Effects of a School-Based Social-Emotional and Character Development Program on Health Behaviors: A Matched-Pair, Cluster-Randomized Controlled Trial

Niloofar Bavarian, Ph.D.[1], **Kendra M. Lewis, Ph.D.**[2], **Alan Acock, Ph.D.**[3], **David L. DuBois, Ph.D.**[4], **Zi Yan, Ph.D.**[5], **Samuel Vuchinich, Ph.D.**[3], **Naida Silverthorn, Ph.D.**[4], **Joseph Day, Ph.D.**[6], and **Brian R. Flay, D.Phil.**[3]

[1]College of Health & Human Services, California State University, Long Beach Long Beach, CA

[2]4-H Youth Development Program, University of California, Division of Agriculture and Natural Resources, Davis, CA

[3]College of Public Health and Human Sciences, Oregon State University, Corvallis, Oregon

[4]School of Public Health, University of Illinois at Chicago, Chicago, IL

[5]Health Sciences, Merrimack College, North Andover, MA

[6]College of Health and Human Services, Governors State University, Chicago, IL

Abstract

There is considerable research that suggests that school-based social-emotional programs can foster improved mental health and reduce problem behaviors for participating youth; in contrast, much less is known about the impact of these programs on physical health, even though some of these programs also include at least limited direct attention to promoting physical health behaviors. We examined the effects of one such program, *Positive Action* (*PA*), on physical health behaviors and body mass index (BMI), and tested for mediation of program effects through a measure of social-emotional and character development (SECD). Participating schools in the matched-pair, cluster-randomized trial were 14 low-performing K-8 Chicago Public Schools. We followed a cohort of students in each school from grades 3 to 8 (eight waves of data collection; 1,170 total students). Student self-reports of health behaviors served as the basis for measures of healthy eating and exercise, unhealthy eating, personal hygiene, consistent bedtime, and SECD. We collected height and weight measurements at endpoint to calculate age- and gender-adjusted BMI *z*-scores. Longitudinal multilevel modeling analyses revealed evidence of favorable program effects on personal hygiene (effect size [ES] = 0.48), healthy eating and exercise (ES = 0.21), and

unhealthy eating (ES = −0.19); in addition, BMI *z*-scores were lower among students in *PA* schools at endpoint (ES = −0.21). Program effects were not moderated by either gender or student mobility. Longitudinal structural equation modeling demonstrated mediation through SECD for healthy eating and exercise, unhealthy eating, and personal hygiene. Findings suggest that a social-emotional and character development program without a primary focus on health behavior promotion can have a modest impact on outcomes in this domain during the childhood to adolescence transition.

Keywords

Introduction

A growing body of research points to the relationship between health and student learning. For example, a longitudinal study by London and Castrechini (2011) found that youths' cardiovascular fitness predicted better academic performance (e.g., standardized math scores). Similarly, a literature review by Basch (2011) concluded that dietary behaviors (e.g., consuming a nutrient-rich breakfast) are associated with indicators of cognition (e.g., problem solving). Furthermore, a meta-analysis by Dewald and colleagues (2010) found evidence of an association between various indicators of sleep and school performance. As such, addressing health behaviors within the school setting may be of significant benefit to youths' academic- *and* health-related development. With respect to primary prevention, it therefore may be beneficial to implement evidence-based programs that can affect both health and learning. This may be especially the case in settings known to be affected by disparities in both educational and health outcomes, such as low-income urban communities (Aud, Fox & KewalRamani, 2010; Orsi et al., 2010).

Traditional approaches to health promotion have focused on addressing the most proximal potential causes of behavior (e.g., promoting exercise by addressing attitudes towards exercise). However, as proposed by a number of theoretical perspectives (e.g., Flay et al., 2009), more distal skills and attitudes (e.g., self-control) may affect a wide range of behaviors, including those that are likely to promote or compromise physical health and well-being. As such, school-based programs that are designed to facilitate the development of core social-emotional competencies (e.g., self-awareness, self-control, social awareness, prosocial skills, responsible decision making; Collaborative for Academic, Social, and Emotional Learning, 2005) could also potentially have an influence on physical health behaviors and related outcomes, through improvements in social-emotional skills (Durlak et al., 2011). Given that schools are increasingly tasked with addressing overall health and well-being, social-emotional programs implemented within the school context should be evaluated for their impact on health-related outcomes.

Positive Action (*PA*) is one example of a school-based program with a curriculum designed to foster the social-emotional, intellectual-, and physical-self (Flay & Allred, 2010). The *PA* program aims to develop these social-emotional skills, and positive actions for the physical

self (e.g., health-promoting behaviors) represent a very small percentage (~8%) of the *PA* curriculum. Cluster-randomized controlled trials of *PA* in Hawai'i, Chicago, and a southeastern rural school district showed positive program effects on a measure of social-emotional and character development (Washburn et al., 2011). The program has also been shown to improve emotional health (Lewis et al., 2013a) and engagement with learning (Bavarian et al., 2013; Snyder et al., 2010), and to reduce such health-compromising behaviors as substance use (Beets et al., 2009; Li et al., 2011; Lewis et al., 2012), violence (Beets et al., 2009; Lewis et al., 2013b), and sexual activity (Beets et al., 2009). In the Chicago trial of *PA*, improvements in substance use and emotional health were mediated through changes in social-emotional and character development (SECD; Lewis et al., 2012; Lewis et al., 2013a). These findings provide a basis for hypothesizing improved physical health behaviors and status among students exposed to the *PA* program, as well as mediation through relative improvements in social-emotional competence. To date, however, the program's impact on physical health outcomes, as mediated by social-emotional improvements, has not been examined.

We used data from a Chicago cluster-randomized controlled trial of *PA* to examine the impact of the program on several health-related behaviors and on one physical health outcome, body mass index; our secondary goal was to test for mediation through changes in a measure of social-emotional and character development (SECD). Our central hypothesis was that students in schools implementing *PA*, compared to those in control schools, would exhibit more favorable change over the course of the study on behaviors relating to eating, exercise, personal hygiene, and sleep, and at the study's end-point would have lower body mass index, and that these effects would be mediated through relative improvements in SECD.

Methods

Design and Sample

For the Chicago trial of *Positive Action*, we received approval of study procedures by the Institutional Review Boards of the University of Illinois at Chicago and Oregon State University, the Research Review Board at Chicago Public Schools (CPS), and the Public/Private Ventures Institutional Review Board for Mathematica Policy Research (MPR). MPR was contracted to collect data during grades 3–5 as part of the Chicago trial's involvement in a seven-site Social and Character Development Research Consortium (SACD) that was funded by the Institute of Education Sciences.

We drew the study sample from the 483 K-6 and K-8 schools in the CPS system. We excluded schools from involvement in the trial if they: 1) were non-community schools (e.g., magnet schools), 2) already had *PA* or a similar intervention, 3) had an enrollment rate 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, and 6) had fewer than 50% of students who received free or reduced-price lunch; these criteria ensured selection of youth who could be considered "high-risk" (Ji et al., 2008). Of the 68 schools meeting eligibility criteria, we successfully contacted 36 who were also able to attend

information sessions; of these 36, 18 agreed to participate given the understanding that they would be randomly assigned to *PA* or a control condition (Ji et al., 2008).

From these 18 schools, we created 9 best-matched pairs (matched on demographic and performance variables such as ethnicity, attendance rate, truancy rate, percentage of students on free or reduced-price lunch, and achievement test scores) using a SAS computer program provided by MPR; within each pair, we randomly assigned schools to *PA* or control using a random-number generation function (Ji et al., 2008). Given the nature of the intervention, blinding of condition status (e.g., by schools, students, teachers) was not possible. Funding allowed for 7 of the 9 pairs of schools to be included in the study; we gave preference for inclusion to pairs that best represented the ethnic diversity within the CPS system (Ji et al., 2008). The 7 pairs of schools we selected did not significantly differ from the remainder of the 68 schools eligible for the study (Ji et al., 2008). Matching and randomization were successful, as evidenced by baseline equivalency between *PA* and control schools on both the matching variables and aggregate and survey data (Lewis et al., 2012).

We implemented the program school-wide beginning in the fall of 2004 and continued doing so throughout the six-year study (the maximum amount of time possible within funding limits); we asked control schools to conduct "business-as-usual." Periodic monitoring of control schools indicated that although their implementation of activities directed toward character and social-emotional development was commonplace (SACD, 2010), there was no implementation of programs similar in scope or intensity to that of *PA*. Students in the treatment schools had their first exposure to *PA* lessons during the 2004/2005 academic year; students in the study cohort received developmentally appropriate *PA* lessons (i.e., those designed for students' particular grade level) from grade 3 to grade 8. We gave teachers and other school staff flexibility to adapt lessons and other program activities for purposes such as cultural appropriateness and tailoring to the school environment (e.g., changing names of story characters to be more culturally relevant), as long as the core content of lessons was maintained. Implementation, as measured by different indices (e.g., trainings attended, teacher description of *PA* activities during classroom time; Bickman et al., 2009), varied across schools and improved across time; by the final year of the study, key program benchmarks were being implemented at moderate (50%–59%), moderate-to-high (60–69%), and high (above 70%) levels by one, three, and three schools, respectively (Li et al., 2011).

Students completed surveys in the fall of 2004 (beginning of grade 3), spring 2005, fall 2005, spring 2006, spring 2007, fall 2008, spring 2009, and spring 2010 (end of grade 8). Our longitudinal trial used a cluster-focused, intent-to-treat design with a dynamic grade cohort (Brown et al., 2008; Vuchinich et al., 2012) at the student level. That is, we followed a cohort of students starting in grade 3 through to the end of grade 8, we included students who entered study schools after baseline in the analysis and we excluded students who left participating schools.

We obtained parental consent before students completed surveys; we obtained parental consent from all students upon entry into the study, and 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 58% to 64% for Waves 6 through 8. The total number of students enrolled in the study across all eight waves of data collection was 1,170, of whom approximately 53% were female; approximately 48% were African American, 27% Hispanic and 19% were classified as Other (e.g., White, Asian, Native American, and "Other"). Demographic indicators (e.g., % male students, % African American students, % students receiving free or reduced price lunch, and school attendance rate) were comparable between control and PA schools at both Wave 1 and Wave 8 (Lewis et al., 2012). There was attrition at the student level due to student mobility (Figure 1; Lewis et al., 2012), which is common in studies conducted in low-income, urban settings (Tobler & Komro, 2011). Specifically, 131 of the original 624 wave 1 students were still present at wave 8; 64 of the 316 *PA* students present at Wave 1 were present at all 8 waves of the study, and 67 of the 308 Control students present at Wave 1 were present at all 8 waves of the study. Table 1 compares student-level demographics at baseline (Wave 1) and endpoint (wave 8). Mobility was more likely among African-American students, as compared to White, Hispanic, and Asian Students (Lewis et al., 2012). Also, mobility was more likely among older students (Lewis et al., 2012). Still, mobility patterns did not differ between control and treatment schools (Lewis et al., 2012). In addition, we conducted attrition analyses to determine if attrition was associated with the outcome variables, and we found no significant differences (results not shown). Lastly, the average number of waves of data collected per student was 3.1. Nonetheless, all 14 schools stayed in the study and in their assigned condition.

Intervention

Positive Action (Flay & Allred, 2010) is a school-based SECD program that includes PreK-12 curricula, of which the K-8 portion was implemented by intervention schools in this study. This portion of the curriculum consists of 140 15–20 minute, age-appropriate, interactive lessons per grade, taught 4 days per week for grades K-6, and 70 20-minute interactive lessons per grade, taught 2 days per week for grades 7 and 8. Lessons are delivered by the classroom teacher and involve both group work and homework. The program curriculum includes six units: self-concept is addressed in Unit 1, health behaviors are directly addressed in half of Unit 2, and Units 3 through 6 focus on social and emotional positive actions.

Measures

Outcomes—We collected student self-reports of health behaviors at each wave using a set of 10 items considered appropriate for grade 3 students. These items drew from the Youth Risk Behavior Surveillance System, and assessed how often (1 = "none of the time" to 4 = "all of the time") students engaged in different behaviors. The correlation analyses we conducted suggested the ten individual items grouped as three multi-item measures focused on healthy eating and exercise, unhealthy eating, and personal hygiene, and one single-item measure pertaining to sleep.

We used Stata's "sem" command to conduct confirmatory factor analysis on the 9 items tapping the 3 types of health behavior. Using data from wave 1 (baseline), the comparative fit index (CFI) and root mean square of approximation (RMSEA) for the model were 0.93 and 0.05, respectively; at wave 8 (endpoint), the CFI and RMSEA for the model were 0.99

and 0.02, respectively. For both waves, items loaded significantly onto each specified factor. Using these suggested groupings, we computed average scores for each type of behavior, with higher scores reflecting engagement in more of that type of behavior.

We calculated student-level stability reliabilities using two-week test-retest data available at Wave 5 for approximately 60 students. Test-retest correlations were 0.68 for healthy eating & exercise, 0.74 for unhealthy eating, 0.75 for consistent bedtime, and 0.76 for personal hygiene, reflecting the moderate to strong stability of the measures. As such, we proceeded with using the three multi-item measures and one single-item measure to assess program effects.

Healthy eating & exercise: Students indicated how much of the time they "eat fresh fruits and vegetables," "drink or eat dairy products," and "exercise hard enough to…sweat and breathe hard." At wave 8, standardized factor loadings for the three items ranged from 0.47 to 0.55.

Unhealthy eating: Students indicated how much of the time they "eat junk food (chips and candy)," "eat fast food," and "drink soda pop." At wave 8, standardized factor loadings for the three items ranged from 0.71 to 0.90.

Personal hygiene: We examined hygienic habits using the following items: "I wash my hands after using the toilet," "I brush my teeth at least twice a day," and "I cover my nose and mouth when I sneeze." At wave 8, standardized factor loadings for the three items ranged from 0.46 to 0.57.

Consistent bedtime: We assessed sleep-related behavior using a single item in which students rated how much of the time they "go to bed by 9:00 PM on school nights." Due to the bimodal distribution of responses, we categorized this measure as dichotomous, with $0 =$ Not all of the time and $1 =$ All of the time; we chose this classification based on research suggesting school-age children need 10–11 hours of sleep each night (National Sleep Foundation, 2012).

Body-mass-index z-scores: Height and weight measurements were taken by trained data collectors at Wave 8 (grade 8). Students were asked to remove hats, coats, and shoes and were assessed one at a time in a private room using the Cardinal® Detecto (model number 6129) scale with a vertical ruler to simultaneously record height and weight. We used height, weight, gender, and age data to calculate body mass index (BMI) scores, and we converted these scores into BMI z-scores (zBMI) using United States reference data collected by the Centers for Disease Control and Prevention in 2000. This method has been used in past studies (Economos et al., 2013).

Mediator

We measured the hypothesized mediator, *SECD*, using the 28-item Social-Emotional and Character Development Scale (Lewis et al., 2012; Ji et al., 2013). Example items, which encompass 6 subscales (i.e. prosocial interactions, honesty, self-development, self-control, and respect for teachers and parents), were: "I play nicely with others," "I apologize when I

have done something wrong," "I make myself a better person," "I keep my temper when I have an argument with other kids," "I speak politely to my parents," and "I speak politely to my teacher." Students indicated how often they engaged in each behavior (1 = None of the time to 4 = All of the time). We created an average of scores on the 6 subscales, with higher scores indicating higher SECD skills. Cronbach's alpha for this composite scale ranged from .92 to .95 across the 8 waves.

Data Analyses

In our *preliminary analyses,* we used Stata v13 to assess the distributions of each outcome, calculate intraclass correlations, and determine correlations between the measures at Waves 1 and 8, respectively.

We investigated program effects on the health behavior measures by estimating multilevel growth-curve models using Stata v13. We estimated models based on their distribution; healthy eating & exercise, unhealthy eating, and personal hygiene were negatively skewed, consistent bedtime was bimodal, and zBMI was normally distributed. We handled missing data using maximum likelihood estimation (Rabe-Hesketh & Skrondal 2008). We used endpoint regression analyses to assess program effects on the zBMI measure. We used two-tailed p-values for the condition by time interaction (for longitudinal data) and condition (for endpoint data) parameter estimates to assess statistical significance of program effects, with $.10 > p > .05$ considered marginal and $p \leq .05$ considered significant.

For the longitudinal models, we first estimated a random-intercept model including quadratic terms for time and the interaction of condition by time (c × t). Nonsignificant higher order terms were dropped from the model for parsimony. When applicable (i.e., for the consistent bedtime outcome), we also estimated a random-coefficient model and we performed a likelihood-ratio test to determine which model (random-intercept or random-coefficient) was more appropriate (Rabe-Hesketh & Skrondal, 2008). We calculated effect sizes for most outcomes using the method described by Lipsey and Wilson (2001). We used the Cox transformation to calculate an effect size for consistent bedtime (a binary outcome; Sánchez-Meca et al., 2003), and we calculated Hedges g effect size for endpoint zBMI.

We used two approaches to assess the robustness of results related to program effects (i.e., sensitivity analyses). First, we estimated a model including pairs as the cluster variable and compared results for the c × t interaction to the main analyses with school as the cluster variable. In addition, we compared the test statistic for the c × t interaction in the multilevel model to 2.18, the $p \leq .05$ critical value for a two-tailed t distribution with 12 degrees of freedom (N=14 schools; Raudenbush & Bryk, 2002).

We examined the roles of gender and student mobility in moderating program effects in supplemental analyses; the confounding nature between race/ethnicity and school in this trial precluded our testing for moderating effects by race/ethnicity. Given the mobility that occurred within this study, it is plausible that students who received the most exposure to the *PA* program (i.e., students who remained in a treatment school throughout the study's duration) would have more favorable change across time. To characterize student mobility patterns, we used latent class analysis to group underlying patterns of mobility. In prior

analyses (Lewis et al., conference presentation), a 5-class solution was found to be the most appropriate fit for the data. In these moderation analyses, students with the greatest time in the study (i.e., "stayers") served as the reference group.

Upon completion of these analyses, we estimated longitudinal mediation models for outcomes with marginal or significant program effects. Prior to running the mediation models, we tested for the presence of a significant quadratic trajectory for measures collected across time (i.e., all measures excepting zBMI). Because our mediator (SECD) included a quadratic, the effect of the linear slope on the change in each outcome is interpretable at the intercept (Gordon, 2012); thus, our time scales were adjusted to run from −4.00 years to 1.58 years (as opposed to the real time of 0.00 to 5.58 years), with the intercept being set at wave 6, the first data collection time after the important transition to middle-school grades.

To test for mediated effects we used M$plus$ v6.12. We included the mediator (i.e., change in SECD) in the model to calculate direct ($c\acute{}$) and indirect ($a \times b$) effects (Figure 2). Given our adjusted time scales, the effect of the mediator on each outcome is based on the point in time when students had started grade 7 (wave 6). Indirect effects (i.e., the effect of PA on the slope of each outcome as mediated by the slope of SECD) were computed as described by MacKinnon (2008). The small number of clusters (i.e., 14 schools; Hox & Maas 2001), low intra-class correlations across schools at baseline (discussed below), and the non-normality of the outcome variables, all in combination with the technical complexities of mediation testing in a multilevel modeling framework, precluded a multi-level structural equation modeling analysis (Hox & Maas, 2001; Singer & Willett, 2003; Zhang, Zyphur, & Preacher, 2009). Instead, we used bootstrap estimation with 1,000 replications to address issues of non-normality (William & MacKinnon, 2008), and used maximum likelihood to handle missing data. Mediation was classified as complete, partial, or non-significant (Baron & Kenny, 1986; MacKinnon et al., 2002).

Results

The baseline intraclass correlations (ICCs) across schools for healthy eating & exercise, personal hygiene, and consistent bedtime (prior to categorizing the measure as binary) were low (0.02, 0.01, and 0.01, respectively). In contrast, the ICC for unhealthy eating was moderately high at baseline (0.14). The ICC across schools for endpoint zBMI was also low (0.01).

Table 2 shows the correlations among the health behavior outcome measures and SECD mediator at Wave 1 and Wave 8. At Wave 1, the healthy eating & exercise measure was significantly and positively correlated with personal hygiene and consistent bedtime, and personal hygiene was significantly and positively correlated with consistent bedtime. At Wave 8, the healthy eating & exercise measure was again positively associated with personal hygiene and consistent bedtime. Scores on zBMI were inversely associated with healthy eating & exercise, unhealthy eating, personal hygiene, and consistent bedtime, although only the association with unhealthy eating, which was in an unexpected direction, was significant. With respect to the mediator, at both points in time SECD had a significant (p<0.01) and

positive association with measures of healthy eating, personal hygiene, and consistent bedtime.

Findings from our analyses testing for program effects and associated effect sizes are presented in Tables 3 and 4, respectively. We observed a significant program effect for personal hygiene (c×t b=0.05(0.01, 0.08), p=0.02, ES = 0.48), indicating that *PA* students had a more favorable trajectory for this behavior than students in control schools; this finding was maintained in both sensitivity analyses. The estimated program effect for healthy eating & exercise was marginally significant (c×t b-coefficient=0.03 (−0.00, 0.06), p=0.09, ES = 0.21), suggesting students in *PA* schools had more favorable growth in these health-promoting behaviors. Results for this measure were not significant in sensitivity analyses. For unhealthy eating, the program effect was not significant in the primary analyses (b=−0.03(−0.06, 0.01), p=0.12, ES = −0.19), but was marginally so in the sensitivity analysis that included pair as the cluster variable (b=−0.03(−0.07, 0.00), p=0.08), suggesting that growth in unhealthy food intake was attenuated by participation in the *PA* program. The program effect was not significant for consistent sleep in any analyses, although the effect size (Cox-d = 0.35) was moderate in magnitude. Considering that a consistent 9:00 PM bedtime may be appropriate only for younger students, we conducted supplemental analyses using wave 5 as the endpoint; these analyses also showed a non-significant but moderate effect size (Cox-d = 0.33). Lastly, our endpoint analysis on the *z*BMI outcome indicated a marginally significant program effect in the hypothesized direction (b=−0.22(−0.46, 0.02), p=0.07, ES =−0.21). For all outcomes, we did not observe moderation by gender or mobility group.

We estimated longitudinal mediation models (Table 5) for all outcomes except consistent bedtime (as it did not have a marginal or significant program effect). In all models, the "a" path (path from program condition to change in SECD) was significant at the 0.01 level. The path from SECD to each outcome was significant at the 0.01 level for healthy eating and hygiene, and at the 0.05 level for unhealthy eating. The calculated indirect effect (a×b) was significant at the 0.01 level for healthy eating and hygiene, and at the 0.05 level for unhealthy eating. As the c' path remained significant in the model for healthy eating, the mediation observed was classified as partial. The c' path was reduced to non-significance in the hygiene and unhealthy eating models. As such, the mediation observed for these two models was classified as complete. Mediation was not observed for the *z*BMI outcome.

Discussion

Our study is the first to use a randomized design to examine the longitudinal impact of a social-emotional program, *Positive Action*, on youths' health behaviors. We observed significant program effects across time for personal hygiene, and marginal condition × time interactions for healthy eating & exercise and unhealthy eating. We also observed a marginal program effect on endpoint BMI *z*-scores. Although the magnitude of effects was modest, they demonstrate the potential of a social-emotional program to impact health behaviors, even though a direct focus on physical health was a relatively small component (approximately 8% of lessons) of the *PA* intervention. Specifically, while only half of Unit 2 covered health behaviors, we observed relative improvements in health behaviors among

students receiving *PA*. Results are particularly encouraging because the trial's setting (i.e., low-income communities) is one known to be impacted by health disparities (Orsi et al., 2010).

We also examined one proposed mechanism of program action. Our cross-sectional correlation analyses showed an association between SECD and health-promoting behaviors. Moreover, our longitudinal mediation analyses showed evidence of complete or partial mediation through changes in SECD for three of our outcomes. This finding supports the literature (Flay, 2009) we reviewed that suggests ultimate- and distal-level predictors' impact various types of behaviors.

Our results also provide direction for future research. Unhealthy eating was the only measure which had an intraclass correlation above 0.02, suggesting that there is variation between schools in this behavior; it is possible that the extent of unhealthy snack availability (e.g., vending machines that dispense sugar-sweetened beverages) may vary by school. Future studies may consider accounting for the obesogenic nature of the school environment. Interestingly, we observed a significant inverse cross-sectional association between zBMI and unhealthy eating, whereby greater unhealthy food consumption was correlated with a lower zBMI. This finding contradicts results from a systematic review of 30 studies showing overweight and obesity to be associated with greater consumption of sugar-sweetened beverages (Malik et al., 2006). Our post-hoc analyses showed the observed inverse correlation at Wave 8 was significant only for girls ($r = -0.25$ for girls; $r = -0.15$ for boys). Given the cross-sectional nature of the relationship, further examination of this observed correlation is warranted.

Limitations

Student self-reports are subject to social desirability bias; this bias, however, should affect students equally in both treatment and control schools, yet program effects were observed nonetheless. Additionally, the sole researcher-collected measure (i.e., zBMI) was available only at endpoint; conclusions related to the program's influence on weight across time, therefore, cannot be drawn. Also, the measure of consistent bedtime did not account for student wake time, limiting the ability to measure sleep duration. Future evaluations of similar programs should, therefore, collect height and weight data across time, as well as include more items related to sleep behaviors.

Because students completed a survey that assessed a broad range of outcomes, the number of items specific to health was limited. Future research should include a more in-depth assessment of health behaviors and related outcomes, using standardized instruments with well-established psychometric properties. Future research could also consider modifying the content of certain questions related to health behaviors across time points in order to be sensitive to developmental differences (e.g., sleep-related items should be adjusted as youth transition into adolescence, as we witnessed marked declines across time points in the percentage of youth who always sleep by 9:00 PM). Even so, we found the items had moderate to strong stability reliabilities.

Another limitation relates to our lack of parent or teacher reports of student health behavior to substantiate our findings. Additionally, data on school-level barriers to healthy decision making (e.g., fast food outlet density) that may confound results were not available. Future research should substantiate student reports with parent- and teacher-reports of health behaviors. Archival data should also be used in future studies to control for the availability of healthy and unhealthy food options. Funding constraints limited our ability to rigorously examine the role of fidelity of implementation on program effects. Future evaluations of *PA* and *PA*-like programs should, therefore, aim to prioritize measurement of implementation.

With respect to study design, the small number of schools (*N*=14) could limit statistical power. However, that some significant findings were found in spite of this limitation suggests that our findings are robust. Also, findings are generalizable only to similar schools (i.e., low-income, urban schools) that would self-select to participate in a trial of this nature. Nonetheless, we were able to focus on a high-risk population affected by disparities in health.

Lastly, the multi-faceted nature of the intervention under study also presented analytical limitations. Ideally, we would have liked to conduct a dose-response analysis, but using a measure of years of classroom exposure to assess dosage appeared problematic for a number of reasons. First, the classroom was only one component of the *PA* intervention. Specifically, the program includes teacher and counselor training in addition to the classroom student curricula. Second, fidelity of implementation within the classroom improved over time (Li et al., 2011) and it is also likely the school context was influenced by the presence of the intervention, thus altering the context for newly arriving students. In other words, fewer years of exposure due to late arrival was confounded with factors (e.g., co-occurring changes in fidelity and school context) in ways that would have prevented us from confidently attributing any differences in outcomes relating to exposure to a dose effect of the intervention as opposed to these other influences. In view of such concerns, we instead examined differences in intervention effects in association with student mobility classes that take into account both timing of entry (or departure) to the school and number of years of exposure. Although it could be useful in future research to attempt to disaggregate these types of influences, attempting to do so reliably was beyond the scope of the present study.

Conclusions

Strengths of this study included its longitudinal nature, experimental design, successful matching, lack of attrition at the unit of randomization (i.e., the school), and use of sophisticated techniques to handle the hierarchical nature of the data, student mobility and mediation. As such, we were able to use a methodologically strong study to examine the impact of a social-emotional intervention on physical health behaviors in a setting where the need to impact both social-emotional skills and physical health is present. We observed small-to-modest effects on health behaviors, which is understandable yet encouraging as only a small percentage of the *PA* curriculum focused on health behaviors. Results from the mediation analyses suggest that the program's focus on ultimate causes of behavior representing core skills (e.g., SECD) was beneficial in addressing health behaviors. Moreover, the findings reported here parallel, albeit to a lesser extent, the trend of positive

impact observed in the Chicago RCT of *PA* on emotional (Lewis et al., 2013a) and behavioral (Lewis et al., 2012; Lewis et al., 2013b) outcomes. Given the interrelatedness of these outcomes, findings suggest that social-emotional and character development programs should be considered by schools seeking to prevent a wide range of detrimental adolescent behaviors. Moreover, the results indicate that schools aiming to address a specific health behavior could see a greater impact by implementing a program that incorporates the development of social-emotional skills. Promoting these skills during childhood and adolescence can have longer-term prevention implications, as health-promoting behaviors can be developed during youth and maintained into young adulthood. Although the intensive nature of *PA* and similar programs may be intimidating, buy-in from key administrative staff, funding for incentives and the ability to readily provide technical support will help schools overcome the challenges of program implementation (Lewis et al., in preparation). The findings of this study suggest that SECD programs may be a useful supplement to other physical health instruction and activities already being implemented in the school setting.

Acknowledgments

References

Aud, S.; Fox, M.; KewalRamani, A. Status and Trends in the Education of Racial and Ethnic Groups (NCES 2010-015). U.S. Department of Education, National Center for Education Statistics; Washington, DC: U.S. Government Printing Office; 2010.

Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology. 1986; 51:1173–1182. [PubMed: 3806354]

Basch CE. Breakfast and the achievement gap among urban minority youth. Journal of School Health. 2011; 81(10):635–640. [PubMed: 21923876]

Bavarian N, Lewis KM, DuBois DL, Acock A, Vuchinich S, Silverthorn N, et al. Using social-emotional and character development to improve academic outcomes: A matched-pair, cluster-randomized trial in low-income, urban schools. Journal of School Health. 2013; 83(11):771–779. [PubMed: 24138347]

Beets MW, Flay BR, Vuchinich S, Snyder FJ, Acock A, Li KK, 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. American Journal of Public Health. 2009; 99(8):1438–1445. [PubMed: 19542037]

Bickman L, Riemer M, Brown JL, Jones SM, Flay BR, Li KK, Massetti G. Approaches to measuring implementation fidelity in school-based program evaluations. Journal of Research in Character Education. 2009; 7(2):75–101.

Brown CH, Wang W, Kellam SG, Muthen BO, Petras H, Toyinbo P, et al. Methods for testing theory and evaluating impact in randomized field trials: Intent-to-treat analyses for integrating the perspectives of person, place, and time. Drug and Alcohol Dependence. 2008; 95(S1):S74–S104. [PubMed: 18215473]

Collaborative for Academic, Social, and Emotional Learning. Safe and sound: An educational leader's guide to evidence-based social and emotional learning programs—Illinois edition. 2005.

Dewald JF, Meijer AM, Oort FJ, Kerkhof GA, Bogels SM. The influence of sleep quality, sleep duration and sleepiness on school performance in children and adolescents: A meta-analytic review. Sleep Medicine Reviews. 2010; 14(3):179–189. [PubMed: 20093054]

DuBois, DL.; Flay, BR.; Fagen, MC. Self-esteem Enhancement Theory: An emerging framework for promoting health across the life-span. In: DiClemente, RJ.; Kegler, MC.; Cosby, RA., editors. Theories in health promotion practice and research. 2. San Francisco: Jossey-Bass; 2009. p. 93-130.

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 Development. 2011; 82:405–432. [PubMed: 21291449]

Economos CD, Hyatt RR, Must A, Goldberg JP, Kuder J, Naumova EN, et al. Shape Up Somerville two-year results: A community-based environmental change intervention sustains weight reduction in children. Preventive Medicine. 2013; 57(4):322–327. [PubMed: 23756187]

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.; Tommey, R.; Clement, N., editors. International Research Handbook on Values Education and Student Wellbeing. Dordrecht: Springer; 2010. p. 471-501.

Flay, BR.; Snyder, F.; Petraitis, J. The theory of triadic influence. In: DiClemente, RJ.; Kegler, MC.; Crosby, RA., editors. Emerging theories in health promotion practice and research. 2. San Francisco CA: Jossey-Bass; 2009. p. 451-510.

Gordon, RAA. Applied Statistics For The Social And Health Sciences. New York, NY: Routledge; 2012.

Hox JJ, Maas CJM. The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples. Structural Equation Modeling. 2001; 8(2):157–174.

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. Journal of School Health. 2008; 78(3):131–139. [PubMed: 18307608]

Ji P, DuBois DL, Flay B. Social-Emotional and Character Development Scale: Development and initial validation with urban elementary school students. Journal of Research on Character Education. 2013; 9(2):121–147.

Lewis KM, Bavarian N, Snyder FJ, Acock A, Day J, DuBois DL, et al. Direct and mediated effects of a social-emotional and character development program on adolescent substance use. The International Journal of Emotional Education. 2012; 4(1):56–78. [PubMed: 24308013]

Lewis KM, DuBois DL, Bavarian N, Acock A, Silverthorn N, Day J, et al. Effects of *Positive Action* on the emotional health of urban youth: A cluster-randomized trial. Journal of Adolescent Health. 2013a; 53(6):706–711. [PubMed: 23890774]

Lewis KM, Schure MB, Bavarian N, DuBois DL, Day J, Ji P, et al. Problem behavior and urban, low-income youth: A randomized controlled trial of Positive Action in Chicago. American Journal of Preventive Medicine. 2013b; 44(6):622–630. [PubMed: 23683980]

Li KK, Washburn I, DuBois DL, Vuchinich S, Ji P, Brechling V, et al. Effects of the Positive Action programme on problem behaviours in elementary school students: A matched-pair randomised control trial in Chicago. Psychology & Health. 2011; 26(2):187–204. [PubMed: 21318929]

Lipsey, MW.; Wilson, DB. Practical meta-analysis. Thousand Oaks: Sage; 2001.

London RA, Castrechini S. A longitudinal examination of the link between youth physical fitness and academic achievement. Journal of School Health. 2011; 81(7):400–408. [PubMed: 21668880]

MacKinnon, DP. Introduction to statistical mediation analysis. New York, NY: Lawrence Erlbaum Associates; 2008.

MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. Psychological Methods. 2002; 7:83–104. [PubMed: 11928892]

Malik VS, Shultze MB, Hu FB. Intake of sugar-sweetened beverages and weight gain: A systematic review. The American Journal of Clinical Nutrition. 2006; 84(2):274–288. [PubMed: 16895873]

National Sleep Foundation. Resource document. National Sleep Foundation; 2013. Children and sleep. http://www.sleepfoundation.org/article/sleep-topics/children-and-sleep

Orsi JM, Margellos-Anast H, Whitman S. Black–white health disparities in the United States and Chicago: A 15-year progress analysis. American Journal of Public Health. 2010; 100(2):349–356. [PubMed: 20019299]

Rabe-Hesketh, S.; Skrondal, A. Multilevel and longitudinal modeling using Stata. 2. College Station: Stata Press; 2008.

Raudenbush, SW.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods, Col 1. Thousand Oaks: Sage; 2002.

Sánchez-Meca J, Marín-Martínez F, Chacón-Moscoso S. Effect-size indices for dichotomized outcomes in meta-analysis. Psychological Methods. 2003; 8(4):448–467. [PubMed: 14664682]

Singer, JD.; Willett, JB. Applied Longitudinal Data Analysis: Modeling Change And Event Occurrence. New York: Oxford University Press; 2003.

Snyder F, Vuchinich S, Acock A, Washburn I, Beets M, Li KK, 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. Journal of Research on Educational Effectiveness. 2010; 3(1):26–55. [PubMed: 20414477]

Social and Character Development Research Consortium. Efficacy of Schoolwide Programs to Promote Social and Character Development and Reduce Problem Behavior in Elementary School Children. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education; 2010. NCER 2011-2001

Tobler AL, Komro KA. Contemporary options for longitudinal follow-up: Lessons learned from a cohort of urban adolescents. Evaluation and Program Planning. 2011; 34(2):87–96. [PubMed: 21338843]

Vuchinich S, Flay BR, Aber L, Bickman L. Person mobility in the design and analysis of cluster-randomized cohort prevention trials. Prevention Science. 2012; 13(3):300–313. [PubMed: 22249907]

Washburn IJ, Acock A, Vuchinich S, Snyder F, Li KK, Ji P, 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. Prevention Science. 2011; 12(3):314–323. [PubMed: 21720782]

Williams J, MacKinnon DP. Resampling and distribution of the product methods for testing indirect effects in complex models. Structural Equation Modeling. 2008; 15(1):23–51. [PubMed: 20179778]

Zhang Z, Zyphur MJ, Preacher KJ. Testing multilevel mediation using hierarchical linear models. Organizational Research Methods. 2009; 12(4):695–719.

Control Participants   Positive Action Participants

308	Wave 1	316
Joiners: 27 Leavers: 36		Joiners: 39 Leavers: 49
299	Wave 2	306
Joiners: 0 Leavers: 79		Joiners: 0 Leavers: 62
220	Wave 3	244
Joiners: 55 Leavers: 31		Joiners: 77 Leavers: 24
244	Wave 4	297
Joiners: 60 Leavers: 51		Joiners: 47 Leavers: 82
253	Wave 5	262
Joiners: 0 Leavers: 161		Joiners: 0 Leavers: 158
92	Wave 6	104
Joiners: 81 Leavers: 1		Joiners: 83 Leavers: 0
172	Wave 7	187
Joiners: 27 Leavers: 31		Joiners: 50 Leavers: 42
168	Wave 8	195

Figure 1.
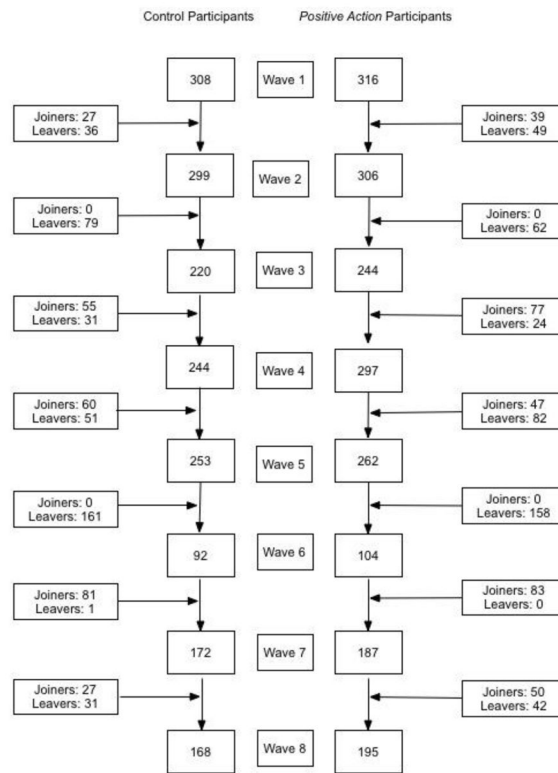
Flow diagram of students.

Reprinted, with permission, from the following publication:

Lewis, K.M., Schure, M.B., Bavarian, N., DuBois, D.L., Day, J., Ji, P., Silverthorn, N., Acock, A., Vuchinich, S., & Flay, B. R. (2013). Problem behavior and urban, low-income youth: A randomized controlled trial of Positive Action in Chicago. *American Journal of Preventive Medicine, 44*(6), 622–630.
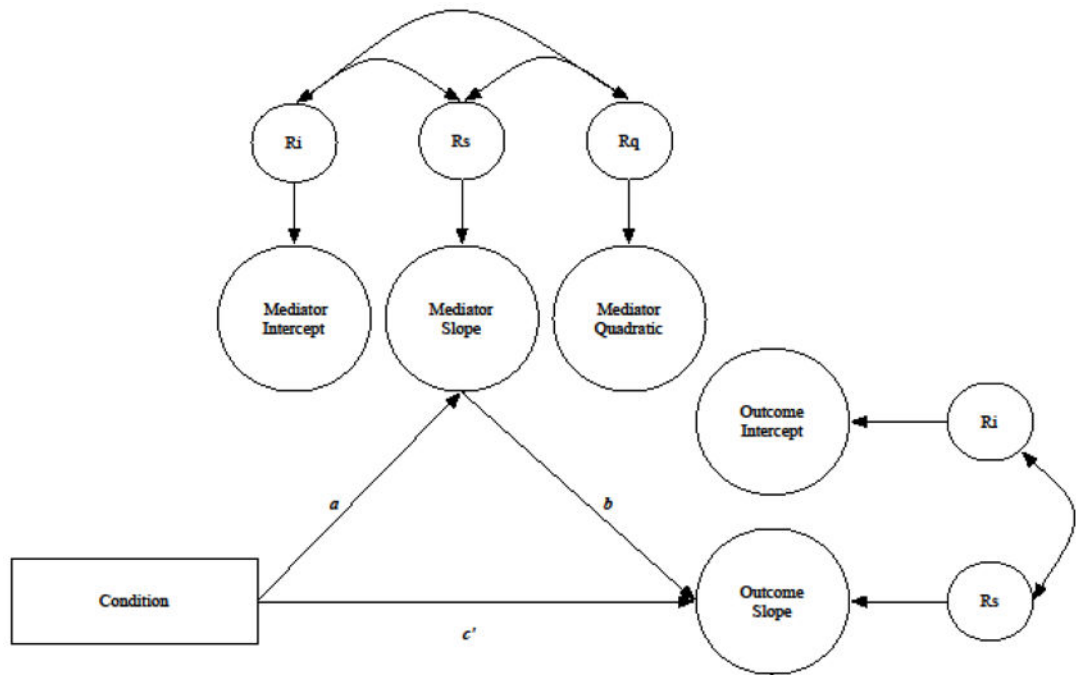
Figure 2.
Longitudinal Mediation Models

Note. In the mediation model, paths were also included from *Positive Action* (Condition) to the intercept of the mediator and the intercept of the outcome. The variance of the quadratic for the mediator (i.e., social-emotional and character development) was set at zero, and the intercepts of the errors of the mediator and the outcome were correlated. In addition, the time measure was set from –4.00 to 1.58, with the intercept (time 0) set at wave 6 (i.e., the start of grade 7). *Notation Note:* "Condition" represents the *Positive Action* program; Ri, Rs, and Rq represent the variances in intercept, slope, and quadratic growth factors, respectively; *a*, *b*, and *c*′ represent the path coefficients where the *a* path represents the direct effect of the independent variable on the mediator, the *b* path represents the direct effect of the mediator on the dependent variable, and the *c*′ path represents the direct effect of the independent variable on the dependent variable with the mediator included.

Author Manuscript   Author Manuscript   Author Manuscript   Author Manuscript

Table 1

Attrition analysis on student demographics

	2004			2010						
	Control Schools	PA Intervention Schools	test statistic	Control Schools			PA Intervention Schools			test statistic
$N=$			p-value	S	L	J	S	L	J	p-value
	299	295		63	237	108	59	235	214	
% Male Students	44.59	47.06	$z=-0.60$ $p=0.55$	39.56	47.12	48.72	47.25	49.36	48.88	$z=.073, p=0.47$
% of White Students	11.37	12.92	$z=0.83$ $p=0.41$	6.45	11.94	7.29	5.26	10.77	6.63	$z=0.91, p=0.36$
% of Black Students	57.52	51.25	NA	45.16	59.70	46.88	42.11	51.79	40.33	NA
% of Hispanic Students	26.27	33.75	$z=1.69, p=0.09$	46.77	22.89	44.79	49.12	35.38	48.07	$z=-4.13, p<.001$
% of Asian Students	4.71	2.08	$z=-1.30, p=0.20$	1.61	5.47	1.04	3.51	2.05	4.97	$z=-0.45, p=0.65$
Age:										
Mean	8.3	8.3	$F=0.03, p=0.60$	13.98	14.34	13.92	13.93	14.3	14.02	$F=0.01$ $p=0.93$
SD	0.55	0.58		0.66	0.55	0.8	0.52	0.6	0.8	

Note: *PA* = Positive Action program; S=stayers; L= leavers; J=joiners; NA= Not Applicable; for ethnicity, "Black" was used as the reference group.

Reprinted with permission from:

Lewis, K.M., Bavarian, N., Snyder, F.J., Acock, A., Day, J., DuBois, D.L., et al. (2012) Direct and mediated effects of a social-emotional and character development program on adolescent substance use. *The International Journal of Emotional Education*, 4(1), 56–78.

Table 2

Correlations Between Health Measures at A) Wave 1 (*N*=593) and B) Wave 8 (*N*=331)

A)

Wave 1 Measures	Healthy Eating & Exercise	Unhealthy Eating	Personal Hygiene	Consistent Bedtime
Healthy Eating & Exercise	---	.		
Unhealthy Eating	0.06	---		
Personal Hygiene	0.25**	−0.03	---	
Consistent Bedtime[a]	0.26**	−0.08	0.20**	---
Social-Emotional Character Development	0.39**	−0.05	0.37**	0.25**

B)

Wave 8 Measures	Healthy Eating & Exercise	Unhealthy Eating	Personal Hygiene	Consistent Bedtime	Body Mass Index z-score
Healthy Eating & Exercise	---				
Unhealthy Eating	0.06	---			
Personal Hygiene	0.34**	0.07	---		
Consistent Bedtime[a]	0.13*	−0.04	0.01	---	
Body Mass Index z-score	−0.05	−0.19**	−0.04	−0.05	---
Social-Emotional Character Development	0.46**	0.01	0.37**	0.18**	−0.06

[a] Measure was re-categorized into a binary variable in the primary analyses

*
p .05.

**
p .01.

Table 3

Multilevel growth-curve model estimates and confidence intervals for student-level measures (N=1,170 students in 14 schools)

Measure	Model Run	Intercept	Time	Time^2	Condition (0 = Non-*PA* 1=*PA*)	Condition × Time
		B (95% CI)	B (95% CI)	B (95% CI)	B (95% CI)	B (95% CI)
Healthy Eating & Exercise	Random Intercept	3.39 (3.31, 3.46)**	−0.13 (−0.19, − 0.08)**	0.01 (0.00, 0.02)*	0.00 (−0.09, 0.10)	0.03 (−0.00, 0.06)+
Unhealthy Eating	Random Intercept	3.174 (2.90, 3.45)**	−0.01 (−0.03, 0.02)	---	−0.03 (−0.41, 0.35)	−0.03 (−0.06, 0.01)[b]
Personal Hygiene	Random Intercept	4.25 (4.11, 4.38)**	−0.09 (−0.11, − 0.06)**	---	−0.10 (−0.29, 0.09)	0.045 (0.01, 0.08)*
Consistent Bedtime[a]	Random Intercept	−0.13 (−0.41, 0.15)	−0.61 (−0.711, − 0.50)**	---	−0.09 (−0.48, 0.30)	0.09 (−0.05, 0.22)
Body Mass Index z-score	Endpoint	1.17 (1.01, 1.32)**			−0.22 (−0.46, 0.02)+	

[a] Logistic Coefficients are provided

[b] In the sensitivity analyses using "pairs" as the cluster variable, the parameter estimate was marginal.

*
p .05.

**
p .01.

PA = Positive Action program

Table 4

Estimated Means and Effect Sizes

| Measure | Response Options | Wave 1 | | Wave 8 | | |
		Control	PA	Control	PA	Effect Size [a]
Healthy Eating & Exercise	1 to 4	3.39	3.39	3.01	3.16	0.21
Unhealthy Eating	1 to 4	3.17	3.14	3.13	2.94	-0.19
Personal Hygiene	1 to 4	4.25	4.14	3.75	3.90	0.48
Consistent Bedtime	0 to 1	0.48	0.47	0.04	0.06	0.35
Body Mass Index z-score	z-scale	--	--	1.17	0.96	-0.21

[a] For continuous measures with data available at all time points, 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 the baseline to calculate effect size. For continuous measures with endpoint data only, Hedges *g* effect size was calculated. For binary measures, the Cox transformation was used to calculate effect size. All calculations were made using real time.

PA = Positive Action program

Table 5

Summary of Unstandardized Mediation Results, Performed Using M*Plus* v6.12 (*N*=1,170 students)[a]

Measure	Mediated Effect			Indirect Effect[b]	Type of Mediation Observed[c]
	c′ path B (SE)	*a* path B (SE)	*b* path B (SE)		
Healthy Habits (Food and Exercise)	−.02(.00)[**]	.04(.01)[**]	1.02(.03)[**]	.05(.00)[**]	Partial
Unhealthy Food	−.01(.02)	.04(.01)[**]	−.40(.16)[*]	−.02(.01)[*]	Full
Personal Hygiene	.00(.03)	.05(.01)[**]	.39(.09)[**]	.02(.01)[**]	Full
Body Mass Index *z*-score	−.20(0.13)	.05(.02)[**]	−.20(.82)	−.01(.04)	None

[a]Analyses were performed only for outcomes with significant or marginal program effects. Models were run using bootstrap estimation with 1,000 replications. Time was set from −4.00 to 1.58. For all analyses except body mass index *z*-score we used 8 waves of mediator, SECD (social-emotional character development), and 8 waves of outcome.

[b]The indirect effect refers to the path from condition to the path from the outcome through the slope of the outcome through the slope of social-emotional and character development. This mediator includes a significant quadratic trajectory with time 0 set at wave 6.

[c]The type of mediation observed was classified using Barron and Kenny (1986) and MacKinnon et al. (2002) as guides.

[*]
p .05.

[**]
p .01.