations of PA. PA was developed by the second author, a public school teacher at the time, over 6 years (1977-83) of planned pilot work, formative evaluation, revision, and further evaluation.[37] These evaluations consistently suggested that the program effectively improved student self-concept, behavior, school involvement, and academic achievement. Using before- and after-PA School Report Card (SRC) data, a wide array of elementary schools have documented strong improvements in achievement and decreases in problem behavior. For example, percentile rankings on standardized tests improved from as low as the 30th percentile to as high as the 90th percentile over the course of only 1 to 3 years. Some schools improved from being the worst in their district to being the best. Admittedly, these are not the average results that might be expected in a more controlled study.

In a more rigorous study,[6] we used a matched-control design and school-level achievement and disciplinary data to evaluate program effects on student performance and behavior in 2 separate school districts. The program improved achievement by 16-52% and reduced disciplinary referrals by 78-85%. The study reported here extends prior work by replicating these results with improved methods in another large school district, and by investigating long-term effects when PA-exposed students graduate into middle and high school.

METHODS

Design

For this study, we chose one large southeastern school district that had school-level archival (SRC) data on student performance and disciplinary referrals/actions easily available for both elementary and secondary (middle and high) schools and that had a significant number of elementary schools that had implemented PA for 4 or more years. Some schools had never used PA or stopped using it 4 or more years before the 1997-98 school year (non-PA, n=28). Others had used it for 4 or more years prior to 1998 (PA-only, n=45), and others had also adopted other supplementary character/behavior programs, such as Skill Streaming, Peace Works, Peace-Able, or combinations of them, in addition to continued use of PA (PA+Other, n=20). We do not have formal data on the elective academic programs (eg, special reading or math programs) used in these schools during this time, but we do know that there was no correlation between whether a school had PA and the special academic programs they used. Each of the latter 2 groups of schools had used PA for an average of 7 years (range = 4-9 years). These 3 groups of schools were compared to assess program effects on elementary school student achievement and behavior.

We used school report card (SRC) data to find matching sets of one PA-only school, one PA+Other school and one non-PA (control) school. In order to find matched sets, we first rank-ordered all schools on percent free/reduced lunch, then on percent mobility (student turnover), and then we selected schools with similar ethnic distributions. These particular variables were chosen because for the non-PA schools in this school district, poverty (percent free/reduced lunch) was the strongest predictor of student performance (accounting for 57% of the variance), and percent African American students was the best predictor of disruptive behavior (accounting for 32% of the variance). Percent mobility was also a strong predictor of both behavior and achievement and the strongest predictor of attendance. These matching variables were not expected to change as a result of PA; therefore, they were presumed to imply pretest matching on the outcome variables of interest (behavior, attendance, and achievement). The PA schools in the resulting matched sets had used PA for 4 or 5 years.

Table 2 shows the comparability of the program schools and their matched control schools compared with all non-PA schools in the district (there were no

Table 2
Differences between PA and non-PA Schools for Total Samples and Matched Sets; 1998 and 1993 Demographic Data and 1993 (pre-PA) Achievement and Behavior Data[a]

	All Schools PA n=65, non-PA n=28				Matched Sets PA n=24, Control n=12			
	PA	NonPA	SD	P	PA	NonPA	SD	P
1998 Demographic Data[b]								
Enrollment	822	690	219.5	0.007	770	700	216.1	0.358
% free/reduced lunch	52.60	69.90	25.03	0.002	62.20	67.60	22.21	0.502
% mobility	41.75	50.78	17.37	0.021	43.83	49.46	15.71	0.318
Student/teacher ratio	10.97	9.22	1.99	0.000	10.71	9.77	1.74	0.129
% White	55.30	44.22	21.54	0.023	50.59	44.66	20.31	0.420
% African American	22.32	28.39	18.98	0.161	24.61	28.48	19.62	0.587
% Hispanic	18.25	24.13	14.91	0.083	20.71	23.23	15.46	0.653
1993 Demographic Data[b]								
Enrollment	760	711	193.80	0.007	770	723	205.90	0.520
% free/reduced lunch	49.90	62.90	21.52	0.002	57.60	59.50	19.92	0.794
% mobility	43.60	53.36	16.67	0.021	47.58	51.83	15.40	0.444
% Minority	22.32	28.39	18.98	0.161	41.58	41.67	16.44	0.989
1993 Achievement and Behavioral Data[b]								
Reading % above median	43.63	35.36	15.32	0.016	37.25	36.50	12.32	0.866
Writing % above 3	19.37	14.46	9.94	0.028	16.46	16.92	7.09	0.858
Math % above median	50.60	44.11	13.68	0.035	46.29	44.92	12.70	0.764
Absentee rate	6.15	6.85	1.18	0.007	6.29	6.77	1.07	0.203
Suspensions	3.49	4.10	2.46	0.279	4.61	4.97	2.35	0.673

Note.
a Means, standard deviations, and P values
b From 3 MANOVAs. There were no significant differences between PA and PA+Other conditions, so the 2 conditions were combined.

significant differences between PA-alone and PA+Other schools). PA schools were substantially different from non-PA schools, being at lower risk because they had lower proportions of students receiving free/reduced lunch, lower mobility rates, and lower proportions of minority students, but at higher risk because they were larger and had a higher student-teacher ratio. As expected, matched control schools were similar to PA schools, including on pre-PA (1993) indicators of achievement and behavior.

For analyses of the sustained effects of *Positive Action* into middle schools, we calculated the proportion of feeder elementary schools that had implemented PA for at least the prior 4 years. For analyses of the sustained effects of PA into high school, we calculated the proportion of feeder schools that had implemented PA for at least the prior 8 years. In each case, we tried to ensure that students in the middle or high schools would have received at least 2 years of PA prior to the year of data available to us. We hypothesized a dose-response relationship, where middle and high schools with more students from elementary schools with PA (ie, PA graduates) would report lower average rates of problem behaviors and higher average achievement.

Measures

Elementary SRC achievement data consisted of mean scores on the Florida Reading Test and the grade 4 Florida Comprehensive Aptitude Test (FCAT) for the 1997-98 school year. Behavioral data consisted of disciplinary referrals for in-

Table 3
Effects of PA on Achievement and Behavior (1998) in Southeastern Elementary Schools (Means, Standard Deviations, P Values,[a] and Percent of Variance Accounted for in Model[b])

	All schools (PA n=65, non-PA n=28)					Matched sets (PA n=24, control n=12)					
	PA	NonPA	SD	P	%diff	PA	NonPA	SD	P	%diff	R^2
Achievement											
Florida Reading Test	110.20	78.00	29.02	0.000	41.30	105.9	73.10	24.80	0.001	**44.90**	0.873
FCAT grade 4 total	295.20	283.10	19.01	0.006	4.30	290.9	278.40	19.30	0.000	**4.50**	0.968
Behavior											
Violence/100 students	5.40	8.74	7.31	0.049	38.20	3.83	12.11	5.94	0.000	**68.40**	0.965
% suspensions	2.52	3.58	2.05	0.057	29.60	2.72	4.09	3.16	0.003	**33.50**	0.836
% absent 21+ days	10.75	12.01	3.91	0.157	10.50	10.79	12.36	3.95	0.179	**12.70**	0.791

Note.
a From ANOVAs for the All vs PA comparisons, and from multivariate GLM fixed effects (PA or not and matched controls) models for all matched controls analyses. Effects were marginally smaller in univariate GLM analyses, but multivariate analyses provide some adjustment for multiple comparisons. There were no significant differences between PA and PA+Other conditions, so the 2 conditions were combined.
b Multivariate GLM fixed effects (PA or not and matched pairs) model.

cidents of violence per 100 students, percent of students who received out-of-school suspensions, and percent of students absent for 21 or more days during the school year. Preliminary analyses found no differences between the PA-alone and PA+Other schools on outcomes; consequently, these 2 conditions were combined for the analyses reported.

Middle-school standardized achievement test data were the percent of students scoring above the median on the 8th-grade norm referenced tests (NRT) of reading and math (1997-98). Available indicators of behavior included incidents per 100 students of substance use (tobacco, alcohol, and illicit drugs), violence, dissing behaviors (disrespect, disobedience, disorderly, and disruptive), and property crimes (larceny, petty theft, and vandalism). All behavioral data were coded disciplinary referrals by school principals or disciplinary officers. Absenteeism data were also available.

High school standardized achievement test data (1997-98) were the percent of 10th-grade students scoring 3 or greater on the Florida Writes test, percent of seniors passing the High School Competency Tests (HSCT) of communications and math, mean Scholastic Aptitude Test (SAT) scores, and mean American College Testing (ACT) composite scores. Percent absent 21 or more days and percent dropout were other indicators of school involvement. Behavioral data (1998-99) included disciplinary referrals for substance use (tobacco, alcohol, and illicit drugs), violence (threatening, fighting, carrying weapons, and battery), dissing behaviors (disrespect, disobedience, disruptive, disorderly, and inappropriate dress), sexual behaviors (sex-related harassment, offences, and battery), property crime (arson, breaking and entering, theft, and vandalism), breaking of school rules, misbehavior on or near school buses, parking violations, and falsification of reports. Data on percent of students suspended (separately for in-school and out-of-school) were also available.

Analyses

All analyses were conducted using SPSS version 10.1.[38] To estimate the effects of PA on elementary school achievement and behavior, we conducted analyses of variance and analyses of covariance (adding the 3 matching variables) for the comparison of all PA schools with all other schools. We conducted multivariate general linear modeling (GLM) with fixed

Figure 1
Effects of the PA Program on Achievement at 3 Levels of Percent Free/reduced Lunch

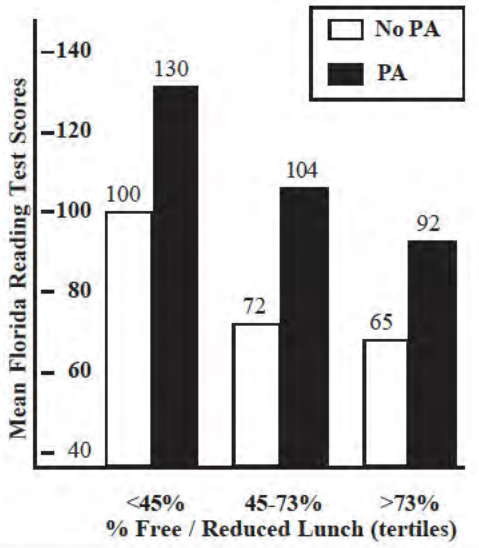
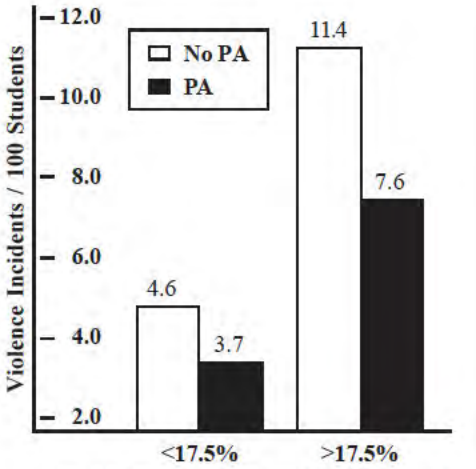


Figure 2
Effects of the PA Program on Violence (incidents per 100 students) at 2 levels of Percent African American Students



effects for condition and pair number (with pairs numbered in order of percent free/reduced lunch for analyses of achievement and in order of percent African American for analysis of behavior) for the comparison of matched PA and non-PA schools.

To estimate the effects of receiving PA in elementary school on achievement and behavior in middle and high school, we conducted multivariate GLMs for each set of outcomes using percent PA students as the independent variable and using percent free/reduced lunch (available for middles schools only), school size, and percent mobility as covariates. For some outcomes, univariate models were necessary because different sets of covariates were significant or interacted with percent of PA.

RESULTS

Elementary School Results

Table 3 shows results for both the all-schools and matched-controls analyses. In the all-schools analysis, scores on the Florida Reading Test were over 40% better in schools with PA compared to schools without PA, and this effect was still significant after adding percent free/reduced lunch as a covariate (P=.003). In the matched-controls analysis, students in PA schools scored an average of 45% better than students in matched control schools. FCAT scores show a less dramatic improvement (4.3% in the all-schools analysis and 4.5% in the matched-controls analysis, not significant after adjusting for percent free/reduced lunch).

In the matched-controls analysis, both main effects (of PA or not and matched set number) were highly significant, and the interaction was not significant. The lack of interaction suggests that PA was equally effective at all levels of school poverty (percent free/reduced lunch). Figure 1 shows that the effects were approximately the same regardless of level of poverty. However, the 27-point improvement for schools with a high proportion of students receiving free/reduced lunch represents a 41% improvement, whereas the 30-point improvement for schools in the lowest tertile of poverty represents only a 30% improvement. Thus, the program

Table 4
Differences between Middle Schools with 3 Levels of PA Graduates; 1998 Demographic Data and 1993 Achievement and Behavior Data

	% of Students PA Graduates				
	<60%[a]	60-79%[b]	80-100%[c]	SD	P
1998 Demographics[d]					
Enrollment	956	1149	1024	286	0.356
% Free/reduced lunch	57.28	47.11	50.41	17.20	0.468
% Mobility	39.59	37.93	36.25	10.12	0.789
% Limited English	7.63	8.13	6.51	5.73	0.814
1993 Outcome Indicators[d]					
Reading % above median	45.20	44.33	44.38	7.57	0.980
Writing % above 3	59.80	60.17	60.38	7.09	0.991
Math % above median	49.60	51.83	50.50	10.39	0.944
Promotion %	68.60	65.45	65.22	16.13	0.934
Suspension %	20.92	19.03	21.94	5.77	0.671
Absenteeism	10.38	11.25	10.86	3.33	0.920

Note.
a 4 schools <50% and 6 schools 50-59%
b 7 schools 60-69%, 5 schools 70-79%
c 5 schools 80-89%, 7 schools 90-100%
d Two multivariate GLM one-way MANOVA

has larger effects for those schools most in need, but still does not close the gap between schools with more versus fewer students receiving free/reduced lunch.

In the all-schools analysis, effects of PA on violence, suspensions, and absenteeism were marginally significant, none of which remained significant after including significant covariates (percent African American). Significant results were found in the matched-controls analyses for violence and suspensions, but not for absenteeism. The number of violence incidents per 100 students was 38% less in PA schools than in control schools in the all-schools analysis and 68% less in the matched-controls analysis. In the multivariate GLM for behavior, the interaction between PA or not and matched pairs was significant for violence, indicating that the program effect was stronger for higher numbered pairs, that is, in schools with higher proportions of African American students. Figure 2 shows the effect – program effects were larger where they were most needed, 33% reduction in schools with higher percentages of African American students compared with a 20% reduction in schools with lower percentages of African American students.

The percentage of students receiving out-of-school suspensions was 29.6% and 33.5% less in PA schools compared to non-PA schools in the all-schools and matched-schools analyses, respectively. The percentage of students reported being absent for 21 or more days was 10.5% and 12.7% less in PA schools compared to non-PA schools in the all-schools and matched-schools analyses respectively. In neither case was the interaction of PA or not and matched set number significant, indicating that the program was equally effective in higher versus lower risk schools.

Middle School Results

For each of the 33 middle schools in the district we calculated the proportion of feeder elementary schools actively implementing PA in 1997-98 and for at least 4 years prior (the percent PA score). The percent PA scores range from 0% to 100% with some skewness toward the high end (Table 4). We compare by tertiles, low-PA middle schools with less than 60% of their students being PA graduates, medium-PA

Table 5
Effects of Elementary PA on Middle School Student Achievement (Multivariate GLM) and Behavior (Univariate GLMs) by 3 Levels of % PA Graduates[a]

	% of Students PA Graduates <60%	60-79%	80-100%	SD	P	Sig Co-variates	Sig Inter-actions	Adj R2
Achievement: % above average grade 8 NRT								
Reading	43.71	48.40	50.89	12.24	0.014	(.001)[b] (.000)[c]	(.012)[b] (.018)[d]	0.918
% change		11	16					
Math	48.14	53.60	58.00	13.93	0.028	(.028)[b] (.000)[d]	(.042)[b] (.058)[d]	0.826
% change		11	20					
Behavior: Incidents per 100 students								
Drug Use[e]	4.09	2.58	1.20	1.65	0.001	(.01)[b]		0.522
% change		37	71					
Violence	39.74	25.44	11.94	16.53	0.047	(.002)[b] (.018)[c]	(.05)[c]	0.665
% change		36	70					
Dis...[f]	322.27	220.27	101.40	120.77	0.047	(.002)[b] (.003)[c]	(.05)[c]	0.767
% change		32	69					
Property Crime[g]	5.52	3.83	2.66	2.12	0.000	.000[b] .000[c] .001[d]	(.000)[c] .001[d]	0.874
% change		31	52					
Days of Absenteeism	72.43	49.49	18.36	31.84	0.000	(.000)[b]		0.750
% change		32	75					

Note.
a Means and percent change shown, as well as pooled standard deviation, P value, significant covariates (with P Value), significant interactions (with P Value) and adjusted R square for the model
b School size
c Mobility
d Lunch
e Tobacco, alcohol and illicit substances. Results for each subcategory parallel those presented for sum.
f Sum of disrespectful, disobedient, and disorderly behaviors.
g Sum of larceny, petty theft, and vandalism.

middle schools with 60-79% PA graduates, and high-PA middle schools with 80-100% PA graduates.

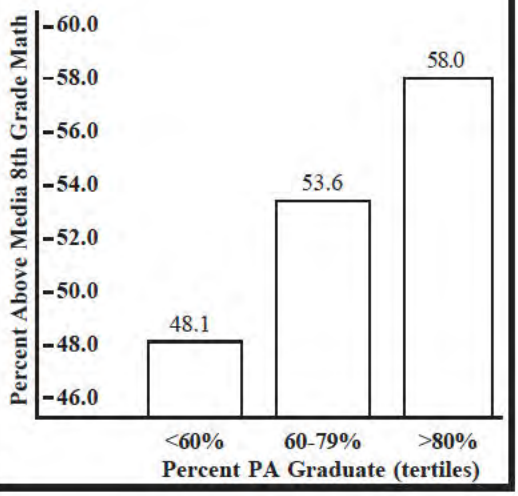
There was no significant relationship between percent PA graduates and available school characteristics: school size (608-1607), percent free/reduced lunch (10.3-75.9), percent mobility (18.3-60.7), percent disabled (60.-21.1), percent limited-English (1.6-26.0), percent gifted (0-18.1), percent teachers with master's degree or higher (19.6-45.8), teachers' average years of experience (7.9-16.5), regular per-pupil expenditures ($2898 to $6284). Data on ethnic distribution were not available, but were estimated from feeder patterns (African American 2-62%, and Hispanic 7-45%). Furthermore, there were no significant differences on pre-PA (1993) indicators of achievement and be-

havior (Table 4).

Table 5 shows a clear dose-response relationship for all outcomes, with middle schools with more PA graduates scoring better than schools with fewer PA graduates. For reading, medium-PA middle schools scored 10.8% better and high-PA schools scored 16.5% better than low-PA schools. For math, medium-PA schools scored 11.4% better and high-PA schools scored 20.6% better than low-PA schools (Figure 3).

Students in medium-PA middle schools had 31-37% less, and students in high-PA schools had 52-75% less problem behaviors than did students in low-PA schools. There were significant interactions with percent mobility for violence, dissing behaviors, and property crimes. For example, as shown in Figure 4, the use of the PA program in elementary schools has larger effects in higher risk middle schools, essentially eliminating the otherwise clear correlation between the predictor covariate and the behavior.

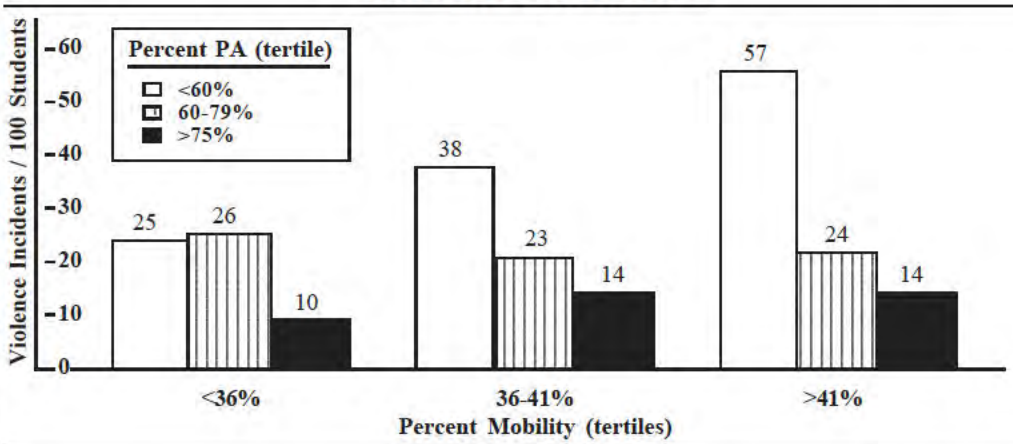
Figure 3
The Dose-response Effect of Percent of Middle School Students from Elementary Schools with the Positive Action Program on Middle School Achievement



High School Results

For each of 18 high schools in the district we calculated the proportion of feeder elementary schools actively implementing the PA program in 1997-98 and for at least 8 years prior. The percent PA scores range from 0% to 50%; 6 low-PA high schools had 0-15% PA graduates, 5 medium-PA schools had 16-26% PA gradu-

Figure 4
Interaction with Percent Mobility for Middle School Behavioral Outcomes[a]



Note.
a Number of drug incidents per 100 students by 3 levels of PA graduates and 3 levels of mobility

Table 6
Significant Effects of Elementary PA on High School Student Achievement (univariate GLM), Employment and Continuing Education (Multivariate GLM), and Behavior (Univariate GLMs) by 3 Levels of % PA Graduates[a]

	% of Students PA Graduates <60%	60-79%	80-100%	SD	P	Sig Co-variates	Sig Inter-actions	Adj R2
Achievement: % above average grade 8 NRT								
% >3 Florida Writes	81.2	85.3	90.1	5.29	0.021		(.057)[b]	0.719
% change		05	11					
% pass HSCT comm	73.7	78.0	81.0	5.98	0.019		(.037)[b]	0.596
% change		06	10					
% pass HSCT math	74.8	78.0	85.7	6.96	0.318			0.330
% change		04	15					
Mean SAT score[c]	951.0	980.8	1046.1	62.85	0.023	(.004)[d]	(.087)[b]	0.767
% change		03	10					
Mean ACT composite	20.22	20.55	21.96	1.44	0.680			0.151
% change		02	09					
Employment and Continuing Education								
% Employed (FT or PT)	63.95	72.88	75.73	7.15	0.183			0.419
% change		14	18					
% Continuing Education	38.75	50.75	53.45	9.98	0.001	(.003)[b]	(.003)[b]	0.870
% change		31	38					
% drop out[c]	6.15	5.49	3.86	1.54	0.001	(.009)[d]		0.623
% change		11	37					
Behavior: Incidents per 100 Students[e]								
Substance use	4.31	3.14	2.20	1.69	0.032			0.289
% change		27	49					
Violence	4.28	2.95	2.16	1.59	0.000	(.000)[d]		0.704
% change		31	50					
Sexual	0.19	0.14	0.07	0.09	0.007	(.023)[d]		0.476
% change		26	63					
Dis…	43.76	35.29	31.48	12.39	0.068	(.073)[d]		0.241
% change		19	28					
Falsify	1.87	0.93	0.80	0.76	0.003	(.033)[d]		0.522
% change		50	57					
Behavior: Percent of Students[f]								
% absent 21+ days	30.53	28.13	26.96	4.23	0.008	(.000)[d]		0.603
% change		08	12					
% in school suspensions	19.71	15.24	13.86	6.28	0.084	(.075)[d]		0.219
% change		23	30					
% out school suspensions	22.18	18.49	16.66	4.22	0.005	(.015)[d]		0.502
% change		17	25					

Note.

a Means and percent change shown, as well as pooled standard deviation, P value, significant covariates (with P Value), significant interactions (with P Value) and adjusted R square for the model.

b School size

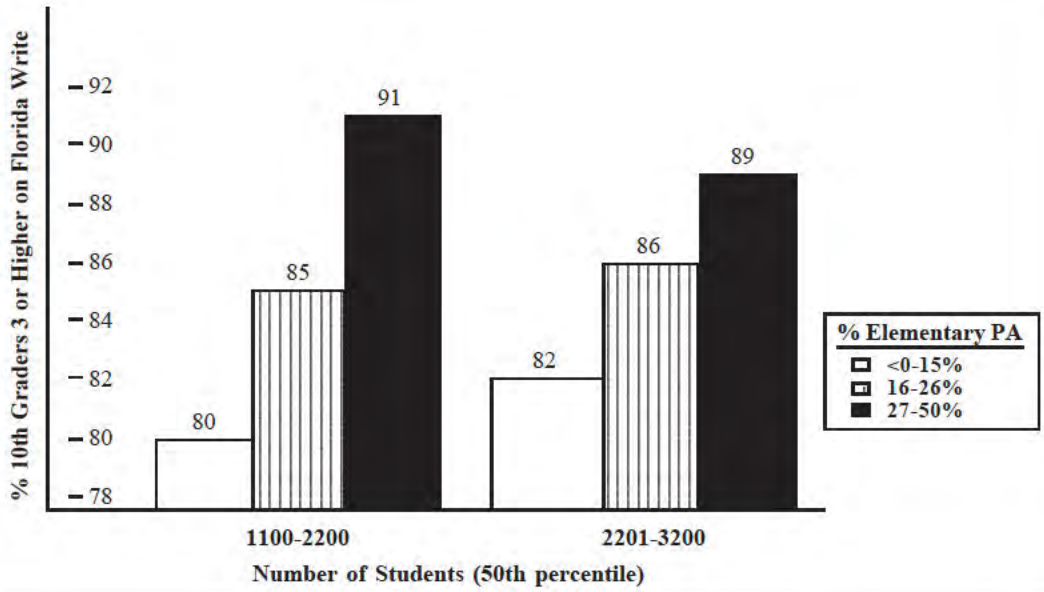
c Results from multivariate GLM with absenteeism and suspensions. All other achievement, employment and continuing education results from one multivariate GLM.

d Mobility

e All results from one multivariate GLM. Substance use = tobacco, alcohol, and illicit substances; Violence = threat, fight, weapon carrying, and battery; Sexual = sex-related harassment, battery and offences; Dis… = disrespect, disobedience, disruptive, and inappropriate dress; and Falsify = falsifying records.

f Absenteeism, suspension and drop out results are from one multivariate GLM.

Figure 5
Interaction with School Size for High School Achievement Outcomes[a]



Note.
a Percent of grade 10 students scoring 3 or better on the Florida Writing Test by 3 levels of PA graduates and 2 levels of high school size.

ates, and 7 high-PA schools had 27-50% PA graduates. No high school had more than 50% of their students coming from elementary schools with PA. The proportions of high school students who were in PA elementary schools were substantially lower than the proportions observed for middle schools because the elementary schools had to have been doing PA for 8 years or more.

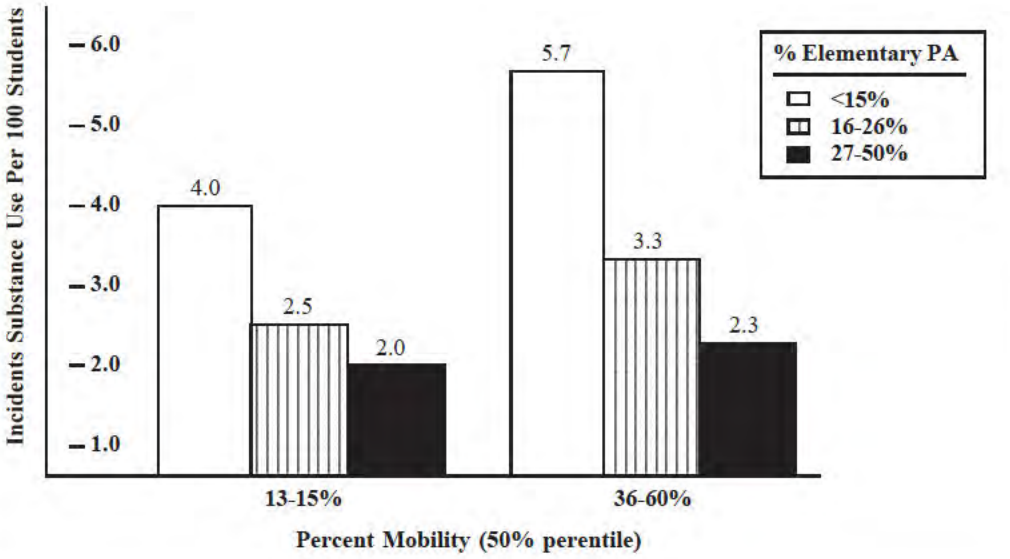
There were no significant relationships between the proportion of PA students and available school characteristics: school size (1121 to 3178 students), percent mobility (18.2-52.1), percent disabled (6.2-12.6), percent limited-English (1.1-16.8), percent teachers with master's degree or higher (31.8-52.7), teachers' average years of experience (9.1-16.5), or school expenditures per students ($3296 to $6064). Data on ethnic distribution or poverty were not available.

Significant effects occurred for a wide range of indicators of achievement and behavior (Table 6). No significant effects occurred for behaviors related to property crime, school rules, bussing, and parking. As with the middle school results, there was a clear dose-response relationship for all significant outcomes.

Medium-PA high schools scored 2-6% better, and high-PA schools scored 9-15% better than low-PA schools on 5 different standardized achievement tests. For 3 outcomes, there was a significant interaction with school size (eg Figure 5), indicating that the endurance of PA program effects on student achievement is better in smaller schools.

Compared with low-PA high schools, the dropout rate is 11% lower from medium-PA high schools, and 37% lower from high-PA high schools. A similar dose-response relationship is observed for the percentage of high school graduates who continue their education (31% and 38%

Figure 6
PA Effects on High School Behavior[a]



Note.
a Number of substance use incidents per 100 students by 3 levels of PA graduates and 2 levels of percent mobility.

improvements respectively). There is a marginal effect of a similar nature for the percentage of high school graduates employed (14% and 18% improvements respectively).

Students in medium-PA high schools were 19-50% less, and students in high-PA schools were 28-63% less likely to engage in problem behaviors; 8% and 12%, less likely to be truant; and 17-23%, and 25-30% less likely to be suspended. There were no significant interactions with covariates, suggesting that effects were equal for high-risk and low-risk schools (eg, Figure 6).

DISCUSSION

Using archival school-level data, we have (a) replicated results of an earlier matched-control study on the effects of the *Positive Action* program on elementary school achievement and behavior;[6] (b) found that adoption of other programs in addition to Positive Action led to no significant improvements; (c) found that the effects endured through middle and high school for a broad array of indicators of both achievement and behavior; (d) found a clear dose-response relationship for most outcomes, such that schools with more PA graduates reported better student behavior, school involvement, and achievement; and (e) found that behavioral effects were as large or larger in higher risk as lower risk schools. These findings provide strong support for (a) the strength of the *Positive Action* program and (b) the idea that a comprehensive program can have broad and long-lasting effects.

This is also the first published evaluation of PA that had pretest data available to establish the equivalence of the matched controls. Non-random assignment to receive PA means that schools that elected to adopt/continue the program may have been different from control schools. By matching on school characteristics normally related to poor academic achievement and problem behavior, we hoped to control for school differences in behavior and achievement prior to the introduction of PA. The availability

of pre-PA achievement and behavioral data (1993), albeit not with the same measures as used in 1998, helped us to establish the statistical comparability of the matched controls. This is a major improvement over the previously published matched control studies.[6]

The above limitation of the elementary school study could have carried over into secondary schools. If so, one would expect that schools with different proportions of PA graduates would differ. We did not find such differences, indicating that the mix of elementary schools feeding students into the secondary schools was not correlated with whether one or more of them had PA. Thus, we can be fairly certain of the pretest comparability of the student bodies in middle and high schools with different proportions of PA graduates.

When students from elementary schools with PA enter a middle or high school with students from an elementary school without PA, we would expect that the more PA graduates there are, the more likely it is that the average behavior of students will be better compared to students in a school with fewer PA graduates. Indeed, we found the hypothesized dose-response relationship. This pattern was very robust, replicating across all measures of achievement and behavior. The enduring effects of PA were especially strong for serious behaviors, high school dropout rates, and the long-term outcome of continuing education after high school. Few other elementary or middle school programs have reported effects enduring through high school.

We found that PA had its multiple effects in all kinds of schools, with equally strong behavioral effects in higher risk schools. Of particular interest is that achievement results do not seem to depend on any particular type of academic program being used with PA.

Our use of school-level archival data may be seen as a limitation or as a strength. Being limited to school-level data did not allow us to investigate program effects on individual students. However, school-level data did allow us to demonstrate that the normative climate was changed sufficiently among a group of elementary school students and their families to carry over into other social environments (their secondary schools) and the rest of their lives. To our knowledge, no other program has reported such results.

The effects of PA support the notion that a comprehensive program can have effects in multiple domains. Only a handful of other programs have also reported effects in multiple domains.[39-45] Many so-called comprehensive character education, social skills development and social-emotional programs have not reported such comprehensive effects. This is because their comprehensiveness is rather limited; eg, social skills training might be expected to improve multiple behaviors, but not all, and not necessarily academic performance.

Most programs are also of limited intensity, duration, and coherence. PA has 4 lessons per week (plus reading of messages from the "ICU Doing Something Good" box on Fridays) for every grade in an elementary school. The program is provided to every grade at the same time. The content of every grade level is parallel but unique, so that as students advance from lower to higher grades they can continue to enhance their PA learning. The curriculum for every grade builds upon what was learned in the previous grade. At the same time, students can start PA at any grade level and still learn the material.

PA offers coherent components for schoolwide climate change, family involvement, and community involvement. Few programs integrate schoolwide, family, and community components as coherently as *Positive Action*. The integration is not just the sum of the classroom activities, but other carefully designed activities are provided to school principals, parents, and community members. These other activities use the same language and follow the same sequence as the classroom curricula, but they provide for an added layer of learning for students, teachers/staff, parents, and community representatives.

PA recognizes the interrelatedness of student character, behavior, and academic achievement. The program impacts school and classroom management, motivation, learning climate, and the skills and knowledge of the core content areas. PA teaches knowledge of, and provides opportunities to practice, skills in these various content areas. The program also teaches thinking skills — reasoning, creativity, problem solving, decision making, higher-order thinking — as

positive actions in the intellectual area. Additionally, PA teaches intrinsic motivation and self-responsibility, thus empowering students to take more initiative for their own learning.

As can be seen, PA is comprehensive in many ways — but another factor may explain the "magic" of PA. The entire focus is positive. Most prevention programs, with a few exceptions, focus on the negative behaviors that they are trying to prevent. Others focus on general health or general social competence development. PA focuses on positive actions — behaviors, thoughts, and feelings – and values. Recent research literature suggests that prevention needs to emphasize youth asset development, resilience, and building on strengths rather than weaknesses or risk factors. PA's approach clearly achieves this. The PA approach brings a comprehensive approach to developing students, school personnel, families, and others through an integrated model that appears to accomplish genuine school reform with one program.

Acknowledgments

The second author developed the program evaluated in the reported study and acquired the school-level data analyzed for this report. The first author conducted the data analyses and interpreted the results. Work on this paper was supported by a grant from the National Institute on Drug Abuse (#DA13474). ■

REFERENCES

1. Flay BR. Positive youth development requires comprehensive health promotion programs. *Am J Health Behav.* 2002;26(6):407-424
2. Jessor R, Jessor SL. Problem Behavior and Psychosocial Development. New York: Academic Press 1977:281.
3. Elliott DS. Health enhancing and health compromising lifestyles. In Milstein SF, Peterson AC, Nightingale EO, (Eds). Promoting Adolescent Health 1993:388.
4. Bryant AL, Schulenberg J, Bachman JG, et al. Understanding the links among school misbehavior, academic achievement, and cigarette use: a national panel study of adolescents. *Prev Sci.* 2000;1(2):71-87.
5. Durlak JA. Common risk and protective factors in successful prevention programs. *Am J Orthopsychiatry.* 1998;68(4):512-520.
6. Flay BR, Allred CG, Ordway N. Effects of the Positive Action program on achievement and discipline: two matched-control comparisons. *Prev Sci.* 2001;2(2):71-90.
7. Slavin RE, Fashola OS. Show Me the Evidence! Proven and Promising Programs for America's Schools. U.S. Dept. of Education, Office of Educational Research and Improvement, Educational Resources Information Center 1998:125.
8. Flay BR. Psychosocial approaches to smoking prevention: a review of findings. *Health Psychol.* 1985;4(5):449-488.
9. Sussman S, Dent CW, Burton D, et al. Developing School-based Tobacco Use Prevention and Cessation Programs. Thousand Oaks, CA: Sage Publications 1995:294.
10. Peters RD, McMahon RJ, (Eds). Preventing Childhood Disorders, Substance Abuse, and Delinquency. Banff International Science Series, Thousand Oaks, CA: Sage 1996;(3):215-240.
11. Tolan P, Guerra N. What works in reducing adolescent violence: an empirical review of the field. The Center for the Study and Prevention of Violence, Institute of Behavioral Sciences, University of Colorado, Boulder 1994:94.
12. Combs A. Perceiving, behaving, becoming: a new focus for educators. ASCD 1962 Yearbook. Washington, DC: Association for Supervision and Curriculum Development 1962:256.
13. Purkey WW. Self-concept and School Achievement. Englewood Cliffs, NJ: Prentice-Hall 1970:86.
14. Purkey W, Novak J. Inviting School Success: A Self-concept Approach to Teaching and Learning. Belmont, CA: Wadsworth 1970:128.
15. Seligman MEP, Csikszentmihalyi M. Positive psychology: an introduction. *Am Psychol.* 2000;55:5-14.
16. Fredrickson BL. Cultivating positive emotions to optimize health and well-being. Prevention and Treatment Volume 3, article 0001a 2000. Available: http://www.journals.apa.org/prevention/volume3/pre0030001a html. Accessed on June 25, 2002.
17. Caine RN, Caine G. Education on the Edge of Possibility. Alexandria,VA: ASCD 1997:279.
18. Bloom BS. All Our Children Learning. New York: McGraw Hill 1981:275.
19. Goleman D. Emotional Intelligence. New York: Bantam Books 1995:352.
20. Gardener H. The Unschooled Mind: How Children Think and How Schools Should Teach. New York: Basic Books 1991:303.
21. Weissberg RP, Gullota TP, Hampton RL, et al. (Eds). Healthy Children 2020: Enhancing Children's Wellness. Thousand Oaks, CA: Sage Publications 1997:278-305.
22. Damasio AR. Descartes' Error: Emotion, Reason, and the Human Brain. New York, NY: Avon Books 1994:312.
23. Akers RL. Deviant Behavior: A Social Learning Approach. 2nd ed. Belmont, CA: Wadsworth Pub. Co. 1977:425.
24. Akers RL. Social Learning and Social Structure: A General Theory of Crime and Deviance. Boston: Northeastern University Press 1998:420.

25.Bandura A. Social Learning Theory. Englewood Cliffs, NJ: Prentice Hall 1977:247.
26.Bandura A. Social Foundations of Thought and Action. Englewood Cliffs, NJ: Prentice Hall 1986:617.
27.Flay BR, Petraitis J. The theory of triadic influence: a new theory of health behavior with implications for preventive interventions. In Albrecht GS (Ed). Advances in Medical Sociology, Vol IV: A Reconsideration of Models of Health Behavior Change. Greenwich, CN: JAI Press 1994:19-44.
28.Petraitis J, Flay BR, Miller TQ. Reviewing the theories of adolescent substance abuse: Organizing pieces of the puzzle. *Psychol Bull.* 1995;117(1):67-86.
29.Battistich V, Hom A. The relationship between students' sense of their school as a community and their involvement in problem behavior. *Am J Public Health.* 1997;87(12):1997-2001.
30.Hawkins JD, Weis JG. The social development model: an integrated approach to delinquency prevention. *Journal of Primary Prevention.* 1985;6:73-97.
31.Comer JP. Educating poor minority children. *Sci Amer.* 1988;259(5):42-48.
32.Haynes NM, Comer JP. Integrating schools, families and communities through successful school reform: The School Development Program. *School Psychology Review*. 1996;4(25):501-506.
33.Haynes NM, Comer JP, Hamilton-Lee M. School climate enhancement through parental involvement. *Journal of School Psychology*. 1989;27:87-90.
34.Allred CG. Positive Action Principal's Kit. Twin Falls, ID: The Positive Action Company 1987.
35.Allred CG. Positive Actions For Living. Twin Falls, ID: Positive Action, Inc. 1995:393.
36.Gorsky RA. Resource reviews: Positive Actions For Living. *Am J Health Promot.* 1996;11(1):15.
37.Allred CG. The development and evaluation of Positive Action: A systematic elementary school self-concept enhancement curriculum [Doctoral thesis]. Provo, UT: Brigham Young University 1984:489.
38.SPSS. SPSS for Windows, Rel. 10.1,.Chicago: SPSS Inc.: 2001.
39.Elias MJ, Gara M, Schuyler T, et al. The promotion of social competence: longitudinal study of a preventive school-based program. *Am J Orthopsychiatry*. 1991;61:409-417.
40.Weissberg RP, Barton HA, Shriver TP. The Social Competence Promotion program for young adolescents. In Albee GW, Gullotta TP, (Eds). Primary Prevention Works. Thousand Oaks, CA: Sage Publications 1997-:268-290.
41.Hawkins JD, Catalano RF, Morrison DM, et al. The Seattle Social Development Project: effects of the first four years on protective factors and problem behaviors. In: McCord J, and Tremblay RE, (Eds). Preventing Antisocial Behavior: Interventions from Birth Through Adolescence. New York: Guilford Press 1992:139-161.
42.O'Donnell J, Hawkins JD, Catalano RF, et al. Preventing school failure, drug use, and delinquency among low-income children: long-term intervention in elementary schools. *Am J Orthopsychiatry*. 1995;65(1):87-161.
43.Battistich V, Schaps E, Watson M, et al. Prevention effects of the Child Development Project: early findings from an ongoing multisite demonstration trial. *Journal of Adolescent Research.* 1996;11(1):12-35.
44.Battistich V, Schaps E, Watson M, et al. Effects of the Child Development Project on students' drug use and other problem behaviors. *Journal of Primary Prevention*. 2000;21(1):75-99.
45.Flay BR, Graumlich S, Segawa E, et al. Effects of two prevention programs on high-risk behaviors among African American youth: a randomized trial. Manuscript, available from author, University of Illinois, March 2003.

JOURNAL OF SCHOOL HEALTH

American School Health Association

RESEARCH ARTICLE

Using Social-Emotional and Character Development to Improve Academic Outcomes: A Matched-Pair, Cluster-Randomized Controlled Trial in Low-Income, Urban Schools*

NILOOFAR BAVARIAN,[a] KENDRA M. LEWIS,[b] DAVID L. DUBOIS,[c] ALAN ACOCK,[d] SAMUEL VUCHINICH,[e] NAIDA SILVERTHORN,[f] FRANK J. SNYDER,[g] JOSEPH DAY,[h] PETER JI,[i] BRIAN R. FLAY[j]

ABSTRACT

BACKGROUND: School-based social-emotional and character development (SECD) programs can influence not only SECD but also academic-related outcomes. This study evaluated the impact of one SECD program, Positive Action (PA), on educational outcomes among low-income, urban youth.

METHODS: The longitudinal study used a matched-pair, cluster-randomized controlled design. Student-reported disaffection with learning and academic grades, and teacher ratings of academic ability and motivation were assessed for a cohort followed from grades 3 to 8. Aggregate school records were used to assess standardized test performance (for entire school, cohort, and demographic subgroups) and absenteeism (entire school). Multilevel growth-curve analyses tested program effects.

RESULTS: PA significantly improved growth in academic motivation and mitigated disaffection with learning. There was a positive impact of PA on absenteeism and marginally significant impact on math performance of all students. There were favorable program effects on reading for African American boys and cohort students transitioning between grades 7 and 8, and on math for girls and low-income students.

CONCLUSIONS: A school-based SECD program was found to influence academic outcomes among students living in low-income, urban communities. Future research should examine mechanisms by which changes in SECD influence changes in academic outcomes.

Keywords: child and adolescent health; emotional health; public health.

Citation: Bavarian N, Lewis KM, DuBois DL, Acock A, Vuchinich S, Silverthorn N, Snyder FJ, Day J, Ji P, Flay BR. Using social-emotional and character development to improve academic outcomes: a matched-pair, cluster-randomized control trial in low-income, urban schools. J Sch Health. 2013; 83: 771-779.

Received on April 16, 2012
Accepted on August 26, 2012

A growing body of research indicates that school-based social-emotional and character development (SECD) and SECD-like programs (eg, social-emotional learning [SEL], positive youth development) can influence health behaviors and academic achievement among low-income minority youth, a population disproportionately affected by disparities in health[1] and education.[2] In their meta-analysis examining the impact of school-based mental health and behavioral programs set in low-income, urban schools, Farahmand et al[3] reported a mean effect size (generally Hedges *g*) on academic outcomes of 0.24. Durlak et al[4] reported a mean effect size (generally Hedges *g*) on academic outcomes of 0.27 in their meta-analysis on school-based SEL programs. With respect to health-related outcomes, the Durlak[4] meta-analysis also showed SEL programs decreased conduct problems (effect size = 0.22) and emotional distress (effect size = 0.24), and improved positive social behaviors (effect size = 0.24). While these findings are

[a]Postdoctoral Fellow, (NBavarian@berkeley.edu), School of Public Health, University of California, Berkeley Prevention Research Center, 1995 University Avenue, Suite 450, Berkeley, CA 94704.

[b]Postdoctoral Researcher, (Kelew@ucdavis.edu), Human Ecology, University of California, Davis One Shields Avenue, Davis, CA 95616.

[c]Associate Professor, (dldubois@uic.edu), School of Public Health, University of Illinois at Chicago, 1747 W. Roosevelt Rd., Chicago, IL 60608.

encouraging, there is a need to accumulate further evidence regarding the capacity of SECD programs to promote academic outcomes, especially when implemented in low-income, urban schools. Accordingly, the primary purpose of this study was to examine the impact of one comprehensive, school-wide SECD program, Positive Action (PA), on academic outcomes using a longitudinal cluster-randomized controlled design in low-income, urban schools.

Positive Action[5] is grounded in theories of self-concept,[6-8] is consistent with social-ecological theories of health behaviors such as the Theory of Triadic Influence (TTI),[9,10] and proposes positive feelings, thoughts, and actions result in fewer negative behaviors and enhanced motivation to learn. The core curriculum is taught through 6 units: self-concept, positive actions for mind and body, positive social-emotional actions focusing on getting along with others, and managing, being honest with, and continually improving oneself. The sequenced classroom curriculum consists of over 140, 15- to 20-minute age-appropriate lessons per grade taught 4 days per week for grades K-6, and 70, 20-minute lessons taught 2 days per week for grades 7 and 8. The PA program also includes teacher, counselor, family, and community training, and school-wide climate development; the school-climate kit, which was used by every school in the trial assigned to the PA condition, focuses on using curriculum lessons and school activities to promote further positive actions amongst students, the school, families, and the community. More information about PA is available at http://www.positiveaction.net.

Prior research has demonstrated that the PA program impacts a range of risk and resilience factors linked to academic outcomes, as well as academic outcomes themselves.[6] In an analysis of 3 longitudinal randomized controlled trials (RCT) of PA involving students aged 6 to 11, PA partially mitigated the decrease in number of positive behaviors endorsed by youth across time.[11] In a matched-pair RCT of PA involving 20 schools in Hawaii, PA was shown to create whole-school contextual change and improve school quality.[12] Students in schools receiving PA were also less likely to engage in substance use, violent behaviors, or sexual activity,[13] and PA schools had significantly higher school-level academic achievement and less absenteeism.[14]

Limitations in prior PA research should be addressed. For example, the academic impact of PA during the middle-school years has not yet been examined. Doing so is critical, as the adolescent years represent a key developmental period with new academic and social demands. Also, the need exists to collect academic-related data from students and teachers so that precursors of academic achievement (eg, engagement with learning) that cannot be measured by school-level archival records alone can be assessed. Finally, the need exists for experimental designs of PA in low-income, urban settings. This study addresses these limitations by (1) following a cohort of students during the elementary- and middle-school years; (2) including student self-reports and teacher ratings of students; and (3) being set in a low-income, urban setting. The purpose was to test the hypothesis that academic performance across time would be better among schools and students receiving PA, than those not receiving PA.

METHODS

Participants

Participating schools were drawn from 483 K-6 and K-8 Chicago Public Schools. Schools were excluded from participation if they (1) were noncommunity schools (eg, charter schools and magnet schools); (2) already had PA or a similar intervention; (3) had an enrollments below 50 or above 140 students per grade; (4) had annual student mobility rates over 40%; (5) had more than 50% of students who passed the Illinois State Achievement Test (ISAT); and (6) had fewer than 50% of students who received free lunch. The latter 2 criteria ensured the selection of high-risk schools.

[d]Professor, (alan.acock@oregonstate.edu), College of Public Health and Human Sciences, Oregon State University, 410 Waldo Hall, Corvallis, OR 97331.

[e]Associate Professor, (vuchinis@oregonstate.edu), College of Public Health and Human Sciences, Oregon State University, 410 Waldo Hall, Corvallis, OR 97331.

[f]Senior Research Specialist, (naida@uic.edu), Institute for Health Research and Policy, University of Illinois at Chicago, 1747 W. Roosevelt Rd., Chicago, IL 60608.

[g]Assistant Professor, (fsnyder@purdue.edu), Department of Health and Kinesiology, Lambert Fieldhouse, 800 West Stadium Avenue, West Lafayette, IN 47907-2046.

[h]Assistant Professor, (jday2@govst.edu), College of Health and Human Services, Governors State University, University Park, IL 60484.

[i]Core Faculty, (pji@adler.edu), Adler School of Professional Psychology, 17 North Dearborn Street, The Adler School—Chicago Campus, Chicago, IL 60602.

[j]Professor, (Brian.Flay@oregonstate.edu), College of Public Health and Human Sciences, Oregon State University, 410 Waldo Hall, Corvallis, OR 97331.

Address correspondence to: Niloofar Bavarian, Postdoctoral Fellow, (NBavarian@berkeley.edu), School of Public Health, University of California, Berkeley Prevention Research Center, 1995 University Avenue, Suite 450, Berkeley, CA 94704.

The findings reported here are based on research funded by grants from the Institute of Education Sciences (IES), U.S. Department of Education:R305L030072, R305L030004 and R305A080253 to the University of Illinois at Chicago (2003-2005) and Oregon State University (2005-2012). The SACD Research Program is a collaboration among IES, the Centers for Disease Control and Prevention's (CDC) Division of Violence Prevention, Mathematica Policy Research Inc. (MPR), and awardees of SACD cooperative agreements (Children's Institute, New York University, Oregon State University, University at Buffalo-SUNY, University of Maryland, University of North Carolina-Chapel Hill, and Vanderbilt University). Moreover, the preparation of this manuscript was supported, in part by the National Institute on Alcohol Abuse and Alcoholism (NIAAA T32 AA014125).

***Indicates CHES continuing education hours are available. Also available at http://www.ashaweb.org/continuing_education.html**

A total of 68 schools met eligibility criteria, of which 18 agreed to participate, and the 7 best-matched pairs (the N that funding would support) were selected for participation; the following variables were used in the matching process: ethnicity, percentage of students who met or exceeded criteria for passing the ISAT, attendance rate, truancy rate, percentage of students who received free lunch, percentage of students who enrolled in or left school during the academic year, number of students per grade, percentage of parents reported to demonstrate school involvement, percentage of teachers employed by the school who met minimal teaching standards, and crime rate for the neighborhood in which the school was located (P. Schochet and T. Novak, unpublished data, 2003).[15-17] A series of t tests revealed that the 7 pairs of schools did not significantly differ from the remainder of the 68 schools eligible for the study, and the PA and control schools were not significantly different from each other on any of the matching variables.[15,16] Throughout the 6 years of the study, 100% of schools were retained.

The total number of students in the analytic sample was 1170, of whom approximately 53% were girls; approximately 48% were African American, 27% Hispanic, and 19% other (eg, White, Asian, Native American, or ''Other''). A total of 247 teachers completed student assessments; 75% of teachers were women; 43% White, 36% African American, 13% Hispanic, and 8% other (eg, Asian and Native American).

Instruments

Student self-report measures. *Disaffection with learning* was assessed using 4 items from a measure of student engagement developed by Furrer and Skinner.[18] Principal components factor analysis on student responses showed this measure loaded strongly onto one factor at both Wave 1 (loadings greater than or equal to 0.66) and Wave 8 (loadings greater than or equal to 0.67). Items were rated on a 4-point Likert scale (''Disagree A LOT'' to ''Agree A LOT'') and included ''When I'm in class, I think about other things'' and ''When I'm in class, my mind wanders.'' A mean of the items was used to create a composite score, whereby higher scores reflected having more disaffection. Cronbach's alpha across the 8 waves of data ranged from 0.64 to 0.71. To assess the impact on *academic grades*, students were asked, ''What grades have you been getting this school year?'' with response options ranging from 1 to 9 (e.g., 1 = Mostly F's, 4 = mix of C's and D's, and 9 = Mostly A's).

Teacher ratings of students. Teachers assessed students using pre-existing measures of academic ability and motivation.[19,20] Each consented student was rated in the areas of reading, mathematics, academic performance, and intellectual functioning using a 5-point Likert scale (1 = Far below grade level to 5 = Far above grade level). Owing to multicollinearity (ie, correlations of 0.84 and higher) between these items, a composite score was created, with higher scores indicating higher teacher ratings of students' *academic ability*. Cronbach's alpha for the composite measure ranged from 0.97 to 0.98. *Academic motivation* was assessed with a single-item measure, with response options ranging from ''Extremely low'' to ''Extremely high.''

School-level archival data. Because state test data provide a policy-relevant measure of achievement,[21] archival reading and math scores of nonEnglish Language Learners on a standardized, school-administered, statewide test (the ISAT) were gathered from the Chicago Public Schools website.[22] The website provided information on the percentages of students tested (all students, grade-specific, and demographic subgroups) whose scores fell into each category (ie, Warning, Not Meeting Standards, Meeting Standards, or Exceeding Standards). A single weighted average of the percentages of students falling into each achievement level was created for each school (ie, [[1 × % of students at Warning level] + [2 × % of students NOT meeting standards] + [3 × % of students meeting standards] + [4 × % of students exceeding standards]]) for both reading and math, overall and by demographic subgroups.

A *value-added metric index of ISAT* performance was also reported by the school district.[23] These indices control for the prior year ISAT scores of students as well as other relevant factors (ie, grade level, gender, race/ethnicity, low income status, English Language Learner status, Individualized Education Plan status, homelessness, and mobility) and are designed to reflect the extent to which scores for a group of students improved (or declined) more than would be predicted based on these factors. Data were available for our student cohort transitioning from grades 7 to 8 (2009 to 2010).

The school district reported average daily attendance rates for each school on a scale from 0 to 100%; these statistics were converted to a measure of average daily *absenteeism* by subtracting 100 from each school's respective year-end attendance.

Procedure

The Chicago trial of PA was longitudinal (ie, 6 years and 8 waves) at the school level and used a place-focused, intent-to-treat design with a dynamic cohort at the student level.[24] Surveys were administered to students beginning in grade 3 (fall 2004), and at 7 additional time points (waves) over 6 years: spring 2005, fall 2005, spring 2006, spring 2007, fall 2008, spring 2009, and spring 2010 (end of grade 8).

Parental consent was obtained before students, parents, or teachers completed surveys when students were in grade 3, with students joining the study at later

waves consented at the time of entry into the study. All students were re-consented for the second phase of funding at Wave 6. At baseline, 79% of parents provided consent; consent rates ranged from 65% to 78% for Waves 2 through 5, and from 58% to 64% for Waves 6 through 8.

The total number of students in the analytic sample across all waves was 1170. Of the original 624 students in grade 3 at the beginning of the trial, only 131 (ie, 21%) remained at grade 8, reflecting the high mobility by low-income urban students. With respect to maintenance of the baseline sample size, 363 students were present at Wave 8 (ie, approximately 61% of the Wave 1 sample size); the decrease in N over time is consistent with the trend among Chicago Public Schools to decrease in size during the study period, together with lower consent rates at Waves 6 through 8.[15]

To substantiate student self-reports, teacher assessments of students and archival data were used. Student assessments were completed by teachers at all waves excepting Wave 6 (the transition from one funding cycle to the next). Percentages of consented students for whom teachers completed ratings for at each wave (excepting Wave 6) ranged from 72% to 93%. Archival ISAT and absenteeism data were collected for the 3 academic years prior to the baseline, as well as throughout the duration of the study.

Data Analyses

Analyses were conducted using Stata version 12.1. Preliminary analyses involved assessing distributions of each outcome and calculating intraclass correlations (ICCs), Cronbach's alphas, and correlations between the student and teacher variables at Waves 1 and 8.

Primary analyses consisted of multilevel growth-curve models to account for all observations and to model school differences. These were 3-level, time within students within schools, analyses for student-level measures, and 2-level, time within schools, analyses for the aggregated school-level data. We used Stata's ''xtmixed'' command for normally distributed outcomes, and ''xttobit'' for outcomes with a positively or negatively skewed distribution (ie, censored below or above, respectively).[25]

A random-intercept model was fitted using the following equations for student- and school-level analysis, respectively:

$$\begin{aligned}\widehat{\mathrm{Y}}_{tij} &= \beta_0 + \beta_1\left(\mathrm{condition}_j\right) \\ &\quad + \beta_2\left(\mathrm{time}_{tij}\right) + \beta_3\left(\mathrm{condition}_j \times \mathrm{time}_{tij}\right) \\ &\quad + \zeta_j + \zeta_{ij} + \epsilon_{tij}\ (\text{Student-level})\end{aligned}$$

$$\begin{aligned}\widehat{\mathrm{Y}}_{tj} &= \beta_{0j} + \beta_1\left(\mathrm{condition}_j\right) \\ &\quad + \beta_2\left(\mathrm{year}_{tj}\right) + \beta_3\left(\mathrm{year}_{tj} \times \mathrm{condition}_j\right) \\ &\quad + \zeta_j + \epsilon_{tj}\ (\text{School-level})\end{aligned}$$

$\hat{\mathrm{Y}}_{tij}$ and $\hat{\mathrm{Y}}_{tj}$ represent the estimated score on a particular outcome at a particular time t (measured as study duration, in years, for student-level models, and as academic year in school-level models). In addition, i represents a student, j represents a school, β_0 represents the mean intercept and ζ_{j} is deviation of a school's mean score from the mean score for all schools. ζ_{ij} is deviation of each student's score from their school's mean, and ϵ_{tij} and ϵ_{tj} are the residual. The original models included quadratic terms for time and the interaction of condition by time. Nonsignificant higher order terms were dropped from the model for parsimony, whereas outcomes with significant quadratic terms (eg, condition $\times$ time^2) were graphed to facilitate interpretation of growth trajectories (not shown).

When applicable, analyses with student-level variables were run using both the fully reduced random-intercept and random-coefficients models, with the former model nested within the latter model. A likelihood-ratio test was performed to determine whether the random-coefficients model was a better fit for the data.[25]

Due to the power and sample size limitations, and because the a priori directional hypothesis was that the PA schools would have greater improvements across time, one-tailed p-values were used in tests of effects of the PA program on school-level outcomes.[26] In the analyses using ISAT weighted averages, 6 matched pairs were retained (for reasons discussed elsewhere);[15] all 7 matched pairs were retained for the end-point value-added ISAT analysis and for the absenteeism growth-curve analysis. For all outcomes (student-level and school-level) analyzed using growth-curve analyses, effect sizes were calculated using the method described by Lipsey and Wilson.[27]

Sensitivity analyses assessed the robustness of results from the primary analyses. A first approach involved including a ''pairs'' variable as an additional level in each of the best-fitting models to determine whether adding a fourth level would affect findings. Second, to provide a more conservative test (from a statistical power perspective) of program effects for each outcome, the test statistic provided by Stata (which assumes a large sample size) in the primary analyses (N = 14 schools) was compared to the critical value for a 2-tailed t-distribution with 12 degrees of freedom at a 95% confidence level (2.18).[28]

For student-level data, the possible *moderating effects* of sex and student mobility were examined. The effect of student mobility groups was examined using results from a latent class analysis (LCA)[15] in which a 5-class solution was found to be the most appropriate fit for the data: (1) stayers (average study duration of 5.72 years, N = 158); (2) temporary participants (present for grade 4 and/or 5 only; average study duration of 1.30 years; N = 196); (3) late joiners (average study duration of 1.38 years; N = 308); (4) early leavers

Table 1. **Youth and Teacher Reports of Academic Outcomes: Correlations at Wave 1 (above the diagonal, N = 603) and Wave 8 (below the diagonal, N = 335)**

Variables	1	2	3	4	5	6	7
Student Self Reports							
1. Disaffection with Learning	—	−0.04	−0.31**	−0.29**	−0.32**	−0.29**	−0.27**
2. Self-Reported Grades	−0.23**	—	0.24**	0.17**	0.21**	0.21**	0.20**
Teacher Ratings of Students							
3. Reading	−0.03	0.33**	—	0.84**	0.89**	0.93**	0.71**
4. Math	−0.06	0.37**	0.93**	—	0.84**	0.87**	0.67**
5. Intellectual Functioning	−0.01	0.29**	0.91**	0.89**	—	0.91**	0.71**
6. Academic Performance	−0.07	0.34**	0.93**	0.93**	0.92**	—	0.73**
7. Academic Motivation	−0.09	0.44*	0.67**	0.67**	0.64**	0.68**	—

*p<0.05; **p<0.01.

(average study duration of 0.94 years; N = 263); and (5) late leavers (average study duration of 3.23 years; N = 287); stayers served as the reference group.

RESULTS

The ICCs for the student-level measures were generally low, with none of the ICCs for student-reported and only 1 of the 14 ICCs for teacher-reported outcomes above 0.10. Scale reliabilities (reported above) were generally high, with a clear increase in Cronbach's alphas as students aged. Table 1 shows the correlations between the student and teacher variables at Waves 1 (beginning of grade 3) and 8 (end of grade 8).

Program effects (significant condition × time and condition × time^2 interactions) were present for disaffection with learning (Table 2). Students in PA schools started off higher than those in control schools (ie, more reported disaffection with learning). There was then an overall trend toward a net increase in disaffection with learning by the end of the study period in both PA and control schools; the pattern of change was linear in control schools and curvilinear within PA schools.

As shown in Table 2, there was evidence of a program effect on teacher ratings of student academic motivation in the form of significant condition × time and condition × time^2 interactions. For students in *PA* schools, after an initial period of modest decline there was an accelerating increase, whereas for control school students there was a gradually decreasing trend. The net result was notably higher predicted levels of teacher-rated academic motivation for students in PA schools. Sensitivity analyses at the pair level supported this finding (results not shown).

With respect to teacher-rated academic ability, a significant condition × time interaction was found in the random-intercept model. In the random-coefficients model, which provided a better fit, the condition × time interaction was not significant (B = 0.03, p < .05 in random-intercept model; B = 0.02, p > .05 in random-coefficients model). For both teacher-rating measures, there was no evidence of moderation of program effects by mobility group; gender moderation was observed for academic ability, with PA boys being rated higher by teachers than control boys.

Growth-curve analyses for the weighted composite measure of ISAT scores for all students in PA and non-PA schools did not reveal evidence of a program effect for Reading. There was, however, evidence of marginal program effects for Math (Table 3). When ''pairs'' was included in the random-intercept model, this finding remained marginal (results not shown). With respect to demographic subgroups, significant

Table 2. **Multilevel Growth-Curve Model Estimates for Student-Level Measures (N = 1170 students) and Aggregated School-Level (N = 14 schools) Archival Measures**

Measure	Model Run	Intercept B (SE)	Time B (SE)	Time^2 B (SE)	Condition (0 = Non-PA; 1 = PA) B (SE)	Condition × Time B (SE)	Condition × Time^2 B (SE)
Student Self-Reports							
Disaffection with Learning	Random Intercept	1.69 (0.06)**	0.03 (0.04)	0.01(0.01)	0.15 (0.08)*	−0.20(0.06)**	0.03 (0.01)**
Self-Reported Grades	Random Intercept	7.89 (0.12)**	−0.81(0.07)**	0.11(0.01)**	0.10(0.17)	0.01(0.03)	—
Teacher Ratings of Students							
Academic Performance†	Random Coefficients	2.62 (0.06)**	−0.05(0.03)*	0.02(0.005)**	−0.06(0.08)	0.02(0.02)	—
Academic Motivation	Random Coefficients	3.01(0.07)**	0.04(0.04)	−0.01(0.01)	0.05(0.10)	−0.12(0.06)*	0.03(0.01)**
School-Level Archival Data‡							
Absenteeism	Random Intercept	6.76 (0.56)**	0.03 (0.05)	—	0.43 (0.65)	−0.16 (0.07)*	—

+p < .10; *p < .05; **p < .01.
†For the random-intercept model, the condition × time interaction is significant at the .05 level (B = 0.03, p < .05).
‡For school-level measures, time variable created using academic year, rather than time since implementation of intervention. Also, the one-tailed p-value is reported for school-level measures.

Table 3. **Multilevel Random-Intercept Growth-Curve Model Estimates for Standardized Academic Test Scores *(N = 12 Schools)**

Variables	Intercept† B (SE)	Time† B (SE)	Time²† B (SE)	Condition (0 = Non-PA; 1 = PA) B (SE)	Condition × Time B (SE)	Condition × Time One-tailed *p*-value
Reading						
All Students (Grades 3 to 8 Combined)	2.26 (0.07)	0.17 (0.01)	−0.02 (0.002)	0.04 (0.10)	0.01 (0.01)	0.16
Subgroups						
Boys	2.22 (0.07)	0.16 (0.02)	−0.02 (0.003)	−0.001 (0.10)	0.01 (0.01)	0.10
Girls	2.30 (0.07)	0.17 (0.02)	−0.02 (0.003)	0.07 (0.10)	0.004 (0.01)	0.35
African Americans	2.20 (0.06)	0.15 (0.02)	−0.01 (0.003)	0.05 (0.08)	0.01 (0.01)	0.10
African American Girls	2.21 (0.05)	0.17 (0.02)	−0.02 (0.003)	0.13 (0.07)	−0.01 (0.01)	0.23
African American Boys	2.17 (0.07)	0.16 (0.03)	−0.02 (0.005)	−0.02 (0.10)	0.03 (0.01)	0.02
Free or Reduced-Price Lunch	2.25 (0.07)	0.17 (0.01)	−0.02 (0.002)	0.03 (0.09)	0.01 (0.01)	0.18
Math						
All Students (Grades 3 to 8 Combined)	2.15 (0.08)	0.24 (0.02)	−0.03 (0.003)	0.04 (0.12)	0.01 (0.01)	0.07
Subgroups						
Boys	2.12 (0.09)	0.24 (0.02)	−0.03 (0.004)	0.04 (0.12)	0.01 (0.01)	0.12
Girls	2.18 (0.08)	0.24 (0.02)	−0.03 (0.004)	0.04 (0.11)	0.02 (0.01)	0.09
African Americans	2.06 (0.06)	0.23 (0.02)	−0.02 (0.004)	0.06 (0.08)	0.02 (0.01)	0.11
African American Girls	2.09 (0.07)	0.23 (0.03)	−0.02 (0.004)	0.10 (0.09)	0.02 (0.01)	0.11
African American Boys	2.02 (0.07)	0.25 (0.03)	−0.03 (0.005)	0.04 (0.09)	0.02 (0.01)	0.12
Free or Reduced-Price Lunch	2.15 (0.08)	0.24 (0.02)	−0.03 (0.003)	0.04 (0.11)	0.01 (0.01)	0.07

*The average of values from 2000/2001 through 2002/2003 was used as the estimate of baseline levels.
†The coefficients for Intercept, Time, and Time² were all significant at the .01 level, except the time² coefficient for African American boys, which was significant at the .05 level.

condition × time interactions were seen in Reading performance for African American boys (B = 0.03, one-tailed $p < .05$). The condition × time interaction remained significant in the pair-level analysis (results not shown). Marginal results (p-values less than or equal to .10) indicative of favorable growth in PA schools as compared to control schools, were observed for Reading performance for boys and African American students, and for Math performance for girls and students receiving free or reduced-price lunch.

End-point regression analyses for our study cohort, using the value-added metric of the same standardized test, showed significant results in Reading, but not Math. As compared to students in control schools making the grade 7 to 8 transition, students in PA schools performed significantly better in reading (B = 1.26, *one-tailed* p = 0.013, effect size = 0.83, results not shown).

As shown in Table 2, growth-curve analyses showed there was lower absenteeism at PA schools than control schools (B = −0.16, *one tailed* p = 0.015). Sensitivity analyses using the pair-level variable and the adjusted degrees of freedom supported these findings (results not shown).

Table 4 shows the estimated means of our outcomes at baseline and end point, as well as the effect sizes for each outcome. The largest effect sizes for school-level measures were for absenteeism (effect size = −0.78) and reading performance on the ISAT for African American boys (effect size = 1.50). With respect to student-level measures, the largest effect size was observed for teacher ratings of academic motivation (effect size = 0.39).

DISCUSSION

In the Chicago trial of PA, the intervention had a positive impact on absenteeism, mitigated a natural increase in disaffection with learning, and PA teachers rated their students as experiencing greater growth in academic motivation and ability; these findings are encouraging, as these outcomes are predictors of long-term academic achievement and school completion.[29-31] Favorable growth was also observed with respect to ISAT Reading and Math performance, particularly for African American boys and students receiving free or reduced-price lunch. Socioeconomic background (ie, low-income), sex (ie, being male) and ethnicity (ie, African American, Hispanic, and Native American youth) are known predictors of school dropout, and school dropout is associated with a multitude of negative outcomes.[30] As prevention programs can only influence those factors amenable to change (eg, motivation to learn), it is encouraging that this trial also demonstrated improvements in test scores for these high-risk groups.

The impact on academic-related outcomes observed in this study may be attributed to a number of factors. For example, the skills fostered by the PA program (eg, problem solving, self-control, emotional regulation, and attention), and lesson plans focusing on improving motivation to learn and do well in school, may in part explain the observed results.[5] In addition, the

Table 4. **Estimated Means and Effect Sizes for Student- and School-Level Data**

Measure	Response Options	Model Run	Wave 1 Control	PA	Wave 8 Control	PA	Effect Size*
Student Self-Reports							
Disaffection with Learning	1 to 4	Random Intercept	1.69	1.85	2.19	2.19	−0.19
Self-Reported Grades	1 to 9	Random Intercept	7.89	7.98	6.67	6.81	0.02
Teacher Ratings of Students							
Academic Ability	1 to 5	Random Coefficients	2.63	2.57	2.84	2.91	0.14
Academic Motivation	1 to 5	Random Coefficients	3.01	3.06	2.80	3.24	0.39
School-Level Archival Data†							
Absenteeism	0 to 100	Random Intercept	6.76	6.33	6.95	5.58	−0.78
ISATs-Reading	1 to 4						
All Students (Grades 3 to 8 Combined)	1 to 4	Random Intercept	2.26	2.29	2.64	2.72	0.22
Boys	1 to 4	Random Intercept	2.22	2.22	2.60	2.66	0.33
Girls	1 to 4	Random Intercept	2.30	2.37	2.68	2.78	0.11
African Americans	1 to 4	Random Intercept	2.20	2.25	2.62	2.74	0.50
African American Girls	1 to 4	Random Intercept	2.21	2.34	2.66	2.74	−0.54
African American Boys	1 to 4	Random Intercept	2.17	2.15	2.57	2.72	1.50
Free or Reduced-Price Lunch	1 to 4	Random Intercept	2.25	2.28	2.63	2.70	0.23
ISATs-Math	1 to 4	Random Intercept					
All Students (Grades 3 to 8 Combined)	1 to 4	Random Intercept	2.15	2.19	2.67	2.79	0.38
Boys	1 to 4	Random Intercept	2.12	2.17	2.67	2.79	0.31
Girls	1 to 4	Random Intercept	2.18	2.22	2.68	2.81	0.41
African Americans	1 to 4	Random Intercept	2.06	2.12	2.62	2.77	0.55
African American Girls	1 to 4	Random Intercept	2.09	2.19	2.61	2.80	0.69
African American Boys	1 to 4	Random Intercept	2.02	2.07	2.62	2.76	0.63
Free or Reduced-Price Lunch	1 to 4	Random Intercept	2.15	2.19	2.67	2.79	0.42

*Effect size calculations made using estimated means. Namely, the estimated mean difference at the baseline was subtracted from the estimated mean difference at the end point to obtain the difference of differences, and this value was then divided by the pooled standard deviation at baseline.
†For school level measures, time variable created using academic year, rather than time since implementation of the Positive Action (PA) intervention.

promotion of positive behaviors may have resulted in less time being spent by teachers on classroom management and, subsequently, more time devoted to interactive strategies that create an intellectually stimulating environment.[5] Moreover, the impact on academics may have been mediated through improvements in attachment to school and teachers.

This study is the first to examine the academic impact of PA in a low-income, urban setting, and supplements Snyder et al's[14] findings on the academic impact of PA in Hawaii by including data from students and teachers of students in the elementary- and middle school grades. The study also adds to the research of Madsen et al,[32] who evaluated the impact of a physical-activity focused, school-based, Positive Youth Development program in low-income Bay Area California schools using a quasi-experimental time series design; namely, the researchers found that each additional year of exposure to the program resulted in significantly higher scores in meaningful participation in school and academic-related goals and aspirations of youth. In this study, for those measures with significant program effects, the effect size for disaffection with learning (effect size = −0.19) was smaller than the effect sizes for academic outcomes reported by the research teams led by Farahmand[3] and Durlak.[4] However, other measures in this study (eg, academic motivation, absenteeism, ISAT Math results) had larger effect sizes than those observed in the aforementioned studies.

Limitations

This study is not without its limitations. Student and teacher-reports on academic measures are subject to social desirability bias; this potential bias was addressed by supplementing student and teacher reports with archival measures representing the actual performance of students on standardized tests. Another possible limitation of the study is that students in the intervention group may have acted differently because they knew they were receiving the PA program, a phenomenon known as the Hawthorne effect. This limitation was addressed through the trial's use of a control group of students and teachers who were also aware they were being observed as part of a study. With respect to external validity, the findings are generalizable only to similar schools (ie, low-income, urban schools) that would self-select to participate in a trial of this nature. The small number of pairs and schools (ie, 7 and 14, respectively) could influence statistical power; however, that significant findings were found in primary and sensitivity analyses suggest that our findings are robust. In addition, student mobility led to high turnover of students, which is problematic as it can become difficult to determine whether observed effects can be attributed to the

intervention or differential attrition.[24] One approach to analyzing mobility patterns is LCA,[33,34] and this study contributes to the LCA literature by examining students who enter a study, not just those who exit;[15] program effects were not found to differ by mobility class.

Limitations notwithstanding, the present study has several strengths. The longitudinal nature of this RCT allowed examination of school performance across 6 years, encompassing both elementary- and middle-school grades. The data from multiple sources as well as the sensitivity analyses provide confidence in study findings. In addition to standardized test performance, our study also reported on theoretically expected mediators of academic success (eg, disaffection with learning). Moreover, this study involved a sample of students in a high-risk setting. Thus, policy makers aiming to alleviate educational disparities should use scientific data from this and other evidence-based studies to advocate for comprehensive school-based SECD programming.

Conclusions

Findings from this study reinforce prior findings that SECD-like programs can improve academic achievement as well as improve student behavior and health. Future studies should determine the mechanism by which SECD programs such as Positive Action improve academic outcomes (eg, mediation through factors that SECD programs seek to foster, such as attachment with teacher and school, improved school climate, emotional regulation, attention, executive function, and increased self-control). Future research could also supplement student and teacher reports by gathering data from parents that may influence academic performance (eg, parent's highest level of education).

IMPLICATIONS FOR SCHOOL HEALTH

In an era where increased pressures to ''teach to the test'' may lead school officials to feel as though they have neither the time nor money to invest in evidence-based prevention programming,[35] there is an increasing need to demonstrate the impact that multifaceted prevention programs can have on academic performance and student and community wellness.[36] When taken together with preliminary research showing the impact of this trial on health behaviors,[37] results from this study demonstrate the possibility of addressing the proverbial ''2 birds'' (ie, health and academics) with ''1 stone'' (ie, school-based SECD programs).

Human Subjects Approval Statement

The research presented herein was approved by the institutional review boards of Oregon State University and the University of Illinois at Chicago, the Research Review Board at Chicago Public Schools and the Public/Private Ventures Institutional Review Board for Mathematica Policy Research Inc.

The SACD research program includes multi-program evaluation data collected by MPR and complementary research study data collected by each grantee. The findings reported here are based only on the Chicago portion of the multi-program data and the complementary research data collected by the University of Illinois at Chicago and Oregon State University (Brian Flay, Principal Investigator) under the SACD program.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Institute of Education Sciences, CDC, MPR, or every Consortium member, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

We are extremely grateful to the participating CPS schools, their principals, teachers, students, and parents. We also thank the CPS Research Review Board and Office of Specialized Services, especially Drs. Renee Grant-Mitchell and Inez Drummond, for their invaluable support of this research.

CONFLICT OF INTEREST

The research described herein was done using the program, the training, and technical support of Positive Action, Inc. in which Dr. Flay's spouse holds a significant financial interest. Issues regarding conflict of interest were reported to the relevant institutions and appropriately managed following the institutional guidelines.

REFERENCES

1. Braveman PA, Cubbin C, Egerter S, Williams DR, Pamuk E. Socioeconomic disparities in health in the United States: what the patterns tell us. *Am J Public Health*. 2010;100(S1):S186-S196.
2. Aud S, Fox MA, Kewal Ramani A. *Status and Trends in the Education of Racial and Ethnic Groups. NCES 2010-015*. National Center for Education Statistics: Washington, DC; 2010.
3. Farahmand FK, Grant KE, Polo AJ, Duffy SN. School-based mental health and behavioral programs for low-income, urban youth: a systematic and meta-analytic review. *Clin Psychol*. 2011;18(4):372-390.
4. Durlak JA, Weissberg RP, Dymnicki AB, Taylor RD, Schellinger KB. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Dev*. 2011;82(1):405-432.
5. Flay BR, Allred CG. The positive action program: improving academics, behavior, and character by teaching comprehensive skills for successful learning and living. In: Lovat T, Toomey R, Clement N, eds. *International Research Handbook on Values Education and Student Wellbeing*. Dordrecht, the Netherlands: Springer; 2010:471-501.

6. DuBois DL, Flay BR, Fagen MC. Self-esteem enhancement theory. In: DiClemente RJ, Crosby RA, Kegler MC, eds. *Emerging Theories in Health Promotion Practice and Research*. 2 ed. San Francisco, CA: Jossey-Bass; 2009:97-130.
7. Purkey WW. *Self-concept and School Achievement*. Englewood Cliffs, NJ: Prentice-Hall; 1970.
8. Purkey WW, Novak J. *Inviting School Success: A Self-concept Approach to Teaching and Learning*. Belmont, CA: Wadsworth; 1970.
9. Flay BR, Petraitis J. The theory of triadic influence: a new theory of health behavior with implications for preventive interventions. In: Albrecht G, ed. *Advances in Medical Sociology*, Vol. 4. Greenwich, CT: JAI Press; 1994:19-44.
10. Flay BR, Snyder F, Petraitis J. The theory of triadic influence. In: DiClemente RJ, Crosby RA, Kegler MC, eds. *Emerging Theories in Health Promotion Practice and Research*. 2 ed. San Francisco, CA: Jossey-Bass; 2009:451-510.
11. Washburn IJ, Acock A, Vuchinich S, et al. Effects of a social-emotional and character development program on the trajectory of behaviors associated with social-emotional and character development: findings from three randomized trials. *Prev Sci*. 2011;12(3):314-323.
12. Snyder FJ, Vuchinich S, Acock A, Washburn IJ, Flay BR. Improving elementary school quality through the use of a social-emotional and character development program: a matched-pair, cluster-randomized, controlled trial in Hawaii. *J Sch Health*. 2012;82(1):11-20.
13. Beets MW, Flay BR, Vuchinich S, et al. Use of a social and character development program to prevent substance use, violent behaviors, and sexual activity among elementary-school students in Hawaii. *Am J Public Health*. 2009;99(8):1438-1445.
14. Snyder F, Vuchinich S, Acock A, et al. Impact of the Positive Action program on school-level indicators of academic achievement, absenteeism, and disciplinary outcomes: a matched-pair, cluster randomized, controlled trial. *J Res Educ Eff*. 2010;3(1): 26-55.
15. Lewis KM, DuBois DL, Ji P, et al. Design, sample and planned analysis of the Chicago trial of the Positive Action program. Oregon State University. In review.
16. Ji P, DuBois DL, Flay BR, Brechling V. ''Congratulations, you have been randomized into the control group!(?)'': issues to consider when recruiting schools for matched-pair randomized control trials of prevention programs. *J Sch Health*. 2008;78(3):131-139.
17. Li K-K, Washburn I, DuBois DL, et al. Effects of the Positive Action programme on problem behaviours in elementary school students: a matched-pair randomised control trial in Chicago. *Psychol Health*. 2011;26(2):187-204.
18. Furrer C, Skinner E. Sense of relatedness as a factor in children's academic engagement and performance. *J Educ Psychol*. 2003;95(1):148-162.
19. Gresham FM, Elliot SN. *Social Skills Rating System*. Circle Pines, MN: American Guidance Service; 1990.
20. Achenbach TM. *Manual for the Teacher's Report Form and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry; 1991.
21. Somers M, Zhu P, Wong E. *Whether and How to Use State Tests to Measure Student Achievement in a Multi-State Randomized Experiment: An Empirical Assessment Based on Four Recent Evaluations (NCES 2012-4015). US Department of Education, National Center for Education Statistics*. Washington, DC: US Government Printing Office; 2011.
22. Chicago Public Schools. Chicago Public Schools. Available at: http://www.cps.edu/Pages/home.aspx. Accessed March 6, 2012.
23. Chicago Public Schools. FAQ on the value-added metric. Available at: http://research.cps.k12.il.us/export/sites/default/accountweb/Research/ValueAdded/valueadd_faq.pdf. Accessed March 6, 2012.
24. Vuchinich S, Flay BR, Aber L, Bickman L. Person mobility in the design and analysis of cluster-randomized cohort prevention trials. *Prev Sci*. 2012;13(3):300-313.
25. Rabe-Hesketh S, Skrondal A. *Multilevel and Longitudinal Modeling Using Stata*. 2 ed. College Station, TX: Stata Press; 2008.
26. Knottnerus JA, Bouter LM. The ethics of sample size: two-sided testing and one-sided thinking. *J Clin Epidemiol*. 2001;54(2):109-110.
27. Lipsey MW, Wilson DB. *Practical Meta-analysis*. Thousand Oaks, CA: Sage; 2001.
28. Raudenbush SW, Bryk AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2 ed. Newbury Park, CA: Sage; 2002.
29. Suh S, Suh J, Houston I. Predictors of categorical at-risk high school dropouts. *J Couns Dev*. 2007;85(2):196-203.
30. Lehr CA, Johnson DR, Bremer CD, Cosio A, Thompson M. *Essential Tools: Increasing Rates of School Completion: Moving From Policy and Research to Practice*. Minneapolis, MN: ICI Publications Office; 2004.
31. Caraway K, Tucker CM, Reinke WM, Hall C. Self-efficacy, goal orientation, and fear of failure as predictors of school engagement in high school students. *Psychol Schools*. 2003;40(4):417-427.
32. Madsen KA, Hicks K, Thompson H. Physical activity and positive youth development: impact of a school-based program. *J Sch Health*. 2011;81(8):462-470.
33. Beunckens C, Molenberghs G, Verbeke G, Mallinckrodt C. Latent-class mixture model for incomplete longitudinal Gaussian data. *Biometrics*. 2008;64(1):96-105.
34. Roy J. Modeling longitudinal data with nonignorable dropouts using a latent dropout class model. *Biometrics*. 2003;59(4):829-836.
35. Kaftarian S, Robinson E, Compton W, Davis BW, Volkow N. Blending prevention research and practice in schools: critical issues and suggestions. *Prev Sci*. 2004;5(1):1-3.
36. Greenberg MT. Current and future challenges in school-based prevention: the researcher perspective. *Prev Sci*. 2004;5(1):5-13.
37. Bavarian N, Lewis KM, Acock A, et al. Effects of a social-emotional and character development program on adolescent health behaviors and outcomes: a matched-pair, cluster-randomized controlled trial. Oregon State University. In review.