Prev Sci (2011) 12:314–323
DOI 10.1007/s11121-011-0230-9

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

Isaac J. Washburn · Alan Acock · Sam Vuchinich · Frank Snyder · Kin-Kit Li · Peter Ji · Joseph Day · David DuBois · Brian R. Flay

Published online: 1 July 2011

Abstract The effects of a school-based social-emotional and character development program, *Positive Action*, on the developmental trajectory of social-emotional and character-related behaviors was evaluated using data from three school-based randomized trials in elementary schools. Results come from 1) 4 years of data from students in 20 Hawai'i schools, 2) 3 years of data from students in 14 schools in Chicago and 3) 3 years of data from students in 8 schools in a southeastern state. Random intercept, multilevel, growth-curve analyses showed that students in both control and *Positive Action* schools exhibited a general decline in the number of positive behaviors associated with social-emotional and character development that were endorsed. However, the *Positive Action* intervention significantly reduced these declines in all three trials. Taken together, these analyses 1) give insight into the normative trajectory of behaviors associated with social-emotional and character development and 2) provide evidence for the effectiveness of *Positive Action* in helping children maintain a relatively beneficial developmental trajectory.

Keywords School-based intervention · Social-emotional and character development · Randomized trial · Trajectory of behaviors

I. J. Washburn (✉)
Oregon Social Learning Center,
10 Shelton McMurphey Blvd,
Eugene, OR 97401, USA
e-mail: Isaacw@oslc.org

A. Acock · S. Vuchinich · F. Snyder · B. R. Flay
Oregon State University,
Corvallis, OR, USA

K.-K. Li
City University of Hong Kong,
Kowloon, Hong Kong

P. Ji · J. Day · D. DuBois
University of Illinois at Chicago,
Chicago, IL, USA

Introduction

Catalano et al. (2004) described the movement over the last 30 years from single-domain intervention programs focused on individual problem behaviors (drugs, delinquency, sexuality) to multiple-domain prevention programs that focus on both problem behaviors and what Catalano et al. (2004) and Flay (2002) called positive youth development. This movement from single-domain intervention to multi-domain prevention is consistent with the latest theoretical thinking in developmental (Lerner et al. 2005) and health promotion/prevention (Flay 2002; Flay et al. 2009) literatures. In particular, there has been an increased 1) interest in social-emotional and character development (SECD; Elias 2009) programs that focus on a child's social, emotional, and character development; and 2) recognition that optimal development in these areas may provide the best protective factors against health-compromising and high-risk behaviors. SECD programs claim they can do more than just give a one-time boost to the number of positive behaviors associated with SECD development; they also claim that they can change the trajectories of SECD of children.

The study of SECD covers multiple disciplines and is of interest to both theorists and prevention scientists. Recent developments include efforts to define SECD within the framework of positive psychology (Parks 2004), encouragement of SECD in adolescents (Catalano et al. 2004; Lerner et al. 2005) and applications of the scientific method to better inform prevention programs that include promotion of character education (Berkowitz and Bier 2004; CASEL 2003). Reviews containing literature on SECD are found in several domains, including the academic (Eisenberg et al. 2006; Lerner et al. 2005), the applied (Benninga et al. 2006; CASEL 2003), and the governmental (Catalano et al. 2010; Ferber et al. 2005).

These advances have stimulated increased interest among prevention scientists in the developmental pathway of positive behaviors associated with SECD from childhood into adolescence. The current research examines the developmental trajectory of positive behaviors from age 6 to 11. Previous theory and research are considered. The study reports the results of three randomized trials that test whether the *Positive Action* intervention can beneficially effect the development of positive behaviors related to SECD.

Given the varied terminology across these different disciplines, the focus of this paper is on the behaviors that overlap in these different disciplines and not on defining any specific new construct. Eisenberg and Morris (2004) provided a review of what was known about prosocial behavior development (a component of SECD development) at that time. Their main point was that, based on the increasing acquisition of cognitive tools, children should be increasing in behaviors associated with moral reasoning and prosocial development into and through adolescence. In general, the studies they reviewed (mainly cross-sectional) showed this increase of positive social behavior into and through adolescence. However, as described below, four subsequent studies using longitudinal data reported contrary results.

Kokko et al. (2006) examined the trajectories of prosocial behavior among a sample of 1,025 boys as reported by teachers between the ages of 6 and 12. The authors were surprised that the boys were found to decline or, at best, stay stable in prosocial behaviors into adolescence. Nantel-Vivier et al. (2009) extended the age of observation with a cross-national comparison of 1,037 boys from 10 to 15, with mother and teacher reports of prosocial behaviors for a Canadian sample and teacher and student reports for an Italian sample. Although differences existed between samples and by who reported (students, mothers, or teacher), similar declines in trajectories were observed. In 2007, Carlo, Crockett, Randall, and Roesch found strikingly similar results from high school student self-reports of their prosocial behavior assessed between grades 7 and 12 with a sample of about 657 youth.

Although the three previous studies looked at behaviors specific to Eisenberg and Morris (2004), Lerner and colleagues (2005) discussed the more broad set of behaviors inherent in their idea of positive youth development, which is roughly analogous to the view found in Catalano and colleagues (2004) and the impetus behind SECD programs. Lerner and colleagues present five Cs (competence, confidence, connection, character, and caring) that make up their larger concept of positive youth development and which can be directly related to the aims of the *Positive Action* program (self-concept, physical and intellectual actions, social/emotional actions for managing oneself responsibly, getting along with others, being honest with yourself and others, and continuous self-improvement). Lerner, Lerner, Phelps, and Colleagues at Tufts University, Institute for Applied Research in Youth Development (2008) measured the five Cs in order to construct a single global measure of positive youth development. Using this global measure with a sample of 1,912 youth assessed between 5th and 8th grade, they found that the general developmental trend was negative; that is, youth endorsed fewer positive outcomes as they got older.

The results of the preceding studies (Carlo et al. 2007; Kokko et al. 2006; Nantel-Vivier et al. 2009; Tufts University, Institute for Applied Research in Youth Development 2008) all indicate a decline in positive behaviors associated with SECD from middle childhood through the end of adolescence. There has been little discussion of reasons for this decline in the literature. Phelps and colleagues (2009) discussed the decline but state that they plan on watching it for a few more years in order to understand it better. Others (Kokko et al. 2006; Nantel-Vivier et al. 2009) have suggested that more work is needed to understand the decline. Carlo et al. (2007) gave several possible reasons for the decline, all of which focused on the environment surrounding boys in school.

Positive Action (PA) is one widely implemented SECD program for school-age youth. Grounded in a broad theory of self-concept (Purkey and Novak 1984), *PA* was developed and revised by educational psychologist Carol Gerber Allred from 1977 to the present, with frequent additions and revisions over the years based on formative and monitoring data. Most recently, the application of the Theory of Triadic Influence (Flay et al. 2009) has moved *PA* in the direction of characterization as an SECD program (Flay and Allred 2010). The combination of the work by Purkey and Novak and Flay et al. posits the *PA* program's influence on SECD as being through the reinforcement of positive behaviors associated with SECD. Children begin a cycle of reinforcement in which positive thoughts lead to positive behaviors that generate positive feelings about self, which, in turn, lead to more positive thoughts and behaviors. More detailed descriptions of the *Positive Action*

program are available at the program website (www.positiveaction.net) and in Flay et al. (2001) and Flay and Allred (2010). Evidence of the effectiveness of the *PA* program for enhancing academic achievement and school involvement, while reducing disciplinary referrals, substance use, risky sexual behavior and violence has been reported previously in quasi-experimental and experimental evaluations of the program (Beets et al. 2009; Flay and Allred 2003; Flay et al. 2001; Li et al. 2011; Snyder et al. 2010). However, the effects of *PA* on SECD are as yet unreported.

The full *PA* program consists of K-12 classroom curricula, of which only the elementary curriculum was used in the present randomized trial; a school-wide 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 present evaluations of PA focus on the school based components and did not include the more intensive family kit or the community-development component of *PA*.

The sequenced elementary curriculum consists of 140 15–20 min lessons per grade, per academic year, provided by classroom teachers. When fully implemented, the total time students are exposed to the program during a 35-week academic year is approximately 35 h. Lessons cover six major units on topics related to self-concept, physical and intellectual actions, social/emotional actions for managing oneself responsibly, getting along with others, being honest with yourself and others, and continuous self-improvement. The classroom curricula, school-climate kit, and other components of the program each encourage and reinforce the six units of *PA*.

The goals of the present study are two-fold: 1) to provide new information that can help reconcile conflicting findings regarding normative changes in positive behaviors associated with SECD during middle childhood (ages 6 to 11 years), and 2) examine the effects of *PA* on behaviors associated with SECD. Based on the recent longitudinal studies reviewed above, we expected to find decreasing reports of positive behaviors in each of three randomized trials of the *PA* program. However, we also expected that children receiving the *PA* program would have significantly mitigated declines in positive behavior.

In the three different randomized trials, a global measure of positive behaviors associated with SECD was collected across varying lengths of times. These are the only school-level randomized trials of *PA* of which the authors are aware. Two of the trials, referred to herein as the Chicago Trial and Hawai'i Trial, had one of the co-authors of this report (Flay) as principal investigator, whereas the third, which we refer to as the Southeastern State Trial, was conducted solely under the auspices of the implementing school district. We present the method, results, and discussion for each trial separately, followed by a final conclusion. Random-intercept multilevel growth models were performed to establish the normative developmental trajectory of positive behavior in each of the three trials as well as to provide tests of the effectiveness of the *PA* program. Although a single model was hypothesized for all three sites, each site was run separately so that any differences in the linear model for each site might be explored without an overabundance of interactions.

Hawai'i Trial

Method

The Hawai'i Trial of the *PA* program took place in 20 public elementary schools on three islands in the unified Hawai'i school district that encompasses the entire state. Student self-reports of their behavior were collected at five time points, on each of two cohorts (first graders and second graders at the start of the project). Data were collected for baseline at the end of the academic school year in half of both the control and PA schools and at the beginning of the next school year in the others. The remaining four waves of data were collected at the next four springs. Data were collected by research, rather than school, staff. The teachers within each of the ten program schools received the *PA* training from the program developer. Brief update trainings were repeated at the start of each subsequent year in the program schools. The teachers in the ten control schools received no *PA* training and were asked to not implement the program.

The schools receiving the *PA* program were randomly assigned from matched pairs based on a multivariate index of factors related to academic risk (cf., Dent et al. 1993). This matching utilized the following school-level characteristics: proportion of students receiving free or reduced price lunches, percentage stability of student enrollment, achievement scores on standardized tests, ethnic distribution, student–teacher ratios, school size, and school-level problem behaviors such as suspensions (see Beets et al. 2009, for details). There were no significant differences between *PA* schools and control schools at baseline on any of the variables used to match schools (Beets et al. 2009).

Students were asked a series of questions about their behaviors and feelings associated with those behaviors. For this analysis, only the behavioral questions were considered. The purpose of these items was to collect information on the positive behaviors associated with SECD and defined by the *PA* program. These behaviors are consistent with Lerner and colleague's (Tufts University, Institute for

Applied Research in Youth Development 2008) conceptualization of positive youth development, the five sets of skills that the Collaborative for Academic, Social and Emotional Learning (CASEL) (2003) conceptualize as social-emotional learning, and the skills presented by Elias (2009) for SECD programs. In fact the *PA* curriculum, which serves as the basis for the index, was given a favorable rating by CASEL (2003) as reflecting a focus on social-emotional learning. These behaviors were also found to be positively associated with teacher reports of the same behaviors (DuBois et al. 2010)

The same 16 behavior items were asked of students across all waves of data collection in Hawai`i. However, the response options differed between years. In grades 1–3 the items had three response options: "*no,*" "*sometimes,*" and "*yes.*" For grades 4 and 5, four response options were used: "*none of the time,*" "*some of the time,*" "*most of the time,*" or "*all of the time.*" To obtain consistent response options across time for the student reports, we coded the items as 1 if they answered *yes* for the three response-option scales, or *all of the time* for the four response-option scales and 0 otherwise. The sum of these items was then transformed into a POMP (percent of maximum possible; Cohen et al. 1999) score so that student reports from the three studies would have the same range of 0–100. This transformation allows the results to be directly compared even with different numbers of items across sites and provides a global measure of positive behavior.

We utilized a longitudinal growth-curve model with a random intercept at the student and school level. This model takes into account similarity of scores within children and within schools. Given our three-level model (observations nested in students nested in schools), two ICC values were obtained for student reports. One ICC is the proportion of variance due to schools and the other is the proportion of variance within children across time (see Table 1). As is typically the case, variance due to child (148.85; s.e. = 8.23) was much larger than variance due to school (4.36; s.e. = 2.15). However, the school ICC (see Table 1) was large enough to justify a three-level model of observations nested in children nested in schools.

The multilevel model adjusting for both individual- and school-level effects was estimated with Stata's xtmixed command using a full-information, maximum-likelihood estimator (Rabe-Hesketh and Skrondal 2008). Because age and gender each have been indicated to be important predictors of behavior at this age (Eisenberg et al. 2006), cohort and gender were included in the model. Our random intercept model can be expressed as

$$\begin{aligned}\hat{Y} = {} & \beta_0 + \beta_1(\text{condition}_k) + \beta_2(\text{gender}_{jk}) + \beta_3(\text{cohort}_{jk}) \\ & + \beta_4(\text{year}_{ijk}) + \beta_5(\text{year}^2_{ijk}) + \beta_6(\text{year}_{ijk}{*}\text{condition}_k) \\ & + \beta_7(\text{year}^2_{ijk}{*}\text{condition}_k) + \zeta_k + \zeta_{jk} + \varepsilon_{ijk}\end{aligned}$$

where $\hat{Y}$ is the estimated POMP score for each child and i represents an observation at wave i (waves 0–4), j represents a child, and k represents a school. ζ_k represents the variance of the random intercept for each school or the deviation of the score for each school from the overall mean score as represented by the intercept, β_0, and ζ_{jk} represents the variance of the random intercept for children in each school and the ε_{ijk} represents the residual at each wave (thus allowing transitory deviations at each wave from the predicted value of Y).

We hypothesized that β_1 for condition, which represents the initial difference between the children in the program and control conditions (condition was coded 0 for control and 1 for *PA*), would not be significant because of randomization. It is included to verify the randomization and to adjust for any possible baseline differences. We hypothesized that β_2 would be significant and negative because boys were expected to report fewer positive behaviors and gender was coded as a 1 for boys and a 0 for girls (Eisenberg et al. 2006). β_3 was expected to be significant and negative because children in the older cohort were expected to have fewer positive behaviors initially than children in the younger cohort and cohort was coded 0 or 1 for the younger or older cohort. β_4 and β_5 test for the normative trajectory of positive behaviors. We hypothesized that β_4 would be significant; that is, we expected there to be a negative trajectory overall. We also included the quadratic term, β_5, to test whether the linear trajectory accelerated or leveled off significantly over time.

The inclusion of the interaction between β_6, (condition and year) and β_7 (condition and the quadratic term) allow our models flexibility in estimating possible differences in linear and quadratic components of trajectories associated with whether a student was attending a school implementing the *PA* program. Our study hypotheses predicted a positive interaction between year and condition as well as between year and condition squared (i.e., those children who were in the *PA* condition would report greater increases in positive behaviors relative to the children in the control condition over time and the rate of this increase would grow over time).

Because some children changed schools, were sometimes absent for an administration of the questionnaire, or refused to answer selected items, there were missing data at all waves. For student reports of their own behavior, 1,544 students responded at the first wave, 2,116 at the second wave, 1,498 at the third wave, 1,493 at the fourth wave and 696 at the final wave. The sharp drop at the final wave was because 6 of the 20 schools (3 control and 3 *PA*) did not contain sixth grade and the entire older cohort in those schools was lost to follow-up. We had a total of 7,347 observations from 2,646 children distributed over 20 schools, with an average of 2.8 waves of data for each

Table 1 Results for multilevel growth model in three trials

	Hawai'i		Chicago		Southeastern state	
Predictor	*B*	*t*	*B*	*t*	*B*	*t*
Intercept	74.23	56.45***	69.61	35.33***	104.48	46.53***
Condition (PA = 1)	3.44	2.35*	−5.64	−2.16*	−2.18	−1.19
Gender (boy = 1)	−8.14	−12.19***	−5.91	−4.24***	−7.77	−11.89***
Cohort (older = 1)	−7.03	−10.44***	Na	Na	−4.51	−17.82***
Year	0.61	0.56	−11.20	−16.81***	−13.32	−6.60***
Year^2	−1.58	−5.99***	Na	Na	1.69	3.43**
Year × condition	−3.99	−3.59***	3.99	4.31***	2.15	3.40**
Year^2 × condition	1.36	4.75***	Na	Na	Na	Na
Random effects	Variance	S.E.	Variance	S.E.	Variance	S.E.
School	4.36	2.13	13.43	8.33	2.81	1.95
Individual	148.85	8.23	298.71	20.77	127.35	8.52
Residual	351.77	7.13	339.98	11.27	270.33	7.53
ICC (school level)	0.01		0.03		0.03	
ICC (between scores)	0.29		0.43		0.43	
PA & control mean difference[a]	11.63		10.48		10.48	
Cohen's d	0.46		0.41		0.41	

[a] At final wave accounting for baseline differences

student. To deal with missing data, full-information, maximum-likelihood estimation was used with the xtmixed command. Given that parents, not students, usually decide if a student is in a school or not and, therefore, missingness is not related to the student behavior outcome (in only two cases was missingness significantly correlated with the outcome and in both cases the correlation was small, only −0.22 and −0.10), it is likely that the missing at random assumption of full-information, maximum-likelihood estimation was met (Brown et al. 2008).

Results

The random intercept multilevel model of student reports of behavior had an overall Wald χ^2 (7)=1,227.55, $p<.001$. The variance at the school level and the individual level, as reported in Table 1, were both substantial and the likelihood ratio χ^2 for the multilevel model versus an OLS regression with 2° of freedom was 666.17, $p<.001$. The Wald χ^2 test is similar to an overall model *F*-test and gives an idea of overall model fit. The likelihood ratio χ^2 supports the use of a multilevel model.

The main effect of condition (at baseline) was significant (B_1=3.44, $p<0.05$) (see Table 1). The significant baseline difference in reports of positive behavior is most likely an artifact of fitting a quadratic model, as the baseline differences in behavior were not different using a simple *t*-test ($M_{control}$=67.57, M_{PA}=68.07, $t(1,576)$=−0.29, *ns*) and in a simple linear growth model. As hypothesized, boys reported significantly fewer positive behaviors than girls (B_2=−8.14, $p<0.001$). The children from the older cohort also endorsed fewer positive behaviors (B_3=−7.03, $p<0.001$). Contrary to our hypotheses, the main effect of year was not significant (B_4=−0.61, $p>0.05$). The year squared term was significant and negative (B_5=−1.58, $p<0.001$) indicating an accelerating decline in the endorsement of positive behaviors during elementary school.

Findings further revealed, as hypothesized, a significant interaction between year and condition (B_6=−3.99, $p<.05$) and between year square and condition (B_7=1.36, $p<.001$). The year by condition interaction is the slope of the curve at year zero, with the year square by condition interaction slowing that decline over time.

The baseline mean on the endorsement of positive behaviors in *PA* and control schools was a close match (68.07 and 67.57, respectively). By the fifth wave there had been a substantial reduction in the number of positive items the children in both sets of schools endorsed with sample means of 50.88 and 37.23 for the children in the *PA* and control schools, respectively. An estimate of the size of the program effect (a simple Cohen's *d* for the final wave controlling for baseline differences) was 0.46 (see Table 1).

Discussion

Figure 1 (first panel) shows these findings graphically. As hypothesized, the number of positive behaviors endorsed decreased from year to year and this decrease was partially

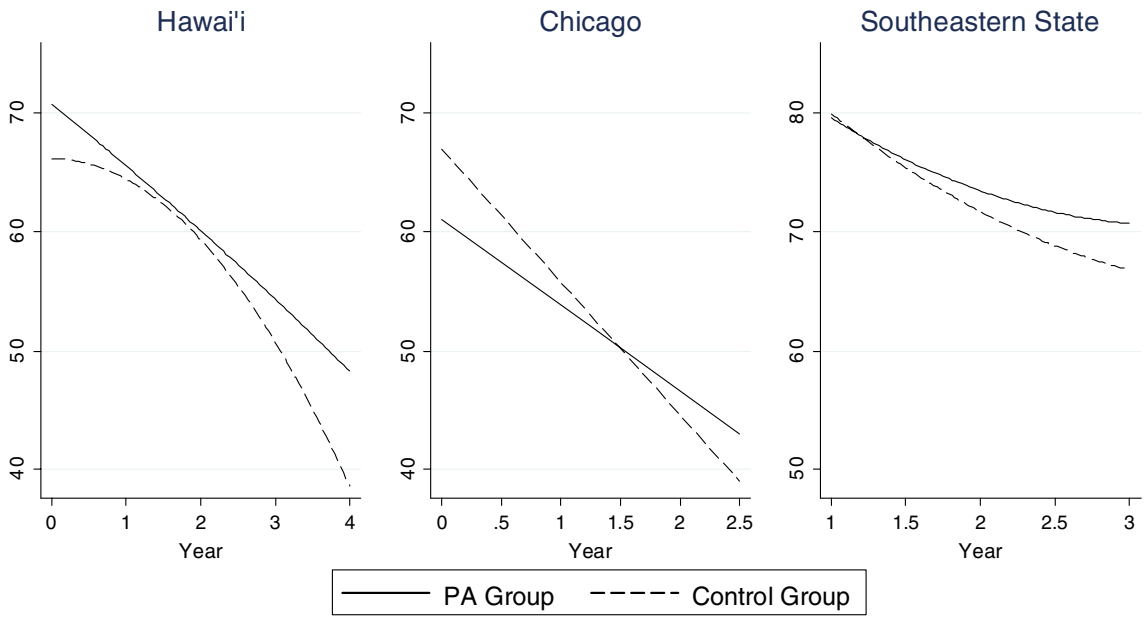

Fig. 1 Estimated trajectories for behaviors by condition for all analyses

mitigated by the significant positive effect of the *PA* program. We also see that, even though the main effect of condition was significant, the accelerating decline of the control group, coupled with the effects of *PA*, created a much larger gap by the end of the study. It appears that the effect of *PA* here was to eliminate the acceleration in the decline of positive behaviors. This is best seen in the figure as the linear nature of the *PA* line compared with the quadratic curve of the control line. Even with *PA*, there was a steady decline with fewer positive behaviors expected each year.

The Hawai'i student reports of positive behaviors support the most current research on positive behaviors associated with SECD; namely, that the prevalence of these behaviors falls as children enter adolescence (Carlo et al. 2007; Kokko et al. 2006). They also show that a global measure of positive behaviors declines from 1st grade to 6th grade, extending backwards the work done by Lerner and Colleagues (Tufts University, Institute for Applied Research in Youth Development 2008) on older children. These data also provide new evidence for the effectiveness of *PA*. In particular, results indicated that, in addition to reducing health-compromising and high-risk behaviors (Beets et al. 2009), *PA* also mitigates the decline of positive behaviors associated with SECD.

Chicago Trial

Method

A second matched-pair randomized study, conducted in 14 elementary schools in the Chicago Public School system, provided five data points across 3 years where data were collected from a single cohort: beginning and end of grades 3 and 4 and the end of grade 5. As in the Hawai'i Trial, the teachers within each of the seven program schools received the *PA* training from the program developer and brief update trainings were repeated at the start of each subsequent year. The teachers in the seven control schools received no *PA* training and were asked not to implement the program. The schools in Chicago were matched and randomized in a similar fashion to those in the Hawai'i Trial (Li et al. 2011). As in Hawai'i there were no significant differences at baseline on any of the variables used for matching (Ji et al. 2008; Li et al. 2011).

The same method of creating a global measure of positive behaviors using a POMP score was used as in the Hawai'i Trial. A total of 51 behavior items were asked that included 16 items common to the Hawai'i Trial, each with the same four response options: "*none of the time,*" "*some of the time,*" "*most of the time,*" or "*all of the time.*" Although the problem of different response options over time did not exist in Chicago as it did Hawai'i, to stay consistent across trials the 51 items were coded 1 for or "*all of the time*" and 0 otherwise. The 51 items were then summed and transformed into a POMP score. Chicago had more items because the children in the study started at an older age and therefore were given more items.

We used the same multilevel growth model with the Chicago data as the Hawai'i data, but without the cohort variable (as only a single cohort was followed in Chicago). β_5, the year square parameter, and β_7, the year square by condition parameter, were not significant, and the model was rerun without the quadratic effects. A log-likelihood

test showed that the model did not significantly degrade with the omission of these two parameters (log-likelihood $\chi^2(2)=0.98$, *ns*).

As in Hawai'i, missing data were handled through use of full-information, maximum-likelihood estimation. Chicago had 593 students at the first wave of data collection, 557 at the second wave, 547 at the third wave, 512 at the fourth wave and 497 at the final wave. For the positive behaviors in Chicago, we had a total of 2,704 observations from 936 children distributed over 14 schools, with an average of 2.9 waves of data for each student. Since the data were collected annually in Hawai'i, but biannually in Chicago, the time variable for Chicago was changed to reflect the difference (0, 1, 2, 3, 4 years in Hawai'i, and 0, .5, 1, 1.5, 2.5 years in Chicago). The same missing at random argument holds in the Chicago Trial as it did in the Hawaii Trial, only more so as at no time was the outcome correlated with missingness.

Results

The results for positive behavior had an overall model Wald $\chi^2(4)=425.83$, $p<0.001$, supporting the overall significance of the model. The multilevel model also fit better than an OLS regression model, $\chi^2(2)=650.30$, $p<.001$. Taken together, we have a multilevel model that reduces variation in the outcome and performs better than an OLS regression.

The main effect of condition was negative ($B_1=-5.64$, $p<.05$) indicating that, despite random assignment of schools, students in *PA* schools started with a lower POMP score at baseline than students in control schools. As hypothesized, boys reported significantly fewer positive behaviors than girls ($B_2=-5.91$, $p<0.001$). The main effect of year was negative and significant ($B_4=-11.20$, $p<.001$), indicating the predicted general decline in positive behaviors. The year by condition interaction was significant and positive ($B_6=3.99$, $p<0.001$) as predicted as well, indicating a positive program effect.

Children in *PA* schools had a mean score of 63.53 at baseline and children in control schools had a mean of 67.55. By the final wave of the study, children in control schools had a mean score of 39.71, and children in *PA* schools had a higher mean of 43.52. An estimate of the size of the program effect (a simple Cohen's *d* at the final wave, controlling for baseline differences) was 0.41.

Discussion

In the Chicago Trial, children in *PA* schools started on average lower than children in control schools but, over the course of the study, these children surpassed the control children and, at the end of the study, had a higher mean score than the control children, overall replicating the findings reported from the Hawai'i Trial. We see in the bottom middle panel of Fig. 1 that, as in Hawai'i, the student data showed that children had a negative trajectory of positive behaviors into the beginning of adolescence. The figure also shows the ability of *PA* to change the trajectory substantially. Unlike in the Hawai'i Trial, neither group was accelerating in their decline.

Southeastern State Trial

Methods

The Southeastern State Trial was conducted in eight rural public elementary schools, with five age cohorts that ranged from children in kindergarten to fourth grade at wave 1. The data were collected at the end of each of three consecutive academic years. A limitation of the Southeastern State data was that the first measurement occurred at the end of the first year of implementation, so no direct baseline comparison was possible. Nonetheless, we can compare the trajectories of children in the *PA* program for 3 years, from the end of the first year of intervention through the end of the third year of intervention. The teachers within each of the four program schools received the *PA* training from the program developer at the beginning of the project, but no follow-up trainings. The teachers in the four control schools received no *PA* training and were asked not to implement the program.

The schools in the Southeastern State were matched and randomized, but this was done by the school district, which has not released details of how they matched or randomized. We do have a set of baseline characteristics for the Southeastern State schools (school-level variables similar to those used in Chicago and Hawai'i Trials) and Table 2 indicates that the control and *PA* schools were not statistically different at the school level on any of the seven variables tested. Unlike in the Hawai'i and Chicago trials, student data were collected by school district personnel rather than research staff.

As in Chicago and Hawai'i, a set of items asking about frequency of positive behaviors associated with SECD were asked each year; 14 items were asked, all of which were in the Hawai'i measure. In grades 2–4 the items had three response options: "*no,"* "*sometimes,"* and "*yes."* For fifth grade, the items had four response options: "*none of the time,*" "*some of the time,*" "*most of the time,"* or "*all of the time."* To allow comparison across time and consistent with the other trials, the items were dichotomized, summed and converted into a POMP score to provide a global measure of positive behavior.

The same initial multilevel model was utilized for the analysis of the data as in Hawai'i; however, β_7 (Year^2 × condition interaction parameter) was not significant so the

Table 2 Baseline equivalence for schools in a southeastern state

	Control schools		Positive Action schools		*p*
	Mean	*SD*	*Mean*	*SD*	
Enrollment	532.50	147.15	543.25	113.60	0.91
Percentage free/reduced lunch	61.08	10.30	70.98	14.93	0.32
Percentage special education programs	5.40	1.64	4.58	0.97	0.42
Percentage limited english proficiency	31.03	4.23	32.58	4.18	0.62
Average daily attendance rate	97.33	1.78	96.93	2.31	0.79
Suspensions	0.18	0.13	0.10	0.20	0.55
Retentions	3.05	1.59	2.93	2.94	0.94

model was re-estimated without this interaction term. To test the effect of these changes on model fit, a log-likelihood test was run and the model fit did not degrade significantly with the above omission (log-likelihood $\chi^2(2)=0.23$, $p>.05$).

The Southeastern State Trial had 1,652 students at the first wave, 1,944 students at the second wave, and 1,504 students at the third wave. There were a total of 5,100 observations distributed over eight schools for 2,610 children with an average of 2.0 waves of data for each student. As in the other trials, full maximum-likelihood estimation was used to account for missing data. Because the data from this trial did not include a baseline measurement, any results reported for baseline are extrapolated by the model. The same missing at random argument holds in the Southeastern State Trial as it did in the Hawai'i Trial except that, with a greater diversity of ages, we found both the outcome (always lower than −0.25) and age were correlated with missingness. The strong correlation between age and outcome could be driving the correlation between the outcome and missingness and, therefore, by controlling for age in the model (through cohort) the assumptions of MAR are likely met.

Results

The results for positive behavior in the Southeastern State had a significant overall model (Wald $\chi^2(6)=624.97$, $p<001$). Again, the multilevel model fit better than an OLS regression model, $\chi^2(2)=317.63$, $p<.001$.

Because students in the Southeastern State Trial were first measured at the end of the first year, but not at the baseline, the intercept is an extrapolated value. The main effect of condition (at extrapolated baseline) was not significant ($B_1=-2.18$, $p>0.05$). Boys reported significantly fewer positive behaviors than girls ($B_2=-7.77$, $p<0.001$). The children from the older cohorts endorsed fewer positive behaviors ($B_3=-4.51$, $p<0.001$). The year term was significant and in the hypothesized negative direction ($B_4=-13.32$, $p<0.001$) and the year squared term was significant and positive ($B_5=1.69$, $p<0.01$), indicating a decelerating rate of decline in the endorsement of positive behaviors during elementary school. The interaction of year by condition was, as hypothesized, positive and significant, $B_6=2.15$, $p<0.01$. This is the intervention effect, as the year square by condition interaction was not included in the final model.

At the end of the first year, our proxy for the baseline mean, reported levels of positive behavior were similar for children in the control and *PA* schools (77.83 and 77.51, respectively). At the end of the study (third wave), the mean number of positive behaviors reported by children in the control schools dropped to 69.89 compared to 74.23 in the *PA* schools. An estimate of the size of the program effect, Cohen's *d* at the final wave, controlling for baseline differences, was 0.22.

Discussion

The mitigation of the decline in endorsement of positive behaviors by students in the Southeastern State Trial is illustrated in the lower right panel of Fig. 1 and is consistent with the results from Hawai'i and Chicago. We also have a replication of the decline in positive behaviors over time. The students in this trial started at much higher levels than in either Hawai'i or Chicago and saw decreasing declines across time. Although this trial is not as methodologically strong as the other two, the greater number of cohorts and the replication of results strengthen both the arguments that behaviors decline into adolescence and that *PA* mitigates this decline.

Conclusion

Even though each trial had its own racial and socioeconomic demographics, students in each trial responded similarly to the *PA* program. In each case, the children in *PA* schools showed smaller declines compared with children in control schools. Thus, the *PA* program prevented a significant reduction in positive behaviors.

Overall, the normative declining developmental trajectories that we found are consistent with those found by Lerner and Colleagues (Tufts University, Institute for Applied

Research in Youth Development 2008) with his global measure of positive behavior, and the literature on prosocial behavior (Carlo et al. 2007; Kokko et al. 2006; Nantel-Vivier et al. 2009). Given that each of these trials focused on schools in high risk areas (i.e., poverty), it is possible that the declines we found were driven by a combination of exposure to high-risk conditions and a lack of access to protective resources (i.e. positive role models, opportunities for constructive interactions, emotional support) as suggested (Carlo et al. 2007). Developmental theory (e.g., Bronfenbrenner and Morris 2006; Wiesner et al. 2007) generally acknowledges the impact of environments in shaping behaviors. This does not, however, mean that the more cognitive-centered theory presented by Eisenberg et al. (2006) is incorrect; it may be that the lack of resources in these areas overpowers any gains through cognitive development. More work on this hypothesis is needed.

Future research also should examine other possible causes of the variability in developmental trajectories for positive behaviors found in this and other studies. The present study showed that the *PA* intervention beneficially influenced the trajectories in diverse contexts. Further work should consider other factors that may shape the levels and slopes of positive behavior. In addition, it would be important to determine if there are subpopulations of children who respond differently to the intervention utilizing growth-mixture modeling to consider variation in developmental trajectories in the context of evaluating interventions such as *PA* (Segawa et al. 2005).

The combination of three trials in three geographically dispersed school districts, each with diverse populations of students, provides strong evidence that the *PA* program significantly reduced the normative decrease in positive behaviors associated with SECD as children develop from age 6 to 11 years. The demographic and cultural differences between the Hawai'i Trial, the urban setting of the Chicago Trial, and the more rural Southeastern State Trial could account for observed differences in levels and slopes of positive behaviors associated with SECD. Such differences in trajectories are common in the available literature on the normative trajectories of positive behaviors (Carlo et al. 2007; Kokko et al. 2006; Nantel-Vivier et al. 2009), but analysis of reasons for the differences is beyond the scope of this paper.

There was variation in the implementation of the intervention that may have influenced its effects. Teachers in all three trials received initial training, but the Southeastern State teachers did not receive the subsequent annual refresher training. Program effects were still evident (though smaller) in the Southeastern State, suggesting that this program's fidelity may depend, to some extent, on the consistent retraining of school staff by program staff. However, all schools that began the trial were included in the analysis regardless of level of implementation, following intent-to-treat criteria at the school level. Another limitation was the use of only student self-reports in the outcome. However, because across the sites the student self-report was the only consistent measure it was a necessary limitation to get comparable results.

Overall, this research shows the effectiveness of *PA* in mitigating the decrease in self-reported positive behaviors. This study adds to the literature on preventive interventions that actively support the development of positive behavior and SECD. Past reports of the beneficial effects of *PA* have shown effects on school-level variables, such as academic achievement (Flay et al. 2001; Flay and Allred 2003; Snyder et al. 2010) and negative behavioral outcomes (Beets et al. 2009; Li et al. 2011). However, the theoretical basis (Flay and Allred 2010) and day-to-day protocol of the intervention focuses on promoting positive attitudes and behaviors. The reduction in negative behaviors and improvement in academic achievement is considered as a result of more positive individual development. It is clear that this approach is effective. But to better understand and demonstrate the mechanism of these beneficial effects it is essential to show that the intervention influences the development of positive behavior. This report partially fills that gap by showing the effects of *PA* on positive behaviors related to SECD. Further research is needed to clarify the role of positive behavior as a potential mediator for the effects of this intervention on negative outcomes and academic achievement. Nevertheless, this study provides new insights on preventive interventions that apply a comprehensive approach that includes the development of positive cognitive, emotional and behavioral characteristic in school-aged children.

Acknowledgments This project was funded by the National Institute on Drug Abuse (R01-DA13474), the Institute of Education Sciences (R305L030072) and the Department of Education (Partnerships in Character Education grants). 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 Hawai'i, Chicago and Southeastern State school districts 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 in the Hawai'i dataset.

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

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.

Benninga, J. S., Berkowitz, M. W., Kuehn, P., & Smith, K. (2006). Character and academics: What good schools do. *Phi Delta Kappan, 87*, 448–452.

Berkowitz, M. W., & Bier, M. C. (2004). Research-based character education. *The Annals of the American Academy of Political and Social Science, 591*(1), 72–85.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Series Eds & Vol. Eds.), *Handbook of child psychology Vol 1. Theoretical models of human development* (6th ed., pp. 793–828). New York: Wiley.

Brown, C. H., Wang, W., Kellam, S. G., Muthén, B. O., Petras, H., Toyinbo, P., et al. (2008). 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, 95*, 74–104.

Carlo, G., Crockett, L. J., Randall, B. A., & Roesch, S. C. (2007). A latent growth curve analysis of prosocial behavior among rural adolescents. *Journal of Research on Adolescence, 17*, 301–324.

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., Gavin, L. E., & Markham, C. M. (2010). Future directions for positive youth development as a strategy to promote adolescent sexual and reproductive health. *Journal of Adolescent Health, 46*, S92–S96.

Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research, 34*, 315–346.

Collaborative for Academic, Social, and Emotional Learning. (2003). *Safe and sound: An educational leader's guide to evidence-based social and emotional learning programs*. Chicago, IL: Author.

Dent, C. W., Galaif, J., Sussman, S., Stacy, A., Burton, D., & Flay, B. R. (1993). Demographic, psychosocial and behavioral differences in samples of actively and passively consented adolescents. *Addictive Behaviors, 18*, 51–56.

DuBois, D. L., Ji, P., Flay, B. R., Day, J., & Silversthorn, N. (2010, June) *Further validation of the youth social and character development scale. Poster session presented at the annual research conference of the Institute of Educational Sciences*, Washington, DC.

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In W. Damon & R. M. Lerner (Series Eds.), N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 646–718). New York: Wiley.

Eisenberg, N., & Morris, A. S. (2004). Moral cognitions and prosocial responding in adolescence. In R. M. Lerner & L. D. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 155–188). Hoboken, NJ: Wiley.

Elias, M. J. (2009). Social-emotional and character development and academics as a dual focus of educational policy. *Education Policy, 23*, 831–846.

Ferber, T., Gaines, E., & Goodman, C. (2005). Positive youth development: State strategies. Retrieved March 3, 2010 from http://www.ncsl.org/Portals/1/documents/cyf/final_positive_youth_development.pdf

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

Flay, B. R., & Allred, C. G. (2003). Long-term effects of the positive action program—a comprehensive, positive youth development program. *American Journal of Health Behavior, 27*, S6–S21.

Flay, B. R., & Allred, C. G. (2010). 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 research handbook on values education and student well-being* (pp. 471–501). New York: 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*, 71–89.

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

Ji, P. Y., DuBois, D. L., Flay, B. R., & Brechling, V. (2008). "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, 78*, 131–139.

Kokko, K., Tremblay, R. E., Lacourse, E., Nagin, D. S., & Vitaro, F. (2006). Trajectories of prosocial behavior and physical aggression in middle childhood: Links to adolescent school dropout and physical violence. *Journal of Research on Adolescence, 16*, 403–428.

Lerner, R. M., Almerigi, J. B., Theokas, C., & Lerner, J. V. (2005). Positive youth development: A view of the issues. *Journal of Early Adolescence, 25*, 10–16.

Li, K.-K., Washburn, I., DuBois, D. L., Vuchinich, S., Ji, P., Brechling, V., et al. (2011). Effects of the *Positive Action* program on problem behaviors in elementary school students: A matched-pair randomized control trial in Chicago. *Psychology & Health, 26*, 187–204.

Nantel-Vivier, A., Kokko, K., Caprara, G. V., Pastorelli, C., Gerbino, M. G., Paciello, M., et al. (2009). Prosocial development from childhood to adolescence: A multi-informant perspective with Canadian and Italian longitudinal studies. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 50*, 590–598.

Park, N. (2004). Character strengths and positive youth development. *The ANNALS of the American Academy of Political and Social Science, 591*, 40–54.

Phelps, E., Zimmerman, S., Warren, A., Jelicic, H., von Eye, A., & Lerner, R. M. (2009). The structure and developmental course of positive youth development (PYD) in early adolescence: Implications for theory and practice. *Journal of Applied Developmental Psychology, 30*(5), 571–584.

Purkey, W., & Novak, J. (1984). *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*. College Station, TX: Stata Press.

Segawa, E., Ngwe, J. E., Li, Y., & Flay, B. R. (2005). Evaluation of the effects of the Aban Aya Youth Project in reducing violence among African American adolescent males using latent class growth mixture modeling techniques. *Evaluation Review, 29*, 128–148.

Snyder, F., Flay, B., Vuchinich, S., Acock, A., Washburn, I., Beets, M., et al. (2010). 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. *Journal of Research on Educational Effectiveness, 3*, 26–55.

Tufts University, Intitute for Applied Research in Youth Development. (2008). *The 4-H study of positive youth development: Report of the findings from the first four waves of data collection: 2002–2003, 2003–2004, 2004–2005, and 2005–2006*. Retrieved March 3, 2010 from http://ase.tufts.edu/iaryd/documents/4HStudyAnnualReport2008.pdf

Wiesner, M., Capaldi, D. M., & Patterson, G. R. (2007). Development of antisocial behavior and crime across the life-span from a social interactional perspective: The coercion modell. In R. L. Akers & G. F. Jensen (Eds.), *Social learning theory and the explanation of crime: A guide for the new century* (pp. 317–338). New Brunswick, NJ: Transaction.