

AN ABSTRACT OF THE DISSERTATION OF

Kendra M. Lewis for the degree of Doctor of Philosophy in Human Development and Family Studies presented on April 20, 2012.

Title: Evaluation of a Social-Emotional and Character Development Program: Methods and Outcomes

Abstract approved:

Samuel Vuchinich

Schools are increasingly expected to prevent and decrease violence, substance use, and other problem behaviors linked to academics and prepare students to be contributing members of society. One approach with proven success in promoting positive outcomes related to a broad range of student behaviors and personal characteristics is social-emotional learning, also known as positive youth development or social-emotional and character development (SECD) programs. Currently, little is known about the relationship between SECD and these outcomes in low-income, urban, minority populations. The present studies (a) examine key methodological design issues in conducting a cluster-randomized trial with such populations, and (b) assess whether an intervention designed to promote SECD was effective in improving the SECD developmental status of children from this population. Data for this study come from the Chicago cluster-randomized controlled trial (CRCT) of *Positive Action (PA)*. The trial was longitudinal at the school level with a place-focused intent-to-treat design at the student level. This CRCT collected data on children in grades 3 through 8, for a total of 6 years and 8 data collection points. Manuscript #1 focuses on the

design, sample, planned analyses, and a latent class analysis (LCA) of mobility patterns. Specifically, the setting and recruitment of schools is described, as well as the process by which schools were matched into pairs and randomized into *PA* or control, including the list of criteria for school eligibility and variables used for matching. Additionally, this paper thoroughly describes the primary analyses to test for program effects using three-level growth curve models (time nested within students nested within schools), as well as several sensitivity analyses that will also be conducted when evaluating this program. Further, this manuscript discusses secondary tests of mediation and moderation, which will assist in the understanding of how the program works (mediation) and for whom (moderation). Finally, this paper also provides several descriptive statistics and characteristics of the students and teachers in this sample. In terms of baseline equivalency, *PA* and control schools did not significantly differ on matching variables before or during the trial. Minimal differences were found on baseline reports from students, teachers, and parents; half of these differences favored *PA* students and half favored control students. Manuscript #2 focuses on the intervention effects on the student-, teacher-, and parent-reported social-emotional outcomes assessed during the Chicago CRCT, following the analytic procedures outlined in the first paper and focusing on the effectiveness of *PA* on social-emotional outcomes. Results indicate that *PA* had a significant effect on student self-reports of prosocial interactions, honesty, self-development, self-control, respect for parents and teachers, empathy, altruism, positive actions/feelings, negative moral center, and aggressive problem solving. Additionally, it was found that *PA* had

marginal effects on teacher-reported responsibility. Minimal differences by gender were found; no differences by mobility status were found. Together, these two papers involved a sample of students in a high-risk setting; generating improvements can be particularly difficult in urban areas. The empirical evidence of effectiveness of a SECD program in a high-risk population, as demonstrated in the present study, should serve as a call to action for policymakers and school officials who are increasingly challenged to positively impact not only academic achievement, but also behavior and character development.

© Copyright by Kendra M. Lewis
April 20, 2012
All Rights Reserved

Evaluation of a Social-Emotional and Character Development Program: Methods and
Outcomes

by
Kendra M. Lewis

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented on April 20, 2012
Commencement June 2012

Doctor of Philosophy dissertation of Kendra M. Lewis presented on April 20, 2012

APPROVED:

Major Professor, representing Human Development and Family Studies

Co-Director of the School of Social and Behavioral Health Sciences

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Kendra M. Lewis, Author

ACKNOWLEDGEMENTS

I would like to thank my parents, Sharon and George, who have provided me with unconditional love and support throughout this journey. They've helped me through my toughest lessons and showed me that I have the strength to achieve this goal.

My committee provided me with guidance and advice throughout this process. Samuel Vuchinich assisted me in the development and conceptualization of my research work, and provided guidance in coursework and my dissertation. Alan Acock fueled my interest in statistics and was always willing to talk about research methods and statistics. Brian Flay provided me with the opportunity to work with his data, and through this work I learned so much more than I ever imagined. I am grateful for this opportunity. Kate MacTavish always supported my decisions and helped me stay focused on my goals.

I would also like to acknowledge my mentor at CSU Chico, Dr. Brad Glanville. From the moment I met him, Brad has supported me in my pursuit of a PhD, and has always provided me with advice along the way. I could not have done this without his guidance and encouragement.

My friends and colleagues at OSU have been fundamental to this process. Whether it was editing drafts of papers, giving feedback on presentations, or having drinks to relax and share a laugh, their help and encouragement helped me through each day. My best friend, Heather Wilhelm, was always ready to listen to me, whether I needed to cry, share exciting news, or just to talk. Her support has been invaluable.

CONTRIBUTION OF AUTHORS

Kendra M. Lewis outlined and performed all data analyses and completed the initial draft of the papers presented.

Dr. Brian Flay provided support in the initial conceptualization as well as editorial comments and advice on both manuscripts.

Drs. Alan Acock , Samuel Vuchinich, and David DuBois provided advice regarding the design, analyses, interpretation of results, and editorial comments on both manuscripts.

Drs. Joseph Day, Peter Ji, and Naida Silverthorn implemented the program and collected data.

Niloofar Bavarian provided editorial comments and advice on both manuscripts.

TABLE OF CONTENTS

	<u>Page</u>
Introduction.....	1
Design, Sample, and Planned Analysis of the Chicago Cluster- Randomized Controlled Trial of the <i>Positive Action</i> Program.....	13
Improving Social-Emotional and Character Development in Children and Adolescents: A Matched-Pair, Cluster-Randomized Controlled Trial in Low-Income, Urban Schools.....	47
Conclusions.....	79
Bibliography.....	86

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
2.1 Diagram of participating schools and students.....	44
3.1 Growth curve graphs for significant outcomes that also have a significant gender difference.....	76

LIST OF TABLES

<u>Table</u>	<u>Page</u>
2.1. Mobility of students by condition	39
2.2. (Online Supplement) Student, Parent, and Teacher-reported Scale Properties.....	40
2.3. Baseline and follow-up equivalence on school demographics.....	43
3.1 Correlations between Student-, Teacher-, and Parent-Reported Outcomes.....	70
3.2 Growth model estimates for Student-, Teacher-, and Parent- Reported outcomes.....	73
3.3 Wave 1 and Wave 8 predicted (model) means and probabilities by condition.....	75

DEDICATION

This is dedicated to my nephews and niece Chase, Cole, Cooper, Connor, and Mallory.

Evaluation of a Social-Emotional and Character-Development Program: Methods and Outcomes

CHAPTER 1: INTRODUCTION

Social emotional and character development (SECD; Elias 2009) is a necessary prerequisite for academic achievement (Ciarrochi, Heaven, & Davies, 2007; Elias 2006; 2009) and the prevention of negative outcomes, such as substance use and delinquency (Hawkins, Kosterman, Catalano, Hill, & Abbott, 2005; O'Donnell, Hawkins, Catalano, Abbott, & Day, 1995). Currently, little is known about the relationship between SECD and these outcomes in low-income, urban, minority populations. The present studies will (a) examine key methodological design issues in conducting a cluster-randomized trial with such populations, and (b) test whether an intervention designed to promote SECD was effective in improving the SECD developmental status of children from this population.

SECD and related skills and outcomes are essential for establishing the kind of social, emotional, and behavioral environments in schools and classrooms needed to promote academic achievement (Coalition for Evidence Based Policy, 2002; Coleman et al., 1966). Schools cannot attain their goals if students lack honesty, respect, discipline, empathy, and other aspects of SECD (Elias, 2009). Additionally, low-income minority youth are particularly affected by the disparities of the educational system (Aud, Fox, M., & KewalRamani, 2010). Given the promise of improved SECD leading to improved academic achievement and decreased negative behaviors, SECD and related skills should be included in the learning process. In response to

these findings there has been a movement toward comprehensive, multi-modal, and multi-level programs that address a range of child and adolescent behaviors often involving families and community. Such programs appear to be more effective than programs that focus on only one behavior or domain (Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2004; Flay, 2002; Flay & Allred, 2010; Flay, Snyder, & Petraitis, 2009). To date, however, it is not clear how well these findings generalize to low-income, urban settings (Farahmand, Grant, Polo, Duffy, & DuBois, 2011). This introduction outlines the development and importance of social-emotional skills and behaviors, followed by an overview of SECD programming. Next, details regarding the evaluation of the program assessed in the present study, *Positive Action*, will be presented, including the rationale, methodology, and challenges faced. Finally, the specifics of each manuscript will be outlined.

Overview of social-emotional outcomes

Emotion plays a crucial role in the development of prosocial behavior, defined as a "voluntary behavior intended to benefit another" (Eisenberg, Fabes, & Spinrad, 2006, pg. 646). Emotional development begins early in a child's life. Once a child is able to internally regulate his or her own emotions, he or she starts to feel and understand new and more ambiguous emotions, such as empathy, guilt or shame. Emotional development provides a basis for many skills such as understanding the meanings and causes of emotions, recognizing the consequences of emotions in certain situations, and the development of empathy and prosocial behavior (Thompson & Lagattuta, 2006). Closely related to such behaviors is SECD. Several concepts are

included in definitions of SECD, such as positive interactions with peers, teachers, and parents (Selman, 2003; Shultz, Selman, & LaRusso, 2003); being honest with peers, teachers, and parents (Park, Peterson, & Seligman, 2004); emotional awareness and self-regulation (Eisenberg, Champion, & Ma, 2004); self development (King et al., 2005; Lerner et al., 2005); positive traits, such as kindness and hope (Park, 2004); and moral functioning, such as moral values and reasoning (Berkowitz & Bier, 2004). Prosocial behaviors and SECD are associated with a variety of outcomes such as increased self-esteem, social competency, academic achievement, well-being, life satisfaction, as well as decreased impulsivity, aggression, externalizing behaviors, depression, and anxiety (for reviews, see: Eisenberg, Fabes, & Spinrad, 2006; Park & Peterson, 2009; Spinrad & Eisenberg, 2009). Moreover, Park and Peterson (2009) suggest that promoting SECD could possibly prevent students' social problems and increase opportunities to build healthy relationships.

In addition to prosocial skills and SECD, this study examines program effects on social competency and related constructs. Social competence is difficult to define, but is generally known as a person's ability to establish and maintain quality relationships (Fabes, Gaertner, & Popp, 2006). In order to establish quality relationships, children first need to develop emotional regulation and understanding (as discussed above), as well as social competency. Social competency is predicted by several factors, including a child's temperament, socio-cognitive skills, and communication skills (Fabes, Gaertner, & Popp, 2006). Without these skills, children are unable to effectively communicate their thoughts and feelings, nor can they

understand the thoughts and feelings of others (which is essential in constructs such as empathy or altruism). These skills are crucial in learning how to socially interact with other children (Fabes, Gaertner, & Popp, 2006) and solve problems in a positive manner, which can improve social behaviors and academic success (Rose-Krasnor & Denham, 2009). These examples and associations emphasize that investing in programs that can increase social-emotional development can also increase academic achievement, problem solving, and other positive outcomes while decreasing negative outcomes.

Students living in low income, urban settings face a multitude of barriers to success (Elias & Haynes, 2008), as well as negative consequences such as poorer academic outcomes (Aud et al., 2010), life event stressors (e.g., child abuse or divorce; Conger, Ge, Elder, Lorenz, & Simons, 1994), increased exposure to crime and violence (Bell & Jenkins, 1994), and increased risk for psychological problems (Grant et al., 2004). Therefore, promoting social-emotional skills in this population is particularly important to help these students succeed. Related approaches with evidence of success in promoting positive outcomes related to a broad range of student behaviors and personal characteristics are social-emotional learning (SEL; Weissberg & O'Brien, 2004), positive youth development (PYD; Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2004; Flay, 2002; J.V. Lerner, et al., 2009; R. M. Lerner, Dowling, & Anderson, 2002; R. M. Lerner et al., 2005; Snyder & Flay, in press), character education (Berkowitz & Bier, 2004) and SECD programs. There is,

however, a lack in the current literature regarding the effectiveness of these programs in low-income, urban youth.

Social-Emotional and Character Development programming

SECD programs use interactive approaches that are designed to be both developmentally and culturally appropriate. They emphasize building students' skills to make responsible decisions, solve problems effectively, recognize and manage their emotions, appreciate the perspectives of others, handle interpersonal situations effectively, be honest with themselves and others, and establish positive goals (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). Importantly, many of these interventions are school-wide in addition to including curricula focused on supporting overall student development and academic achievement (Durlak et al., 2011).

Berkowitz and Bier (2004) conducted a review of a broad selection of school-based character education programs to determine what these programs achieve and how they do so. They concluded that these programs succeeded when implemented comprehensively and with fidelity. Additionally, these programs have an impact on a variety of outcomes such as prosocial behaviors, substance use, violence, academic achievement, and problem solving skills.

In a meta-analysis of 213 SEL programs, Durlak and colleagues (2011) found that such programs significantly increased social and emotional skills; improved students' attitudes about themselves, others, and school; improved positive social behaviors; improved academic achievement; and decreased conduct problems and emotional distress (Durlak et al., 2011). Additionally, in their meta-analyses of 23

articles examining the impact of school-based mental health and behavioral programs set in low-income, urban schools, Farahmand and colleagues (2011) reported a mean effect size (generally Hedges g) of 0.31 on competence and social skills, 0.28 on internalizing behaviors, 0.24 on academic outcomes, and 0.02 on externalizing and substance use behaviors, suggesting that some programs are effective in this population. More research, however, is needed on the effectiveness of programs in this setting. Further, in their review of 77 programs focused on positive youth development, Catalano et al. (2004) found that youth participating in these programs had a wide variety of improved outcomes compared to those youth not in these programs, including greater assertiveness, sociability, problem solving, interpersonal skills, self-control, empathy, and health management, as well as reduced problem behaviors such as substance use and aggression.

Description of Positive Action and the underlying theory

One example of a SECD program being implemented throughout the United States is *Positive Action* (PA; <http://www.positiveaction.net>; Flay & Allred, 2010), a comprehensive, school-wide program. The program posits that children who engage in positive behaviors will have more positive feelings about themselves and subsequent positive thoughts, leading to more positive behaviors. Moreover, PA (as well as other SECD programs) proposes a link between positive and negative behaviors, with increased positive feelings, thoughts, and actions resulting in fewer negative behaviors (Durlak et al., 2011; Flay & Allred, 2010).

The first core element of the program is the *Positive Action* philosophy, which is grounded in theories of self-concept and self-esteem (Purkey, 1970; Purkey & Novak; 1970), particularly Self-esteem Enhancement Theory (SET), a broad theory of self-concept and self-esteem (DuBois, Flay, & Fagen, 2009). This theory assumes that motivation to feel good about oneself is prevalent throughout the life span and that people will use a wide range of cognitive, affective, and behavioral strategies to help acquire and sustain feelings of self-worth.

The second core element of the program is the Thoughts-Actions-Feelings About Self circle. The components of the program (including the content of the classroom curricula) are based on the idea that “You feel good about yourself when you do positive actions and there is always a positive way to do everything.” This circle can also be seen as teaching some basics about self-regulation. Self-regulation includes control over cognition ("thoughts"), behaviors ("actions"), and emotions ("feelings"), and is related to a variety of outcomes, such as school readiness, social development, and academic success (McClelland, Ponitz, Messersmith, & Tominey, 2010). The Thoughts-Actions-Feelings About Self circle illustrates this self-reinforcing process that is taught to students; showing them that thoughts lead to actions, actions lead to feelings about self, and feelings about self lead to more thoughts. The aim of *PA* is to get everyone into the positive cycle by consciously making positive choices. This is intrinsically motivated change (Deci, 2009; Deci, Koestner, & Ryan, 1999, 2001; Deci & Ryan, 1985; Gottfried, Marcoulides, Gottfried, & Oliver, 2009; Ryan & Deci, 2000, 2006), where a person chooses to do positive

actions to feel good about his or her self. Children and youth must not only be taught what is right versus wrong, but also need to be supported and reinforced for acting in such a way. The *PA* program does this in a way that is effective for students, as well as their instructors, parents, and community members.

The third core element of the *PA* program is the classroom curriculum, which consists of the following six components: self-concept, positive actions for body and mind, social and emotional positive actions focusing on getting along with others, and managing, being honest with, and continually improving oneself. Together, these make up the comprehensive set of necessary skills for successful learning and living. The program trains teachers and parents to identify, teach, and reinforce positive thoughts, actions, and feelings about themselves by students and adults in the school. This leads to continual reinforcement of positive actions and enhanced student bonding with parents and school, and is consistent with ecological theories of behavior. For example, although the Theory of Triadic Influence (TTI; Flay & Petraitis, 1994; Flay et al., 2009) did not guide the development of *PA*, the outcomes assessed by *PA* and the focus of the program can clearly be mapped onto the TTI

The TTI consists of three streams of influence: intrapersonal, social, and cultural. These streams of influence are similar to the ecological rings of the Bioecological Model (Bronfenbrenner & Morris, 2006). Within these streams exist multiple levels of causation, ranging from proximal to distal. Cognitive and affective processes are also included within each stream of influence (as substreams). The TTI can also be framed as a developmental theory, as it considers development over time

and the impact this may have on behaviors. Further, once a behavior is performed, reactions to that behavior feed back into the different streams and levels of influence, as the experience of a behavior may change a person's self-control, self-efficacy, attitudes, bonds with others, beliefs about norms, etc. The TTI provides an inclusive perspective for understanding the mechanisms of behaviors and details the multiple influences that are a part of even a single act or behavior. The TTI can provide a more comprehensive explanation of behaviors than other theories that focus on just an individual, society, cultural, or some limited combination to explain behaviors. The TTI provides a theoretical foundation for *PA* and its effects on different domains. Using the TTI, we can identify potential mediators of program effects on a variety of outcomes, giving us a clearer understanding of how the program works; that is, the mechanisms underlying *PA*.

Previous Research on PA

Previous quasi-experimental and experimental evaluations have found significant effects of *PA* on several outcomes (Flay & Allred, 2010). In quasi-experimental studies in Nevada, Hawai'i, and Florida school districts, schools receiving *PA* had higher achievement scores and fewer violent incidents, disciplinary referrals and suspensions than schools not receiving *PA* (Flay, Allred, & Ordway, 2001; Flay & Allred, 2003). Additionally, in a cluster-randomized controlled trial (CRCT) of *PA* in Hawai'i, which followed students in 20 schools from grade 1 or 2 through grade 5 or 6, students receiving *PA* were less likely to engage in substance use, violent behaviors or sexual activity (Beets et al., 2009). At the school level, *PA*

schools had higher academic achievement on standardized test scores, lower rates of absenteeism, fewer disciplinary referrals and suspensions (Flay & Allred, 2010; Snyder et al., 2010), and improved school quality (Snyder, Vuchinich, Acock, Washburn, & Flay, 2012).

The current paper focused on a CRCT of *PA* in Chicago schools, conducted in a high-minority, low-income (i.e., more than 80% of students receiving free or reduced lunch), urban population. This CRCT involved collecting data eight times on children in grades 3 through 8. Prior papers from this trial found that students receiving *PA* reported less SU, violence and bullying behaviors than control students not receiving *PA* by grade 5 (Li et al., 2011) and less substance use by grade 8 (Lewis et al., 2012a), and that *PA* students reported higher SECD-related skills that mediated the decreased substance use (Lewis et al., 2012a). Using data from both the Hawai'i and Chicago trials, as well as an RCT from a southeastern state, Washburn et al. (2011) found that while SECD-related skills decreased over time for both *PA* and control students, *PA* significantly mitigated this decline; that is, *PA* students showed less of a decline in SECD than those not receiving *PA*.

Methodological issues in longitudinal and cluster-randomized trials

Longitudinal studies of SECD programs face the unavoidable issue of attrition. This issue is especially prevalent in the Chicago CRCT of *PA*, in which the student mobility rates were very high. Few students remained in the study across the eight waves; many students left the study schools or joined them during later waves. This challenge created the opportunity to utilize a novel approach to assessing mobility

effects. In order to test these potential moderating effects in future analyses, a latent class analysis (LCA; Marsh, Ludtke, Trautwein, & Morin, 2009) was conducted to identify groups of children with similar patterns of dropout and entry. . This allowed for identification of patterns of students based on their mobility in and out of the trial, and the possible impact of these mobility patterns on outcomes. Other methodological issues include small sample size and power, recruitment of clusters and handling the disappointment of those randomized to the control condition, receiving consent from parents, as well as parent reports on children, and measurement challenges (e.g., "floor" or "ceiling" effects).

The present study

Overall, previous studies on *PA* have shown the program to be effective in influencing multiple behaviors. Only one study, however, has assessed the effectiveness of *PA* on social-emotional outcomes (Washburn et al., 2011); this paper analyzed only one outcome (rather than multiple, as done in the present study) and through grade 5, whereas the present study extends into grade 8. In addition, there have been calls to make public the methodology of a trial before the data is analyzed; therefore, there is an increasing number of articles that describe the design and methodology of trials of various interventions (for example, see Morone et al., 2012; Winters-Stone, Lyons, Nail, & Beer, 2012). This particular trial was set in a highly mobile population, emphasizing the need to take a closer look at how to analyze these data and account for mobility.

This dissertation had two aims: to (a) describe the design, sample, methods, and planned analyses of the Chicago CRCT of *PA*, and present a new approach to analyzing attrition in a highly mobile population, and (b) test the long-term effects of *PA* on social-emotional outcomes in a low-income, urban setting. It was hypothesized that students in *PA* schools will show significantly better change in outcomes than students attending control schools.

These aims were addressed in two manuscripts. The first manuscript describes the design on the Chicago CRCT, including the process of matching schools (clusters), the sample included in analyses, baseline equivalency, and planned analyses to evaluate the effectiveness of *PA* on a variety of outcomes. This manuscript reviews the design utilized, lessons learned from challenges both before and during the trial, and describes the most appropriate analytic techniques given the design, strengths, and limitations of the trial.

The second paper provides an example of the application of the analytic techniques in evaluating the effectiveness of *PA* on social-emotional outcomes in a low-income, urban population. Students in these settings face a variety of barriers to social-emotional, physical, and mental health, and this program may serve as a buffer to some of these barriers. Together, these papers can provide new information on conducting randomized trials on school-based preventive interventions in urban minority populations, and new empirical evidence on the impact of the *PA* intervention on the development of social-emotional outcomes.

Design, Sample, and Planned Analysis of the Chicago Cluster-Randomized Controlled Trial of the *Positive Action* Program

Kendra M. Lewis^{a,d}, David L. DuBois^b, Peter Ji^c, Joseph Day^b, Naida Silverthorn^b, Niloofar Bavarian^a, Samuel Vuchinich^a, Alan Acock^a, and Brian R. Flay^{a,d}

^aOregon State University, Corvallis, OR, 97330, USA

^bUniversity of Illinois at Chicago, Chicago, IL, 60608, USA

^cAdler School of Professional Psychology, Chicago, IL, 60602, USA

^dCorresponding author: Oregon State University, College of Public Health and Human Sciences, School of Social and Behavioral Health Sciences, 410 Waldo Hall, Corvallis OR 97331-5102; Telephone: 541-737-4765; Fax: 541-737-1076; Email:

lewiske@onid.orst.edu.

Brian Flay email: brian.flay@oregonstate.edu

ABSTRACT

Positive Action (PA) is a social-emotional and character development program designed to improve academics, behavior and character. This paper describes the design, planned analyses, and baseline comparability of a longitudinal, matched-pair, cluster-randomized controlled trial of *PA* conducted in 14 low-income Chicago public schools. Additionally, this paper introduces a novel approach, latent class analysis, to handle student mobility. The trial was conducted during the 2004-05 to 2009-10 academic years and incorporates examination of student, parent, and teacher reports on student outcomes as well as aggregated data from school records across elementary- and middle-school grades. The sample consists of a cohort of 1170 students measured up to eight times from grades three through eight. We found evidence of baseline equivalency between *PA* and control schools both on aggregate school-level data and survey data obtained from the student cohort and their parents and teachers, suggesting successful matching and randomization. Future papers will evaluate program effects on social-emotional outcomes, internalizing and externalizing behaviors, self-esteem, academics, and environment. Limitations and strengths, as well as lessons learned during the trial, are discussed.

Keywords: Longitudinal Design, Cluster Randomized Trial, Methods, Mobility, Latent Class Analysis.

Design, Sample, and Planned Analysis of the Chicago Cluster-Randomized Controlled Trial of the *Positive Action* Program

Since the passing of the No Child Left Behind Act, K-12 education in the United States has primarily emphasized core academic content areas, such as reading and math (Hamilton et al., 2007). It is increasingly recognized, however, that in order for students to be successful in school and in life, they also need curricula to help them acquire social skills, build character, improve mental and physical health, and alleviate problem behaviors (Coalition for Evidence-based Policy, 2002). Growing evidence indicates that programs designed to address these types of skills and behaviors can help cultivate learning environments that both prevent violence (Eron, Gentry, & Schlegel, 1994) and other problem behaviors such as bullying (Afshar & Kenny, 2008) and substance use (Greenberg et al., 2003), and promote academic achievement (Coalition for Evidence-based Policy, 2002; Coleman et al., 1966).

Low-income minority youth are particularly affected by the disparities of the educational system. Research on this population is important, because students living in low-income, urban settings face a multitude of barriers to success (Elias & Haynes, 2008), poorer academic outcomes (Aud, Fox, & KewalRamani, 2010), life event stressors such as child abuse and divorce (Conger, Ge, Elder, Lorenz, & Simons, 1994), increased exposure to crime and violence (Bell & Jenkins, 1994), and increased risk for psychological problems (Grant et al., 2004). One approach with evidence of success in promoting positive outcomes related to a broad range of student behaviors and personal characteristics is social-emotional learning (SEL; Weissberg & O'Brien,

2004) which overlaps greatly with positive youth development (PYD; Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2004; Flay, 2002; Lerner et al., 2009; Lerner, Dowling, & Anderson, 2002; Lerner et al., 2005; Snyder & Flay, in press) and social-emotional and character development (SECD) programs (Elias, 2009). To date, however, it is not clear how well these encouraging findings generalize to low-income, urban settings (Farahmand, Grant, Polo, Duffy, DuBois, 2011). In view of this concern, the present trial seeks to extend the knowledge base of the effectiveness of SECD programs for urban minority students. The study is a matched-pair, cluster-randomized controlled trial (CRCT; Murray, 1998) of *Positive Action (PA)*, an evidence-supported SECD program, as it was implemented in low-performing, high-poverty (K-8) schools in Chicago, with a largely minority (87%) student population.

There have been calls to make public the methodology of a trial before the data are analyzed; therefore, there are increasing numbers of articles that describe the design and methodology of trials of various interventions (for example, see Morone et al, 2012; Winters-Stone, Luyons, Nail, & Beer, 2012). The purpose of the present paper is to describe the design and methodology of the Chicago CRCT of *PA*, including the process through which the schools were matched and participants recruited. We follow the consolidated standards of reporting trial (CONSORT) statement checklist as outlined in Walleser, Hill, and Bero (2011). Included are the results of tests of baseline equivalency on both aggregated school archival data and survey data collected from students, parents, and teachers, as well as an overview of planned analyses for the assessment of program effects, including tests of mediation

and moderation. Additionally, we introduce a novel approach, latent class analysis, to address mobility in CRCTs.

METHOD

Positive Action

Positive Action (<http://www.positiveaction.net>) (Flay & Allred, 2010) is a comprehensive, school-wide, SECD program grounded in theories of self-concept (DuBois, Flay, & Fagen, 2009; Purkey, 1970; Purkey & Novak, 1970) and consistent with integrative and social-ecological theories of health behaviors such as the Theory of Triadic Influence (TTI; Flay & Petraitis, 1994; Flay, Snyder, & Petraitis, 2009). The program posits that children who engage in positive behaviors will have more positive feelings about themselves and subsequent positive thoughts, leading to more positive behaviors. Moreover, *PA* proposes that more positive feelings, thoughts, and actions will result in fewer negative behaviors and improved motivation to learn (Flay & Allred, 2010).

The *PA* program includes PreK-12 curricula, of which the K-8 portion was tested in this trial. The scoped and sequenced classroom curricula consists of over 140 age-appropriate lessons per grade taught for 15-20 minutes 4 days per week for grades K-6, and 70 lessons taught 2-3 days per week for grades 7 and 8. The core curricula consist of the following six units: self-concept, positive actions for body and mind, social and emotional positive actions focusing on getting along with others, and managing, being honest with, and continually improving oneself. In addition to the

student core curricula, the program includes teacher, counselor and family training, and school-wide climate development.

Previous Research on PA

Previous quasi-experimental and experimental evaluations have found significant effects of *PA* on several outcomes (Flay & Allred, 2010). In quasi-experimental studies in Nevada, Hawai'i, and Florida school districts, schools implementing *PA* had higher achievement scores and fewer violent incidents, disciplinary referrals, and suspensions than schools not implementing *PA* (Flay, Allred, & Ordway, 2001; Flay & Allred, 2003). Additionally, in a CRCT of *PA* in Hawai'i, which followed students in 20 schools from grade 1 or 2 through grade 5 or 6, students receiving *PA* were less likely to engage in substance use, violence, or sexual activity (Beets et al., 2009). At the school level, *PA* schools had less absenteeism, fewer disciplinary referrals and suspensions, higher academic achievement (Flay & Allred, 2010; Snyder et al., 2010) and improved school quality (Snyder, Vuchinich, Acock, Washburn, & Flay, 2012)

The Chicago trial is the first matched-pair, CRCT investigating the effects of *PA* in a low-income, urban environment and also the first to include middle-school grades. Because the unit of randomization was the school and because *PA* is a school-wide intervention, the appropriate unit of inference is the school (Donner & Klar, 2004). A repeated cross-sectional approach (i.e., analyzing just school-level means of all student scores) would have serious limitations. It would ignore change at the student level and variation in student change, and students could not “serve as their

own controls,” thus precluding important statistical controls for measured and unmeasured student-level confounds. An approach that overcomes these disadvantages is the cluster-focused intent-to-treat (ITT) analysis (Vuchinich, Flay, Aber, & Bickman, 2012). It acknowledges the focus on schools and follows all schools randomized to condition to trial endpoint, regardless of whether the school continues the *PA* intervention, and regardless of how well *PA* is implemented. It also collects data from all students who are in the appropriate grade cohort in the schools when the assessments occur. Typically students in low-income urban areas stay in the same school a few years, thus providing a basis for the repeated measures advantages listed above. However, student mobility is also high. Because the trial was cluster-focused, we assessed students who entered schools after the beginning of the trial (joiners), but did not follow individual students who stopped attending the study schools (leavers). From the standpoint of students, across time they could be considered a “dynamic” (i.e. changing) cohort. Multilevel models will be used to take into account variation at the school and student levels. Missing data will be addressed using the missing-at-random assumption, as it is unlikely that a single unmeasured variable or set of variables would predict missingness for all students who left or joined the trial schools after randomization (Brown et al., 2008).

Students in the seven matched pairs of schools were followed, beginning in grade 3 (fall 2004 and spring 2005), and at seven additional time points (waves) over six years: beginning and end of grade 4 (fall 2005 and spring 2006), end of grade 5 (spring 2007), beginning and end of grade 7 (fall 2008 and spring 2009), and end of

grade 8 (spring 2010). At the beginning of the study, nine schools were K-8, and five were K-6. By the second phase of the study (grades 6-8) 13 schools were K-8, one was K-5, and one was 6-8 that students from the one K-5 school matriculated into. Thus, 14 schools and seven pairs of schools were included in the study at each wave.

Setting, Design, School Recruitment, and Sample

Sampling and recruitment of schools took place during spring 2004 and were described in detail elsewhere (Ji, DuBois, Flay, & Brechling, 2008). Briefly, the participating schools were low-performing, high-poverty elementary to middle schools within the Chicago Public Schools (CPS) system that were drawn from the population of 483 K-6 and K-8 schools. Schools were excluded from participation if they: 1) were non-community schools (e.g., charter and magnet schools), 2) already had *PA* or a similar SEL/SECD intervention (so that the *PA* program effects would not be confounded with other programs), 3) had an enrollment below 50 or above 140 students per grade (so that at least two classrooms per grade were available for the study, but an enrollment cap was in place for cost and management reasons), 4) had annual student mobility rates under 40% (to ensure some stability in the student population across adjacent years), 5) had more than 50% of students who passed the Illinois State Achievement Test (ISAT), and 6) had more than 50% of students who received free lunch. The latter two criteria ensured the selection of high-risk schools. Sixty-eight schools met all of the above criteria, and 18 of those agreed to participate in the study with the understanding that they would have to be successfully matched with another school and randomly assigned to condition.

The following variables from the 2003-2004 CPS data were used as matching variables: percentage of White, African American, Hispanic, and Asian students; percentage of students who met or exceeded criteria for passing the state achievement test; attendance rate; truancy rate; percentage of students who received a free or reduced-price lunch; percentage of students who enrolled or left school during the school year; number of students per grade; percentage of parents who were involved with school activities; and percentage of teachers employed by the school who met minimal teaching standards. We also used information about the crime rate in each school's neighborhood (Chicago Police Department, n.d.). Using these variables and an additional requirement that each school in a pair be located in the same region of the city, the 18 schools were successfully matched into nine pairs. Schools were matched using a SAS program provided by Mathematica Policy Research (MPR; Schochet & Novak, 2003). The matching variables listed above were entered into the SAS matching program, as well as the school's region code so that schools in similar regions could be matched. The SAS program matched schools into pairs based on a "distance matrix method", and the best fit-squared distance matrix was utilized. The request for application (RFA; #NCER-03-06) requested five pairs for the trial, and final funding allowed for seven pairs. The seven best-matched pairs, which also happened to provide the most accurate representation of the ethnic composition at CPS schools, were recruited for participation. A random-number generation function in the Microsoft Excel program was used by the principal investigator (P.I) to randomize the schools within each pair to control and treatment conditions. Allocation to condition

was blinded until assigned by the P.I. Because of the nature of the intervention, schools, students, those delivering the intervention, and the outcome assessors could not be blinded to condition. See Figure 2.1 for a diagram of school recruitment.

<Figure 2.1 about here>

Student-Level Recruitment

Parental consent was obtained before students, parents or teachers completed surveys when students were in grade 3. Seventy-nine percent of parents provided consent at baseline and ranged from 65% to 78% for Waves 2-5. Students joining the study at later waves were consented at that time. All students were re-consented for the second phase of funding at Wave 6 (beginning of grade 7). Consent rates were lower at Waves 6 through 8 (~58 to 64%) than Waves 1 through 5. This is consistent with previous studies that have found that consent rates drop as grade levels increase (Ji, Pokorny, & Jason, 2004; Thompson, 1984).

Students signed assent forms at the point of entry to the study. The assent forms described that the student was agreeing to participate in the study for multiple years; therefore, signed assent forms were considered to be agreement to participate for the entire (or remaining) study period.

The percentages of consenting parents who provided reports on their children were 72.3% at Wave 1, 58.9% at Wave 2, 52.2% at Wave 4, 50.5% at Wave 5, and 72.9% at Wave 8. Two factors that likely increased parent response rate at Wave 8 were (1) an increase in the financial incentive for completing the parent report and (2) an intensive period of phone outreach to families to note the incentive increase and to

encourage survey completion. Percentages of consented students for whom teachers completed ratings were 74.6%, 74.8%, 72.4%, 78.3%, 74.4%, and 92.7% for Waves 1, 2, 4, 5, 7, and 8, respectively. At Wave 8, we introduced an additional school-level incentive for 100% rates of teacher survey completion, which likely resulted in the increase in completed teacher ratings.

Some adjustments to the data set were necessary before analyses. Students who were new at Waves 3 and 6 (at beginning of grade 4 or 7) are considered new at the following wave (i.e., wave 4 or 7, at the end of grade 4 or 7, respectively) because they could not have been exposed to *PA* the prior year. Additionally, students who, during Waves 6, 7, and 8, attended the middle school, but had not attended the feeder elementary school that was in the trial, were not included in the analysis sample for two reasons: the elementary schools these children did attend (a) were not study schools and (b) were demographically different and higher-risk than study schools. Finally, a small number of students (n=30) on whom only height and weight data were collected at Wave 8 were removed from the analysis sample for all other analyses, as they did not have data on any other covariates or outcomes.

Measures

A total of 56 student, 19 teacher, and 13 parent (or primary caregiver) scales were used to assess a variety of outcomes throughout the duration of the CRCT. A portion of the first five waves of data were collected as part of a multisite study on social and character development programs conducted by the Institute of Education Sciences (IES) and MPR. The research team at the University of Illinois at Chicago

(UIC) collected data on other outcomes for waves 1-5 and on all of the scales for the final 3 waves.

Collectively, student-report measures included questions about social-emotional and character development (e.g., prosocial interactions, honesty, self-control, respect for teachers and parents), internalizing behaviors (e.g., anxiety, positive and negative affect), externalizing behaviors (e.g., bullying, substance use, delinquent behaviors, beliefs about aggression), self-esteem, academic achievement, and perceptions of school climate. Questions about substance use, violence, depression, and anxiety were asked of students starting in grade 5 (Waves 5 through 8). Parent and teacher reports included measures of their perceptions of a student's SECD, level of responsibility, social competency, aggression, conduct, and academic performance and motivation. Additional measures assessed parental involvement with student's teacher and school, parenting skills, household confusion/disorder and perceptions of community risks and resources. Teacher reports were collected at all waves except Wave 6 and parent reports were collected at all waves except Waves 6 and 7. Questionnaires were written for students in a range of grades (from 3rd to 8th); therefore, some censoring or "floor" and "ceiling" effects (restricted range) may be observed for some of these measures. We will use different estimators to appropriately address outcomes that violate the assumption of normality. This will be accounted for in the analytic model (to be described in detail below).

In addition to survey data from students, parents, and teachers, this CRCT will examine intervention effects on school archival records, including absenteeism,

disciplinary referrals and suspensions, and reading and math standardized-test scores. These data, reported at the end of each academic year, were obtained from the CPS website once they became public. For this study, we are utilizing data from all study years, as well as several academic years prior to the present trial to establish a reliable baseline.

Data Collection Procedures

Student surveys were administered during class time. During Waves 1 through 5, research staff read survey instructions and items aloud; beginning at Wave 6, students read the individual survey items themselves, although complicated items (e.g., items with double negatives) were explained. Research staff members were also available to assist students who had questions or who appeared to have difficulty taking the survey. Student ID numbers were pre-printed on each page of the survey; but student names were printed only on the front page. To maintain confidentiality, this front-page sheet was removed upon return of completed surveys by students, and eventually discarded/shredded.

Parents completed surveys at home and returned them to the school in a sealed envelope. For surveys not returned to the school within 3 or 4 weeks, the MPR research team attempted to get the parent to complete the survey via telephone interview. Teacher surveys were also self-administered. Parents and teachers, who were offered a small financial incentive to compensate them for their time, returned their surveys to a central point in the school for collection by the research team.

Human Subjects Approvals

This trial was approved by Institutional Review Boards at the University of Illinois at Chicago and Oregon State University, the Research Review Board at Chicago Public Schools, and the Public/Private Ventures Institutional Review Board for MPR.

Planned Statistical Analyses

Scale Properties. Scale alphas at all waves and test-retest reliability at Wave 5 were run for each outcome. Indices of clustering were also computed; intraclass correlations (ICC) for scales and median odds ratios (MOR; Merlo et al., 2006) for dichotomous measures were calculated.

Baseline equivalency. A series of independent-groups *t*-tests were conducted for each school demographic variable to determine whether the treatment and control schools were significantly different from each other. These tests were conducted pre-baseline, at baseline, and twice more during the trial. We also tested for differences in baseline scores on the 50 student-, 11 teacher- and 13 parent-reported scales.

Growth-Curve Models. We propose a three-level (occasions of measurement nested within students nested within schools) growth-curve model for analyzing treatment effects on various outcomes; analyses will be conducted using Stata. These models will account for all observations and model school differences (Rabe-Hesketh & Skrondal, 2008). This approach allows for a complete analysis of the multiple waves of available data and takes into account the patterns of change over time. When outcomes consist of more than one item, an average of the scores will be used to create a composite score, with higher scores reflecting having more of an outcome. The

distribution of each scale will be examined in order to determine the most appropriate model. Outcomes may have a normal distribution, be censored (skewed right or left, indicative of floor or ceiling effects), be a count, or be bimodal, or a dichotomy. Bimodal outcomes will be transformed into a dichotomy. Each outcome will be assessed using the most appropriate estimator to fit the distribution.

Random-intercept, linear growth-curve models will first be estimated. This will be followed by tests for whether there is a significant quadratic trend over time and whether there is significant variation in the time slope across students (i.e., a random-coefficient model). The basic random-intercept linear model for normally distributed outcomes can be expressed as:

$$\hat{Y}_{tij} = \beta_0 + \beta_1(\text{condition}_j) + \beta_2(\text{time}_{tij}) + \beta_3(\text{condition}_j \times \text{time}_{tij}) + \zeta_j + \zeta_{ij} + \varepsilon_{tij}$$

The subscripts represent the three-level structure of the data. Since these models for cluster-randomized trials are well-known (e.g., Murray, Feldman, McGovern, 2000; Raudenbush & Bryk, 2002), the presentation here is streamlined. \hat{Y}_{tij} is the estimated score on the outcome at time t (Waves 1-8), for student i , in school j . Condition is a binary variable at the school level (level-3), with 0 indicating control status and 1 indicating receipt of *PA*. Time is measured as years of implementation, a total of 5.58 years by the end of grade 8. The effect of the intervention will be tested with the statistical significance of β_3 . This will indicate if the students in the *PA* schools change differently over time than those in the control schools in the predicted direction. The ζ_j represents the deviation of a school's mean score from the mean score for all schools. The ζ_{ij} represents the deviation of each student's score from their school's mean. These

two random effects represent the unmeasured tendencies for schools and students to score higher or lower on the outcome. The ε_{tij} represents the residual at each wave. An additional term, $\zeta_{1ij}(\text{time}_{tij})$ will be added in the model to test whether there is significant variation in student change across time, rather than all students in each condition having the same change pattern (i.e., a random-coefficient model). Similarly, additional terms, $\beta_4(\text{time}_{tij}^2)$ and $\beta_5(\text{condition} \times \text{time}_{tij}^2)$, will be used to test whether there is a quadratic time trend, indicating acceleration or deceleration in change over time, and whether *PA* influenced that pattern. These terms will be removed from the final model in the interest of parsimony if insignificant. If the $\text{condition} \times \text{time}_{tij}^2$ term is significant, we will rerun the model with the intercept set at Wave 8 to test for a significant condition effect. This will tell us if the difference at endpoint (Wave 8) is significantly different between *PA* and Control. Since the random-intercept model is nested within the random-coefficients model, a Likelihood Ratio Chi-square (LR) test will be used to compare model fit (Rabe-Hesketh & Skrondal, 2008) with and without the random coefficient. If a model with a random time coefficient provides a significantly better fit for a given outcome, it will be reported as the final model. The multilevel models for outcomes that are binary or count have a similar structure but, due to their nature, have no residual at level 1 (Brown et al., 2009).

Because only 14 schools are in this trial, and the *PA* effect is tested at the school level in a cluster-randomized trial, we will conduct several sensitivity analyses. First, we will assess the statistical significance of the *PA* coefficient estimate and its standard error using the *t*-distribution with 12 degrees of freedom: 14 schools – 1 (the

condition effect) – 1 = 12 df (Raudenbush & Bryk, 2002) providing for a more conservative approach (from the perspective of statistical power) than the z -distribution provided by Stata. A second approach will be the pair-level analysis, estimated as a four-level model: occasions of measurement nested within students, nested within schools, nested within matched pairs (e.g., Brown et al., 2009), with the linear random-intercept model expressed in simplified notation as:

$$\hat{Y}_{tijk} = \beta_0 + \beta_1(\text{condition}_{jk}) + \beta_2(\text{time}_{tijk}) + \beta_3(\text{condition}_{jk} \times \text{time}_{tijk}) + \zeta_k + \zeta_{jk} + \zeta_{ij} + \varepsilon_{tijk}$$

where the t , i , and j subscripts representing the time, student, and school clusters as in Equation 1 above, with k representing the 7 matched pairs of schools. The ζ_k term represents deviations of each matched pair from the overall mean. As described above, additional terms will be tested to determine if there is significant variation in the time trend across students (random-coefficient model) and whether there is a significant quadratic pattern.

Intervention effects on scales collected only at later waves (Waves 5 or 6 onwards) will be tested in two ways. First, the models will be ran with the intercept set at the first time the scale was measured (Wave 5 or 6) to test if program effects had already occurred by Wave 5 or 6. Second, the model will be run again with the intercept set at Wave 8 to test if program effects are (still) significant at Wave 8. In either analysis, a significant ($p < 0.05$) condition term will indicate a *PA* effect as measurement occurs long after implementation, and a significant condition by time interaction will indicate whether the effect changed in magnitude between the first and

later measures. Each of these analyses will include a related school-level covariate that was assessed at baseline, which will be grand mean-centered.

In addition to the student-level survey data, several school-level archival measures will be analyzed. These data will be analyzed similarly to cohort data but time will be measured in equal, one-unit increments (e.g., 1, 2, 3, etc.) rather than in years since implementation, as these data were collected at the same time each year – and summarize performance for the whole school year. Additionally, because these data are at the school level, the growth-curve models will be two-level (observations within schools) rather than three-level. Because of the small amount of data (the number of schools times the number of waves) and the resulting power limitations, these analyses will use the random-intercept model only. Due to the power and sample size limitations, and because the *a priori* hypothesis was that the *PA* schools would have greater improvements across time, *one-tailed p-values* were used in tests of effects of the *PA* program on school-level outcomes (Knotterus & Bouter, 2001).

Mediation and moderation. Mediation analyses will allow us to examine the *PA* program's mechanisms of action. To test for mediation, we will use the framework described by Baron and Kenny (1986) and MacKinnon (2000; 2002; 2008). We will first estimate the bivariate effect of X on Y without the mediator included in the model. Then, we will simultaneously estimate the direct effect of X on Y with the mediator included in the model, as well as the mediated effect, which consists of the effect of X on M \times M on Y (MacKinnon, 2008).

In addition, we will test for moderation (Baron & Kenny, 1986) by gender and by student mobility. The moderation tests will reveal for whom the program has its effects; that is, these tests will allow us to assess whether program effects differ by gender or a child's mobility. We will not test for moderation by ethnicity because it is highly confounded with school, with 3 pairs of schools having a mostly African-American enrollment and 2 pairs of schools having a mostly Hispanic enrollment.

While all 14 schools were retained throughout the CRCT, the student population in this trial was highly mobile. Mobility poses a challenge to efficacy trials of school-based programs in these kinds of schools, and it can be difficult to infer whether the observed effects are because of the intervention or differential attrition (Vuchinich et al., 2012). Thus, it is important to test for potential moderating effects of student mobility patterns. A recent approach to analyzing mobility patterns is latent class analysis (LCA; Beunckens, Molenberghs, Verbeke, & Mallinckrodt, 2008; Lin, McCulloch, & Rosenheck, 2004; Roy, 2003). LCA will be used to identify subgroups of individuals on unobservable factors (Marsh, Lüdtke, Trautwein, & Morin, 2009) and will allow for the identification of classes of students with similar mobility patterns. These patterns can then be tested as a moderator of program effects; that is, examining whether students with different mobility patterns have different program effects.

RESULTS

Sample Characteristics

Table 2.1 shows the number of students present at each wave, as well as the number of students that remained in later waves of the trial. One hundred percent of the schools remained throughout the study; that is, there was no attrition at the school level. At the student-level, the total sample size across all eight waves was 1170, with approximately 21% (131) of the original 624 grade 3 cohort remaining at grade 8, illustrating the high mobility by low-income urban students, a trend that has been noted in prior studies in this population (e.g., Tobler & Komro, 2011). Additionally, parental consent rates were lower during the middle-school grades and Chicago school enrollment was decreasing; therefore the student sample size at Wave 8 (363) was smaller than at Wave 1 (624). The average number of waves of data per child was 3.1.

<Table 2.1 about here>

Of the students, 53% were female, 48% were African American, 27% Hispanic, and 19% “Other” (e.g., White, Asian, and Native American, as well as self-described responses). Approximately 6% of the sample never reported, nor had a parent report, their ethnicity. At Wave 1, students were, on average, 8.39 years of age ($SD=0.61$). Fifty-two percent of the students over the eight waves were in treatment schools, and 48% were in control schools.

Of the teachers, 75% were female; 43% were White, 36% African American, 13% Hispanic, and 8% “Other” (e.g., Asian and Native American). Fifty-six percent had a Master's degree, 37% a Bachelor's degree, 5% a specialist degree, 1% a doctorate degree, and 1% listed "other" as their highest level of education. For teacher certification, 83% had a regular certificate, 7% a probationary certificate, 3% a

provisional certificate, 1% a temporary certificate, 1% had no certification, and 5% had other certifications (such as administrative, behavior specialist, special education, or counseling).

Scale Properties

Table 2.2 (Online Supplement) provides a list of scales, internal consistency reliability (using Cronbach's alpha), stability reliability (using Pearson's r correlation for test-retest data) at Wave 5, and ICC or MOR for each scale. Alpha for all outcomes ranged from 0.53 to 0.98. For student-reported outcomes, alphas generally increased as students got older, and were relatively high for most measures. Alphas for teacher- and parent-reported outcomes were also consistently high. There was no discernable pattern of increase or decrease over time for ICCs or MORs. ICCs and MORs were generally low (ICCs mostly $<.05$, MORs mostly <1.5) for student reports and even lower for parent reports, and highest for teacher reports.

Baseline Equivalency

Baseline school equivalency tests found no statistically significant differences between the treatment and control group schools on any of the matching variables at any of the four times tested (see Table 2.3). Only four student-reported scales (out of the 50 available at baseline) showed significant differences, all of which favored control school students. Students in control schools had higher baseline scores than students in *PA* schools on three scales: Belief in Moral Center-Positive Values ($t=2.20$, $p<0.05$), perceived Rewards for Prosocial Behaviors from parents ($t=3.50$, $p<0.001$), and perceived Rewards for Prosocial Behaviors from teachers ($t=2.22$, $p<0.05$); and

significantly lower scores than students in *PA* schools on Peer Group Affiliation with Bad Friends ($t=-2.96, p<0.01$). Of the 11 teacher-reported outcomes available at baseline, three showed significant differences, all favoring *PA* students. Teachers in *PA* schools reported students as having fewer conduct problems ($t=2.90, p<0.01$), higher altruism ($t=-5.13, p<0.001$), and higher parent communication ($t=-2.52, p<0.05$). Of the 13 parent-reported outcomes, one showed significant differences by condition favoring *PA* students; parents of children in *PA* schools reported fewer community risks than parents of children in control schools ($t=4.39, p<0.001$). Thus, of 74 baseline measures, only 8 showed significant differences and, of these, half favored control students and half favored *PA* students.

<Table 2.3 about here>

DISCUSSION

PA is a comprehensive, school-wide, SECD program designed to increase positive youth development outcomes such as self-esteem, character, morality, problem solving, and empathy, and to also decrease negative externalizing behaviors such as substance use and violence, and internalizing symptoms such as depression and anxiety. The Chicago trial of *PA* utilized a longitudinal design that aimed to assess the effectiveness of *PA* in a low-income, urban setting. Additionally, it is the first trial of *PA* that included students in middle-school grades. Effectiveness of the program will be assessed using multilevel, growth-curve analyses. Mediation analyses will investigate the mechanisms by which the program works and moderation models will assess whether program effects differ by gender or mobility class.

Equivalency tests at the school-level revealed no significant differences at four time points before and during the trial. At the student level, 74 student-, parent-, and teacher-reported scales were tested for baseline equivalency, with only 4 of 50 student-report scales favoring control students, and 3 of 11 teacher-report scales and 1 of 13 parent-report scales favoring *PA* students. The facts that the number of statistically significant differences was low, and that they were in opposite directions, suggest minimal threats to internal validity. Random assignment (from matched pairs) of clusters (schools) cannot guarantee equivalence of the nested subjects (in this case, students; Giraudeau & Rivaud, 2009), so these minimal baseline differences that did not consistently favor one group or the other indicate that the *a priori* matching and randomization were successful.

The design, methodology, and results of this trial should be viewed in the context of some limitations. Most outcomes were assessed through student self-report, potentially leading to a method bias (Podsakoff, MacKenzie, & Podsakoff, 2003), which can inflate the observed relationships between the variables. Self-reports are also susceptible to social desirability bias; students may exaggerate their participation in high-risk behaviors in order to feel as if they fit in with their peers, or underreport such behaviors knowing society's negative views on behaviors such as substance use, violence, and bullying. However, the collection of teacher and parent ratings of student behavior should mitigate this limitation. Additionally, the sample for this study was U.S. youth from low-income, urban environments; therefore, the results cannot be generalized beyond this demographic group. Demonstrating program

effects in a high-risk population, however, would be beneficial, given that these are the populations most in need of such programs.

Lessons Learned

We learned some lessons in the process of designing this trial, collecting the data, and planning the analyses. Recruiting schools to participate required a large amount of work (e.g., obtaining archival data for eligibility, visiting schools, holding meetings) before the trial could begin. Investigators need to be prepared to handle the disappointment of principals assigned to the control condition. One way to reduce tension would be to reiterate the benefits of the research agreement, such as receiving *PA* (including training) in three years and an annual stipend (Ji et al., 2008). In order to get high parental consent rates, the use of incentives was extremely useful, encouraging parents to respond who may not have otherwise. In addition, future investigators should be aware that extra efforts (time and money) became necessary as the students got older.

This trial was conducted in a low-income, urban setting with a highly mobile population. This population challenged us to pay close attention to possible mobility effects - do program effects differ by students' mobility patterns? This obstacle gave us the opportunity to explore methods for dealing with mobility and to propose the use of advanced statistical techniques such as latent class analysis to identify such mobility patterns and test for differences in program effectiveness among these patterns (Vuchinich et al., 2012).

The young age of the students at the beginning of the trial and the longitudinal nature of the study presented some measurement challenges, in that some outcomes have restricted ranges or show "floor" or "ceiling" effects. To account for this challenge, we will employ different estimators that are most appropriate for the distribution of each outcome. Finally, it has been acknowledged from the start of this study (Social and Character Development Research Consortium, 2010) that this trial was underpowered at the school-level. That is, current literature suggests that a sample size of seven school pairs would usually not be adequate for multilevel modeling. Funders should be aware of this issue when planning future funding opportunities.

Despite the issues we faced with this trial, there are several strengths. This trial utilizes a "cluster-randomized with intent-to-treat" design, allowing for the assessment of program effects on students within clusters (schools). This analysis is the most appropriate for school level intervention, particularly in schools with naturally occurring mobility (Vuchinich et al., 2012). Additionally, the longitudinal nature of this trial allowed examination of students across elementary and secondary grades, and will provide the first experimental test of *PA* effects in middle-school grades. The design will also provide for temporal ordering of *PA* as a cause on various outcomes when testing for mediation effects. This design also avoids many issues found in earlier school-based RCT's (Flay & Collins, 2005) and meets most of the standards put forth by the Society of Prevention Research (Flay et al., 2005). This paper proposed a novel approach to handling mobility in trials with such designs,

using latent class analysis to identify patterns of mobility, then to test these patterns for moderation effects of the program. Moreover, this trial involves a sample of students in a high-risk setting; generating improvements can be particularly difficult in urban areas facing rising poverty rates (DeNavas-Walt, Proctor, & Smith, 2009), health disparities (Braveman & Egerter, 2008) and cuts in social and educational programs (Johnson, Oliff, & Williams, 2010).

The current paper described the CRCT of *PA* in Chicago schools. Future papers forthcoming from this trial will examine the impact of *PA* on academic, social-emotional, problem behaviors, physical and mental health, and environmental outcomes, allowing for the examination of long-term program effects on such outcomes. Additionally, other papers will test for mediation in order to gain a better understanding of the mechanism of *PA*; that is, how does *PA* work. Future trials by independent evaluators of equal or better quality are needed for this program.

Table 2.1: Mobility of students by condition

	Wave	1	2	3	4	5	6	7	8
	Season/Year	Fall 2004	Spring 2005	Fall 2005	Spring 2006	Spring 2007	Fall 2008	Spring 2009	Spring 2010
	Grade	3	3	4	4	5	7	7	8
Treatment									
# of Participating Students		316	306	244	297	262	104	188	195
# of Joiners			39	0	77	47	0	83	50
# of Leavers			49	62	24	82	158	0	43
Control									
# of Participating Students		308	299	220	244	253	92	172	168
# of Joiners			27	0	55	60	0	81	26
# of Leavers			36	79	31	51	161	1	30
Totals		624	605	464	541	515	196	360	363
# of Joiners			66	0	132	107	0	164	76
# of Leavers			85	141	55	133	319	1	73

Note: The increase in mobility rates after Wave 5 may be partially explained by the time difference between Wave 5 and Wave 6 representing one school year plus two summer breaks of mobility (end of grade 5 to beginning of grade 7) and the transition from elementary to middle-school grades. Joiners in the Fall of 2005 and 2008 were considered as joiners at the end of the school year (Spring 2006 and 2009, respectively). Table shows the analysis sample. There may be small variations between this table and tables in papers (and reports from IES and the SACD Research Consortium) because some students may have only been present for the "multisite" or "site-specific" days of data collection. Additionally, some students may not have been present for any data collection and therefore only have teacher-reported data.

Table 2.2: (Online Supplement) Student, Parent, and Teacher-reported Scale Properties

<i>Student Self-Report Scale</i>	<i>Source</i>	<i># of items</i>	<i>Coefficient alphas: w1/w2/w3/w4/w5/ w6/ w7/w8</i>	<i>Test-retest reliability:w5</i>	<i>ICCs or MORs^a: w1/w2/w3/w4/w5/w6/ w7/w8</i>
Multisite Scales					
Positive School Orientation	Kaminski et al., 2009	10	.87 / .86 / .89 / .88 / .89 / .86 / .87 / .88	--	.08 / .16 / .26 / .25 / .30 / .28 / .18 / .16
Feelings of Safety	IES ^b	4	.72 / .77 / .80 / .83 / .84 / .85 / .83 / .86	--	.03 / .01 / .04 / .04 / .06 / .00 / .01 / .01
Normative Beliefs about Aggression ^a	Huesmann & Guerra, 1997	8	.81 / .89 / .92 / .92 / .90 / .93 / .91 / .93	--	1.47 / 1.18 / 1.08 / 1.37 / 1.16 / 1.85 / 1.73 / 1.72
Children Empathy Questionnaire	Funk et al., 2003	16	.79 / .86 / .86 / .88 / .87 / .87 / .89 / .87	--	.02 / .05 / .07 / .09 / .10 / .19 / .06 / .13
Engagement vs. Disaffection with Learning – Positive	Furrer & Skinner, 2003	5	.81 / .72 / .80 / .71 / .82 / .86 / .81 / .83	--	.04 / .00 / .01 / .02 / .07 / .05 / .04 / .05
Engagement vs. Disaffection with Learning – Negative		5	.53 / .54 / .55 / .56 / .55 / .64 / .62 / .68	--	.01 / .01 / .02 / .04 / .05 / .01 / .04 / .00
Self-Efficacy for Peer Interaction	Wheeler & Ladd, 1982	12	.81 / .84 / .85 / .86 / .87 / .90 / .89 / .89	--	.00 / .06 / .09 / .06 / .03 / .08 / .05 / .07
Altruism Scale	Solomon et al., 2000	8	.83 / .86 / .86 / .87 / .86 / .81 / .86 / .87	--	.02 / .02 / .03 / .04 / .00 / .03 / .00 / .02
Victimization Scale	Orpinas & Kelder, 1995	6	.83 / .85 / .86 / .85 / .86 / .86 / .87 / .84	--	.01 / .01 / .02 / .06 / .03 / .00 / .00 / .03
Frequency of Disruptive Behavior ^a	Loeber & Dishion, 1983	6	.77 / .77 / .79 / .77 / .81 / .77 / .78 / .76	--	1.00 / 1.57 / 1.50 / 1.82 / 1.48 / 1.69 / 1.42 / 1.38
Bullying Scale ^a	Orpinas & Frankowski, 2001	6	.81 / .86 / .90 / .89 / .88 / .86 / .87 / .83	--	1.47 / 1.66 / 1.90 / 1.96 / 1.96 / 2.36 / 1.93 / 1.38
Site-Specific Scales					
Inventory of School Climate ^c	Brand, 2003	9	.76 / .87 / .88	--	.00 / .06 / .07
MESA (Stress) ^c	Gonzales et al. (n.d.)	33	.73 / .80 / .76	--	.11 / .01 / .04
Emotional Intelligence ^d	Petrides et al. 2009	30	.69 / .84	--	.01 / .02
Rewards for Prosocial Behavior - Parent	Arthur et al. 2002	3	.72 / .76 / .79 / .82 / .84 / .87 / .89 / .91	.68	.02 / .01 / .03 / .03 / .03 / .02 / .06 / .04
Rewards for Prosocial Behavior - Teacher		3	.80 / .83 / .85 / .87 / .89 / .88 / .92 / .92	.67	.05 / .06 / .07 / .09 / .07 / .00 / .07 / .09
Student Perception of Neighborhood Context	Chipuer et al. 1999	9	.70 / .77 / .74 / .77 / .79 / .75 / .73 / .76	.86	.05 / .05 / .07 / .08 / .06 / .09 / .05 / .06
Student Attachment to School ^a	Cook et al. 1995	4	.75 / .81 / .85 / .89 / .83 / .81 / .82 / .79	.80	1.48 / 1.33 / 1.68 / 2.32 / 2.11 / 2.11 / 1.62 / 1.85
Student Attachment to Parents ^a		4	.62 / .76 / .77 / .79 / .77 / .86 / .85 / .88	.76	1.18 / 1.45 / 1.20 / 1.00 / 1.26 / 1.20 / 1.40 / 1.10
Student Attachment to Teacher ^a		4	.71 / .80 / .87 / .86 / .86 / .87 / .87 / .84	.78	1.48 / 1.55 / 1.66 / 1.56 / 1.44 / 2.02 / 1.76 / 1.80
Student Attachment to Friends ^a		4	.74 / .79 / .79 / .81 / .83 / .87 / .81 / .83	.76	1.03 / 1.00 / 1.49 / 1.35 / 1.47 / 1.44 / 1.38 / 1.32
Peer Group Affiliation – Good Friends	Elliott et al. 1996	3	.60 / .60 / .60 / .68 / .64 / .69 / .69 / .73	.50	.05 / .06 / .05 / .04 / .06 / .07 / .01 / .05
Peer Group Affiliation – Bad Friends		4	.84 / .85 / .86 / .85 / .83 / .87 / .85 / .81	.60	.08 / .12 / .13 / .18 / .13 / .22 / .07 / .09
Belief in the Moral Order - Positive Values	Arthur et al. 2002	6	.70 / .74 / .74 / .76 / .81 / .81 / .83 / .81	.55	.02 / .02 / .02 / .03 / .06 / .15 / .03 / .02
Belief in the Moral Order - Negative Values		5	.61 / .75 / .75 / .77 / .75 / .73 / .75 / .79	.62	.04 / .07 / .14 / .14 / .10 / .27 / .15 / .07
Self-Esteem Questionnaire (SEQ)	DuBois et al., 1996	4	.72 / .77 / .80 / .75 / .76 / .80 / .80 / .79	.72	.02 / .01 / .00 / .04 / .03 / .10 / .03 / .03

Peer					
SEQ – School		4	.69 / .76 / .79 / .77 / .78 / .84 / .85 / .85	.85	.01 / .02 / .02 / .03 / .05 / .09 / .05 / .05
SEQ – Family		4	.76 / .76 / .77 / .77 / .82 / .82 / .85 / .86	.74	.00 / .02 / .01 / .03 / .02 / .08 / .06 / .04
SEQ – Appearance		4	.70 / .73 / .76 / .77 / .79 / .81 / .79 / .82	.76	.02 / .04 / .05 / .03 / .05 / .13 / .11 / .13
SEQ – Sports		4	.73 / .72 / .78 / .79 / .82 / .86 / .88 / .88	.87	.01 / .02 / .01 / .02 / .00 / .08 / .02 / .02
SEQ – Global		8	.68 / .71 / .77 / .76 / .82 / .85 / .83 / .83	.81	.01 / .02 / .04 / .04 / .04 / .13 / .12 / .13
Self-Esteem Motivation	DuBois & Flay, 2003a	3	.73 / .79 / .82 / .82 / .82 / .86 / .84 / .88	.75	.01 / .01 / .01 / .03 / .05 / .09 / .05 / .08
Self-Esteem Formation-Adaptive	DuBois et al., 2003	7	.74 / .79 / .76 / .79 / .77 / .83 / .82 / .84	.81	.02 / .03 / .02 / .03 / .06 / .06 / .03 / .05
Self-Esteem Formation-Maladaptive		13	.78 / .79 / .81 / .81 / .81 / .76 / .81 / .82	.53	.02 / .04 / .04 / .04 / .06 / .06 / .08 / .00
Aggressive Problem Solving ^a	Aber et al. 1995	5	.79 / .84 / .83 / .84 / .81 / .78 / .69 / .65	.66	1.69 / 1.73 / 2.32 / 2.28 / 1.89 / 2.46 / 1.59 / 1.67
Competence Problem Solving ^a		5	.59 / .69 / .69 / .71 / .67 / .67 / .62 / .57	.81	1.30 / 1.90 / 1.78 / 1.71 / 1.63 / 1.78 / 1.00 / 1.71
Social-Emotional and Character Development Scale (SECD) ^e	Ji et al. 2011	6	.88 / .90 / .90 / .91 / .90 / .90 / .90 / .92	.79	.04 / .05 / .03 / .05 / .07 / .10 / .03 / .04
SECD – Prosocial Interactions ^e		6	.77 / .82 / .83 / .83 / .85 / .85 / .83 / .85	.76	.02 / .04 / .02 / .09 / .08 / .18 / .02 / .04
SECD – Self Development ^e		4	.64 / .74 / .77 / .79 / .78 / .86 / .81 / .83	.62	.01 / .05 / .00 / .00 / .04 / .06 / .06 / .06
SECD – Honesty with self and others ^e		5	.71 / .79 / .82 / .80 / .81 / .82 / .80 / .80	.75	.05 / .03 / .01 / .05 / .05 / .10 / .03 / .04
SECD – Self-Control ^e		4	.64 / .67 / .69 / .74 / .76 / .82 / .72 / .73	.63	.01 / .05 / .04 / .08 / .05 / .13 / .04 / .05
SECD – Respect for Teachers ^e		5	.71 / .83 / .84 / .86 / .88 / .90 / .90 / .90	.78	.04 / .06 / .06 / .08 / .08 / .11 / .04 / .05
SECD – Respect for Parents ^e	Laurent et al. 1999	4	.61 / .71 / .73 / .78 / .81 / .86 / .87 / .89	.73	.02 / .02 / .03 / .02 / .06 / .02 / .07 / .05
Positive Affect Scale for Children		6	.75 / .70 / .80 / .83 / .84 / .87 / .87 / .86	.80	.02 / .07 / .07 / .03 / .04 / .06 / .06 / .07
Negative Affect Scale for Children		6	.74 / .80 / .76 / .74 / .76 / .72 / .71 / .74	.70	.01 / .02 / .04 / .01 / .01 / .00 / .01 / .03
Student's Life Satisfaction Scale	Huebner, 1991	3	.71 / .79 / .77 / .79 / .80 / .83 / .82 / .84	.85	.00 / .02 / .02 / .00 / .03 / .01 / .01 / .03
BASC Depression Scale ^{af}	Reynolds & Kamphaus, 2002	6	.79 / .70 / .78 / .78	.76	1.21 / 1.00 / 1.22 / 1.20
BASC Anxiety Scale ^{af}		6	.77 / .81 / .75 / .79	.60	1.25 / 1.47 / 1.00 / 1.00
Positive Health Behavior, Positive Food & Exercise Index	Positive Action Inc., 2007	3	N/A	.68	.02 / .00 / .00 / .01 / .01 / .05 / .02 / .01
Positive Health Behavior-Negative Food Index		3	N/A	.74	.14 / .11 / .13 / .18 / .16 / .22 / .14 / .07
Positive Health Behavior, Hygiene Index		3	N/A	.76	.01 / .02 / .01 / .01 / .02 / .01 / .01 / .02
Positive Health Behavior, Sleep Index		1	N/A	.75	.01 / .04 / .03 / .04 / .03 / .01 / .02 / .03
Substance Use Index ^{a,f,g}	CDC, 2004	5	N/A	.58	.04 / .09 / .05 / .03
Violence Index ^{a,f,g}		6	N/A	.99	1.76 / 2.55 / 1.54 / 2.03
Positive Action/Positive Feeling Scale	Positive Action Inc., 2007	8	.71 / .67 / .73 / .75 / .78 / .82 / .82 / .82	.86	.02 / .02 / .03 / .04 / .04 / .03 / .09 / .08
Parent-Report Scales			Coefficient alphas: w1/w2/w3/w4/w5/w8		ICCs or MORs^d: w1/w2/w3/w4/w5/w8
Parent and Teacher Involvement	CPPRG, 1991	9	.63 / .64 / .69 / .75 / .73 / .75	--	.00 / .04 / .04 / .05 / .02 / .10
Social Competence	CPPRG, 1999	19	.89 / .90 / .91 / .90 / .90 / .92	--	.05 / .03 / .04 / .06 / .07 / .00
Responsibility	IES ^b	8	.80 / .82 / .82 / .81 / .82 / .83	--	.05 / .04 / .06 / .06 / .05 / .04
Bullying ^a		10	.70 / .79 / .83 / .81 / .83 / .84	--	1.00 / 1.00 / 1.00 / 1.22 / 1.49 / 1.00

Conduct Problems ^a	Reynolds & Kamphaus, 1998	7	.75 / .74 / .74 / .81 / .74 / .81	--	1.64 / 1.74 / 1.33 / 1.54 / 1.53 / 1.16
Altruism	Solomon et al., 2000	8	.86 / .87 / .90 / .91 / .88 / .91	--	.03 / .00 / .00 / .00 / .05 / .00
Alabama Parenting Questionnaire-Positive	Shelton et al., 1996	6	.79 / .79 / .82 / .81 / .83 / .84	--	.01 / .01 / .00 / .00 / .00 / .01
Alabama Parenting Questionnaire-Negative		10	.66 / .54 / .60 / .71 / .67 / .76	--	.00 / .02 / .00 / .02 / .00 / .00
Confusion, Hubbub, and Order	Matheny et al., 1995	14	.79 / .79 / .81 / .81 / .81 / .81	--	.00 / .04 / .00 / .00 / .01 / .00
Community Risks	Forehand et al., 2000	7	.89 / .87 / .88 / .87 / .86 / .86	--	.20 / .16 / .17 / .18 / .24 / .09
Intergenerational Closure	Sampson et al., 1999	3	.73 / .77 / .74 / .74 / .75 / .69	--	.01 / .03 / .02 / .03 / .02 / .10
Community Resources	IES ^b	5	.74 / .76 / .75 / .76 / .75 / .72	--	.06 / .10 / .06 / .08 / .10 / .11
Child-Centered Social Control	Sampson et al., 1999	5	.85 / .86 / .86 / .84 / .85 / .89	--	.01 / .00 / .05 / .01 / .03 / .00
Teacher-Report Scales			Coefficient alphas: w1/w2/w3/w4/w5/w7/w8		ICCs or MORs^d: w1/w2/w3/w4/w5/w7/w8
Social Competence	CPPRG, 1999	19	.96 / .97 / .97 / .96 / .97 / .97 / .96	--	.10 / .17 / .22 / .16 / .17 / .25 / .15
Responsibility	IES ^b	8	.91 / .93 / .92 / .90 / .92 / .91 / .91	--	.10 / .16 / .22 / .13 / .13 / .20 / .19
Bullying ^a	Reynolds & Kamphaus, 1998	13	.94 / .95 / .94 / .93 / .95 / .94 / .92	--	1.70 / 1.89 / 2.28 / 1.98 / 1.79 / 1.52 / 2.57
Conduct Problems ^a		10	.74 / .86 / .78 / .81 / .84 / .88 / .84	--	1.56 / 2.31 / 2.47 / 2.16 / 2.03 / 1.77 / 3.33
Altruism	Solomon et al., 2000	8	.91 / .94 / .93 / .95 / .95 / .95 / .96	--	.23 / .31 / .34 / .57 / .45 / .42 / .32
Academic Performance	IES ^b	4	.97 / .97 / .97 / .97 / .97 / .97 / .98	--	.02 / .02 / .06 / .00 / .07 / .05 / .06
Motivation	IES ^b	1	N/A	--	.01 / .07 / .07 / .03 / .07 / .20 / .06
Attendance & Delinquency ^e	IES ^b	5	.71 / .71 / .76	--	.06 / .07 / .08
Social-Emotional and Character Development Scale (SECD) ^{df}	Ji et al. 2011	6	.96 / .98	--	.12 / .17
SECD-Prosocial Interactions ^{df}		1	N/A	--	.01 / .18
SECD-Honesty ^{df}		1	N/A	--	.12 / .16
SECD-Self-Development ^{df}		1	N/A	--	.11 / .13
SECD-Self-Control ^{df}		1	N/A	--	.13 / .17
SECD-Respect for Teacher ^{df}		1	N/A	--	.12 / .15
SECD-Respect for Parents ^{df}		1	N/A	--	.10 / .23
Parent Actions	CPPRG, 1991	6	.70 / .73 / .77 / .73 / .82 / .89 / .86	--	.08 / .13 / .22 / .22 / .40 / .07 / .23
Parent Communication		1	N/A	--	.07 / .07 / .18 / .24 / .12 / .18 / .19
Parent Involvement		1	N/A	--	.04 / .03 / .15 / .18 / .10 / .15 / .20
Parent Encouragement		1	N/A	--	.04 / .00 / .12 / .09 / .10 / .11 / .28

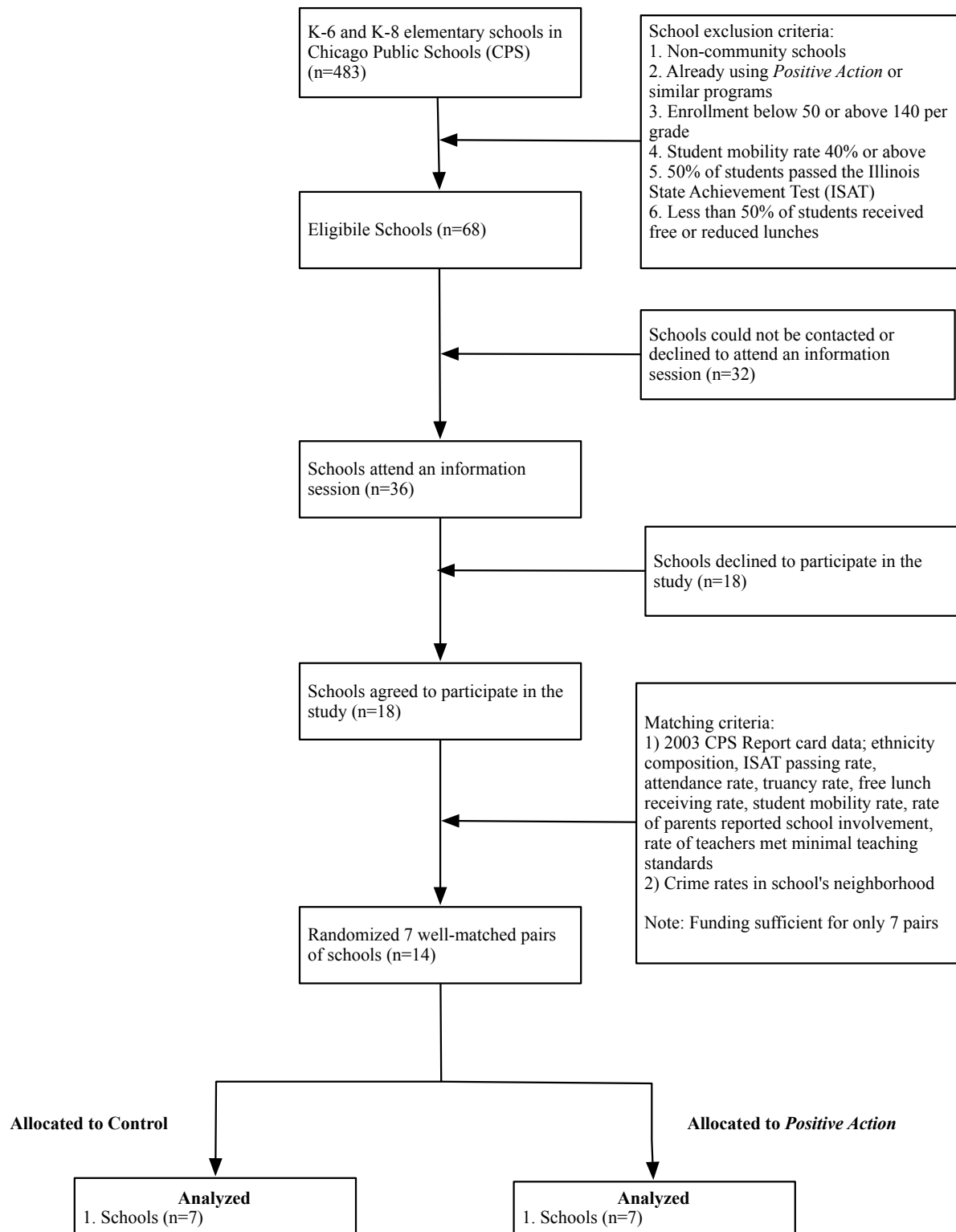
Note: The outcomes assessed by MPR during Waves 1-5 (and by UIC research staff during Waves 6-8) are labeled as "Multisite Scales" and outcomes assessed by UIC project staff during Waves 1-8 are labeled as "Site-Specific Scales." N/A indicates that alphas could not be calculated because the scale consisted of one item. Test-retest reliability available at Wave 5 for Site-Specific Scales only. ^aMOR reported rather than ICC. ^bThis scale was created by IES for the SACD multiprogram evaluation. ^cThis scale was administered starting at wave 6. ^dThis scale was administered starting at wave 7. ^eThis scale is a 2nd-order factor. For student reports, this scale was created as an average of the 6 subscales. For the teacher reports, this scale was created as an average of the 6 single items. ^fThis scale was administered starting at wave 5. ^gThis is an index of behavior rather than a scale, so alpha is an inappropriate measure and therefore not reported.

Table 2.3: Baseline and follow-up equivalence on school demographics

	2001					2004					2007					2010				
	Control Schools		PA Schools		df=12 t	Control Schools		PA Schools		df=12 t	Control Schools		PA Schools		df=12 t	Control Schools		PA Schools		df=12 t
	M	SD	M	SD		M	SD	M	SD		M	SD	M	SD		M	SD	M	SD	
% Male Students	52.70	1.98	52.00	0.98	0.84	52.64	2.89	52.47	2.11	0.13	52.03	2.35	50.90	2.17	0.93	52.09	2.42	52.10	2.22	0.00
% of White Students	10.51	16.75	9.58	13.07	0.12	9.38	14.80	9.07	12.68	0.04	7.93	13.18	7.00	10.62	0.15	8.57	13.81	7.49	11.53	0.16
% of African American Students	55.21	43.91	52.07	48.01	0.13	56.49	43.35	53.64	47.35	0.12	56.73	42.67	55.31	47.10	0.06	56.20	41.64	55.57	47.79	0.03
% of Hispanic Students	27.27	32.10	32.23	34.23	-0.28	31	35.16	32.79	36.28	-0.09	31.04	33.76	31.33	35.71	-0.02	31.8	34.07	32.63	38.94	-0.40
% of Asian American Students	4.03	6.17	4.69	7.05	-0.19	2.91	4.30	4.21	6.57	-0.44	2.74	4.45	3.91	7.00	-0.37	3.36	5.94	4.17	7.61	-0.22
% of "Other" Students	3.91	4.60	4.64	4.92	-0.29	4.46	5.03	4.72	5.21	-0.10	4.63	5.05	4.45	5.19	.063	0.04	0.03	0.02	0.05	-0.40
%Students with LEP	14.27	16.13	17.50	16.94	-0.37	11.41	14.10	17.04	17.20	-0.67	12.50	13.82	14.61	16.32	-0.26	10.87	11.92	12.87	14.78	-0.28
%Students with an IEP	13.00	6.37	8.70	2.31	1.67	12.84	5.33	9.46	2.36	1.53	14.08	6.46	9.86	3.23	1.55	14.76	6.01	12.06	3.73	1.01
% of Students Receiving a Free Lunch	83.59	3.39	83.09	6.85	0.17	81.46	3.81	85.51	4.56	-1.81	85.86	7.86	84.97	7.36	0.22	94.60	3.92	92.70	6.30	0.68
School Attendance Rate	93.01	1.17	93.38	1.28	-0.57	93.54	1.09	93.74	1.79	-0.25	92.87	2.09	93.71	1.78	-0.84	93.27	1.87	95.03	1.52	-2.01

Notes: 2001 data were used for the matching before randomization, 2004 was the first year of the trial, 2007 was when cohort students ended grade 5, and 2010 was the last year of the study. Other= %Native American, % Multirace, % Mexican, % Puerto Rican, % Cuban, % Other Hispanic, and % Multiethnic. At 2010, only information for % Native Americans was available. LEP= Limited English Proficiency, IEP= Individualized Education Plan. Degrees of Freedom for School Attendance Rate is 13. None of the above t-test were significant at the $p < 0.05$ level

Figure 2.1: Diagram of participating schools and students



Acknowledgements

We would like to thank Marc Schure for comments on previous drafts of this paper. This project was funded by grants from the Institute of Education Sciences, US Department of Education: R305L030072, R305L030004 and R305A080253 to the University of Illinois at Chicago (2003-05) and Oregon State University (2005-12). The SACD Research Program was a collaboration among IES, the Centers for Disease Control and Prevention's (CDC) Division of Violence Prevention, Mathematica Policy Research Inc. (MPR), and awardees of SACD cooperative agreements (Children's Institute, New York University, Oregon State University, University at Buffalo-SUNY, University of Maryland, University of North Carolina-Chapel Hill, and Vanderbilt University).

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

Brian Flay and David DuBois conceived the study and obtained funding, David DuBois and UIC staff oversaw program implementation, the program developer (Carol G. Allred) provided teacher/staff training, UIC and MPR staff collected all data, Brian Flay and OSU co-investigators and staff conducted data analysis, Kendra Lewis wrote the first draft of the paper, and investigators and staff participated in paper revisions.

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

Correspondence concerning this article should be addressed to lewiske@onid.orst.edu or brian.flay@oregonstate.edu

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

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

Improving Social-Emotional and Character Development in Children and
Adolescents: A Matched-Pair, Cluster-Randomized Controlled Trial in Low-Income,
Urban Schools

Kendra M. Lewis^{a,d}, Samuel Vuchinich^a, Peter Ji^b, David L. DuBois^c, Alan Acock^a,
Niloofar Bavarian^a, Joseph Day^c, Naida Silverthorn^c, and Brian R Flay^a

^aOregon State University, Corvallis, OR, 97330, USA

^bAdler School of Professional Psychology, Chicago, IL, 60602, USA

^cUniversity of Illinois at Chicago, Chicago, IL, 60608, USA

^dCorresponding author: Oregon State University, College of Public Health and Human
Sciences, School of Social and Behavioral Health Sciences, 410 Waldo Hall, Corvallis
OR 97331-5102; Telephone: 541-737-4765; Fax: 541-737-1076; Email:

lewiske@onid.orst.edu.

Brian Flay email: brian.flay@oregonstate.edu

ABSTRACT

Social-emotional development has been linked with higher levels of academic achievement and lower levels of negative behaviors such as substance use and violence. This study evaluated the impact of a school-based social-emotional and character development intervention, *Positive Action (PA)*, on social-emotional outcomes in children ages 8 to 14. A matched-pair, cluster-randomized controlled trial was conducted in a sample of urban, largely minority schools in Chicago. Data included student self-reports, as well as teacher and parent ratings of students. Multilevel growth curve analyses revealed that students who received *PA* had higher scores of student-reported empathy, altruism, positive moral beliefs, social-emotional and character development, positive actions/feelings, and social competency, and lower scores of negative moral beliefs and aggressive problem solving. There were no significant program effects for teacher- or parent-reported items. There were minimal differences by gender and no differences by patterns of student mobility. Implications of the results are discussed, with special interest in the mechanisms involved in obtaining beneficial effects in school-based interventions.

Keywords: social-emotional character development, adolescence, longitudinal

Improving Social-Emotional and Character Development in Children and
Adolescents: A Matched-Pair, Cluster-Randomized Controlled Trial in Low-Income,
Urban Schools

It has become increasingly evident that domains of social-emotional development such as discipline, morality, emotional regulation, social skills, and character are prerequisites for academic achievement in reading, writing, math and science (Ciarrochi, Heaven, & Davies, 2007; Elias, 2006; 2009). These skills are essential for establishing the kind of social, emotional, and behavioral environments in schools and classrooms needed to promote academic achievement (Coalition for Evidence Based Policy, 2002; Coleman et al., 1966). Schools cannot attain their goals if students lack honesty, respect, discipline, empathy, and other aspects of social-emotional and character development (SECD; Elias, 2009). Nevertheless, schools are expected to prevent and decrease violence, substance use, and other problem behaviors linked to academics (Fleming, 2005; Malecki & Elliot, 2002; Wentzel, 1993) and prepare students to be contributing members of society (Elias, 2009). This has led to increased interest in interventions that promote child development in these areas (e.g., Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011).

Given that addressing and promoting SECD is crucial for promoting academic achievement and preventing negative behaviors, and that emotions are a critical part of learning and development, SECD and related skills and outcomes should not be excluded from the learning process. Fortunately, recent decades have seen an increase in SECD programs that are comprehensive in nature, promoting positive behaviors

and reducing negative behaviors. SECD programs aim to not only change the behaviors associated with positive behavior development, but also the developmental trajectories. For example, the Promoting Alternative THinking Strategies (PATHS) program has been found to improve students' ability to understand, discuss, and manage emotions, as well as decrease aggression and disruptive, externalizing, and internalizing behaviors (Riggs, Greenberg, Kusché, & Pentz, 2006). Additionally, in a meta-analysis of 213 social and emotional learning (SEL) programs, Durlak and colleagues (2011) found that SEL programs significantly increased social and emotional skills; improved students' attitudes about themselves, others, and school; improved positive social behaviors; improved academic achievement; and decreased conduct problems and emotional distress. Although these programs have been shown to succeed when implemented comprehensively and with fidelity (Berkowitz and Bier, 2004; 2007; Durlak et al., 2011), to date it is not clear how well these encouraging findings generalize to low-income, urban settings (Farahmand, Grant, Polo, Duffy, DuBois, 2011). This paper reports results of an evaluation of the effectiveness of such a program on multiple indicators of the social-emotional development of children and adolescents in a low-income, urban setting to gain a better understanding of the effectiveness of SECD programs in this population. Although the term "social-emotional" encompasses several constructs, we focus on SECD, prosocial behaviors such as empathy, altruism, morality/moral behaviors, and social problem solving.

Development of SECD and prosocial behaviors in adolescence

The study of SECD and prosocial behaviors has occurred in numerous research disciplines, including general education, moral education, citizen education, and positive psychology (Althof & Berkowitz, 2006; Berkowitz & Bier, 2004; Park, Peterson, & Seligman, 2004). Current literature suggests that SECD and related prosocial behaviors may decrease as children move into adolescence. Washburn et al. (2011) and Lewis et al. (in press) found that SECD reduced over time, but that *PA* significantly mitigated this decline; this indicated that despite an overall decline in SECD, students in *PA* had less of a reduction than control students. These findings support the work of Lerner et al. (2008) who used the 5 Cs (competence, confidence, connection, character, and caring) to create a global measure of positive youth development, and found a similar negative trajectory; that is, youth endorsed fewer positive outcomes as they grew older. Similarly, Carlo and colleagues (2007) found comparable results from high school student self-reports of their positive social behavior assessed between grades 7 and 12 in a sample of about 657 youth. The results of the preceding studies indicate a decline in positive behaviors associated with SECD from middle childhood through the end of adolescence.

Studies have shown that physical aggression in reaction to conflict declines in early childhood, and may be replaced with verbal and relational aggression as children get older (Laursen & Pursell, 2009). Children tend to use coercion (similar to what we call "aggressive problem solving" below) to resolve conflicts (rather than disengagement or negotiation), whereas adolescents resort to disengagement first, followed by coercion and negotiation. Negotiation (similar to what we call

"competent problem solving" below) does not appear as a preferable method of resolving peer conflict until young adulthood (Laursen & Pursell, 2009). In terms of the development of empathy, altruism, and morality, reviews by Spinrad and Eisenberg (2009) and Eisenberg and Morris (2004) suggest that adolescents have more empathy and altruism and higher morality than children, though it should be noted that these studies were mostly cross-sectional. Rosenblum and Lewis (2003) point out that even though the development of empathy (which requires the shift from a self-focused view of emotions to also understanding the emotions of others) occurs between ages seven and 13, having the ability to be aware of another's emotions does not necessarily mean that a person will respond supportively or appropriately (Rosenblum & Lewis, 2003). Additionally, Eisenberg and Morris (2004) mention that while self-reflective and internalized prosocial reasoning increases, stereotypic and expected or normative behavior reasoning ("It's nice to help other people") decreases from childhood until late teens. Additionally, hedonistic reasoning (using one's own desires to justify actions) also decreases during this time (Eisenberg & Morris, 2004).

Research has found that children who are more prosocial in their behaviors are more socially competent. That is, these children are more likely to have positive and supportive peer relationships and interactions, tend to be popular and sociable, and are viewed positively by peers and adults (Spinrad & Eisenberg, 2009). Moreover, prosocial behaviors (such as empathy, altruism, and morality or prosocial reasoning) have been shown to be related to one another, as well as to more classroom cooperation and appropriate classroom behaviors, higher academic achievement, more

assertiveness, higher self-esteem, higher self-regulation, positive emotionality, and lower levels of adult-reported externalizing behaviors (Eisenberg & Fabes, 1998; Spinrad & Eisenberg, 2009). Additionally, children with low social competency will likely not have effective problem solving strategies or ways to deal with conflict (Laursen & Pursell, 2009). This review suggests that investing in SECD programs can influence multiple domains of children's lives, including improving their prosocial behaviors, social competency, and academic achievement, and reducing internalizing and externalizing outcomes and behaviors. One example of such a program being implemented throughout the U.S. is *Positive Action (PA)*; <http://www.positiveaction.net>; Flay and Allred, 2010).

The *Positive Action* Program

PA is a comprehensive, school-wide, SECD program grounded in theories of self-concept (DuBois, Flay, & Fagen, 2009; Purkey, 1970; Purkey & Novak, 1970) and consistent with social-ecological theories of health behaviors such as the Theory of Triadic Influence (TTI; Flay & Petraitis, 1994; Flay, Snyder, & Petraitis, 2009). The program posits that children who engage in positive behaviors will have more positive feelings about themselves and subsequent positive thoughts, leading back to more positive behaviors. Moreover, *PA* proposes a link between positive and negative behaviors, with positive feelings, thoughts, and actions resulting in fewer negative behaviors (Flay & Allred, 2010).

The *PA* program consists of a K-12 curriculum, of which the K-8 portion was used for this study. The sequenced classroom curriculum includes of over 140 15-

minute, age-appropriate lessons taught 4 days per week for grades K-6, and 70 lessons taught 2 days per week for grades 7 and 8. In addition to the student core curriculum, the *PA* program includes teacher training; counselor, family, and community training; and school-wide climate development. The core curriculum consists of the following six components: self-concept, positive actions for body and mind, social and emotional positive actions focusing on getting along with others, and managing, being honest with, and continually improving oneself.

Previous quasi-experimental and experimental evaluations have found schools receiving *PA* to have higher academic achievement and performance on standardized tests (Bavarian et al., 2012; Flay, Allred, & Ordway, 2001; Snyder et al., 2010), greater growth in academic motivation and ability, and less disaffection with learning (Bavarian et al., 2012), less absenteeism (Bavarian et al., 2012; Flay & Allred, 2010; Snyder et al., 2010), less violent behavior (Beets et al., 2009; Flay, Allred, & Ordway, 2001; Li et al., 2011), and fewer disciplinary referrals and suspensions (Flay, Allred, & Ordway, 2001; Snyder et al., 2010). *PA* has also been shown to create whole-school contextual change and improve school quality (Snyder, Vuchinich, Acock, Washburn, & Flay, 2012). Recent research has reported the effect of *PA* on reducing substance use among students (see Beets et al., 2009; Lewis et al., 2012a; Li et al., 2011), and has identified SECD as a mediator (Lewis et al., 2012a) through which this is accomplished.

To date, the effects of *PA* in middle school grades have been examined in few studies (Bavarian et al., 2012; Lewis et al., 2012a). Additionally, program effects on

most of the social-emotional outcomes in this paper have not yet been examined. Further, these outcomes were assessed in a new context, involving high-risk, inner-city students in urban Chicago. Therefore, the aim of the present study is to examine the effects of *PA* on the social-emotional outcomes assessed during this trial. We used multi-level growth curve models to test the hypotheses that students in *PA* would have better change in SECD and related outcomes than students in control schools.

METHOD

Design and Sample

The Chicago trial of *PA* is the first matched-pair, cluster-randomized controlled trial investigating the effects of the program in a low-income, urban environment. Schools participating in the study were drawn from 483 K-6 and K-8 Chicago Public Schools. Sixty-eight schools met eligibility criteria (Ji, DuBois, Flay, and Brechling, 2008; Lewis et al., 2012b), of which 18 agreed to participate, and the seven best-matched pairs were selected for participation (Ji et al., 2008; Schochet & Novak, 2003). A series of *t*-tests revealed that the *PA* and control schools were not significantly different from each other on any of the matching variables, and that the seven pairs of schools did not significantly differ from the remainder of the 68 schools eligible for the study (Ji et al., 2008; Lewis et al., 2012b).

The trial was longitudinal at the school-level with a place-focused intent-to-treat design at the cohort-level (Vuchinich, Flay, Aber, & Bickman, 2012). Specifically, a cohort of students in the seven matched pairs of schools were followed, beginning in grade 3 (fall 2004), and at seven additional times (waves) over six years:

spring 2005, fall 2005, spring 2006, spring 2007, fall 2008, spring 2009, and spring 2010 (end of grade 8). Survey data were also collected about students from their teachers (at all waves except Wave 6) and parents (at six waves). Throughout the six years of the study, 100% of schools were retained.

Parental consent was obtained before students, parents or teachers completed surveys when students were in grade 3. Seventy-nine percent of parents provided consent at baseline and ranged from 65% to 78% for Waves 2-5. Students joining the study at later waves were consented at that time. All students were re-consented for the second phase of funding at Wave 6 (beginning of grade 7); consent rates were lower at Waves 6 through 8 (\approx 58 to 64%) than Waves 1 through 5. The percentages of consenting parents who provided reports on their children were 72.3%, 58.9%, 52.2%, 50.5%, and 72.9% at Waves, 1, 2, 4, 5 and 8, respectively. Percentages of consented students for whom teachers completed ratings were 74.6%, 74.8%, 72.4%, 78.3%, 74.4%, and 92.7% for Waves 1, 2, 4, 5, 7, and 8, respectively (Lewis et al., 2012b).

The total number of students enrolled in the study across all eight waves was 1,170, of whom approximately 53% were female; approximately 48% were African American, 27% Hispanic and 19% other (e.g., White, Asian, and Native American, and “Other”). A total of 247 teachers completed student assessments; 75% of teachers were female; 43% White, 36% African American, 13% Hispanic and 8% other (Asian and Native American). Of the original 624 students in grade 3 at the beginning of the trial, only 131 (i.e. 21%) remained at grade 8, reflecting the high mobility by low-

income urban students, a pattern that has been documented with this population (Tobler & Komro, 2011). With respect to maintenance of the baseline sample size, 363 students were present at Wave 8 (i.e., approximately 58% of the Wave 1 sample size); the decrease in N over time is consistent with the trend among Chicago Public Schools to decrease in size during the study period. The decrease in sample sizes across time is also reflective of the lower consent rates at Waves 6 through 8.

Measures

The data for the study came from student self-reports, as well as teacher and parent ratings of students. For each measure, an average composite was created, with higher scores reflecting having more of the construct being assessed.

Student Self-report Measures

SECD. SECD was measured using the 28-item Child SECD Scale (DuBois, Flay, Day, & Silverthorn, 2010; Ji, DuBois, & Flay, 2011). This scale was adapted from multiple existing measures of social skills (Achenbach, 1991; Bar-On, 2002; Elliott, et al., 1988; Goodman & Goodman, 2009; Leffert et al., 1998; Smart & Sanson, 2003; Walker & McConnell, 1995; Wilson, O'Brien, & Sesma, 2009; see Ji et al., 2011). Ji et al. (2011) found six first-order factors of these 28 items: Prosocial Interactions (α range=0.77-0.85; ICCs range=0.02-0.18) Honesty (α range=0.71-0.82; ICCs range=0.01-0.10), Self-Development (α range=0.64-0.83; ICCs range=approximately 0.00-0.06), Self-Control (α range=0.64-0.82; ICCs range=0.01-0.13), Respect for Teacher (α range=0.71-0.90; ICCs range=0.04-0.10), and Respect for Parents (α range=0.61-0.89; ICCs range=0.02-0.07). These scales load on a

second-order factor, SECD, previously assessed by Lewis et al. (in press). Responses were on a four-point Likert scale that allowed students to indicate how often they performed each character-related behavior ranging from "None of the time" to "All of the time".

Self-Efficacy for Peer Interaction. Self-Efficacy for Peer Interaction was assessed using 12-item, modified version of the Self-Efficacy for Peer Interaction Scale (Wheeler & Ladd, 1982). Students were asked to rate how hard it is for them to interact with peers in different scenarios. Example items include "Some kids are teasing your friend. How easy or hard would it be for you to tell them to stop a kid does not like your friend?", "How easy or hard would it be for you to tell the kid to be nice to your friend?" and "A kid is yelling at you. How easy or hard would it be to tell the kid to stop?". Responses were on a four-point Likert scale ranging from "Really hard" to "Really easy" (α range=0.81-0.90; ICCs range=approximately 0.00-0.09).

Empathy. Empathy was measured using the Children's Empathic Attitudes Questionnaire (Funk, Fox, Chan, & Curtiss, 2008). The scale included 16 items; example items include: "When I'm mean to someone, I usually feel bad about it later", "I understand how other kids feel", and "Seeing a kid who is crying makes me feel like crying". Responses were on a three-point scale, "No", "Sometimes", and "Yes" (α range=0.79-0.89; ICCs range=0.02-0.19).

Altruism. Altruism was assessed by 8 items (Solomon, Battistich, Watson, Schaps, & Lewis, 2000). Students were asked to report how often they helped another person at school or elsewhere. Example items include "I helped someone who was

hurt", "I helped an older person", and "I stopped a kid from hurting another kid".

Responses were on a four-point Likert scale ranging from "Never" to "Many times" (α range=0.81-0.87; ICCs range=approximately 0.00-0.04).

Belief in Moral Order. Students' morals and endorsement of socially desirable and undesirable behaviors were assessed using a modified version of the Belief in Moral Order scale (Arthur, Hawkins, Pollard, Catalano, & Baglioni, 2002; Glaser, Van Horn, Arthur, Hawkins, & Catalano, 2005). The scale includes 11 items, six assessing positive values (α range=0.70-0.83; ICCs range=0.02-0.15) and five assessing negative values (α range=0.61-0.79; ICCs range=0.04-0.27). Students were asked to indicate if they agreed with the statements; responses were on a four-point Likert scale ranging from "NO!" to "YES!" Example positive items include "It is important to be honest with your parents, even if they get upset with you or you get punished", "It is important to play by the rules, even if no one is watching" and "It is important to control your temper, even when something happened that you don't like". Example of negative items include "It is OK to beat up people if they start a fight", "It is OK to cheat on tests at school" and "It is OK to hit someone if they hit your first".

Social Problem Solving. student's ability to solve interpersonal social problem was assessed using five of the eight scenarios from The Social Skills Problem Solving Measure (Aber, Brown, Jones, & Samples, 1995). These scenarios depict an interpersonal problem, and students chose which of five possible responses they felt would help solve the problem. The questionnaire yields a score for students' selection of aggressive social problem solving responses (α range=0.65-0.84; ICCs range=0.05-

0.18) and a score for students' selection of competent social problem solving responses (α range=0.57-0.71; ICCs range=0.02-0.11). These scores were then dichotomized into whether the child chose any responses to reflect the specific problem solving method (for a score of 1) or no responses reflecting that method (for a score of 0).

Positive Actions/Feelings. The Positive Actions/Feelings scale (*Positive Action*, 2007) was created to measure the theoretical construct of Unit 1 (self-concept) of the *PA* program. The items pertain to feelings about oneself and one's actions that are expected to be fostered by *PA*; items include "I feel good about my future", "I feel happy about my life", "I feel good when I do good things", and "I feel good about what I am doing". Responses were on a four-point Likert scale ranging from "None of the time" to "All of the time" (α range=0.73-0.90; ICCs range=0.02-0.09).

Teacher and Parent Reported Outcomes

In addition to the child self-report outcomes, three related scales reported by parents and teachers were also measured. Parent reports were assessed at Waves 1-5 and Wave 8; teacher reports were assessed at all waves except Wave 6.

Social Competency. Social competency was assessed using a modified version of the Social Competence Scale (Conduct Problems Prevention Research Group, 1999). Teachers and parents were asked to rate the child's typical behavior in the past 30 days. Responses to the 19 items were on a four-point Likert scale ranging from "Never" to "Almost always" (teacher-report: α range=0.96-0.97; ICCs range=0.15-0.25; parent-report: α range=0.89-0.92; ICCs range=approximately 0.00-0.07).

Example items include "Expresses needs and feelings appropriately", "Is very good at understanding other people's feelings", "Works well in a group", "Cooperates with peers without prompting", "Listens to others' points of view".

Responsibility. Responsibility was assessed using a scale created by IES for the multisite trial (Social and Character Development Research Consortium, 2010).

Teachers and parents were asked to rate the child's typical behavior in the past 30 days. Responses were on a four-point Likert scale ranging from "Never" to "Almost always" (teacher-report: α range=0.90-0.93; ICCs range=0.10-0.22; parent-report: α range=0.80-0.83; ICCs range=0.03-0.06). This scale included eight items; example items are "Asks before borrowing or taking something", "Apologizes when he/she has done something wrong", and "Returns borrowed belongings or materials".

Altruism. Altruism was assessed with eight items using an adapted version of the student-report altruism scale (Solomon, Battistich, Watson, Schaps, & Lewis, 2000), reporting how often children showed altruistic behaviors such as helping someone who was hurt, cheering up someone who was sad, and helping an older person (teacher-report: α range=0.91-0.96; ICCs range=0.23-0.57; parent-report: α range=0.86-0.91; ICCs range=approximately 0.00-0.05).

Analysis

Primary analyses consisted of multi-level (observations nested within students nested within schools) growth curve models using Stata's (version 12) "xtmixed" (for normally distributed outcomes) "xttobit" (for outcomes with a skewed distribution) and "xtlogit" (for binary outcomes) commands to account for all observations and to

model school differences (Rabe-Hesketh & Skrondal, 2008). This approach allows for a complete analysis of the multiple waves of available data and takes into account the pattern of change over time.

For the growth-curve analysis of each outcome measure, we first fit a full random-intercept model including condition (i.e., *PA* or non-*PA* school), time (measured as years of exposure to *PA*), condition by time (condition \times time), and quadratic terms for time and the interaction of condition by time (time² and condition \times time²). Higher order terms found to lack statistical significance were then dropped from the model for parsimony in a step-wise fashion, until a fully-reduced model was achieved. This was followed by the fully reduced random-coefficients models (when applicable), with the former model nested within the latter model. A likelihood-ratio test was performed to determine which model was a better fit for the data (Rabe-Hesketh & Skrondal, 2008); model estimates for the better fitting model are reported in the appropriate tables. Additionally, effect size calculations made using estimated means (Lipsey and Wilson, 2001). Effect sizes for binary outcomes were calculated using the Cox transformation (Sánchez-Meca, Marín-Martínez, & Chacón-Moscoso, 2003).

Supplementary Analyses

Sensitivity analyses assessed the robustness of results from the primary analyses. One approach involved including a “pairs” variable as an additional level in each of the best-fitting models to determine whether adding a fourth level would affect findings. Second, to provide a more conservative test (from a statistical power

perspective) of program effects for each outcome, the test statistic provided by Stata (which assume a large sample size) in the primary analyses ($N=14$ schools) was compared to the critical value for a two-tailed t -distribution with 12 degrees of freedom at a 95% confidence level (2.18; Raudenbush & Bryk, 2002).

Possible moderating effects of gender and student mobility were examined. The effect of student mobility groups was examined using results from a latent class analysis (Lewis et al., 2012b) in which a 5-class solution was found to be the most appropriate fit for the data: 1) stayers (average study duration of 5.72 years, $N = 158$), (2) temporary participants (present for grade 4 and/or 5 only; average study duration of 1.30 years; $N= 196$), 3) late joiners (average study duration of 1.38 years; $N= 308$); 4) early leavers (average study duration of 0.94 years; $N= 263$), and 5) late leavers (average study duration of 3.23 years; $N= 287$); stayers served as the reference group.

RESULTS

Table 3.1 shows the correlations between the student, teacher, and parent variables at Waves 1 (beginning of grade 3) and 8 (end of grade 8), and program effects on student-, teacher-, and parent-reported social and emotional outcomes are presented in Table 3.2. Table 3.3 reports the baseline and endpoint predicted means by condition and effect sizes. Figure 3.1 presents trajectories of outcomes with significant condition \times time² interactions or gender interactions.

<Table 3.1 about here>

Student Self-Report

Students in *PA* schools reported higher scores on all first-order factor SECD outcomes:

prosocial interactions ($B=0.06, p<0.001; ES=0.56$), honesty ($B=0.05, p<0.01; ES=0.49$), self-development ($B=0.04, p<0.05; ES=0.43$) self-control ($B=0.08, p<0.001; ES=0.67$), respect for teachers ($B=0.08, p<0.001; ES=0.82$), and respect for parents ($B=0.07, p<0.001; ES=0.74$). These findings remained significant in both the pair-level and adjusted degrees of freedom (12) analyses (results not shown).

Additionally, students receiving *PA* reported significantly higher scores on empathy ($B=0.02, p<0.05; ES=0.29$), altruism ($B=0.03, p<0.05; ES=0.22$), and positive actions/feelings ($B=0.05, p<0.05; ES=0.47$) over time compared to students in control schools. Empathy and altruism remained significant in the pair-level analysis and positive actions/feelings remained significant using the adjusted degrees of freedom (12; results not shown). Finally, students in *PA* schools reported significantly lower negative morals ($B=-0.06, p<0.001; ES=-0.67$), and aggressive problem solving ($OR=0.75, p<0.001; ES=0.30$), than students in control schools. These findings remained significant in both the pair-level and adjusted degrees of freedom (12) analyses (results not shown).

All outcomes except for negative moral center and aggressive problem solving showed a decline over time. *PA*, however, was shown to significantly mitigate this decline; that is, students receiving *PA* showed less of a decline than students in the control group. Additionally, although children increased in self-report of negative moral center and aggressive problem solving, *PA* significantly mitigated this increase,

such that students in *PA* schools were lower than students in control schools on these outcomes.

Teacher-and Parent-Report

Students in *PA* were reported by teachers to have marginally higher responsibility ($B=0.03$ $p=0.08$; $ES=0.25$) than students in control schools. There were no other significant effects on teacher or parent reports of student behavior.

<Table 3.2 about here>

<Table 3.3 about here>

Supplementary Analyses

Three outcomes showed significant gender moderation. Specifically, the positive effect of *PA* is significant for girls and less pronounced for boys for measures of honesty (condition \times time \times gender interaction $B= -0.07$ $p<0.05$), self-control (condition \times time \times gender interaction $B= -0.07$, $p<0.05$), and respect for teacher (condition \times time \times gender interaction $B= -0.08$, $p<0.05$). No differences were found by mobility status.

<Figure 3.1 about here>

DISCUSSION

There is a need for schools to address SECD outcomes that are essential to the development of academic achievement and preventing negative behaviors. Findings from the present study lend support to the ability of social-emotional and character development programs to impact such SECD outcomes. Specifically, students in *PA* reported higher levels of various social-emotional outcomes, including empathy,

altruism, prosocial interactions, honesty, self-control, self-development, respect for teachers and parents, and positive actions and feelings. Additionally, students in *PA* also reported lower aggressive problem solving and negative morality. Teachers reported students in *PA* to have marginally higher responsibility.

This is the first study to examine the impact of *PA* on social-emotional outcomes in a low-income, high minority urban population. *PA* significantly mitigated the decline in a variety of positive outcomes, as well as the increase in negative outcomes. Additionally, supplemental analyses revealed few differential effects by gender; program effects were significant for girls, but not boys, on honesty, self-control, and respect for teachers. This might be because girls tend to be more socially-emotionally skilled and attuned and thus may be better able to take advantage of SECD curriculum; however, this is a post-hoc explanation than needs to be researched. There were no differences found by mobility status, suggesting that despite the high mobility of this population, limited exposure to *PA* can still be influential on students.

Our finding of a general decline in related SECD outcomes is consistent with previous studies (see Carlo et al., 2007; Lerner et al., 2008; Kokko et al., 2006; Nantel-Vivier, 2009). Washburn et al. (2011) and Lewis et al. (in press) found SECD, a second-order factor of six outcomes assessed in this paper, to decline over time. There has been little discussion of reasons for this decline in the literature; Kokko et al. (2006) and Nantel-Vivier et al. (2009) suggested that more work is needed to understand the decline, while Carlo et al. (2007) and Washburn et al. (2011) suggested

that environmental factors such as exposure to high risk conditions and lack of access to protective factors (e.g., positive role models, opportunities for constructive interactions, and emotional support) could influence these negative trajectories. Additionally, Wigelsworth, Humphrey, Kalambouka, and Lendrum (2010) point out that older children and adolescents have a heightened sense of self-awareness compared to younger children, and it is possible that this increased self-awareness contributes to lower (and more accurate) reports of SECD and related behaviors.

There was also a general decline in empathy and altruism, which is inconsistent with previous findings. It should be noted, however, that most previous studies on these prosocial behaviors have been cross-sectional. Additionally, Rosenblum and Lewis (2003) suggested that the development of empathy does not equate with the use of empathetic responses. This study provides a longitudinal assessment of these prosocial behaviors, as well as problem solving, which is lacking in the current developmental literature (Laursen & Pursell, 2009). Our finding of higher aggressive problem solving over time is consistent with cross-sectional studies of problem solving, such that aggressive problem solving techniques were used over competent problem solving until late teens or early adulthood. *PA*, however, was shown to significantly mitigate the increase in aggressive problem solving.

The findings of this study should be viewed in the context of several limitations. Student outcomes were assessed through student self-report, potentially leading to a method bias (Podsakoff, MacKenzie, Lee, and Podsakoff, 2003) that could inflate the observed relationships between the variables. Self-reports are also

susceptible to social desirability such that students may overstate their positive thoughts, feelings, and behaviors or understate their negative thoughts, feelings, and behaviors to feel as if they "fit in" with their peers and society. Additionally, significant effects were not found for parent or teacher reports. Wigelsworth et al. (2010) point out, however, that parents have a restricted frame of reference compared to teachers; that is, parents only see SECD and related behaviors of their child, whereas a teacher sees many children. Teachers, however, fill out surveys for many children, and measures may lose their sensitivity with the burden of completing so many surveys (Wigelsworth et al., 2010). Additionally, the teachers rating the students changed every year; that is, several different teachers rated the same students through the trial. These issues and challenges could contribute to our lack of significant findings for the adult-reported outcomes.

With respect to external validity, the findings are generalizable only to low-income, inner-city schools that would self-select to participate in a trial of this nature. With respect to internal validity, the small number of pairs (i.e., seven) and schools (i.e., 14) led to limited statistical power; however, that significant findings were found in spite of this limitation suggest that our findings may be conservative. Additionally, as has been seen in other studies within low income, urban school settings (Tobler & Komro, 2011), student mobility led to high turnover of students. However, the use of latent class analysis is an innovative method we have used to address this issue (Lewis et al., 2012b). One final limitation is the decreasing school size during the time of the study; nonetheless, no schools left the study.

The present study has several strengths. The longitudinal nature of this cluster-randomized controlled trial allowed for the examination of children across elementary and middle school grades. The sensitivity analyses serve to support the study findings, and the use of latent class analysis to examine the differences in study duration (and program exposure for students in *PA* schools) due to the high mobility of the student population is a novel approach that is an additional strength of the study. This paper included new measures that have not been previously assessed in evaluations of the *PA* program, and showed significant effects for a majority of these outcomes. Moreover, this study involved a sample of students in a high-risk setting; generating improvements can be particularly difficult in urban areas facing rising poverty rates (DeNavas-Walt, Proctor, & Smith, 2009), health disparities (Braveman & Egerter, 2008), and cuts in social and educational programs (Johnson, Oliff, & Williams, 2010). The empirical evidence of effectiveness of a SECD program in a high-risk population, as demonstrated in the present study, should serve as a call to action for policymakers and school officials who are increasingly challenged to positively impact not only academic achievement, but also behavior and character development (Elias, 2006; Flay & Allred, 2010).

Table 3.1: Correlations between Student-, Teacher-, and Parent-Reported Outcomes (N=1,170 students). Student-reported outcomes; Wave 1 above the diagonal, Wave 8 below.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.Prosocial Interactions	--	.66*	.56*	.62*	.57*	.46*	.07†	.37*	.32*	.20*	-.32*	-.32*	.21*	.38*
2.Honesty	.79*	--	.52*	.58*	.57*	.52*	.05	.33*	.25*	.26*	-.25*	-.21*	.19*	.33*
3.Self-Development	.67*	.59*	--	.56*	.53*	.50*	.03	.29*	.23*	.23*	-.20*	-.16*	.13*	.39*
4.Self-Control	.72*	.70*	.47*	--	.59*	.43*	.08†	.33*	.24*	.26*	-.33*	-.36*	.27*	.38*
5.Respect for Teachers	.77*	.67*	.59*	.71*	--	.51*	.10*	.21*	.18*	.16*	-.25*	-.26*	.18*	.36*
6.Respect for Parents	.60*	.57*	.60*	.56*	.62*	--	.14*	.20*	.15*	.21*	-.13*	-.12*	.08*	.39*
7.Self-Efficacy for Peer Interactions	.21*	.10†	.20*	.05	.04	.11*	--	-.03	-.01	-.04	.00	.15*	-.01	.10*
8.Empathy	.53*	.46*	.34*	.45*	.50*	.17*	-.00	--	.32*	.20*	-.25*	-.28*	.22*	.18*
9.Altruism	.28*	.26*	.26*	.17*	.11*	.16*	.06	.22*	--	.03	-.11*	-.10*	.06	.08†
10.Positive Moral Center	.49*	.47*	.46*	.47*	.48*	.38*	.14*	.33*	.18*	--	-.11*	-.09*	.17*	.16*
11.Negative Moral Center	-.34*	-.31*	-.30*	-.40*	-.45*	-.27*	.06	-.37*	-.04	-.24*	--	.42*	-.31*	-.19*
12.Aggressive Problem Solving	-.36*	-.36*	-.23*	-.44*	-.41*	-.21*	.16*	-.43*	-.13*	-.22*	.46*	--	-.57*	-.23*
13.Competent Problem Solving	.42*	.38*	.29*	.42*	.41*	.28*	.06	.41*	.14*	.25*	-.41*	-.74*	--	.23*
14.Positive Actions/Feelings	.48*	.41*	.61*	.32*	.42*	.46*	.24*	.16*	.19*	.40*	-.21*	-.12*	.23*	--

Table 3.1 continued. Teacher- and parent-reported outcomes with student-reported outcomes, Wave 1.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
15.Responsibility (Teacher)	.15*	.08†	.08†	.12*	.19*	.09*	-.00	.05	-.03	.09*	-.23*	-.16*	.18*	.13*
16.Social Competency (Teacher)	.15*	.09*	.07†	.14*	.17*	.07	-.00	.04	-.04	.13*	-.24*	-.15*	.17*	.11*
17.Altruism (Teacher)	.01	-.01	.04	.01	.03	.01	.04	.03	.03	.04	-.06*	-.03	.09*	.04
18.Responsibility (Parent)	.12*	.14*	.04	.04	.14*	.13*	-.00	.02	-.06	.07	-.14*	-.14*	.15*	.15*
19.Social Competency (Parent)	.14*	.15*	.10*	.09*	.17*	.15*	.02	.03	-.05	.08†	-.13*	-.10*	.16*	.13*
20.Altruism (Parent)	.02	-.04	.03	.00	-.00	.05	.02	.01	.00	-.03	-.03	-.00	.03	.05

Table 3.1 continued. Teacher- and parent-reported outcomes with student-reported outcomes, Wave 8.

	15	16	17	18	19	20
1.Prosocial Interactions	.10†	.16*	.09	.16*	.21*	-.01
2.Honesty	.09	.12*	.08	.12†	.16*	.04
3.Self-Development	.10	.13*	.08	.05	.18*	.04
4.Self-Control	.21*	.27*	.11†	.13*	.18*	.03
5.Respect for Teachers	.18*	.23*	.11†	.23*	.27*	-.03
6.Respect for Parents	.04	.07	-.02	.12†	.19*	.10
7.Self-Efficacy for Peer Interactions	-.01	.01	-.03	-.02	.05	.14*
8.Empathy	.20*	.22*	.10†	.09	.15*	.02
9.Altruism	-.12*	-.14*	-.03	-.05	.04	.26*
10.Positive Moral Center	.13*	.17*	-.06	.13*	.20*	.02
11.Negative Moral Center	-.11†	-.12*	-.10	-.17*	-.15*	-.00
12.Aggressive Problem Solving	-.19*	-.17*	-.07	-.10	-.12†	-.06
13.Competent Problem Solving	.10	.10†	.08	.11	.15*	.12†
14.Positive Actions/Feelings	-.01	.01	-.06	.07	.19*	-.02

Table 3.1 continued. Teacher- and parent-reported outcomes with teacher- and parent-reported outcomes; Wave 1 above the diagonal, Wave 8 below.

	15	16	17	18	19	20
15. Responsibility (Teacher Rating)	--	.90*	.14*	.31*	.27*	.00
16. Social Competency (Teacher Rating)	.89*	--	.16*	.27*	.27*	-.01
17. Altruism (Teacher Rating)	.16*	.24*	--	.11*	.07	.08†
18. Responsibility (Parent Rating)	.15*	.19*	.06	--	.73*	.13*
19. Social Competency (Parent Rating)	.09	.16*	.07	.76*	--	.15*
20. Altruism (Parent Rating)	-.02	-.06	-.07	.20*	.24*	--

Note: † $p < .10$; * $p < .05$.

Table 3.2: Growth model estimates for student-, teacher-, and parent-reported outcomes (N=1,170 students)

Scale	Intercept	Condition	Time	Condition × Time	Time ²	Condition × Time ²
Student-Report	B (SE)					
Prosocial Interactions ^a	3.64 (0.06)	-0.08 (0.09)	-0.30 (0.03)***	0.06 (0.01)***	0.03 (0.00)***	--
Honesty ^a	3.69 (0.06)	-0.09 (0.08)	-0.36 (0.03)***	0.05 (0.02)**	0.03 (0.00)***	--
Self-Development ^a	3.83 (0.07)	-0.15 (0.10)	-0.11 (0.01)***	0.04 (0.02)*	N/A	--
Self-Control ^a	3.55 (0.07)	-0.09 (0.09)	-0.42 (0.03)***	0.08 (0.02)***	0.04 (0.01)***	--
Respect for Teachers ^a	4.05 (0.08)	-0.14 (0.11)	-0.33 (0.03)***	0.08 (0.02)***	0.02 (0.01)***	--
Respect for Parents ^a	4.09 (0.07)	-0.13 (0.09)	-0.19 (0.01)***	0.07 (0.02)***	--	--
Self-Efficacy for Peer Interaction ^a	2.90 (0.06)	-0.29 (0.07)	0.24 (0.02)***	-0.02 (0.01)	-0.02 (0.00)***	--
Empathy ^b	2.44 (0.03)	0.04 (0.04)	-0.19 (0.01)***	0.02 (0.01)*	0.02 (0.00)***	--
Altruism ^b	1.64 (0.05)	-0.08 (0.06)	-0.26 (0.02)***	0.03 (0.01)*	0.03 (0.00)***	--
Positive Moral Center ^a	3.60 (0.06)	-0.09 (0.08)	-0.02 (0.03)	0.02 (0.02)	-0.01 (0.01)*	--
Negative Moral Center ^a	1.21 (0.10)	-0.06 (0.13)	0.20 (0.01)***	-0.06 (0.02)***	--	--
Positive Actions/Feelings ^a	4.10 (0.09)	-0.13 (0.12)	-0.15 (0.02)***	0.05 (0.02)*	--	--
Teacher-Report						
Responsibility ^a	3.21 (0.10)	-0.10 (0.14)	-0.02 (0.01)	0.03 (0.02) [†]	--	--
Social Competency ^b	2.87 (0.08)	-0.07 (0.11)	-0.05 (0.03) [†]	0.12 (0.04)**	0.01 (0.01)	-0.02 (0.01)**
Altruism ^a	1.15 (0.10)	0.28 (0.13)*	-0.02(0.03)	-0.02 (0.02)	0.01 (0.01)***	--
Parent-Report						

Responsibility ^a	3.16 (0.06)	0.07 (0.08)	0.02 (0.01) *	-0.00 (0.01)	--	--
Social Competency ^b	2.85 (0.05)	0.03 (0.07)	0.01 (0.07)	0.01 (0.01)	--	--
Altruism ^a	2.44 (0.05)	-0.06 (0.07)	-0.07 (0.03)*	0.12 (0.05)*	0.01 (0.01)*	-0.02 (0.01)*
Student-Report	OR (95% CI)					
Aggressive Problem Solving ^a	0.15 (0.05)	0.73 (0.35)	4.33 (0.53)***	0.75 (0.05)***	0.88 (0.02)***	--
Competent Problem Solving ^a	37.79 (13.28)	1.30 (0.55)	0.55 (0.09)***	1.07 (0.12)	1.12 (0.03)***	--

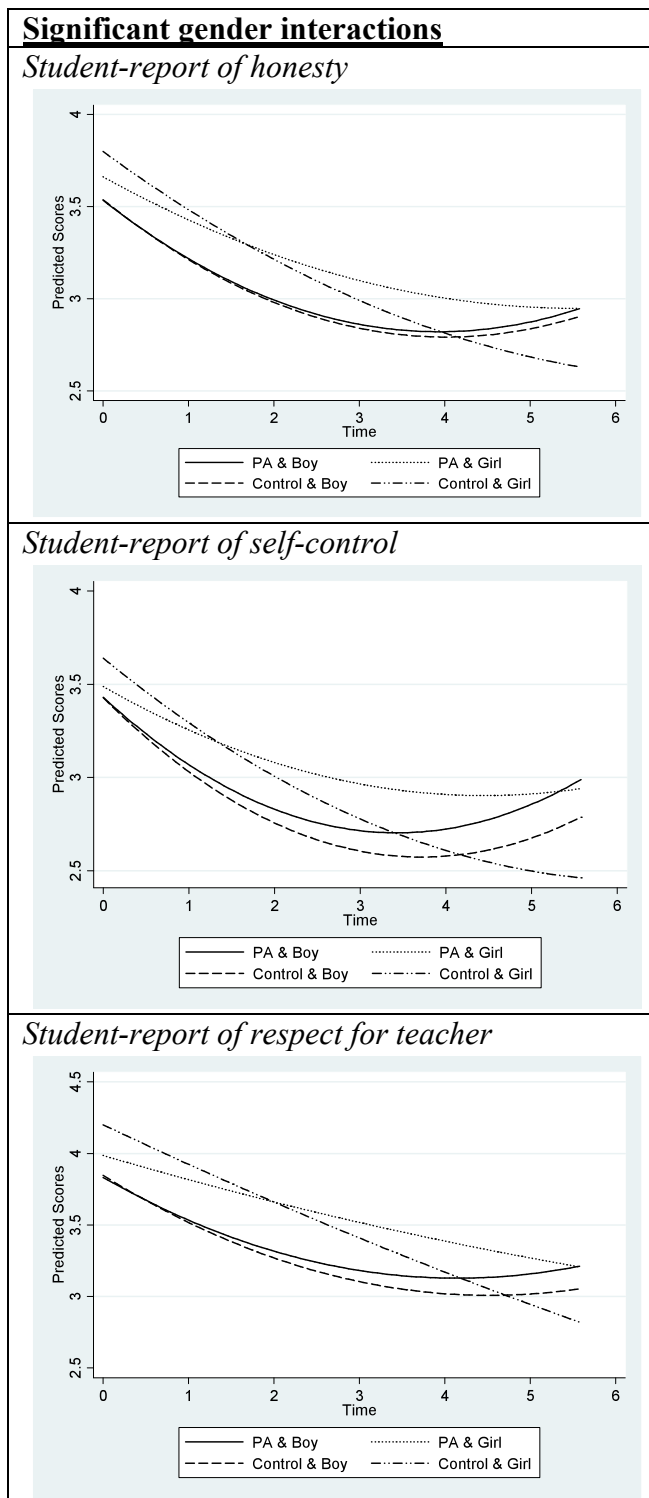
Note: [†] $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$. Insignificant quadratic terms were removed from model in the interest of parsimony. ^aResults are for a random intercept, fixed slope model. ^bResults are for a random coefficient (random intercept and random slope) model; LR tests revealed these models to be preferable over random intercept, fixed slope models.

Table 3.3: Wave 1 and Wave 8 predicted (model) means and probabilities by condition (N=1,170 students)

Variables	Wave 1		Wave 8		Effect Size
	<i>PA</i> Mean	Control Mean	<i>PA</i> Mean	Control Mean	
Prosocial Interactions	3.57	3.64	3.14	2.90	0.56
Honesty	3.60	3.69	2.95	2.74	0.49
Self-Development	3.68	3.83	3.33	3.23	0.43
Self-Control	3.46	3.55	2.96	2.61	0.67
Respect for Teachers	3.91	4.05	3.21	2.92	0.82
Respect for Parents	4.00	4.13	3.30	3.06	0.74
Self-Efficacy for Peer Interaction	2.87	2.90	3.49	3.62	-0.17
Empathy	2.48	2.44	2.14	2.00	0.29
Altruism	1.56	1.64	1.33	1.23	0.22
Positive Moral Center	3.52	3.60	3.14	3.10	0.18
Negative Moral Center	1.09	1.15	1.82	2.24	-0.67
Positive Actions/Feelings	3.97	4.10	3.42	3.29	0.41
Aggressive Problem Solving ^a	0.10	0.13	0.63	0.92	0.12
Competent Problem Solving ^a	0.98	0.97	0.99	0.98	0.22
Responsibility (Teacher Rating)	3.11	3.22	3.20	3.13	0.25
Social Competency (Teacher Rating)	2.80	2.87	2.80	2.79	0.11
Altruism (Teacher Rating)	1.43	1.15	1.48	1.31	-0.24
Responsibility (Parent Rating)	3.23	3.16	3.34	3.29	-0.04
Social Competency (Parent Rating)	2.87	2.85	2.98	2.88	0.14
Altruism (Parent Rating)	2.38	2.44	2.43	2.42	0.09

Note: Means for censored outcomes may fall outside of the range as an effect of the estimator. Baseline means were significantly different for Altruism (Teacher Rating), favoring *PA* students. ^aIndicates predicted probabilities.

Figure 3.1. Growth curve graphs for significant outcomes that also have a significant gender difference.



Acknowledgements

We would like to thank Marc Schure and Margaret Malloy for comments on previous drafts of this paper. This project was funded by grants from the Institute of Education Sciences, US Department of Education: R305L030072, R305L030004 and R305A080253 to the University of Illinois at Chicago (2003-05) and Oregon State University (2005-12). The SACD Research Program was a collaboration among IES, the Centers for Disease Control and Prevention's (CDC) Division of Violence Prevention, Mathematica Policy Research Inc. (MPR), and awardees of SACD cooperative agreements (Children's Institute, New York University, Oregon State University, University at Buffalo-SUNY, University of Maryland, University of North Carolina-Chapel Hill, and Vanderbilt University).

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

Brian Flay and David DuBois conceived the study and obtained funding, David DuBois and UIC staff oversaw program implementation, the program developer (Carol G. Allred) provided teacher/staff training, UIC and MPR staff collected all data, Brian Flay and OSU co-investigators and staff conducted data analysis, Kendra Lewis wrote the first draft of the paper, and investigators and staff participated in paper revisions.

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

Correspondence concerning this article should be addressed to lewiske@onid.orst.edu or brian.flay@oregonstate.edu

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

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

CONCLUSIONS

A limited, but growing, body of research indicates that school-based SECD and SECD-like programs can influence social outcomes, health behaviors, academic achievement, and problem behaviors among low-income minority youth, a population disproportionately affected by disparities in health (Braveman, Cubbin, Egerter, Williams, & Pamuk, 2010) and education (Aud et al., 2010). The current paper extends this research in two ways: (a) by showing how it is possible to conduct cluster-randomized prevention trials on complex social-psychological outcomes in urban, minority populations, and (b) by showing that the *PA* intervention had the predicted beneficial effects on the development of social-emotional from age 8 to 14.

Summary of findings from the first paper

Following the recent trend to publicize the methodology of a trial (see Morone et al, 2012; Winters-Stone et al., 2012, for examples), this manuscript focused on providing details about the Chicago CRCT of *PA* that will be useful for prevention scientists, educators, researchers, and policy makers interested in promoting positive outcomes and behaviors and outcomes, academic achievement, reducing negative outcome behaviors. Specifically, the setting and recruitment of schools was described, as well as the process by which schools were matched into pairs and randomized into *PA* or control, including the list of criteria for school eligibility and variables used for matching. Additionally, this paper described the primary analyses to test for program effects, as well as several sensitivity analyses that will also be conducted when evaluating this program. Further, this manuscript described secondary tests of

meditation and moderation. Finally, this paper also provided several descriptive statistics and results regarding the sample. For all scales that are to be assessed, alpha reliability and intraclass correlations (ICC) or median odds ratio (MOR) at all eight waves were presented, as well as test-retest reliability at Wave 5. This manuscript also presented characteristics of the students and teachers in this sample. In terms of baseline equivalency, *PA* and control schools did not significantly differ on matching variables before or during the trial. Minimal differences were found on baseline reports from students, teachers, and parents; half of these differences favored *PA* students and half favored control students. This suggests that the a priori matching and randomization were successful, as randomization does not guarantee that nested subjects (i.e., students within schools) will be equivalent (Giraudeau & Rivaud, 2009). Finally, this manuscript provided some lessons learned from designing, conducting, and analyzing this trial. The lessons described in the present study regarding the design and implementation of the trial were derived from discussing the issues with research staff at UIC who undertook these steps. Challenges regarding the analyses of program effects resulted from conducting preliminary analyses and diagnostics on each outcome assessed.

Summary of findings from the second paper

The second paper conducted an evaluation of *PA*, following the analytic procedures outlined in the first paper and focusing on the effectiveness of *PA* on social-emotional outcomes. Results indicated that *PA* had a significant effect on student self-reports of prosocial interactions, honesty, self-development, self-control,

respect for parents and teachers, empathy, altruism, positive actions/feelings, negative moral center, and aggressive problem solving. Additionally, it was found that *PA* had marginal effects on teacher-reported responsibility. Minimal differences by gender were found; these moderation analyses revealed significant effects for girls but not boys on honesty, self-control, and respect for teachers. No differences were found by mobility status. These findings suggest, that in comparison to "stayers" (or students who participated in a majority of the study), outcome trajectories are similar for students who left or joined the study, and were there for a shorter amount of time. This supports the theory of the program that *PA* changes the school climate in addition to student behavior; therefore students entering the trial are influenced immediately by the school climate. Additionally, exposure to the program at the beginning of the trial is influential on trajectories as well, such that students in the early part of the study are positively effected by *PA*.

Limitations

There are a few limitations to this study that should be mentioned. Most outcomes were measured by student self-report, potentially leading to a method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), which can inflate the observed relationships between the variables. Self-reports are also susceptible to social desirability bias; students may exaggerate their participation in high-risk behaviors in order to feel as if they fit in with their peers, or underreport such behaviors knowing society's negative views on behaviors such as substance use, violence, and bullying.

By the use of incentives and extensive reminders and incentives, we obtained Implementation Reports from an average of 75% of teachers at the end of each content unit, and up to 79% of teachers and 100% of school-based *PA* coordinators for the end-of-year implementation survey. These data revealed wide variability between schools in implementation fidelity, especially in early years, with improvements over time. It should also be noted that while schools with SECD-like interventions were excluded from participation in the study, some control schools did implement SECD-like activities after baseline (Social and Character Development Research Consortium, 2010). This makes our estimates of effect sizes conservative (Hulleman and Cordray, 2009).

There are several threats to validity (Shadish, Cook, and Campbell, 2002), and this present study was susceptible to some of these threats. With respect to statistical conclusion validity, restricted range of outcomes can weaken the relationship between variables (Shadish et al., 2002). The young age of the students at the beginning of the trial and the longitudinal nature of the study presented some measurement challenges, in that some outcomes have restricted ranges or show "floor" or "ceiling" effects. These restricted ranges violate the assumption of normality; therefore, it is inappropriate to run a model utilizing an estimator that assumes normality. To account for this challenge, we employed different models that were most appropriate for the distribution of each outcome. Additionally, low statistical power can lead one to conclude that there is no relationship between treatment and the outcome (Shadish et al., 2002). The small number of pairs (seven) and schools (14) could influence the

statistical power; however, that significant findings were found in spite of this limitation suggest that our findings may be conservative. With respect to external validity, it is possible that had the study been conducted in a different setting, with different schools, or under different conditions that we would not have found the same results. Therefore, the findings presented here are generalizable only to low-income, inner-city schools that would self-select to participate in a trial of this nature (Shadish et al., 2002). Finally, as has been seen in other studies within low income, urban school settings (Tobler & Komro, 2011), student mobility led to high turnover of students. However, the use of latent class analysis is an innovative method we have used to address this issue. One final limitation is the decreasing school size during the time of the study; nonetheless, no schools left the study.

Connections between the papers and future directions

The first paper in this study gave thorough detail about the Chicago CRCT of *PA*. This description consisted of the protocol used to analyze for program effects, including the primary analyses of three-level growth curves, secondary (moderating) analyses, and sensitivity analyses. This CRCT had the challenge of being conducted in an urban, low-income setting, creating the additional hurdle of dealing with a highly mobile population. In developing the design of the trial and reviewing this paper, this obstacle allowed for the development of several other papers. One paper focused on describing four approaches to dealing with mobility and addresses who to collect data from, who should be included in analysis, and to whom results can be generalized in cluster-randomized cohort trials (Vuchinich et al., 2012). The second is a paper on a

latent class analysis, of which preliminary methodology and results are presented in the current study. The lack of differences by mobility status suggests that despite the high mobility of this population, limited exposure to *PA* can still be influential on students.

The second paper uses the protocol outline in the first paper to analyze program effects on various social-emotional outcomes, and finds *PA* to be effective on these outcomes. Strong, positive effects on these outcomes support the theory of the program, and the effectiveness of SECD programs. Given the effectiveness of *PA* to influence social-emotional constructs, these can be tested as mediators of other outcomes. While SECD programs focus on the development of social-emotional skills, these programs also aim to improve other outcomes such as academic achievement, externalizing behaviors (e.g., violence, aggression, substance use), internalizing behaviors (e.g., anxiety, depression, negative affect), and health behaviors (e.g., exercising, eating fruits and vegetables). Future evaluations of *PA* should test SECD as a mediator of these outcomes. This will allow for a greater understanding of the mechanisms underlying *PA*; that is, how does *PA* work?

To date, only one paper has explored this mechanism (Lewis et al., 2012a). This paper found the second-order factor of SECD (made up of the six first order factors assessed in the current study) to completely mediate the relationship between *PA* and grade 8 substance use (including alcohol, tobacco, marijuana, and other serious drug use), such that students in *PA* had higher SECD scores than students in control schools, and students with higher SECD scores reported significantly less

substance use that those with lower SECD scores. This paper provides the first look at understanding how *PA* works to influence outcomes not directly addressed in program curriculum. That is, *PA*, as implemented in this setting, does not include antidrug messages or content. These preliminary findings support the importance developing these SECD skills in order to potentially have an effect on other domains of a student's life (e.g., academics, externalizing behaviors, internalizing behaviors). The present paper has shown effects on not only the six first-order factors that are included in the overall SECD measure but several other social-emotional constructs as well. These constructs, clearly related to SECD should be tested as potential mediators as well.

Together, these two papers involved a sample of students in a high-risk setting; generating improvements can be particularly difficult in urban areas facing rising poverty rates (DeNavas-Walt, Proctor, & Smith, 2009), health disparities (Braveman & Egerter, 2008), and cuts in social and educational programs (Johnson, Oliff, & Williams, 2010). The empirical evidence of effectiveness of a SECD program in a high-risk population, as demonstrated in the present study, should serve as a call to action for policymakers and school officials who are increasingly challenged to positively impact not only academic achievement, but also behavior and character development (Elias, 2006; Flay & Allred, 2010).

REFERENCES

*References are for scales in Table 2.3 (online supplement) of Manuscript 1 and are not referenced in text of Manuscript 1 or 2.

Aber, J. L., Brown, J. L., Jones, S., & Samples, F. (1995). Adapting measures of children's beliefs, attributions and skills for use in the evaluation of violence prevention projects. New York, NY: National Center for Children in Poverty, Columbia University, 1995. Retrieved from http://www.cdc.gov/ncipc/pub-res/pdf/YV/CDC_YV_SecIII.pdf.

Achenbach, T M. (1991). *Manual for the Child Behavior Checklist/4-J8 and 1991 profile*. Burlington: University of Vermont.

Afshar, P., & Kenny, M. (2008). Violence in schools, Cross-National and Cross-Cultural Perspectives. *Journal of Psychological Trauma*, 6, 87-89.
doi:10.1080/19322880802096624

Althof, W., & Berkowitz, M. W. (2006). Moral education and character education: Their relationship and roles in citizenship education. *Journal of Moral Education. Special Issue: The Moral Roots of Citizenship and Citizenship Education*, 35, 495-518. doi:10.1080/03057240601012204

Arthur, M. W., Hawkins, J. D., Pollard, J. A., Catalano, R. F., & Baglioni, A. J. (2002). Measuring risk and protective factors for substance use, delinquency, and other adolescent problem behaviors: The Communities That Care Youth Survey. *Evaluation Review*, 26, 575-601. doi:10.1177/0193841X0202600601

Aud, S., Fox, M., and KewalRamani, A. (2010). *Status and Trends in the Education of Racial and Ethnic Groups* (NCES 2010-015). U.S. Department of Education,

National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Bar-On, R. (2002). *Bar-On EQ-i:S Technical Manual*. Toronto: Multi-Health Systems.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182. doi:10.1037/0022-3514.51.6.1173

Bavarian, N., Lewis, K. M., DuBois, D. L., Acock, A., Vuchinich, S., Silverthorn, N., . . . Flay, B. R. (2012). Using social-emotional and character development to improve academic outcomes: A matched-pair, cluster-randomized trial in low-income, urban schools. Manuscript submitted for publication.

Beets, M. W., Flay, B. R., Vuchinich, S., Snyder, F. J., Acock, A., Li, K. K., et al. (2009). Use of a social and character development program to prevent substance use, violent behaviors, and sexual activity among elementary-school students in Hawaii. *American Journal of Public Health*, 99, 1438-1445. doi:10.2105/AJPH.2008.142919

Bell, C. C., & Jenkins, E. (1994). Effects of child abuse and race. *Journal of the National Medical Association*, 86, 165– 232. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/journals/655/>

Berkowitz, M. W., & Bier, M. C. (2004). Research-based character education. *The Annals of the American Academy of Political and Social Science*, 591, 72-85. doi:10.1177/0002716203260082

- Berkowitz, M. W., & Bier, M. C. (2007). What works in character education. *Journal of Research in Character Education*, 5, 29-48. Retrieved from <http://www.characterandcitizenship.org/services/testjournal.htm>
- Beunckens, C., Molenberghs, G., Verbeke, G., & Mallinckrodt, C. (2008). A latent-class mixture model for incomplete longitudinal Gaussian data. *Biometrics*, 64, 96-105. doi: 10.1111/j.1541-0420.2007.00837.x
- *Brand, S. A., Felner, R., Shim, M., Seitsinger, A., & Dumas, T. (2003). Middle school improvement and reform: Development and validation of a school-level assessment of climate, cultural pluralism, and school safety. *Journal of Educational Psychology*, 95, 570-588. doi:10.1037/0022-0663.95.3.570
- Braveman, P. A., Cubbin, C., Egerter, S., Williams, D. R., & Pamuk, E. (2010). Socioeconomic disparities in health in United States: What the patterns tell us. *American Journal of Public Health*, 100, S186-S196. doi:10.2105/AJPH.2009.166082
- Braveman, P., & Egerter, S. (2008). Overcoming obstacles to health: report from the Robert Wood Johnson Foundation to the Commission to Build a Healthier America. Princeton, NJ: Robert Wood Johnson Foundation.
- Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Series Eds.) & D. Kuhn (Vol. Ed.), *Handbook on child psychology Vol 1: Theoretical models of human development* (6th ed.; pp. 793-828). New York: Wiley.

- Brown, E. C., Graham, J. W., Hawkins, J. D., Arthur, M. W., Baldwin, M. M., Oesterle, S., . . . Abbott, R. D. (2009). Design and analysis of the Community Youth Development Study longitudinal cohort sample. *Evaluation Review*, 33, 311-334. doi:10.1177/0193841X09337356
- Brown, C. H., Wang, W., Kellam, S. G., Muthen, B. O., Petras, H., Toyinbo, P., . . . Windham, A. (2008). Models for testing and evaluating impact in randomized field trials: Intent-to-treat analyses for integrating the perspectives of person, place, and time. *Drug and Alcohol Dependence*, 95S, S74-S104. doi:10.1016/j.drugalcdep.2007.11.013
- Carlo, G., Crockett, L. J., Randall, B. A., & Roesch, S. C. (2007). A latent growth curve analysis of prosocial behavior among rural adolescents. *Journal of Research on Adolescence*, 17, 301-324. doi:10.1111/j.1532-7795.2007.00524.x
- Catalano R. F., Berglund M. L., Ryan J. A. M., Lonczak H. S., & Hawkins J. D. (2004). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *Annals of the American Academy of Political and Social Science*, Vol. 591, 98-124. doi:10.1177/0002716203260102
- *Centers for Disease Control and Prevention (2004). Methodology of the Youth Risk Behavior Surveillance System. RR-12. Retrieved from <http://www.cdc.gov/HealthyYouth/yrbs/index.htm>

Chicago Police Department (n.d.). Chicago Police Department Clear Map Crime

Incidents. Retrieved from

<http://gis.chicagopolice.org/CLEARMap/startPage.htm>.

*Chipuer, H., Pretty, G. H., Delorey, E., Miller, M., Powers, T., Rumstein, O., Barnes,

A., Cordasic, N., Laurent, K. (1999). The Neighborhood Youth Inventory:

Development and validation. *Journal of Community & Applied Social*

Psychology, 9, 355-368. doi:10.1002/(SICI)1099-

1298(199909/10)9:5<355::AID-CASP511>3.0.CO; 2-6

Ciarrochi, J., Heaven, P. C. L., & Davies, F. (2007). The impact of hope, self-esteem,

and attributional style on adolescents' school grades and emotional well-being:

A longitudinal study. *Journal of Research in Personality*, 41(6), 1161-1178.

doi:10.1016/j.jrp.2007.02.001

Coalition for Evidence Based Policy. (2002). *Bringing evidence-driven progress to*

education: A recommended strategy for the U.S. Department of Education.

Washington, DC: Coalition for Evidence-Based Policy.

Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F. D., et

al. (1966). *Equality of educational opportunity*. Washington, DC: U.S.

Government Printing Office.

*Conduct Problems Prevention Research Group (1991). Parent-Teacher Involvement

Measure (Online). Retrieved from

<http://www.fasttrackproject.org/techrept/p/ptp/index.php>

- Conduct Problems Prevention Research Group (1999). Initial impact of the Fast Track prevention trial for conduct problems I: The high-risk sample. *Journal of Consulting and Clinical Psychology*, 67, 631-647. doi:0022-006X/99/J3.00
- Conger, R. D., Ge, X., Elder, G. H., Jr., Lorenz, F. O., & Simons, R. L. (1994). Economic stressors, coercive family process, and developmental problems of adolescents. *Child Development*, 65, 541-561. doi:10.1111/j.1467-8624.1994.tb00768.x
- *Cook, E. T., Greenberg, M. T., & Kusche, C. A. (1995, March). People In My Life: Attachment relationships in middle childhood. Paper presented at the Society for Research in Child Development, Indianapolis, IN.
- Deci, E. L. (2009). Large-scale school reform as viewed from the self-determination theory perspective. *Theory and Research in Education*, 7, 244-253. doi:10.1177/1477878509104329
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum Press.
- Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125, 627-668. doi:10.1037/0033-2909.125.6.627
- Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research*, 71, 1-27. doi:10.3102/00346543071001001

- DeNavas-Walt, C., Proctor, B., & Smith, J. (2009). Income, Poverty, and Health Insurance Coverage in the United States, 2006. Washington, DC: US Census Bureau; August 2007.
- Donner, A., & Klar, N. (2004). Pitfalls and controversies in cluster randomization trials. *American Journal of Public Health, 94*, 416-422.
doi:10.2105/AJPH.94.3.416
- *DuBois, D. L., & Flay, B. R. (2003). *Self-Esteem Motivation Scale for Youth*. Unpublished measure, Institute for Health Research and Policy, University of Illinois at Chicago.
- *DuBois, D. L., Ji, P., & Flay, B. R. (2003). *Self-Esteem Processes Scale for Youth*. Unpublished measure, Institute for Health Research and Policy, University of Illinois at Chicago.
- *DuBois, D. L., Felner, R. D., Brand, S. Phillips, R. S. C., & Lease, A. M. (1996). Early adolescent self-esteem: A developmental-ecological framework and assessment strategy. *Journal of Research on Adolescence, 6*, 543-579.
Retrieved from <http://www.tandfonline.com/toc/hzzg20/current>
- 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.

- DuBois, D. L., Ji, P., Flay, B. R., Day, J. & Silverthorn, N. (2010). Further validation of the youth social and character development scale. Institute of Educational Sciences annual meeting, Washington, DC, June 28, 2010.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82, 405-432. doi:10.1111/j.1467-8624.2010.01564.x
- Eisenberg, N., Champion, C., & Ma, Y. (2004). Emotion-related regulation: An emerging construct. *Merrill-Palmer Quarterly*, 50, 236–259.
doi:10.1353/mpq.2004.0016
- Eisenberg, N., & Fabes, R.A. (1998). Prosocial development. In W. Damon (Ed.), *Handbook of Child Psychology* (5th Ed., pp. 701-778). New York: John Wiley & Sons.
- Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In W. Damon (Series Ed.), & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 646-718). New York: Wiley.
- Eisenberg, N., & Morris, A. S. (2004). Moral cognitions and prosocial responding in adolescence. In R. M. Lerner & L. D. Steinberg (Eds.), *Handbook of Adolescent Psychology* (2nd Ed.). (155-188). Hoboken, New Jersey: John Wiley & Sons.
- Elias, M. J. (2006). The connection between academic and social and emotional

learning. In M. J. Elias & H. Arnold (Eds.), *The educator's guide to emotional intelligence and academic achievement* (pp. 4–14). Thousand Oaks, CA: Corwin Press.

Elias, M. J. (2009). Social-emotional and character development and academics as a dual focus of educational policy. *Educational Policy*, 23(6), 831-846.
doi:10.1177/0895904808330167

Elias, M. J., & Hayes, N. M. (2008). Social competence, social support, and academic achievement in minority, low-income, urban elementary school children. *School Psychology Quarterly*, 23, 474-495. doi:10.1037/1045-3830.23.4.474

Elliott, S. N., Gresham, F. M., Freeman, T., & McCloskey, G. (1988). Teacher and observer ratings of children's social skills: Validation of the Social Skills Rating Scales. *Journal of Psychoeducational Assessment*, 6, 152-161.
doi:10.1177/073428298800600206

*Elliott, D. S., Wilson, W. J. Huizinga, D., Sampson, R. J., Elliott, A., & Rankin, B. (1996). The effects of neighborhood disadvantage on adolescent development. *Journal of Research in Crime and Delinquency*, 33, 389-426.
doi:10.1177/0022427896033004002

Eron, L. D., Gentry, J. H., and Schlegel, P. "*Reason to hope: A psychosocial perspective on violence and youth*". Washington, DC: American Psychological Association, 1994.

Fabes, R. A., Gaertner, B. M., & Popp, T. K. (2006). Getting along with others: Social competence in early childhood. In K. McCartney & D. Phillips (Eds.),

Blackwell handbook of early childhood development. (pp.297-316). Malden, MA: Blackwell Publishing.

Farahmand, F. K., Grant, K. E., Polo, A. J., Duffy, S. N., & DuBois, D. L. (2011).

School-based mental health and behavioral programs for low-income, urban youth: A systematic and meta-analytic review. *Clinical Psychology: Science and Practice*, 18, 372-390. doi:10.1111/j.1468-2850.2011.01265.x

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

Retrieved from <http://www.ingentaconnect.com/content/png/ajhb>

Flay, B.R., & Allred, C.G. (2003) Long-term effects of the *Positive Action* program.

American Journal of Health Behavior, 27, S6-S21. Retrieved from <http://www.ingentaconnect.com/content/png/ajhb>

Flay, B.R., & Allred, C.G. (2010). The Positive Action Program: Improving academics, behavior and character by teaching comprehensive skills for successful learning and living. In T. Lovat & R. Toomey (Eds.), *International Handbook on Values Education and Student Well-Being* (pp. 471-501).

Dordrecht: Springer.

Flay, B. R., & Collins, L. M. (2005). Historical review of school-based randomized trials for evaluating problem behavior prevention programs. *The ANNALS of the American Academy of Political and Social Sciences*, 599, 115-146.

doi:10.1177/0002716205274941

- Flay, B. R., & Petraitis, J. (1994). The theory of triadic influence: A new theory of health behavior with implications for preventive interventions. *Advances in Medical Sociology*, 4, 19-44. Retrieved from <http://www.emeraldinsight.com/books.htm?issn=1057-6290>
- Flay, B. R., Allred, C. G., & Ordway, N. (2001). Effects of the *positive action* program on achievement and discipline: Two matched-control comparisons. *Prevention Science*, 2, 71-89. doi:1389-4986/01/0600-0071\$19.50/1
- Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., . . . Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science*, 6, 151-175. doi:10.1007/s11121-005-5553-y
- Flay, B. R., Snyder, F. J., & Petraitis, J. (2009). The theory of triadic influence. In R. J. DiClemente, R. A. Crosby & M. C. Kegler (Eds.), *Emerging Theories in Health Promotion Practice and Research* (Second ed., pp. 451-510). San Francisco, CA: Jossey-Bass.
- Fleming, C. B., Haggerty, K. P., Catalano, R. F., Harachi, T. W., Mazza, J. J., & Gruman, D. H. (2005). Do social and behavioral characteristics targeted by preventive interventions predict standardized test scores and grades? *Journal of School Health*, 75, 342-349. doi:10.1111/j.1746-1561.2005.00048.x
- *Forehand, R., Brody, G.H., & Armistead, L., et al. (2000). The role of community risks and resources in the psychosocial adjustment of at-risk children: An

examination across two community contexts and two informants. *Behavior Therapy*, 13, 395-414. doi:10.1016/S0005-7894(00)80022-2

Funk, J., Fox, C., Chan, M., & Curtiss, K. (2008). The development of the Children's Empathic Attitudes Questionnaire using classical and Rasch analyses. *Journal of Applied Developmental Psychology*, 29, 187-196. doi: 10.1016/j.appdev.2008.02.005

*Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, 95, 148-162. doi:10.1037/0022-0663.95.1.148

Giraudeau, B., & Rivaud, P. (2009). Preventing bias in cluster randomized trials. *PLoS Medicine* 6: e1000065.

Glaser, R., Van Horn, M., Arthur, M., Hawkins, J., & Catalano, R. (2005). Measurement properties of the Communities That Care Youth Survey across demographic groups. *Journal of Quantitative Criminology*, 21, 73-102. doi:10.1007/s10940-004-1788-1

Goodman, R. & Goodman, A (2009). Strengths and difficulties questionnaire as a dimensional measure of child mental health. *Journal of the American Academy of Child and Adolescent Psychiatry*, 48, 400-403. doi:10.1097/CHI.0b013e3181985068

*Gonzales, N. A., Gunnoe, M. L., Jackson, K. M., & Samaniego, R. Y. (n.d). Validation of a multicultural events scale for urban adolescents: Preliminary

strategies for enhancing cross-ethnic and cross-language equivalence.

Unpublished manuscript.

- Gottfried, A., Marcoulides, G., Gottfried, A., & Oliver, P. (2009). A latent curve model of parental motivational practices and developmental decline in math and science academic intrinsic motivation. *Journal of educational psychology, 101*, 729-739. doi:10.1037/a0015084
- Grant, K. E., Katz, B. N., Thomas, K. J., O’Koon, J. H., Meza, C. M., DiPasquale, A. M., et al. (2004). Psychological symptoms affecting low-income urban youth. *Journal of Adolescent Research, 19*, 613–634. doi:10.1177/0743558403260014
- Hamilton, L. S., Stecher, B. M., Marsh, J. A., Sloan McCombs, J., Robyn, A., Russell, J., et al. (2007). *Standards-based accountability under no child left behind: Experiences of teachers and administrators in three states*. Santa Monica, CA: RAND Corporation.
- Hawkins, J. D., Kosterman, R., Catalano, R. F., Hill, K. G., & Abbott, R. D. (2005). Promoting Positive Adult Functioning Through Social Development Intervention in Childhood: Long-term Effects From the Seattle Social Development Project. *Archives of Pediatric and Adolescent Medicine, 159*, 25-31. doi:10.1001/archpedi.159.1.25
- *Huebner, E. S. (1991). Further validation of the Students' Life Satisfaction Scale: The independence of satisfaction and affect ratings. *Journal of Psychoeducational Assessment, 4*, 363-368. doi:10.1177/073428299100900408

- *Huesmann, L.R., & Guerra, N.G. (1997). Children's normative beliefs about aggression and aggressive behavior. *Journal of Personality and Social Psychology*, 72, 408-419. doi:10.1037/0022-3514.72.2.408
- Hulleman, C.S., and Cordray, D.S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness*, 1, 88-110. doi: 10.1080/19345740802539325
- Ji, P. DuBois, D.L., & Flay, B. (2011). Social-Emotional and Character Development Scale: Development and Initial Validation with Elementary School Students. Unpublished manuscript.
- Ji, P., DuBois, D. L., Flay, B. R., & Brechling, V. (2008). "Congratulations, you have been randomized into the control group!(!)": Issues to consider when recruiting schools for matched-pair randomized control trials of prevention programs. *Journal of School Health*, 78, 131-139. doi:10.1111/j.1746-1561.2007.00275.x
- Ji, P.Y., Pokorny, S.B., & Jason, L.A. (2004). Factors influencing middle and high schools' active parental consent return rates. *Evaluation Review*, 28, 578-591. doi:10.1177/0193841X04263917
- Johnson, N., Oliff, P., & Williams, E. (2010). An update on state budget cuts: At least 46 states have imposed cuts that hurt vulnerable residents and the economy. Washington, DC: Center on Budget and Policy Priorities. Retrieved from <http://www.cbpp.org/cms/?fa=view&id=1214>
- *Kaminski, J. W., Battistich, V. A., David-Ferdon, C., & Social and Character

- Development Research Consortium. (2009). Development and validation of outcome measures for the school-based and character development program. *Journal of Research in Character Education*, 7, 51-73. Retrieved from <http://www.characterandcitizenship.org/services/testjournal.htm>
- King, P. E., Dowling, E. M., Mueller, R. A., White, K., Schultz, W., Osborn, P., et al. (2005). Thriving in adolescence: The voices of youth-serving practitioners, parents, and early and late adolescents. *The Journal of Early Adolescence*, 25, 94-112. doi:10.1177/0272431604272459
- Knotterus, J. A., & Bouter, L. M. (2001). The ethics of sample size: Two-sided testing and one-sided thinking. *Journal of Clinical Epidemiology*, 54(2), 109-110. doi:10.1016/s0895-4356(00)00276-6
- Kokko, K., Tremblay, R. E., Lacourse, E., Nagin, D. S., & Vitaro, F. (2006). Trajectories of prosocial behavior and physical aggression in middle childhood: Links to adolescent school dropout and physical violence. *Journal of Research on Adolescence*, 16, 403-428. doi:10.1111/j.1532-7795.2006.00500.x
- *Laurent, J., Catanzaro, S. J., Joiner, T. E., Rudolph, K.D., Potter, K.I., Lambert, S., Osborne, L., Gathright, T. (1999). A measure of positive and negative affect for children: Scale development and preliminary validation. *Psychological Assessment*, 11, 326-338. doi:10.1037/1040-3590.11.3.326

- Laursen, B. & Pursell, G. (2009). Conflict in peer relationships. In K.H. Rubin, W.M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups* (pp.267-286). New York: The Guilford Press.
- Leffert, N., Benson, P. L., Scales, P. C., Sharma, A. R., Drake, D. R., & Blyth, D. A. (1998). Developmental assets: Measurement and prediction of risk behaviors among adolescents. *Applied Developmental Science, 2*, 209-230.
doi:10.1207/s1532480xads0204_4
- Lerner, J. V., Phelps, E., Forman, Y., and Bowers, E. P. (2009). Positive youth development. In R.M. Lerner and L. Steinberg (Eds.). *Handbook of Adolescent Psychology* (3rd ed., Vol. 1), 524-558. Hoboken, NJ: John Wiley and Sons, Inc.
- Lerner, R. M., Almerigi, J. B., Theokas, C., & Lerner, J. V. (2005). Positive youth development: A view of the issues. *Journal of Early Adolescence, 25*, 10- 16.
doi: 10.1177/0272431604273211
- Lerner, R. M., Dowling, E. M., & Anderson, P. M. (2002). Positive youth development: Thriving as the basis of personhood and civil society. *Applied Developmental Science, 7*, 172-180. doi:10.1207/S1532480XADS0703_8
- Lerner, R. M., Lerner, J. V., Phelps, E. et al. (2008). The positive development of youth technical report the 4-h study of positive youth development: Report of the findings from the first four waves of data collection: 2002-2003, 2003-2004, 2004-2005, and 2005-2006. Retrieved from
<http://ase.tufts.edu/iaryd/documents/4HStudyAnnualReport2008.pdf>

- Lewis, K. M., Bavarian, N., Snyder, F. J., Acock, A., Day, J., DuBois, D. L., . . . Flay, B. R. (2012a). Direct and Mediated Effects of a Social-Emotional and Character Development Program on Adolescent Substance Use. *International Journal of Emotional Education*, 4, 56-78. Retrieved from <http://www.enseceurope.org/journal/>
- Lewis, K. M., DuBois, D. L., Ji, P., Day, J., Silverthorn, N., Bavarian, N., . . . Flay, B. (2012b). Design, sample, and planned analysis of the Chicago randomized control trial of the *Positive Action* program. Manuscript submitted for publication.
- Li, K., Washburn, I., DuBois, D. L., Vuchinich, S., Ji, P., Brechling, V., et al. (2011). Effects of the *Positive Action* programme on problem behaviors in elementary school students: A matched-pair randomised control trial in Chicago. *Psychology & Health*, 26, 187-204. doi:10.1080/08879446.2011.531574
- Lin, H., McCulloch, C. E., & Rosenheck, R. A. (2004). Latent pattern mixture models for informative intermittent missing data in longitudinal studies. *Biometrics*, 60, 295-305.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications, Inc.
- *Loeber, R., & Dishion, T.J. (1983). Early predictors of male delinquency: A review. *Psychological Bulletin*, 94, 68-99. doi:10.1037/0033-2909.94.1.68
- MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York, NY: Lawrence Erlbaum Associates.

- MacKinnon, D. P., Krull, J. L., and Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1*, 173-181. doi:1389-4986/00/1200-0173
- MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., and Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83-104. doi:10.1037/1082-989X.7.1.83
- Malecki, C. K., & Elliott, S. N. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly, 17*, 1-23. doi:10.1521/scpq.17.1.1.19902
- Marsh, H. W., Ludtke, O., Trautwein, U., and Morin, A.J.S. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling: A Multidisciplinary Journal, 16*, 191-225. doi:10.1080/10705510902751010
- *Matheny, A.P., Wachs, T.D., Ludwig, J.L., & Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the Confusion, Hubbub, and Order Scale. *Journal of Applied Developmental Psychology, 16*, 429-444. doi:10.1016/0193-3973(95)90028-4
- McClelland, M. M., Ponitz, C. C., Messersmith, E., & Tominey, S. L. (2010). Self-regulation: The integration of cognition and emotion. In R. M. Lerner (Series

- Ed.) & W. F. Overton (Vol. Ed.), *Handbook of lifespan human development: Vol. 1. Cognition, biology and methods* (pp. 509-553). Hoboken, NJ: Wiley.
- Merlo, J., Chaix, B., Ohlsson, H., Beckman, A., Johnell, K., Hjerpe, P., et al. (2006). A brief conceptual tutorial of multilevel analysis in social epidemiology: Using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *Journal of Epidemiology and Community Health*, 60, 290-297. doi:10.1136/jech.2004.029454
- Morone, N. E., Greco, C. M., Rollman, B. L., Moore, C. G., Lane, B., Morrow, L., et al. (2012). The design and methods of the aging successfully with pain study. *Contemporary Clinical Trials*, 33, 417-425. doi:10.1016/j.cct.2011.11.012
- Murray, D. M. (1998). *Design and Analysis of Group-Randomized Trials*. New York: Oxford University Press.
- Murray, D. M., Feldman, H. A., McGovern, P. G. (2000). Components of variance in a group-randomized trial analysed via a random-coefficients model: the Rapid Early Action for Coronary Treatment (REACT) trial. *Statistical Methods in Medical Research*, 9, 117-133. doi:10.1177/096228020000900204
- Nantel-Vivier, A., Kokko, K., Caprara, G. V., Pastorelli, C., Gerbino, M. G., Paciello, M., et al. (2009). Prosocial development from childhood to adolescence: a multi-informant perspective with Canadian and Italian longitudinal studies. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 50, 590-598. doi:10.1111/j.1469-7610.2008.02039.x

O'Donnell, J., Hawkins, J. D., Catalano, R. F., Abbott, R. D., & Day, L. E. (1995).

Preventing school failure, drug use, and delinquency among low-income children: Long-term intervention in elementary schools. *American Journal of Orthopsychiatry*, 65, 87-100. doi:10.1037/h0079598

*Orpinas, P., & Frankowski, R. (2001). The Aggression scale: A self-report measure of aggressive behavior for young adolescents. *Journal of Early Adolescence*, 21, 50-67. doi:10.1177/0272431601021001003

*Orpinas, P., & Kelder, S. (1995). Students for Peace Project: Second student evaluation. Houston, TX: University of Texas Health Science Center at Houston, School of Public Health.

Park, N. (2004). Character strengths and positive youth development. *The ANNALS of American Academy of Political and Social Science*, 591, 40-54. doi:10.1177/0002716203260079

Park, N., & Peterson, C. (2009). Strengths of character in schools. In R. Gilman, E.S. Huebner, & M.J. Furlong (Eds.) *Handbook of Positive Psychology in Schools* (pp. 65-76). New York: Routledge.

Park, N., Peterson, C., & Seligman, M. E. P. (2004). Strengths of character and well being. *Journal of Social & Clinical Psychology*, 23, 603-619. Retrieved from http://www.guilford.com/cgi-bin/cartscript.cgi?page=pr/jnsc.htm&dir=periodicals/per_psych&cart_id=

- *Petrides, K. V., Sangareau, Y., Furnham, A., & Frederickson, N. (2006). Trait emotional intelligence and children's peer relations at school. *Social Development, 15*, 537-547. doi:10.1111/j.1467-9507.2006.00355.x
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*, 879-903. doi:10.1037/0021-9010.88.5.879
- Positive Action, Inc. (2007). Evaluation services. Retrieved from <http://www.positiveaction.net/support/index.asp?ID1=2&ID2=215>
- Purkey, W. W. (1970). *Self-concept and school achievement*. Englewood Cliffs, NJ: Prentice-Hall.
- Purkey, W. W., and Novak, J. (1984). *Inviting school success: A self-concept approach to teaching and learning*. Belmont, CA: Wadsworth.
- Rabe-Hesketh, S. and Skrondal, A. (2008). *Multilevel and Longitudinal Modeling using Stata (Second Edition)*. College Station, TX: Stata Press.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. (2nd ed.). Newbury Park, CA: Sage.
- *Reynolds, C., & Kamphaus, R. (2002). *The clinician's guide to the Behavior Assessment System for Children (BASC)*. New York, NY: Guilford Press.
- *Reynolds, C.R., & Kamphaus, R.W. (1998). *Behavioral Assessment System for Children*. Circle Pines, MN: American Guidance Service Inc.

- Riggs, N. R., Greenberg, M. T., Kusché, C. A., & Pentz, M. A. (2006). The mediational role of neurocognition in the behavioral outcomes of a social-emotional prevention program in elementary school students: Effects of the PATHS curriculum. *Prevention Science*, 7, 91-102. doi:10.1007/s11121-005-0022-1
- Rose-Krasnor, L., & Denham, S. (2009). Social-emotional competence in early childhood. In K.H. Rubin, W.M. Bukowski, & B. Laursen (Eds.) *Handbook of Peer Interactions, Relationships, and Groups* (pp.162-179). New York: Guilford Press.
- Rosenblum, G.D., & Lewis, M. (2003). Emotional development in adolescence. In G.R. Adams & M.D. Berzonsky (Eds.), *Blackwell handbook of adolescence* (pp. 269-289). Malden, MA: Blackwell.
- Roy, J. (2003). Modeling longitudinal data with nonignorable dropouts using a latent dropout class model. *Biometrics*, 59, 829-836. doi:10.1111/j.0006-341X.2003.00097.x
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68-78. doi:10.1037/0003-066X.55.1.68
- Ryan, R. M., & Deci, E. L. (2006). Self-regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will? *Journal of Personality*, 74, 1557-1586. doi:10.1111/j.1467-6494.2006.00420.x

- *Sampson, R.J., Morenoff, J.D., & Earls, F. (1999). Beyond social capital: Spatial dynamics of collective efficacy for children. *American Sociological Review*, 64, 633-660. Retrieved from <http://www.jstor.org/action/showPublication?journalCode=amersocirevi>
- Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, 8, 448-467. doi:10.1037/1082-989X.8.4.448
- Schochet, P., & Novak, T. (2003). Computer program to construct school pairs. Unpublished manuscript.
- Schultz, L. H., Selman, R. L., & LaRusso, M. D. (2003). The assessment of psychosocial maturity in children and adolescents: Implications for the evaluation of school-based character education programs. *Journal of Research in Character Education*, 1, 67-87. Retrieved from <http://www.characterandcitizenship.org/services/testjournal.htm>
- Selmen, R. L., (2003). The promotion of social awareness: Powerful lessons from the partnership of development theory and classroom practice. New York: Russell Sage Foundation.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized casual inference*. Boston: Houghton-Mifflin.

- *Shelton, K.K., Frick, P.J., & Wootton, J. (1996). Assessment of parenting practices in families of elementary school-age children. *Journal of Clinical Child Psychology, 25*, 317-329. doi:10.1207/s15374424jccp2503_8
- Smart, D., & Sanson, A. (2003). Social competence in young adulthood, its nature and antecedents. *Family Matters, 64*, 4–9. Retrieved from: <http://www.aifs.gov.au/institute/pubs/fammats.html>.
- Snyder, F.J., Flay, B. R., Vuchinich, S., Acock, A., Washburn, I. J., Beets, M., & Li, K. (2010). Impact of a social-emotional and character development program on school-level indicators of academic achievement, absenteeism, and disciplinary outcomes: A matched-pair, cluster-randomized, controlled trial. *Journal of Research on Educational Effectiveness, 3*, 26-55.
doi:10.1080/19345740903353436
- Snyder, F. J., Vuchinich, S., Acock, A. C., Washburn, I. J., & Flay, B. R. (2012). Improving elementary-school quality through the use of a social-emotional and character development program: A matched-pair, cluster-randomized, controlled trial in Hawai`i. *Journal of School Health, 82*, 11-20.
doi:10.1111/j.1746-1561.2011.00662.x
- Snyder, F. J., and Flay, B. R. (in press). Positive youth development. In A. Higgins-D'Alessandro, M. Corrigan and P. Brown (Eds.), *The Handbook of Prosocial Education*. Lanham, MD: Rowman and Littlefield Publishing Group.
- Social and Character Development Research Consortium (2010). *Efficacy of Schoolwide programs to promote social and character development and*

reduce problem behavior in elementary school children (NCER 2011-2001).

Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Solomon, D., Battistich, V., Watson, M., Schaps, E., & Lewis, C. (2000). A six-district study of educational change: Direct and mediating effects of the Child Development Project. *Social Psychology of Education*, 4, 3-51.
doi:10.1023/A:1009609606692.

Spinard, T.L., & Eisenberg, N. (2009). Empathy, prosocial behavior, and positive development in schools. In R. Gilman, E.S. Huebner, & M.J. Furlong (Eds.) *Handbook of Positive Psychology in Schools* (pp. 119-130). New York: Routledge.

Thompson, T. L. 1984. A comparison of methods of increasing parental consent rates in social research. *Public Opinion Quarterly*, 48, 779-787. doi:10.1086/268883

Thompson, R. A., & Lagattuta, K. H. (2006). Feeling and understanding: Early emotional development. In K. McCartney & D. Phillips (Eds.), *Blackwell handbook of early childhood development*. (p.317-337). Malden, MA: Blackwell Publishing.

Tobler, A. L., and Komro, K. A. (2011). Contemporary options for longitudinal follow-up: Lessons learned from a cohort of urban adolescents. *Evaluation and Program Planning*, 34, 87-96. doi:10.1016/j.evalprogplan.2010.12.002

- Vuchinich, S., Flay, B., Aber, L., & Bickman, L. (2012). Person mobility in the design and analysis of cluster-randomized cohort prevention trials. *Prevention Science*. Advance online publication. doi:10.1007/s11121-011-0265-y
- Walker, H. M., & McConnell, S. R. (1995). *The Walker-McConnell scale of social competence and school adjustment (SSCSA)*. San Diego, CA: Singular.
- Walleser, S., Hill, S. R., & Bero, L. A. (2011). Characteristics and quality of reporting of cluster randomized trials in children: Reporting needs improvement. *Journal of Clinical Epidemiology*, 64, 1331-1340.
doi:10.1016/j.jclinepi.2011.04.006
- Washburn, I. J., Acock, A., Vuchinich, S., Snyder, F., Li, K.K., Ji, P., et al. (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
- Weissberg, R. P., and O'Brien, M. U. (2004). What works in school-based social and emotional learning programs for positive youth development. *The Annals of the American Academy of Political and Social Science*, 591, 86-97.
doi:10.1177/0002716203260093
- Wentzel, K. R. (1993). Does being good make the grade? Social behavior and academic competence in middle school. *Journal of Educational Psychology*, 85, 357-364. doi:10.1037/0022-0663.85.2.357
- Wheeler, V A., & Ladd, G. W. (1982). Assessment of children's self-efficacy for

social interactions with peers. *Developmental Psychology*, 18, 795-805.

doi:10.1037/0012-1649.18.6.795

Wigelsworth, M., Humphrey, N., Kalamhouka, A., & Lendrum, A. (2010). A review of key issues in the measurement of children's social and emotional skills.

Education psychology in practice, 26, 173-186.

doi:10.1080/02667361003768526

Wilson, D., O'Brien, D., & Sesma, A. (2009). Human prosociality from an evolutionary perspective: Variation and correlations at a city-wide scale.

Evolution and Human Behavior, 30, 190-200. Retrieved from

<http://www.ehbonline.org/>

Winters-Stone, K. M., Luyons, K. S., Nail, L. M., & Beer, T. M. (2012). The Exercising Together project: Design and recruitment for a randomized, controlled trial to determine the benefits of partnered strength training for couples coping with prostate cancer. *Contemporary Clinical Trials*, 33, 342-350. doi:10.1016/j.cct.2011.10.015