

# A Longitudinal Evaluation of the Positive Action Program in a Low-Income, Racially Diverse, Rural County: Effects on Self-Esteem, School Hassles, Aggression, and Internalizing Symptoms

Shenyang Guo<sup>1</sup> · Qi Wu<sup>2</sup> · Paul R. Smokowski<sup>3</sup> · Martica Bacallao<sup>3</sup> · Caroline B. R. Evans<sup>2</sup> · Katie L. Cotter<sup>4</sup>

Received: 31 July 2015 / Accepted: 22 September 2015 / Published online: 28 September 2015  
© Springer Science+Business Media New York 2015

**Abstract** Positive Action is a school-based program that aims to decrease problem behaviors (e.g., violence, substance use) and increase positive behaviors (e.g., school engagement, academic achievement). Although a number of studies have shown that Positive Action successfully achieves these goals, few studies have evaluated the program's effectiveness in rural schools. Given that rural youth are at an increased risk for risky behaviors (e.g., violence, substance use), this is a critical gap in the existing Positive Action research base. The current study assesses the impact of Positive Action on change rates of self-esteem, school hassles, aggression, and internalizing symptoms in a group ( $N = 1246$ , 52 % female) of ethnically/

racially diverse (27 % White, 23 % African American, 12 % mixed race/other, 8 % Latino, 30 % as American Indian) middle school youth (age range 9–20) located in two violent, low-income rural counties in North Carolina. One county engaged in Positive Action over the 3-year study window while the other county did not. Following multiple imputation and propensity score analysis, 4 two-level hierarchical linear models were run using each of the outcome measures as dependent variables. The results indicate that the program generates statistically significant beneficial effects for youth from the intervention county on self-esteem scores and school hassles scores. Although the program generates beneficial effects for intervention youth on the change in aggression scores, the finding is not statistically significant. The finding on the change in internalizing scores shows a non-significant detrimental effect: the youth from the comparison county have lower internalizing scores than those from the intervention county. Implications are discussed.

---

✉ Shenyang Guo  
s.guo@wustl.edu

Qi Wu  
qiwu@email.unc.edu

Paul R. Smokowski  
smokowski@ku.edu

Martica Bacallao  
marticabacallao@gmail.com

Caroline B. R. Evans  
careyroberson@gmail.com

Katie L. Cotter  
katie.cotter@asu.edu

**Keywords** School-based interventions · Positive Action · Self-esteem · Aggression · Rural · Propensity score analysis

## Introduction

Adolescence marks an intense developmental period focused on identity formation (Erikson 1950). As youth develop a sense of self, they face a number of challenges (e.g., bullying, parent–child conflict, friend rejection, delinquent friends, peer pressure) that put them at risk for negative developmental outcomes such as low self-esteem, internalizing symptoms, and aggression (Smokowski et al. 2014; Smokowski et al., in press). Further, youth often negatively impact their development by engaging in risk

<sup>1</sup> Brown School of Social Work, Washington University, CB# 1196, One Brookings Drive, St. Louis, MO 63130, USA

<sup>2</sup> School of Social Work, University of North Carolina, CB# 3550, 325 Pittsboro St, Chapel Hill, NC 27599, USA

<sup>3</sup> School of Social Welfare, University of Kansas, 1545 Lilac Ln, Lawrence, KS 66045, USA

<sup>4</sup> School of Social Work, Arizona State University, 411 N. Central Avenue, Suite 800, Phoenix, AZ 85004, USA

taking behaviors such as violence, aggression, substance use, and dropping out of school. These risk-taking behaviors are a significant problem across the United States, but especially in rural states such as North Carolina. For example, according to the Youth Risk Behavior Surveillance Survey (YRBSS), a national survey of U.S. middle- and high-school students conducted by the Centers for Disease Control (Centers for Disease Control 2013), 63.8 % of middle school males and 50.7 % of middle school females in North Carolina reported having been in a physical fight in the current school year (Centers for Disease Control 2013). Middle school youth in North Carolina are also at considerable risk for substance use and 6.1 % of middle school youth currently smoke cigarettes, 26.2 % have used alcohol, 12.7 % have smoked marijuana, and 2.4 % have used cocaine (Centers for Disease Control 2013). Further, research in rural areas in general highlights that, compared to urban and suburban youth, rural youth are at an increased risk for problem behaviors such as binge drinking (Substance Abuse and Mental Health Services Administration 2014), low school engagement (Witherspoon and Ennett 2011), and dropping out of school (Provasnik et al. 2007). This past research suggests that rural youth in particular would benefit from universal prevention programs, such as Positive Action, that aim to decrease risky behaviors such as violence, dropping out of school, and substance use while simultaneously improving self-esteem and increasing prosocial actions. The Positive Action program is especially well suited for middle school youth who often display high rates of problematic behaviors and are at risk for continuing such negative behaviors in high school.

### Overview of the Positive Action Intervention

Positive Action is a school-based intervention designed to improve academic achievement, school attendance, problem behaviors (e.g., substance use, violence, disruptive behaviors, dropping out of school, sexual behavior), parent–child bonding, family cohesion, and family conflict (National Registry 2014). This intervention is presented as a series of kits consisting of lesson plans and materials that are developmentally appropriate for use with elementary-, middle-, or high-school aged youth. The elementary-school kit consists of 140 brief lessons (approximately 15–20 min) to be used with Kindergarten aged youth through Grade 6; the middle-school curriculum contains two kits of 82 lessons ranging from 15 to 20 min for use with youth in Grades 7 and 8; the high school kit contains four kits for Grades 9 through 12. Every kit has materials for the following six units: (a) Self-Concept, designed to enhance students' self-understanding and identity; (b) Positive Actions for Your Body and Mind, designed to teach proper

hygiene, exercise, and creative thinking skills; (c) Managing Yourself Responsibly, designed to help students identify positive skills and learn skills to effectively manage time and resources; (d) Treating Others the Way You Like to be Treated, designed to teach students positive social skills through skills practice; (e) Telling Yourself the Truth, designed to teach the importance of self-honesty and implementation strategies; and (f) Improving Yourself Continually, designed to teach students learn how to apply positive action skills in all areas of life and convey the idea that self-improvement is an ongoing, continual process (Positive Action, n.d.). Thus, the focus of the Positive Action program is on increasing adolescent positive behaviors and actions with the aim of improving behavioral and developmental outcomes.

Since the creation of Positive Action in 1977, various research studies have established the utility of the program. For example, Blueprints for Healthy Youth Development, a database of violence, delinquency, and drug prevention programs that have met a stringent standard of program effectiveness, has labeled Positive Action as a *model program*. Notably, the model program designation is reserved for programs that have undergone the most rigorous evaluation, including at least two high-quality randomized controlled trials, or one high-quality randomized controlled trial and one high-quality quasi-experimental evaluation. In addition, model programs must demonstrate positive program effects that are sustained a minimum of 12 months following program completion (Blueprints 2012). Positive Action has received a number of other national accolades. For example, the U.S. Department of Education has twice recognized Positive Action: The Department of Education *What Works Clearing House* first gave Positive Action a positive rating in 2007 based on strong evidence that the intervention had positive effects on outcomes (Department of Education 2007). In 2009, the Department of Education's Safe, Disciplined, and Drug-Free Schools Expert Panel labeled Positive Action as a promising program (Department of Education 2009). Similarly, the U.S. Department of Justice has recognized Positive Action as an effective program, citing the program was supported by strong evidence to indicate the intervention achieves its intended outcome (National Institute of Justice, n.d.). To date, the majority of Positive Action research has been conducted in urban areas, leaving a critical gap in the evidence base regarding the effectiveness of this program with rural, racially/ethnically diverse populations. Therefore, the current study sought to evaluate the impact of Positive Action over a 3-year period on the change rates of self-esteem, school hassles, internalizing symptoms, and aggression in a low-income, racially/ethnically diverse, rural county in North Carolina relative to a neighboring county that did not receive the intervention. Middle schools

(Grades 6, 7, 8) in the intervention county implemented Positive Action whereas middle schools in the comparison county did not. Propensity score analysis was used in the evaluation to control for various nonrandom selections and to enhance the rigor of comparison. It is important to note that the current research team has no connection to the Positive Action program and that no conflicts of interest exist.

### Theory of Change: Self-Esteem Enhancement Theory and Organizational Culture Theory

The theory of change underlying the Positive Action program is based on the simple philosophy that individuals feel good about themselves when they engage in positive actions (Positive Action, n.d.). This philosophy is grounded in self-esteem enhancement theory, which posits that individuals desire positive self-worth and that positive outcomes result when individuals are supported in their desire for self-esteem through adaptive thoughts, values, and behaviors (DuBois et al. 2009). This theoretical framework is illustrated through the thoughts-actions-feelings circle (Positive Action, n.d.; see Fig. 1). Given that thoughts lead to actions and actions lead to feelings, the Positive Action program seeks to encourage actions that promote a healthy and positive cycle. Engaging in prosocial actions is expected to positively impact one's thoughts and feelings, especially with respect to self-appraisal. Positive Action encourages youth to engage in prosocial behaviors, to think in a positive manner, and to appreciate themselves for their positive actions; in this regard, the program aims to increase adolescent's self-esteem. Youth who feel positively about themselves (i.e., have high self-esteem) are likely to continue behaving in a positive and prosocial manner and to refrain from engaging in negative

behaviors, suggesting that Positive Action also functions to decrease harmful behaviors such as aggression. Further, as youth's self-esteem is continuously bolstered through ongoing Positive Action lessons and continued engagement in positive actions, this high self-esteem might function to combat poor mental health, such as symptoms of anxiety and depression.

The Positive Action curriculum was designed to first introduce the overarching goal of positive self-appraisal (i.e., feeling good about self), and then increase specific skills needed to achieve this goal (Lewis et al. 2013a). The theory of change guiding the classroom curriculum component can be divided into immediate outcomes, short-term effects (mediators), and long-term impact. The classroom curriculum is hypothesized to improve teacher–student relations, student–student relations, student–parent relations, and student engagement with the community (Flay et al. 2001). In addition to improving the learning environment, short-term program goals include improved self-concept; enhanced study/thinking skills; increased motivation to learn; improved health habits; effective self-management; increased interpersonal/social/emotional skills; increased self-honesty; and improved skill levels in goal setting, problem solving, and decision making (Flay et al. 2001). These short-term program goals are hypothesized to translate into improved school attendance, improved grades, higher test scores, fewer disciplinary problems, reduced substance use, and less aggression and violence (Flay et al. 2001), resulting in overall improvement in the school culture.

If all youth in a middle school are exposed to the tenants of the Positive Action program, the entire culture of the school might be improved, causing youth to perceive the overall school environment more positively (e.g., decreased perception of physical and verbal aggression). According to organizational culture theory, every organization has a unique culture (Geertz 1973) and school culture refers to a school's "...unwritten rules and traditions, norms, and expectations...that seem to permeate everything..." (Deal and Peterson 2009, p. 2). Thus, Positive Action aims to improve the culture of a school by making positive behaviors the norm. Research suggests that a positive school climate and culture is associated with improved student behavior in the form of decreased bullying (Lee and Song 2012), supporting the notion that if Positive Action can universally increase positive behaviors, negative actions will decrease and youth will view school as safer (e.g., decreased school hassles).



**Fig. 1** Positive Action thoughts, actions, feelings circle (Positive Action, n.d.)

### Past Research on the Positive Action Intervention

Past research on Positive Action has indicated improvements in a number of the programs' targeted outcomes. For

example, in a 6 year, eight wave longitudinal study with a sample of 1170 low-income urban youth in Grades 3 through 8, participation in Positive Action had a positive effect on reducing absenteeism, and lessened the natural increase in students' disaffection with learning. In addition, relative to teachers in control schools, teachers in the Positive Action intervention schools rated their students as achieving greater growth in academic motivation and academic ability. Further, all students in Positive Action schools showed an increase in math scores relative to their counterparts in the control schools, and intervention schools also showed significant, positive effects on reading for African American boys (Bavarian et al. 2013).

The research on Positive Action in rural areas is limited, with only two studies that have examined the effects of the program in rural schools. One study of fifth grade youth attending rural and urban schools in Hawaii found that as compared with control schools, students who participated in Positive Action demonstrated significant decreases in self- and teacher-reported violence as well as decreased self-reported substance use and sexual activity (Beets et al. 2009). Another study in eight public elementary schools in rural areas of one state found that after 3 years of the Positive Action intervention, youth in the intervention schools reported significantly higher numbers of positive behaviors as compared to youth in control schools. However, this study design had methodological limitations because the baseline measures were collected 1 year after implementation of the Positive Action program (Washburn et al. 2011).

Other evaluations of Positive Action in urban areas have reported that participation in the program resulted in decreased violence (Flay et al. 2001; Flay and Allred 2003; Lewis et al. 2013b; Li et al. 2011; Snyder et al. 2013), absenteeism (Snyder et al. 2010), suspensions (Snyder et al. 2010), bullying (Li et al. 2011), depression and anxiety (mediated by social-emotional and character development; Lewis et al. 2013a), and substance use (Li et al. 2011; Snyder et al. 2013). In addition, evaluations of Positive Action have also shown program effects of increased academic performance (Flay et al. 2001; Flay and Allred 2003; Snyder et al. 2010, 2013) and improved school involvement (Flay and Allred 2003). To date, few methodologically strong studies of Positive Action have been conducted in rural schools, which is troubling given the unique and elevated risk factors present in rural areas.

### Rural Areas

Rural environments expose residents to stressors absent in urban environments, such as geographic isolation, restricted social networks, and limited community resources (U.S. Department of Justice 2001). Specifically for

adolescents in rural areas, these stressors decrease access to extracurricular activities, mental health services, and social interactions with nonfamily members. The stress of rural living likely contributes to the higher rates of risk-taking behavior displayed by rural youth relative to urban and suburban youth. For example, compared with urban and suburban middle- and high-school aged youth, rural adolescents in this age group have been found to be more likely to engage in sexual intercourse and become pregnant (Atav and Spencer 2002), resulting in a rural teen birth rate that is one third higher compared to the birth rate in metropolitan counties (Finley and Stewart 2013). Further, rural youth are more likely than urban youth to binge drink (Substance Abuse and Mental Health Services Administration 2014) and to use non-medical prescription drugs (Barnett et al. 2015).

Other comparisons of rural and urban youth have reported rural middle- and high-school aged youth experienced decreased school belongingness and increased school misbehavior over a three and a half year study window (Witherspoon and Ennett 2011), which likely contributes to the high drop-out rates of high-school youth in impoverished rural communities (Provasnik et al. 2007). The environment of rural schools is particularly ripe with problematic behaviors; reports of rates of bullying victimization in rural areas range from 33 % (Price et al. 2013) to 82.3 % (Dulmus et al. 2004). Notably, these bullying rates in rural areas far exceed the rates in national bullying surveys ranging from 10.6 % (Health Behavior of School-Aged Children Survey; Nansel et al. 2001) to 27.8 % (School Crime Supplement; Robers et al. 2013). Taken together, the research on rural areas suggests that rural youth are a high-risk population in need of support. Based on previous studies, Positive Action appears to have the potential to provide support for rural adolescents, and ultimately improve their academic achievement, school attendance, problem behaviors, and family relationships. However, additional methodologically rigorous evaluation studies of the Positive Action program in rural areas are necessary.

### Hypothesis for Current Study

Based on the theory of change, organizational culture theory, and past research, the hypothesis guiding the current study was that, relative to youth in the comparison county (Positive Action not implemented), youth in the intervention county who received the Positive Action intervention would display increases in self-esteem and decreases in school hassles, internalizing symptoms, and aggression over a 3-year period. Specifically, past studies of Positive Action indicate increased academic ability, academic motivation, and positive behaviors suggesting

that this program helps youth feel good about themselves and likely increases self-esteem; further, engaging in the positive actions espoused by the program also bolsters self-esteem. This increase in self-esteem might fuel positive behaviors and extinguish negative behaviors such as aggression. A number of studies found that participation in Positive Action was associated with decreases in violence, aggression, bullying, and suspensions, indicating that program participation likely decreases aggression and improves the school atmosphere thus decreasing school hassles. Finally, although less research has focused on Positive Action's effect on internalizing symptoms, one study found decreased depression and anxiety, suggesting that the program might also be associated with improved mental health functioning.

## Methods

The current study was funded through a cooperative agreement between the Centers for Disease Control and the North Carolina Academic Center for Excellence in Youth Violence Prevention project. The study sample came from the North Carolina Academic Center for Excellence in Youth Violence Prevention project's Rural Adaptation Project, a 5-year longitudinal panel study of more than 4000 middle-school students from 27 public middle schools and 11 public high schools in two rural, economically disadvantaged counties in North Carolina. The current study used four waves of Rural Adaptation Project panel data collected between 2011 and 2014 (i.e., Years 1–4 of the Rural Adaptation Project).

## Sample

In Year 1 of the current study, all middle school students (Grades 6 through 8) in the comparison county were included in the sample. However, the intervention county was larger geographically and had a larger student population than the comparison county; therefore, a random sample of 40 % of middle-school students from the intervention county was included as the current study's intervention sample. Students from both counties were tracked longitudinally as they moved through middle school and into high school. At the beginning of each new academic year, the new cohort of sixth graders from the comparison county and a random sample of 500 sixth graders from the intervention county were added to the sample.

### *Analytic Sample and Sample Size*

The analytic sample varied by the propensity score method used to account for the number of participants who

provided non-missing data on the outcome variables at a specific wave. In general, the sample size for the analysis using the inverse probability of treatment weighting ranged from 3715 to 5894, while the sample size for the propensity score matching ranged from 1246 to 1968 participants. Using the baseline data and the imputed files for the inverse probability of treatment weighting analysis, the study found that 27 % of participants identified as White, 23 % as African American, 12 % as mixed race/other, 8 % as Latino, and 30 % as American Indian. About half of the sample (52 %) was female, 88 % of participants received free or reduced price lunch, and 91.52 % of participants lived in a two-parent family. The mean age of the sample was 12.78 years.

### *Intervention and Comparison Counties*

The intervention county is one of the most ethnically/racially diverse rural counties in the nation and, according to the United States Census Bureau (2015) in 2013, 39.5 % of residents identified as American Indian, 27 % as Caucasian, 25 % as African American, 8 % as Hispanic/Latino, and 0.8 % as other. Although slightly less diverse, the comparison county also has a mix of races/ethnicities: 61 % Caucasian, 31 % African American, 5 % Hispanic/Latino, 4 % American Indian, and 0.5 % other. Both counties are low income and from 2009 to 2013, 32 % of residents in the intervention county and 25 % of residents in the comparison county lived below the poverty level, rates almost double that of the state of North Carolina (18 %) and the United States (15 %). Further, during the same time period, the median household income was \$29,806 in the intervention county and \$35,761 in the comparison county, compared to \$46,334 in the state of North Carolina and \$53,046 in the United States (U.S. Census Bureau 2015). In addition to being low income, both counties are quite violent. From 2009 to 2013, the homicide rate per 100,000 inhabitants was 19.3 in the intervention county and 14.3 in the comparison county, much higher than the state average of 5.8 (NC States Center for Health Statistics 2015).

### **Implementation and Fidelity Procedures**

Optimal implementation of Positive Action for middle school (i.e., Grades 6, 7, and 8) consists of brief lessons (approximately 15 min) taught two or three days a week by classroom teachers. During this project, the Positive Action program was delivered to approximately 4700 middle school children in Grades 6, 7, and 8 in 13 middle schools. To launch Positive Action in the intervention county, the North Carolina Academic Center for Excellence in Youth Violence Prevention project (NC-ACE) provided program

materials, training, and supervision for all 13 middle schools. During teacher training each fall, NC-ACE staff modeled the implementation of Positive Action lessons, showing teachers and counselors how to facilitate the program. Most schools implemented Positive Action during social studies or health classes and implemented between two and three lessons per week. NC-ACE staff sometimes assisted teachers in implementing Positive Action lessons to ensure consistency; the role of the NC-ACE staff was to ensure that teachers felt competent following the Positive Action program and to provide teachers with support as they learned how to implement the program. By November in Year 1, 65 teachers took over the program in their 13 middle schools and the NC-ACE staff monitored progress for implementation fidelity and observed teachers and completed rating forms to document that teachers had attained adequate implementation skills. Teachers documented each lesson completed using weekly Implementation Logs provided by NC-ACE staff. These logs were collected by NC-ACE staff and entered into an Excel spreadsheet that allowed NC-ACE to closely monitor progress within and across the schools. Program dosage was assessed by the number of lessons taught and by the duration of lessons. Despite typical implementation barriers encountered (schedule changes, teacher absences), all 65 teachers reached, and most exceeded, their implementation goals for Positive Action lessons, especially during implementation Years 2 and 3. Year 1 was the most difficult because of solving implementation problems (i.e., recruiting the appropriate staff, fostering trust with principals, finding incentives for teachers). During the final implementation year (Year 3) the 16 Grade 6 teachers taught 1193 Positive Action lessons, ranging from 74 to 77 lessons per teacher and exceeding their goal of 73 lessons for the year. In Year 3, the 24 Grade 7 teachers taught 1527 lessons, ranging from 63 to 65 lessons per teacher, and dramatically exceeding their goal of 51 lessons for the year. Also in Year 3, the 25 Grade 8 teachers taught 1509 lessons, ranging from 60 to 64 lessons each, and exceeding their yearly goal of 45 lessons. Given the number of teachers involved and the chaotic, impoverished schools participating in the project, this is an excellent implementation record, and this level of implementation fidelity had not been achieved in previous Positive Action studies. According to the Positive Action program designer, the typical level of implementation is approximately 25 lessons per year (C. Allred, personal communication 2013). NC-ACE staff and the intervention school teachers tripled that number. As expected, some teachers enthusiastically adopted Positive Action whereas other teachers struggled with the lessons. School-level reports from the staff implementing the program indicated that they encountered a number of implementation barriers in the schools,

including difficulties situating Positive Action in the curriculum; unannounced changes in the Positive Action plans put forth by principals or in school schedules; teacher absences; and teacher turnover or reassignment of teachers and substitutes. However, encountering such challenges is not unusual when launching a school-based intervention of this scope. In years 2 and 3, teachers implementing Positive Action received a \$50 incentive each month if lesson goals were met, dramatically increasing teacher motivation.

#### *Additional Positive Action Materials: Climate and Counselor Kits*

The Positive Action program offers supplemental materials to help schools create an environment that reinforces the philosophy and lessons of the program. As described on the program's website, climate kits contain

...items and activities to reinforce positive actions: words-of-the-week, newsletter templates, assemblies, stickers, tokens, positive notes, etc. They reinforce the good feelings that students have when they do positive actions (that is, they help develop intrinsic, rather than extrinsic, motivation to engage in positive actions; Positive Action, n.d.)

As part of the current study, North Carolina Academic Center for Excellence in Youth Violence Prevention project provided the 13 middle schools with Positive Action Climate Kits. Each school's principal selected teachers, staff, or community members to use the climate kit materials in hallways, classrooms, offices and elsewhere to highlight Positive Action program themes. In addition, counselors at each of these 13 schools received Positive Action Counselor Kits that contained "the text, *Positive Actions for Living*, used for additional counseling sessions and with individuals, small groups and families" (Positive Action, n.d.).

#### **Data Collection Procedures**

Nearly identical data collection procedures were used in the two study county sites. In accordance with school district policies, the comparison county adopted the assessment as part of normal school procedures, whereas the intervention county sent a letter home to all parents explaining the study. If parents in the intervention county did not want their child to participate, they returned a letter requesting nonparticipation, and their child was removed from the study roster. Students in both counties were advised that participation was voluntary, and they were free to decline to participate; students assented to participate by reading and electronically signing an assent screen before completing the online assessment. Assessments were

completed in school computer labs, which were closely monitored by research staff. To maintain participant confidentiality, each participant was assigned a unique identification number and assessments did not collect any other identifying information. All study procedures were approved by the Institutional Review Board of the major research university in North Carolina with which the researchers were affiliated.

## Measures

The primary measure used in the current study was a modified version of the School Success Profile (Bowen and Richman 2008). The School Success Profile is a 195-item online, youth self-report survey with 22 subscales that measure perceptions and attitudes about school, friends, family, neighborhood, self, health, and well-being. Since its creation in 1993, the School Success Profile has been administered to tens of thousands of students, and has well-documented reliability and validity (Bowen et al. 2005). The current study used a modified version of the School Success Profile, the School Success Profile Plus, which included 17 of the original School Success Profile subscales and 12 additional subscales.

### Self-Esteem

A five-item modified version of the frequently used Rosenberg (1965) self-esteem scale measured self-esteem. Example items included: “I am able to do things as well as most other people” and “I have confidence in myself.” Each item was rated on a 3-point Likert scale (*Not Like Me, A Little Like Me, A Lot Like Me*).

### Aggression

Twelve items from the externalizing subscale of the Youth Self-Report (Achenbach and Rescorla 2001) assessed aggression. Example items included: “I get in many fights” and “I break rules at home, school, or elsewhere.” Each item was rated on a 3-point Likert scale (*Not Like Me, A Little Like Me, A Lot Like Me*).

### Internalizing Symptoms

Symptoms of depression and anxiety were assessed using seven items from the internalizing subscale from the Youth Self-Report (Achenbach and Rescorla 2001). Example items included: “I often feel nervous or tense” and “I often feel fearful or anxious.” Each item was rated on a 3-point Likert scale (*Not Like Me, A Little Like Me, A Lot Like Me*).

### School Hassles

The amount of verbal and physical harassment youth endured at school was assessed with the 13-item school hassles scale (Bowen and Richman 2008). Example items included: “Someone treated you in a disrespectful way” and “Someone at school pushed, shoved, or hit you.” Each item was rated on a 3-point Likert scale (*Never, Once or Twice, More than Twice*).

### Friend Rejection

The degree to which youth were rejected by their friends was assessed with the three-item friend rejection scale (Bowen and Richman 2008). Example items included: “I am made fun of by my friends” and “I wish my friends would show me more respect.” Each item was rated on a 3-point Likert scale (*Not Like Me, A Little Like Me, A Lot Like Me*).

### Parent–Child Conflict

The amount of conflict present in the parent–child relationship was assessed with 10 items from the Conflict Behavior Questionnaire (Prinz et al. 1979). Following the prompt “at least three times per week” example items included: “My parent(s) and I get angry at each other” and “My parent(s) put me down.” Each item was rated as *true* or *false*.

### Religious Orientation

The importance of religion in participants’ lives was assessed with the three-item religious orientation scale (Bowen and Richman 2008). Example items included: “My religious faith gives me strength” and “My religious faith influences the decisions I make.” Each item was rated on a 3-point Likert scale (*Not Like Me, A Little Like Me, or A Lot Like Me*).

### School Satisfaction

Participants’ positive experiences at school were assessed with the seven-item school satisfaction scale (Bowen and Richman 2008). Example items included: “I enjoy going to this school” and “I get along well with teacher at this school.” Each item was rated on a 3-point Likert scale (*Not Like Me, A Little Like Me, or A Lot Like Me*).

### Future Optimism

Participants’ expectations for future success were measured with the 12-item Future Optimism scale (Bowen and

Richman 2008). Example items included: “When I think about my future, I feel very positive” and “I see myself accomplishing great things in life.” Each item was rated on a 4-point Likert scale (*Strongly Disagree, Disagree, Agree, and Strongly Agree*).

#### *Parent Support*

The five-item Parent Support scale (Bowen and Richman 2008) measured the degree to which an adult caregiver in the participants’ home provided emotional support in the past 30 days. Example items included “How often did the adults in your home let you know that you were loved?” and “How often did the adults in your home tell you that you did a good job?” Each item was rated on a 3-point Likert Scale (*Never, Once or Twice, or More than Twice*).

#### *Teacher Support*

The eight-item Teacher Support scale (Bowen and Richman 2008) assessed participants’ perceptions of their teachers’ supportive behavior. Example items included: “My teachers care about me” and “My teachers give me a lot of encouragement.” Each item was rated on a 4-point Likert scale (*Strongly Disagree, Disagree, Agree, or Strongly Agree*).

#### *Friend Support*

The five-item Friend Support scale (Bowen and Richman 2008) measured participants’ perceptions of how supportive their friends are. Example items included: “I can count on my friends for support” and “I can trust my friends.” Each item was rated on a 3-point Likert scale (*Not Like Me, A Little Like Me, or A Lot Like Me*).

#### *Delinquent Friends*

The nine-item Delinquent Friends subscale (Bowen and Richman 2008) measured the degree to which participants’ friends engaged in delinquent activities. Example items included: “I have friends who get in trouble with the police” and “I have friends who cut classes.” Each item was rated on a 3-point Likert Scale (*Not Like Me, A Little Like Me, or A Lot Like Me*).

#### *Peer Pressure*

The degree to which participants felt their friends negatively pressured them was assessed with a five-item scale (Bowen and Richman 2008). Example items included: “I let my friends talk me into doing things I really don’t want to do” and “I tend to go along with the crowd.” Each item

was rated on a 3-point Likert Scale (*Not Like Me, A Little Like Me, or A Lot Like Me*).

#### *Perceived Discrimination*

The amount of racial discrimination youth endured and witnessed at school was assessed with the three-item perceived discrimination scale (Gil et al. 1994). Example items included: “How often do people dislike you because of your race?” and “How often have you seen friends treated unfairly because of their race or ethnicity?” Each item was rated on a 4-Likert scale (*Never, Sometimes, Frequently, Always*).

#### *School Danger*

The 11-item school danger scale (Bowen and Richman 2008) assessed the frequency of dangerous behaviors at school. Following the prompt “How often does each of the following happen at your school,” example items include: “Fights among students” and “Students carrying weapons.” Each item was rated on a 3-point Likert Scale (*Does Not Happen, Happens Sometimes, Happens A Lot*). See Table 1 for a description of measures, and Table 2 for sample statistics by matched and weighted samples.

### **Analytic Plan**

The current study used a quasi-experimental design with an intervention and comparison county that had similar demographics. Implementing the study as a randomized controlled trial was not feasible, and therefore, the evaluation encountered a fundamental challenge: determining to what extent the differences in youth’s changes on key outcome variables could be attributed to the Positive Action intervention. Further, as well as controlling for numerous selection biases, the data analysis needed to address key methodological issues such as the violation of a normality assumption embedded in the linear model, clustering effects inevitably existing in the study of change trajectories, and missing data imputation. The analytic protocol addressed these methodological challenges.

### **Correcting for Selectivity and Two Propensity Score Models**

Because receiving the Positive Action intervention was not random (i.e., only youth residing in the intervention county were eligible to receive the intervention) and involved many factors affecting selectivity, the samples from the two counties were imbalanced on covariates (a report on an imbalance check of the baseline data between the intervention and control counties is available from the first



**Table 1** Description of measures

Measure	Type	HLM level	Alpha (Year 1; 2; 3; 4)	Response options	Example items
Self-esteem (5 items)	Dependent	1 (Time)	.87; .91; .92; .94	Not like me A little like me A lot like me	I am able to do things as well as most other people I have confidence in myself
Aggression (12 items)	Dependent; time varying	1 (Time)	.86; .87; .86; .90	Not like me A little like me A lot like me	I get in many fights I break rules at home, school, or elsewhere
Internalizing (7 items)	Dependent; time varying	1 (Time)	.89; .90; .91; .95	Not like me A little like me A lot like me	I often feel nervous or tense I often feel fearful or anxious
School hassles (13 items)	Dependent; time varying	1 (Time)	.90; .92; .92; .93	Never Once or twice More than twice	Someone treated you in a disrespectful way Someone at school pushed, shoved, or hit you
Friend rejection (3 items)	Time varying	1 (Time)	.70; .74; .75; .80	Not like me A little like me A lot like me	I am made fun of by my friends I wish my friends would show me more respect
Parent–child conflict (10 items)	Time varying	1 (Time)	.82; .83; .84; .85	True False	At least 3 times a week, my parent(s) and I get angry at each other. My parent(s) put me down
Religious orientation (3 items)	Independent	2 (Student)	.88; .91; .92; .93	Not like me A little like me A lot like me	My religious faith gives me strength My religious faith influences the decisions I make
School satisfaction (7 items)	Independent	2 (Student)	.84; .85; .87; .88	Not like me A little like me A lot like me	I enjoy going to this school I get along well with teachers at this school
Future optimism (12 items)	Independent	2 (Student)	.88; .90; .92; .97	Strongly disagree Disagree Agree	When I think about my future, I feel very positive I see myself accomplishing great things in life
Parent support (5 items)	Independent	2 (Student)	.89; .92; .93; .94	Strongly agree Never Once or twice More than twice	How often did the adults in your home let you know that you were loved? How often did the adults in your home tell you that you did a good job?
Teacher support (8 items)	Independent	2 (Student)	.88; .90; .92; .92	Strongly disagree Disagree Agree Strongly agree	My teachers care about me My teachers give me a lot of encouragement

Table 1 continued

Measure	Type	HLM level	Alpha (Year 1; 2; 3; 4)	Response options	Example items
Friend support (5 items)	Independent	2 (Student)	.89; .91; .92; .94	Not like me A little like me A lot like me	I can count on my friends for support I can trust my friends
Delinquent friends (9 items)	Independent	2 (Student)	.90; .91; .91; .92	Not like me A little like me A lot like me	I have friends who get in trouble with the police I have friends who cut classes.
Peer pressure (5 items)	Independent	2 (Student)	.73; .77; .79; .83	Not like me A little like me A lot like me	I let my friends talk me into doing things I really don't want to do I tend to go along with the crowd
Perceived discrimination (3 items)	Independent	2 (Student)	.71; .75; .76; .76	Never Sometimes Frequently Always	How often do people dislike you because of your race or ethnicity? How often have you seen friends treated unfairly because of their race or ethnicity?
School danger (11 items)	Independent	2 (Student)	.85; .87; .89; .91	Does not happen Happens Sometimes Happens a lot	How often does each of the following happen at your school? Examples items included: Fights among students and Students carrying weapons.
Symptoms of depression (4 items)	Independent	2 (Student)	.86; .84; .88; .90	Not like me A little like me A lot like me	I often feel sad I often feel alone
Symptoms of anxiety (3 items)	Independent	2 (Student)	.76; .79; .81; .85	Not like me A little like me A lot like me	I often feel nervous or tense I often feel fearful or anxious

**Table 2** Baseline sample descriptive statistics based on 15 imputed files sample statistics for matched sample and weighted sample

Variables	Matched sample		Weighted sample	
	% or <i>M</i>	<i>SE</i>	% or <i>M</i>	<i>SE</i>
Aggression (time-varying)	1.31	0.01	1.33	0.006
Internalizing behavior (time-varying)	1.42	0.013	1.43	0.008
School hassles (time-varying)	1.49	0.012	1.49	0.008
Parent–child conflict (time-varying)	1.84	0.06	2.0	0.041
Friend rejection (time-varying)	1.3	0.012	1.29	0.008
Target county	0.5	0.012	0.66	0.008
African American	0.35	0.012	0.23	0.007
Hispanic	0.09	0.007	0.08	0.004
Native American	0.08	0.007	0.3	0.008
Mixed race and other	0.11	0.008	0.12	0.005
Gender female (reference male)	0.52	0.012	0.52	0.008
Age at baseline	12.79	0.026	12.77	0.017
Receipt of free/reduced lunch (reference No)	0.89	0.008	0.88	0.006
Religious orientation	2.34	0.014	2.31	0.009
School satisfaction	2.38	0.012	2.38	0.008
Future optimism	3.46	0.012	3.47	0.008
Parent support	2.7	0.012	2.68	0.008
Teacher support	3.19	0.014	3.17	0.009
Friend support	2.47	0.014	2.49	0.009
Delinquent friends	1.35	0.011	1.38	0.007
Peer pressure	1.3	0.01	1.31	0.007
Perceived discrimination	1.42	0.014	1.43	0.009
School danger	1.74	0.009	1.79	0.007
Anxiety	1.46	0.015	1.47	0.009
Depression	1.43	0.015	1.44	0.009

SEs were estimated by aggregating 15 imputed files using Rubin’s rule

author upon request). Due to these selection issues, running a covariance control, such as a linear regression model or a hierarchical linear model, would inevitably encounter the statistical problem known as endogeneity and results from such models are deemed biased and inefficient (Guo and Fraser 2015; Imbens 2004). To correct for endogeneity, we applied the Neyman-Rubin counterfactual framework as a conceptual model to guide the data analysis.

To analyze the data, we used two propensity score models. A propensity score is a conditional probability of a participant receiving treatment, given observed covariates (Rosenbaum and Rubin 1983). The samples from the two counties were imbalanced on 20 covariates at baseline, including six demographic variables (i.e. *African American, Hispanic, Native American, White, Mixed race and other, and receipt of free/reduced lunch*), and 14 other variables at individual level (i.e., *religious orientation, school satisfaction, future optimism, parent support, teacher support, friend support, delinquent friends, peer pressure, perceived discrimination, school danger, friend rejection, parent–child conflict, anxiety, and depression*) at

baseline. Thus, a binary logistic regression used four social demographic variables (i.e., *gender, race, age, receipt of free or reduced lunch*) and the imbalanced variables at baseline to estimate propensity scores for the probability of receiving the Positive Action treatment. This procedure enabled us to control overt selectivity, and made the two counties balanced on observed covariates. We then used the estimated propensity scores to conduct a growth curve analysis in conjunction with the two propensity score models.

#### *Inverse Probability of Treatment Weighting*

The first propensity score model used inverse probability of treatment weighting. Imbens and Wooldridge (2009) showed that this estimator is equivalent to a sample average from a random sample, consistent for population average treatment effect, and is  $\sqrt{N}$  asymptotically normally distributed. Because the Horvitz-Thompson (Horvitz and Thompson 1952) estimator is based on sample averages and widely employed in weighted analysis to adjust

for unequal probabilities employed in stratified sampling, applying the adjustments to propensity score weighting is analogous to weighted analysis using unequal sampling weights.

The above estimator is used for estimating the average treatment effect. Researchers (Imbens and Wooldridge 2009) have also developed a similar estimator to evaluate the average treatment effect for the treated participants. The average treatment effect for the treated depicts treatment effect for the targeted subpopulation, and is more important than average treatment effect in program evaluation (Guo and Fraser 2015). Denoting the estimated propensity score for the  $i$ th observation as  $\hat{e}(x_i)$ , the propensity score weight for estimating average treatment effect is  $1/\hat{e}(x_i)$  for a treated participant, and is  $[1/(1 - \hat{e}(x_i))]$  for a comparison participant; the propensity score weight for estimating the average treatment effect of the treated participants is 1 for a treated participant, and  $[\hat{e}(x_i)/(1 - \hat{e}(x_i))]$  for a comparison participant. After creating these weights, we ran a growth curve model using one of the two weights to estimate the average treatment effect and the average treatment effect of the treated participants.

#### *The 1-to-1 Nearest-Neighbor Within Caliper Matching*

In practice, researchers often perform several propensity score models to ensure robust findings. Constructing multiple models is a prudent step because the various propensity score models are based on different assumptions about data and it is unknown whether those assumptions are valid in empirical analysis (Guo and Fraser 2014). Such a robust check or sensitivity analysis is important, and particularly important for evaluations using quasi-experimental data. To carry out this task, we used 1-to-1 nearest-neighbor within caliper matching as a second propensity score model in the analysis.

The 1-to-1 nearest-neighbor within caliper matching can be expressed by the following equation:

$$\min_j \|\hat{e}(x_i) - \hat{e}(x_j)\| < \varepsilon$$

where  $\hat{e}(x_i)$  and  $\hat{e}(x_j)$  are the sample estimated propensity scores for the treated participant  $i$  and untreated participant  $j$ , respectively; for each treated participant  $i$ , the algorithm chooses untreated participant  $j$  as a match of  $i$ , if, and only if, the difference on the estimated propensity scores is the minimum among all possible pairs between this participant  $i$  and each of the untreated participants in the potential pool within the predetermined caliper of size  $\varepsilon$  (i.e.,  $j$  is the nearest neighbor of  $i$  within  $\varepsilon$ ). Rosenbaum and Rubin (1985) suggested using a caliper size of a quarter of a standard deviation of the sample estimated propensity

scores (i.e.,  $\varepsilon \leq .25\sigma_P$ , where  $\sigma_P$  denotes standard deviation of the estimated propensity scores of the sample). After matching using the above procedure, we ran a growth curve model based on the matched sample to discern the average treatment effect.

#### **Growth Curve Analysis**

The evaluation aimed to analyze students' change on an outcome variable over a 3-year study period. Therefore, the analytic sample included only those students who provided data for at least two data waves. Students who had data on only one wave were excluded because those data do not meet the definition of outcome change.

The dataset in the current analysis has a typical nesting structure because the data collection points (i.e., four waves) are nested within students, and students are nested within schools. To correct for the clustering effects naturally formed by the data collection, and hence, to address the violation of independent-observation assumption embedded in a linear regression model, we applied a two-level hierarchical linear modeling to the data analysis. The two-level hierarchical linear model is shown by the following equation (i.e., the combined equation):

$$\ln(Y_{it}) = \gamma_{00} + \gamma_{10}(\text{Time})_{it} + \sum_{p=2}^P \gamma_{p0}(TV)_{pti} + \sum_{q=1}^Q \gamma_{0q}(X)_{qi} + r_{0i} + e_{it}$$

where  $\ln(Y_{it})$  is the outcome variable of interest,  $(\text{Time})_{it}$  is the time variable measured in months from baseline or Wave 1,  $(TV)_{pti}$  are  $P - 1$  time-varying variables,  $(X)_{qi}$  are  $Q$  student-level variables,  $r_{0i}$  is a random effect for the  $i$ th student, and  $e_{it}$  is a residual term incorporating temporal random effect for the  $i$ th student at time  $t$ . The data do not show nontrivial clustering of students within a school (i.e., the intraclass correlation coefficient on most outcomes is below .03), and therefore, we did not treat school as a study level. The specification of the above model was based on prior studies using Rural Adaptation Project data to evaluate students' change on key outcomes over time (e.g., Smokowski et al. 2014; Smokowski et al., in press). The current study is distinct from previous evaluations in that our research focused specifically on the evaluation of the Positive Action effects, whereas prior studies had a different aim.

Four specifications of the growth curve model are worth noting: first, all four outcome variables of interest had a skewed distribution and did not meet the normality assumption embedded in hierarchical linear modeling about the outcome variable. Therefore, we followed the

convention in econometrics to take a natural-logarithm transformation of the dependent variable (Greene 2003).

Second, the analysis specified only a linear time variable, which is typical for growth curve analysis using 4- or 5-point panel data where a quadratic or other type of curvilinear model adds unnecessary complexity (Raudenbush and Bryk 2002). A special feature, and perhaps an advantage of the current analytic model, is the use of several time-varying variables [i.e.,  $(TV)_{piti}$  variables] in Level 1. The inclusion of these time-varying variables investigates the relationship between a predictor and the outcome variable from a dynamic point of view, and therefore, makes best use of the rich information offered by the longitudinal data. Based on examination of the dynamic influences of predictors on outcome change, we chose five or four time-varying covariates, depending on the outcome variable. Supported by the literature, the four or five time-varying covariates are the most important predictors of outcome change; our selected time-varying covariates were school hassles scale, aggressive behavior (when not the dependent variable), internalizing behavior (when not the dependent variable), parent–child conflict scale, and friend rejection (Smokowski et al. 2014; Smokowski et al., in press).

Third, we chose  $Q = 19$ , that is, used 19 predictor variables at Level 2. All 19 variables were measured at the study entry. Please see measure descriptions above.

Finally, because the dependent variable in our final model takes a natural-logarithm transformation, we used the exponent of estimated coefficient  $\exp(\beta)$  to ease the burden of interpreting the findings, where  $\beta$  is the estimated coefficient associated with the independent variable  $X$ . As such, we can interpret the finding of the difference between  $X = 1$  and  $X = 0$  on the outcome  $Y$  as: the group of  $X = 1$  on average has an outcome that is  $[100 - 100 \times \exp(\beta)]\%$  lower than the outcome of the group of  $X = 0$  when  $\exp(\beta) < 1$ , and the group of  $X = 1$  on average has an outcome that is  $[100 \times \exp(\beta) - 100]\%$  higher than the outcome of the group of  $X = 0$  when  $\exp(\beta) > 1$ . The  $X$  variable of primary interest is the intervention county: the percentage difference on this variable shows the direction of Positive Action impact, and test of its statistical significance generalizes the effect to the targeted population.

Because this study is a program evaluation that examines whether the Positive Action intervention has beneficial impacts for the participants, we performed directional hypothesis tests of treatment effects based on our hypotheses. For the four different outcome variables, we had a hypothesized sign for each coefficient for a one-tailed test. The hypothesized sign for “Aggression”, “Internalizing” or “School Hassles” is “–”, indicating hypotheses about a negative sign of the coefficient. Specifically, we

hypothesized that students who received the Positive Action intervention would show lower aggression scores, lower internalizing scores, and lower school hassles scores, compared with students who did not receive the intervention. The hypothesized sign for “Self-esteem” is “+”, indicating that we hypothesized that students who received Positive Action would show higher self-esteem scores relative to students who did not engage in the program. For the demographic variables and other covariates, we performed non-directional hypothesis tests with a given level of statistical significance.

### Multiple Imputation and the Aggregated Analysis Applying Rubin’s Rule

Missing data is a typical problem encountered in datasets used to evaluate change in youth’s outcomes over time. To address the problem of missing data, we followed guidelines suggested by Allison (2002) and conducted multiple imputation. We created 15 imputed files. The final results were aggregated from these 15 multiply imputed files using Rubin’s (1987) rule. Few existing studies have shown the analytic procedure that combines the multiple imputation and growth curve modeling with propensity score correction into one procedure. Ultimately, the current investigation sought to incorporate three statistical models into one: the multiple imputation of missing data, the hierarchical linear modeling-type growth curve analysis, and propensity score modeling using either inverse probability of treatment weighting or matching.

Based on a careful study of existing literature and numerous experimental runs, we developed an analytic protocol and applied the protocol to each outcome variable. For the inverse probability of treatment weighting, we first estimated propensity scores for each of the 15 imputed files, then conducted the inverse probability of treatment weighting-growth-curve analysis using the estimated propensity scores for each of the 15 imputed files, and finally used Rubin’s (1987) rule to aggregate the 15 sets of estimated results. For matching, we estimated the propensity scores for each of the 15 imputed files, and then proceeded with the 1-to-1 nearest-neighbor within caliper matching to create a matched sample for each of the 15 imputed files. Next, growth curve analysis was conducted for each of the 15 matched samples. Last, we used Rubin’s rule to aggregate the 15 sets of estimated results into one set.

## Results

Findings indicate that after controlling for selection biases, the Positive Action program as implemented in the intervention county was associated with statistically significant

**Table 3** Estimated HLM coefficients of the change of self-esteem score using three PSA methods

Fixed and random effects	IPTW–ATE	IPTW–ATT	Matching
Fixed effect			
Level 1: time			
Time (months since baseline)	–0.002***	–0.003***	–0.002***
Aggression (time-varying)	0.003	0.010	0.009
Internalizing behavior (time-varying)	–0.093***	–0.090***	–0.081***
School hassles (time-varying)	–0.059*	–0.074*	–0.013
Friend rejection (time-varying)	–0.020	–0.020	–0.037***
Parent–child conflict (time-varying)	–0.006	–0.006	–0.013***
Level 2: individual			
Intervention county (comparison)	0.018*	0.016*	0.021**
Treatment effect: exponent of coefficient	1.018*	1.016*	1.021**
Gender (male)			
Female	–0.029**	–0.029*	–0.018*
Race (White)			
African American	0.062***	0.063***	0.062***
Hispanic	0.050***	0.053***	0.054***
Native American	0.023*	0.025*	0.037*
Mixed race and other	0.057***	0.059***	0.045***
Age at baseline	–0.007*	–0.008*	–0.007*
Receipt of free/reduced lunch at baseline (No)			
Yes	0.014	0.012	0.015
Religious orientation at baseline	0.065***	0.066***	0.064***
School satisfaction at baseline	0.024*	0.016	0.038***
Future optimism at baseline	0.052***	0.059***	0.042***
Parent support at baseline	0.036*	0.031*	0.036***
Teacher support at baseline	0.008	0.013	–0.004
Friend support at baseline	0.027*	0.027	0.016*
Delinquent friends at baseline	0.003	0.000	0.012
Peer pressure at baseline	–0.014	–0.022	0.002
Discrimination at baseline	0.010	0.018	–0.012
School danger at baseline	0.018	0.018	0.024*
Anxiety at baseline	0.018	0.024	0.004
Depression at baseline	0.006	0.004	0.006
Intercept	0.652***	0.668***	0.645***
Random effect (variance component)			
Level 2 intercept	0.010***	0.010***	0.009***
Level 1 residual	0.033***	0.034***	0.030***

Reference group for categorical variables is shown in parenthesis after variable name

One-tailed for “Intervention County”, two-tailed for all others

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$

beneficial effects for youth in self-esteem and school hassles scores. The Positive Action program was also related to beneficial effects for the intervention youth on the change of aggression score, but the finding did not reach statistical significance. The change related to internalizing score showed that Positive Action program exposure had a detrimental effect from the two-tailed tests: as compared with youth in the intervention county, the youth from the

comparison county had lower internalizing scores. However, the detrimental effect was not statistically significant from the one-tailed tests conducted in our analysis.

Specifically, the results (see Table 3) show that, other things being equal and controlling for selection biases, at any point in time during the study period intervention youth had a self-esteem score 1.8 % higher than that of the control youth ( $p < .05$ , estimated by the inverse probability

**Table 4** Estimated HLM coefficients of the change of school hassle score using three PSA methods

Fixed and random effects	IPTW–ATE	IPTW–ATT	Matching
Fixed effect			
Level 1: time			
Time (months since baseline)	−0.004***	−0.004***	−0.003***
Aggression (time-varying)	0.094***	0.089***	0.101***
Internalizing behavior (time-varying)	0.120***	0.126***	0.084***
Friend rejection (time-varying)	0.092*	0.092	0.133***
Parent–child conflict (time-varying)	0.005	0.005	0.007***
Level 2: individual			
Intervention county (comparison)	−0.040***	−0.047***	−0.022**
Treatment effect: exponent of coefficient	0.961***	0.954***	0.978**
Gender (male)			
Female	−0.027*	−0.030	−0.025**
Race (White)			
African American	−0.035***	−0.035**	−0.031**
Hispanic	−0.065***	−0.068***	−0.063***
Native American	−0.009	−0.010	−0.012
Mixed race and other	−0.038**	−0.039**	−0.031
Age at baseline	−0.023***	−0.024***	−0.020***
Receipt of free/reduced lunch baseline (No)			
Yes	0.008	0.006	0.017
Religious orientation at baseline	0.010	0.009	0.016
School satisfaction at baseline	−0.022	−0.020	−0.030**
Future optimism at baseline	0.034	0.037	0.020*
Parent support at baseline	−0.001	−0.004	0.009
Teacher support at baseline	0.001	0.000	0.005
Friend support at baseline	−0.028*	−0.034	−0.019*
Delinquent friends at baseline	0.014	0.021	−0.016
Peer pressure at baseline	0.004	0.006	0.000
Discrimination at baseline	0.064***	0.065***	0.063***
School danger at baseline	0.108***	0.098***	0.113***
Anxiety at baseline	0.000	−0.003	0.013
Depression at baseline	0.015	0.014	0.004
Intercept	−0.059	−0.014	−0.130
Random effect (variance component)			
Level 2 intercept	0.018***	0.019***	0.015***
Level 1 residual	0.037***	0.038***	0.036***

Reference group for categorical variables is shown in parenthesis after variable name

One-tailed for “Intervention County”, two-tailed for all others

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$

of treatment weighting-average treatment effect), 1.6 % higher ( $p < .05$ , estimated by the inverse probability of treatment weighting-average treatment effect of the treated), or 2.1 % higher ( $p < .01$ , estimated by matching). Similarly, as shown in Table 4, with other things being equal and controlling for selection biases, intervention youth at any point in time during the study period had a school hassles score 3.9 % lower than that of the control youth ( $p < .001$ , estimated by the inverse probability of

treatment weighting-average treatment effect), 4.6 % lower ( $p < .001$ , estimated by inverse probability of treatment weighting-average treatment effect of the treated), or 2.2 % lower ( $p < .01$ , estimated by matching). In addition, the results (Table 5), indicate that, at any point in time during the study period, intervention youth had an aggression score that was 0.7 % lower than that of the control youth (not significant, estimated by the inverse probability of treatment weighting-average treatment effect), 0.6 % lower

**Table 5** Estimated HLM coefficients of the change of aggression score using three PSA methods

Fixed and random effects	IPTW-ATE	IPTW-ATT	Matching
Fixed effect			
Level 1: time			
Time (months since baseline)	0.000	0.000	0.000
Internalizing behavior (time-varying)	0.223***	0.223***	0.214***
School hassles (time-varying)	0.065***	0.063***	0.069***
Parent-child conflict (time-varying)	0.007	0.007	0.009***
Friend rejection (time-varying)	-0.002	-0.001	0.021
Level 2: individual			
Intervention County (Comparison)	-0.007	-0.006	-0.007
Treatment effect: Exponent of Coefficient	0.993	0.994	0.993
Gender (male)			
Female	0.019*	0.019	0.023**
Race (White)			
African American	0.019*	0.021*	0.015
Hispanic	-0.024*	-0.023*	-0.027*
Native American	0.011	0.010	0.016
Mixed race and other	0.010	0.010	0.009
Age at baseline	0.005	0.005	0.005
Receipt of Free/reduced lunch at baseline (No)			
Yes	0.000	0.001	-0.002
Religious orientation at baseline	-0.035***	-0.034***	-0.033***
School satisfaction at baseline	-0.052***	-0.052***	-0.055***
Future optimism at baseline	-0.014	-0.022	0.002
Parent support at baseline	0.019*	0.021	0.013
Teacher support at baseline	0.002	0.002	0.002
Friend support at baseline	0.014	0.013	0.024**
Delinquent friends at baseline	0.084***	0.077***	0.098***
Peer pressure at baseline	0.022	0.024	0.010
Discrimination at baseline	-0.004	-0.005	0.000
School danger at baseline	0.006	0.002	0.000
Anxiety at baseline	-0.015	-0.012	-0.017
Depression at baseline	-0.007	-0.006	0.004
Intercept	-0.212*	-0.174	-0.285***
Random effect (variance component)			
Level 2 intercept	0.011***	0.011***	0.011***
Level 1 residual	0.026***	0.028***	0.023***

Reference group for categorical variables is shown in parenthesis after variable name

One-tailed for "Intervention County", two-tailed for all others

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$

(not significant, estimated by inverse probability of treatment weighting-average treatment effect of the treated), or 0.7 % lower (not significant, estimated by matching). Finally, with regard to the change of the internalizing score (Table 6), findings show that, at any point in time during the study period, intervention youth had an internalizing score that was 2.6 % higher than that of the control youth

(estimated by the inverse probability of treatment weighting-average treatment effect), 3.0 % higher (estimated by inverse probability of treatment weighting-average treatment effect of the treated), or 1.8 % (estimated by matching). However, none of these effects were statistically significant in the one-tailed tests, though they were significant in the two-tailed tests.



**Table 6** Estimated HLM coefficients of the change of internalizing score using three PSA methods

Fixed and random effects	IPTW–ATE	IPTW–ATT	Matching
Fixed effect			
Level 1: time			
Time (months since baseline)	0.000	0.000	0.000*
Aggression (time-varying)	0.304***	0.295***	0.306***
School hassles (time-varying)	0.115***	0.120***	0.083***
Friend rejection (time-varying)	0.021	0.021	0.047***
Parent–child conflict (time-varying)	0.011	0.012	0.026***
Level 2: individual			
Intervention county (comparison)	0.025	0.029	0.018
Treatment effect: exponent of coefficient	1.026	1.030	1.018
Gender (male)			
Female	0.068***	0.074***	0.048***
Race (White)			
African American	0.008	0.007	0.008
Hispanic	0.035**	0.031*	0.045***
Native American	–0.005	–0.008	–0.008
Mixed race and other	0.020	0.018	0.017
Age at baseline	0.003	0.003	0.003
Receipt of Free/reduced lunch at baseline (No)			
Yes	0.010	0.011	0.005
Religious orientation at baseline	–0.009	–0.006	–0.002
School satisfaction at baseline	0.011	0.007	0.019
Future optimism at baseline	0.000	0.001	–0.009
Parent support at baseline	–0.005	–0.010	0.019*
Teacher support at baseline	0.018	0.021*	0.012
Friend support at baseline	–0.008	–0.007	–0.006
Delinquent friends at baseline	–0.022	–0.017	–0.033**
Peer pressure at baseline	–0.014	–0.012	–0.022*
Discrimination at baseline	0.000	0.004	–0.010
School danger at baseline	–0.045**	–0.053**	–0.027*
Anxiety at baseline	0.089***	0.084***	0.095***
Depression at baseline	0.128***	0.121***	0.129***
Intercept	–0.640***	–0.624***	–0.698***
Random effect (variance component)			
Level 2 intercept	0.006***	0.006***	0.006***
Level 1 residual	0.040***	0.042***	0.037***

Reference group for categorical variables is shown in parenthesis after variable name  
 One-tailed for “Intervention County”, two-tailed for all others  
 \*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$

**Discussion**

Existing research on Positive Action has largely neglected to examine program effects in rural areas. The current study aimed to fill this gap in the literature by assessing the impact of participation in Positive Action over a 3 year study window on rates of self-esteem, school hassles, aggression, and internalizing symptoms in a racially diverse sample of rural youth. Based on self-esteem

enhancement theory, the theory of organizational culture, and past Positive Action research, it was hypothesized that middle school youth who participated in Positive Action would report increased self-esteem and decreased school hassles, aggression, and internalizing symptoms relative to youth who did not engage in the program. Our hypothesis was partially supported and participation in the Positive Action program was associated with statistically significant beneficial effects in self-esteem and school hassles for

youth from the intervention county. Also in line with our hypothesis, the Positive Action program generated beneficial effects for intervention youth on the change of aggression scores; however, this finding was not statistically significant. In addition, the Positive Action program showed a detrimental effect on youth's internalizing scores, but this effect was not statistically significant in the one-tailed tests.

The current study's examination of the impact of Positive Action on self-esteem is an important contribution to the literature given that no studies have used rigorous methodology to assess the impact that this program has on self-esteem. Specifically, youth who participated in Positive Action had a self-esteem score that ranged from 1.6 to 2.1 % higher than the self-esteem scores of youth in the comparison county; the range of this score depends on the type of propensity score analysis used. Given the lack of research on Positive Action and changes in self-esteem, it is not possible to compare our results to past findings. In addition to the positive effects of the program on outcomes reported in previous studies (e.g., academic motivation, school attendance, problem behaviors; Bavarian et al. 2013; Beets et al. 2009), our finding indicates that Positive Action functioned to increase the self-esteem of participating youth. Positive Action focuses on both developing youth's self-concept, and helping youth to better understand how their actions can have an impact on their lives and others. The program also aims to increase youth's positive, prosocial behaviors. Behaving in a positive way, such as a youth helping friends, teachers, and family members, or a youth taking good care of her- or himself physically and emotionally, likely fosters the self-esteem of that youth. This finding is particularly important given the rural setting where this study took place and the various risk factors in that environment (e.g., high unemployment, high crime and violence). High self-esteem serves as a protective factor against depression (Derdikman-Eiron et al. 2011), delinquency (Donnellan et al. 2005), and bullying victimization (Huitsing et al. 2012). Thus, it is possible that youth who participate in the Positive Action program are less likely to experience low self-esteem and its consequences. Future research using structural equation modeling could explore the extent to which self-esteem plays a mediating role in the influence of Positive Action participation on depression, delinquency, and bullying victimization.

This study also found that participation in the Positive Action program resulted in a significant decrease in school hassles. Specifically, youth who participated in Positive Action had a school hassles score that was 2.2–4.6 % lower than the school hassles scores of youth from comparison schools; again, the results depended upon the type of propensity score analysis used. In the current study, school

hassles were measured by assessing the extent to which youth had experienced adverse treatment in school such as being insulted, excluded, or treated unfairly. The significant decrease in school hassles experienced by youth in the Positive Action schools suggests that the program helped to decrease negative social interactions between students and improve the overall culture of the school, which likely improved the school experiences of the intervention youth. By replacing negative interactions with positive behaviors, the Positive Action program decreased youth's perceptions of school hassles and likely improved the school culture and climate. While other studies of Positive Action have not specifically examined how the program impacts perceptions of school hassles, past studies have indicated that behaviors representing school hassles have decreased dramatically in Positive Action schools. For example, students in urban Positive Action schools reported 37 % less violent behaviors, 41 % fewer bullying behaviors, and a 27 % reduction in disruptive behaviors relative to youth who did not engage in Positive Action (Li et al. 2011). The current decrease in school hassles is much smaller than these previous findings, perhaps reflecting the difficulty of implementing Positive Action in a rural school system.

The finding that participation in Positive Action resulted in decreased aggression is consistent with prior research findings that participation in the program decreased violence (e.g., Lewis et al. 2013b; Snyder et al. 2013) and in line with our hypothesis; however, in our study, the decrease in aggression was not significant. Although aggression and violence are closely related, our study measures assessed a type of aggression consisting of mild forms of violence such as getting in fights, being mean, and arguing, whereas the violent behavior assessed in past studies of Positive Action were more severe forms of violence such as carrying a knife, threatening someone, or stabbing someone. It is possible that the Positive Action program effectively targets and decreases more severe forms of violence, but does not affect aggressive behavior. Alternately, small changes in severe violence measures from prior studies may appear to render large results due to low base rates. Indeed, the decrease in aggression found in the current study is much lower than the changes in violence seen in other studies. For example, a 3-year study of Positive Action in Nevada found that violence decreased in Positive Action schools 87 % as compared to control schools (Flay et al. 2001), a study in rural Hawaii found that the odds of students behaving violently decreased about 58 % in Positive Action schools compared to control schools (Beets et al. 2009), and a second study in Hawaii found that participation in Positive Action was associated with a 75.6 % decrease in violent behavior (Snyder et al. 2013). These large percentages may result from small changes in behaviors that are rare. Alternately, the low-

income rural setting in which the current study took place has a unique racial/ethnic makeup and is exceptionally violent. Perhaps the risk factors present in the study county presented challenges that Positive Action could not combat, resulting in a less significant decrease of aggression as compared with previous studies. These disparate findings underscore the importance of replicating the program in different settings to confirm or counter past results.

Finally, although the detrimental effect of the Positive Action program on internalizing scores was not statistically significant in the one-tailed test, it was significant in the two-tailed test. Contrary to our hypothesis and past research (Lewis et al. 2013a), findings showed that, at any point in time during the study period, intervention youth had internalizing scores that were 1.8–3.0 % higher than the internalizing scores of control youth (with the range of scores depending on the form of propensity score analysis that was used). The difference on this outcome shows that the control youth generally had lower scores than that of the intervention youth. It is worth noting that this difference could be a consequence of a more severe extent of an unknown internalizing problem experienced by the intervention youth—a problem that was so severe that the current intervention could not successfully or significantly reduce the problem. It is possible that the elements of the Positive Action program simply do not affect these mental health symptoms. One study of Positive Action in an urban sample found that participation in the program was associated with decreased depression and anxiety, however this decrease was mediated by social-emotional and character development (Lewis et al. 2013a). Further, to the authors' knowledge, no studies have examined Positive Action's effect on depression and anxiety in a rural sample. Given the multiple risk factors present in the intervention county as compared with the neighboring county, it is possible that youth in the intervention county suffered from significant depression and anxiety, but were unable to seek treatment. For example, the intervention county has an unemployment rate of 12 %, which is five percentage points higher than the national average (Bureau of Labor Statistics 2012); between 2009 and 2013, 31.7 % of residents of the intervention county lived below the poverty threshold, which is more than double the national rate of 15.4 % (U.S. Census Bureau 2015). Further, the intervention county is one of the most violent rural counties in the nation with a homicide rate of 19.3 per 100,000 inhabitants, or more than 3 times the state average (5.8) from 2009 to 2013 (NC State Center for Health Statistics 2015). It is possible that, as compared to their counterparts in the neighboring county, youth in the intervention county suffer from higher rates of depression and anxiety because of the pervasive poverty, poor health, and crime that are the background of their daily lives. Indeed, the internalizing symptoms seen in youth in the

intervention county may be a realistic appraisal of the risks inherent in their environment. Current findings indicate that the depression and anxiety problems in the intervention county were not decreased by participation in the Positive Action program.

The findings of the current study must be considered in light of certain limitations. First, it would have been ideal to randomly assign counties and schools to the intervention or control group; however, random assignment was not feasible. Therefore, propensity score analysis was used to rectify this shortcoming. However, unlike random assignment, propensity score analysis cannot balance data to control for hidden selection biases (Guo and Fraser 2015). Second, the two propensity score methods (i.e., matching and weighting) do not always provide the same effect sizes, which leaves uncertainty for understanding the true treatment effects. Such discrepancies are primarily due to different assumptions embedded in the models and differential extents to which our dataset meets these assumptions. This uncertainty needs to be addressed in future studies. However, for all four outcomes, all propensity score models show a treatment effect in the same direction and are consistent in findings of significance tests at a significance level of .05. This level of convergence is important, and warrants robustness of the current findings. Third, when filling out the online assessment, youth might have been affected by the presence of their peers. It would have been ideal to have youth fill out surveys in private rooms, but this approach was not feasible due to space and time constraints. To address this limitation, trained research staff monitored the process to ensure privacy and confidentiality. Fourth, it would have been interesting to examine the effects of the Positive Action intervention by race/ethnicity, however, this was beyond the scope of the current study. Given the dearth of research on Positive Action in rural areas, the purpose of the current study was to gain a more comprehensive understanding of how the program functions in rural areas in general; later analysis of Positive Action should investigate whether results of the program differ by race/ethnicity. Last, given that multiple teachers taught Positive Action, it is probable that the lessons were taught in slightly different ways, which could have an impact on the outcomes; however, such variation in program delivery is a limitation of all large-scale intervention research.

Future studies using propensity score analysis should complete a balance check after running probability of treatment weighting and matching. We used one of the 15 imputed files to check the imbalance following probability of treatment weighting or matching. Ideally, the imbalance check should be applied to all files, and then be aggregated into one set of results by using Rubin's rule. Although we are confident that the current findings are accurate—

because both inverse probability of treatment weighting and matching methods balanced the two groups (i.e., youth from the intervention and comparison counties) on more than 90 % of covariates shown by one imputed file—an aggregated imbalance check using all 15 imputed files would have enhanced the rigor of this evaluation.

## Conclusion

The current study investigated the effects of Positive Action on self-esteem, school hassles, aggression, and internalizing symptoms in a sample of low-income, racially/ethnically diverse rural youth. To date, the majority of Positive Action research has been conducted in urban areas. The current research adds to the existing Positive Action knowledge base by testing the program in an impoverished, rural, ethnically/racially diverse county. Current findings indicate that participation in Positive Action were related to significant increases in self-esteem and significant decreases in school hassles relative to youth who did not participate in the program. Although not statistically significant, participation in Positive Action was also associated with decreases in aggressive behavior. Contrary to expectations, participation in Positive Action was associated with increased internalizing symptoms, but the effect was not statistically significant in the one-tailed test. Overall, the current study indicates that Positive Action has some positive results when implemented in a low-income, diverse, violent, rural community. At-risk youth in the intervention rural community benefited from increased self-esteem and decreased school hassles, which might ultimately lead to more positive school experiences and improved academic functioning. However, Positive Action was unable to improve the internalizing symptoms of youth residing in the intervention county, suggesting that the program could be supplemented with additional programming when mental health improvement is desired and youth live in a challenging environment with multiple risks factors. No matter how many accolades an evidenced based program has, it remains critical for researchers to continue to evaluate the program in different contexts.

**Acknowledgments** Funding for this research was provided through a cooperative agreement with the U.S. Centers for Disease Control and Prevention's National Center for Injury Prevention and Control (5 U01 CE001948-03).

**Authors' Contributions** S.G. conceived of the study by selecting the independent variables and statistical methodologies to use and oversaw the statistical analysis. Q.W. conducted the statistical analysis and wrote the methods and analysis sections with support and guidance from S.G. P.R.S. obtained the funding to make the current study possible, implemented the data collection, and assisted in drafting and editing the manuscript. M.B. oversaw the implementation of the

program and worked in the field ensuring the fidelity of Positive Action; she also wrote the implementation section. C.B.R.E. and K.L.C. wrote the remaining sections of the manuscript. All authors read and approved the final manuscript.

**Conflict of interest** The authors report no conflicts of interest.

## References

- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for ASEBA school-age forms and profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth & Families.
- Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: Sage.
- Atav, S., & Spencer, G. A. (2002). Health risk behaviors among adolescents attending rural, suburban, and urban schools: A comparative study. *Family and Community Health, 17*(12), 53–64. doi:10.1097/00003727-200207000-00007.
- Barnett, T., Clements-Nolle, K., Lensch, T., Frankenberger, D., Larson, S., & Yang, W. (2015). *Non-medical prescription drug use in a representative sample of high school students: A need for prevention in rural settings*. American Public Health Association. Retrieved from <https://apha-confex.com.libproxy.lib.unc.edu/apha/143am/webprogram/Paper335314.html>
- Bavarian, N., Lewis, K. M., DuBois, D. L., Acock, A., Vuchinich, S., Silverthorn, N., & Flay, B. R. (2013). Using social-emotional and character development to improve academic outcomes: A matched-pair cluster-randomized controlled trial in low-income, urban schools. *Journal of School Health, 83*, 771–779. doi:10.1111/josh.12093.
- Beets, M. W., Ray, B. R., Vuchinich, S., Snyder, F. J., Acock, A., Li, K. K., & Durlak, J. (2009). Use of 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. doi:10.2105/AJPH.2008.142919.
- Blueprints for Healthy Youth Development. (2012). *Program criteria*. Retrieved from <http://www.blueprintsprograms.com/programCriteria.php>
- Bowen, G. L., & Richman, J. M. (2008). *The School Success Profile*. Chapel Hill, NC: University of North Carolina.
- Bowen, G. L., Rose, R. A., & Bowen, N. K. (2005). *The reliability and validity of the school success profile*. Philadelphia, PA: Xlibris.
- Bureau of Labor Statistics. (2012). *Local area unemployment statistics map*. Retrieved from <http://www.bls.gov/lau/>
- Centers for Disease Control and Prevention. (2013). *1995–2013 Middle school youth risk behavior survey data*. <http://nccd.cdc.gov/youthonline/>
- Cohen, J., McCabe, L., Michelli, N. M., & Pickeral, T. (2009). School climate: Research, policy, practice, and teacher education. *Teachers College Record, 111*, 180–213. Retrieved from <http://www.tcrecord.org>
- Deal, T. E., & Peterson, K. D. (2009). *Shaping School Culture: Pitfalls, Paradoxes, and Promises*. San Francisco, CA: Jossey-Bass.
- Derdikman-Eiron, R., Indredavik, M. S., Bratberg, G. H., Taraldsen, G., Bakken, I. J., & Colton, M. (2011). Gender differences in subjective well-being, self-esteem, and psychosocial functioning in adolescents with symptoms of anxiety and depression: Findings from the Nord-Trøndelag health study. *Scandinavian Journal of Psychology, 52*, 261–267. doi:10.1111/j.1467-9450.2010.00859.x.
- Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., Moffitt, T. E., & Caspi, A. (2005). Low self-esteem is related to aggression,

- antisocial behavior, and delinquency. *Psychological Science*, *16*, 328–335. doi:10.1111/j.0956-7976.2005.01535.x.
- DuBois, D. L., Flay, B. R., & Fagen, M. C. (2009). Self-esteem enhancement theory: An emerging framework for promoting health across the life-span. In R. J. DiClemente, M. C. Kegler, & R. A. Crosby (Eds.), *Emerging theories in health promotion practice and research* (2nd ed., pp. 97–130). San Francisco, CA: Jossey-Bass.
- Dulmus, C. N., Theriot, M. T., Sowers, K. M., & Blackburn, J. A. (2004). Student reports of peer bullying victimization in a rural school. *Stress, Trauma, and Crisis*, *7*, 1–16. doi:10.1080/15434610490281093.
- Erikson, E. H. (1950). *Childhood and society*. New York, NY: Norton.
- Finley, C. & Stewart, A. (2013). *Working with rural teens: Adolescent reproductive health in rural America*. The National Campaign to Prevent Teen and Unplanned Pregnancy. Third Annual Teen Pregnancy Prevention Grantee Conference. Retrieved from [http://www.hhs.gov/ash/oah/oah-initiatives/teen\\_pregnancy/training/Assests/adolescent\\_reproductive\\_ruralam.pdf](http://www.hhs.gov/ash/oah/oah-initiatives/teen_pregnancy/training/Assests/adolescent_reproductive_ruralam.pdf)
- Flay, B. R., & Allred, C. G. (2003). Long-term effects of the positive-action program. *American Journal of Health Behavior*, *27*, S6–S21. doi:10.5993/AJHB.27.1.s1.2.
- Flay, B. R., Allred, C. G., & Ordway, N. (2001). Effects of the positive action program on achievement and discipline: Two match-control comparisons. *Prevention Science*, *2*(2), 71–89. doi:10.1023/A:1011591613728.
- Geertz, C. (1973). *The Interpretation of Culture*. New York, NY: Basic Books.
- Gil, A. G., Vega, W. A., & Dimas, J. M. (1994). Acculturative stress and personal adjustment among Hispanic adolescent joys. *Journal of Community Psychology*, *22*, 43–54. doi:10.1002/1520-6629(199401)22:1<43::AID-JCOP2290220106>3.0.CO;2-T.
- Greene, W. H. (2003). *Econometric analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
- Guo, S., & Fraser, W. M. (2015). *Propensity score analysis: Statistical methods and applications* (2nd ed.). Thousand Oaks, CA: Sage.
- Horvitz, D., & Thompson, D. (1952). A generalization of sampling without replacement from a finite population. *Journal of the American Statistical Association*, *47*, 663–685. doi:10.1080/01621459.1952.10483446.
- Huitsing, G., Veenstra, R., Sainio, M., & Salmivalli, C. (2012). “It must be me” or “It could be them?” The impact of the social network position of bullies and victims on victims’ adjustment. *Social Networks*, *34*, 279–386. doi:10.1016/j.socnet.2010.07.002.
- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, *86*, 4–29. doi:10.1162/003465304323023651.
- Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*, 5–86. doi:10.1257/jel.47.1.5.
- Lee, C. H., & Song, J. (2012). Functions of parental involvement and effects of school climate on bullying behaviors among South Korean middle school students. *Journal of Interpersonal Violence*, *27*, 2437–2464. doi:10.1177/0886260511433508.
- Lewis, K. M., DuBois, D. L., Bavarian, N., Acock, A., Silverthorn, N., Day, S., & Flay, B. R. (2013a). Effects of Positive Action on the emotional health of urban youth: A cluster-randomized trial. *Journal of Adolescent Health*, *53*, 706–711. doi:10.1016/j.jadohealth.2013.06.012.
- Lewis, K. M., Schure, M. B., Bavarian, N., DuBois, D. L., Day, J., Ji, P., & Flay, B. R. (2013b). Problem behavior and urban, low-income youth: A randomized controlled trial of Positive Action in Chicago. *American Journal of Preventative Medicine*, *44*, 622–630. doi:10.1016/j.amepre.2013.01.030.
- Li, K. K., Washburn, I., DuBois, D. L., Vuchinich, S., Ji, P., Brechling, V., & Flay, B. R. (2011). Effects of the Positive Action programme on problem behaviors in elementary school students: A match-pair randomised control trial in Chicago. *Psychology and Health*, *26*(2), 187–204. doi:10.1080/08870446.2011.531574.
- Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. *Journal of the American Medical Association*, *285*, 2094–2110. doi:10.1001/jama.285.16.2094.
- National Institute of Justice. (2015, n.d.). Program profile: Positive Action. Retrieved from <http://www.crimesolutions.gov/ProgramDetails.aspx?ID=113>
- National Registry of Evidence-Based Programs and Practices. (2014). Minimum requirements. Retrieved from <http://www.nrepp.samhsa.gov/ReviewSubmission.aspx>
- NC State Center for Health Statistics. (2015). North Carolina statewide and county trends in key health indicators. Retrieved from <http://www.schs.state.nc.us/data/keyindicators/>
- Positive Action (2015 n.d.) *Positive Action overview*. Retrieved from <https://www.positiveaction.net/>
- Price, M., Chin, M. A., Higa-McMillan, C., Kim, S., & Frueh, B. C. (2013). Prevalence and internalizing problems of ethnoracially diverse victims of traditional and cyberbullying. *School Mental Health*, *5*, 183–191. doi:10.1007/s12310-013-9104-6.
- Prinz, R. J., Foster, S., Kent, R. N., & O’Leary, K. D. (1979). Multivariate assessment of conflict in distressed and nondistressed mother-adolescent dyads. *Journal of Applied Behavior Analysis*, *12*, 691–700. doi:10.1901/jaba.1979.12-691.
- Provasnik, S., Ramani, A. K., Coleman, M. M., Gilbertson, L., Herring, W., & Xie, Q. (2007). *Status of education in rural America (NCES 2007-040)*. Washington, DC: National Center for Education Statistics.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.
- Robers, S., Kemp, J., & Truman, J. (2013). *Indicators of school crime and safety: 2012 (NCES 2013-036/NCJ241446)*. Washington, DC: National Center for Education Statistics. Retrieved from <http://nces.ed.gov/pubs2013/2013036.pdf>
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*, 41–55. doi:10.1093/biomet/70.1.41.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician*, *39*, 33–38. doi:10.2307/2683903.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: J. Wiley & Sons.
- Smokowski, P. R., Guo, S., Cotter, K. L., Evans, C. B. R., & Rose, R. A. (2015). in press). Multi-level risk factors and developmental assets associated with aggressive behavior in disadvantaged adolescents: Modeling longitudinal trajectories from the rural adaptation project. *Aggressive Behavior*.
- Smokowski, P. R., Guo, S., Rose, R., Evans, C. B., Cotter, K. L., & Bacallao, M. (2014). Multilevel risk factors and developmental assets for internalizing symptoms and self-esteem in disadvantaged adolescents: Modeling longitudinal trajectories from the Rural Adaptation Project. *Development and Psychopathology*, *26*, 1495–1513. doi:10.1017/S0954579414001163.

- Snyder, F. J., Acock, A. C., Vuchinich, S., Beets, M. W., Washburn, I., & Flay, B. R. (2013). Preventing negative behaviors among elementary-school students through enhancing students' social-emotional and character development. *American Journal of Health Promotion*, 28(1), 50–58. doi:10.4278/ajhp.120419-QUAN-207.2.
- Snyder, F., Flay, B., Vuchinich, S., Acock, A., Washburn, I., Beets, M., & Li, K. K. (2010). Impact of a social-emotional and character development program on school-level indicators of academic achievement, absenteeism, and disciplinary outcome: A matched-pair, cluster-randomized, controlled trial. *Journal of Research on Educational Effectiveness*, 3, 26–55. doi:10.1080/19345740903353436.
- Substance Abuse and Mental Health Services Administration. (2014). *Results from the 2013 national survey on drug use and health: summary of national findings, NSDUH Series H-48, HHS Publication No. (SMA) 14-4863*. Rockville, MD: Substance Abuse and Mental Health Services Administration, 2014.
- U.S. Census Bureau. (2015). State and county quick facts. Retrieved from <http://quickfacts.census.gov/qfd/states/00000.html>
- U.S. Department of Education. (2007). *What works clearing house*. Retrieved from <http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=380>
- U.S. Department of Education. (2009). *Save, disciplined, and drug-free schools*. Retrieved from [http://www2.ed.gov/admins/lead/safety/exemplary01/report\\_pg9.html](http://www2.ed.gov/admins/lead/safety/exemplary01/report_pg9.html)
- U.S. Department of Justice, Office for Victims of Crime. (2001). *Rural victim assistance: A victim/witness guide for rural prosecutors (NCJ No. 211106)*. Retrieved from [http://www.ovc.gov/publications/infores/rural\\_victim\\_assistance/pfv.html](http://www.ovc.gov/publications/infores/rural_victim_assistance/pfv.html)
- Washburn, I. J., Acock, A., Vuchinich, S., Snyder, F., Li, K. K., Ji, P., & Flay, B. R. (2011). 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*, 12, 314–323. doi:10.1007/s11121-011-0230-9.
- Witherspoon, D., & Ennett, S. (2011). Stability and change in rural youths' educational outcomes through the middle and high school years. *Journal of Youth and Adolescence*, 40, 1077–1090. doi:10.1007/s10964-010-9614-6.
- Shenyang Guo, Ph.D.** is the Frank J. Bruno Distinguished Professor of Social Work at the George Warren Brown School of Social Work at Washington University in St. Louis. His research focuses on applying advanced statistical models (e.g., propensity score analysis, survival analysis, hierarchical linear modeling, growth curve analysis, structural equation modeling, cross-classified random effects models) to solving social welfare problems.
- Qi Wu, MM.** is a doctoral student at the University of North Carolina at Chapel Hill. Her research interests include child welfare and applying advanced statistical modeling such as propensity score analysis and hierarchical linear modeling to social welfare problems.
- Paul R. Smokowski, Ph.D.** is a professor and Dean at the University of Kansas School of Social Welfare. His research interests include risk, resilience, acculturation, adolescent mental health, and youth violence prevention.
- Martica Bacallao, Ph.D.** is an associate professor of practice at the University of Kansas School of Social Welfare. Her research interests include the implementation of youth violence prevention programs for community-based research, immigration and acculturation, and community-based mental health.
- Caroline B. R. Evans, Ph.D.** is a research associate at the University of North Carolina at Chapel Hill. Her research interests include bullying perpetration and victimization, bullying bystander behavior, adolescent aggression and violence, and adolescent development.
- Katie L. Cotter, Ph.D.** is an assistant professor at Arizona State University's School of Social Work. Her research interests include female aggression, youth violence prevention, adolescent development, and risk and resilience.