# A Randomized Trial of the Effects of an Enhanced After-School Program for Middle-School Students

June, 2010

Denise C. Gottfredson, Amanda B. Cross, Denise M. Wilson, Nadine Connell, & Melissa Rorie

This work was supported through Grant R305F050069 from the U.S. Department of Education Institute of Educational Sciences to the University of Maryland. We acknowledge the support of the Baltimore County Local Management Board, the Baltimore County Department of Recreation and Parks, and the Baltimore County Public Schools for implementing and managing the after-school programs. We especially acknowledge the assistance of Elise Andrews of the Baltimore County Local Management Board and Beahta Davis of the Baltimore County Department of Recreation and Parks. We would like to thank Gordon Bonham for collecting data from the school system and Stephanie DiPietro, Mathew Gugino, Lynda Okeke, Matthew Brigham, Freshta Rahimi, and Jaynie Trageser for research assistance. Thanks also to Elise Andrews, Shawn Bushway, J. David Hawkins, and Rebecca Maynard for comments on prior drafts of this report. All authors except Connell are with the Department of Criminology & Criminal Justice, University of Maryland, College Park. Connell is with the Department of Law and Justice Studies, Rowan University.

## Table of Contents

## List of Tables

## List of Figures

## List of Appendix Tables

1. Abstract

**Background:** Considerable government funds are invested in after-school programs (ASPs) intended to improve academic performance and to alter related student behaviors (e.g., attendance, drug use, and conduct). Nevertheless, rigorous research on the effects of ASPs on those outcomes is sparse and results are mixed.

**Purpose:** To assess the extent to which the routine practices of an implementing agency can be shifted in the direction of providing more research-based programming, and to measure the effects of doing so on a range of outcomes for middle-school youths, including academic performance, school attendance, and conduct problems, as well as attitudes, beliefs, and behaviors related to these outcomes.

**Setting:** ASPs in five public middle schools in Baltimore County, Maryland. Registration was open to all students who attended the participating schools, but principals were asked to encourage youths whom they considered especially "at risk" for academic or behavioral problems to register. The participating schools served high percentages of minority youth (47–99% minority population) and many students who received subsidized meals (64–67% receiving free or reduced lunch).

**Subjects:** $N = 447$ students in grades 6–8. About half of this sample were males (54%), and 70% were African Americans. The average age for participants was 12 years, and 59% of students received free or reduced meals at school.

**Intervention:** The program, which was free to participants, operated for 9 hours per week for 30 weeks. The plan for the enhanced ASP included attendance monitoring and reinforcement, structured tutoring, and the All Stars curriculum (http://www.allstarsprevention.com/). The attendance monitoring and reinforcement component was intended to provide rewards to individual students with good attendance in school and small groups with good attendance in the ASP. Structured tutoring was to be conducted by teachers, adult volunteers, and older students for 1.5 hours per week. The All Stars curriculum focuses on building attitudes and beliefs that are inconsistent with substance use and other risky behaviors, and on teaching skills necessary for healthy decision making. It was to be delivered by ASP staff for 1.5 hours per week. Previous research showed evidence of effectiveness for each of these components when offered during the school day. These enhancements were to be added to a typical ASP offered in the county, which consisted mainly of sports, crafts, snack, and so on.

**Implementation:** The enhancement plan was partially implemented. Although the planned incentives for attending school were implemented, the group-based incentives for attending the ASP were not. The structured tutoring program was replaced with a group-based academic assistance program that resembled the academic services typically available in ASPs. The All Stars prevention curriculum was implemented as planned, but high dropout and sporadic attendance limited student exposure. In the end, the study provided a rigorous test of a reasonably well-structured program that resembled a typical ASP except that it also included a research-based prevention curriculum, All Stars.

**Research Design:** Students were randomly assigned within school to the enhanced ASP (treatment n = 224) or to a "treatment as usual" control group (n = 223).

**Control Condition:** The services offered to the control group included one after-school activity per month, usually a special event or party. Fifty-two percent of control group members never attended these events. Control group students were free to enroll in any available after-school activity other than the experimental program. Virtually all (96%) members of the control group participated in some organized after-school activity. Nearly 60% participated in an after-school activity at their schools. This level of after-school activity participation is comparable with levels reported by same-aged youths in national surveys.

**Key Measures:** The following measures were collected preprogram and postprogram: school records on attendance, grades, promotion, achievement test scores, discipline records, and youth surveys measuring academic outcomes and student attitudes as well as experiences and behaviors (social skills related to substance abuse, attitudes about substance abuse, substance use, aggression, delinquency, victimization, school conduct, educational plans and aspirations, commitment to academics, and studying behavior). Teacher ratings of conduct, academic competence, and social skills were collected at the end of the program year. A survey measuring exposure to after-school activities was administered midway through the program year. Program observations were conducted twice per month, and implementation data were collected daily.

**Data Analytic Strategy:** Regression models appropriate for each dependent variable treated gender, race, age, and a nominal variable measuring school to correct for the clustering of

individual cases within school as covariates. A covariate for the outcome variable measured at pretest was also included to increase power. Covariate-adjusted posttest means were calculated from these regression models. Standardized mean difference effects size statistics were calculated using the adjusted posttest means. Exploratory analyses also examined the effects of actual ASP participation as well as the possibility that program effectiveness depended on certain characteristics of the participants or the programs.

**Findings:** The program was not fully implemented in ways that would be expected to achieve the desired student outcomes. Assignment to the treatment condition resulted in a substantial increase in the level of participation in both the experimental ASP and the school-based ASPs more generally, relative to the control students. Thus, treatment youths experienced increased exposure to attendance incentives, academic assistance, and the All Stars curriculum. However, nearly all youth in both the treatment and control conditions reported participating in some organized activity during the after-school hours, and the number of different activities in which treatment and control youths reported being involved did not differ. Participation in the treatment reduced time spent with friends with no adults present during the after-school hours. The magnitude of this effect was small relative to our expectation: Youths attending the ASP reported being with their friends with no adults present for approximately half a day per week less than control youths. No significant posttest differences between treatment and control youths were found on measures of conduct problems, academic performance, school attendance, prosocial/antidrug attitudes, social competence, school bonding, or positive peer influence. Frequent attenders did not have different outcomes than infrequent attenders, and the analysis of moderators showed that differential effects for subgroups of youth were detected in fewer

instances than predicted by chance. Although the quality of implementation varied across the five implementing sites, we found no evidence that exposure to higher quality programs resulted in more beneficial outcomes for participating youths. At least within the range of program quality observed in this study, the program did not produce the expected effects.

**Conclusion:** The findings from this study are broadly consistent with the results from the 21st Century Learning Community evaluation (Dynarski et al., 2003) and with the results from smaller randomized controlled trials examining the benefits of after-school programs (e.g., Zief et al., 2006). The findings suggest that it is difficult to achieve high fidelity in the implementation of research-based practices in the typical ASP setting. The modest improvements in access to and quality of after-school options available to treatment youth in this study are unlikely to lead to desired outcomes. Because the program was not implemented as intended, the findings from this study should not be interpreted as estimates of the impacts of the intended program. They should also not be interpreted as indicating that ASPs in general are not beneficial, as youth in the control group also participated in ASPs.

## 2. Background and Purpose

After-school programming has been increasing in the United States. Considerable federal, state, local, and private monies are being invested in these programs. For example, the 21st Century Community Learning Center Program received approximately $1 billion in federal funds annually from 2002 to 2007 to provide before- and after-school enrichment for students in low-performing schools. Estimates of total annual federal investment in out-of-school time have reached as high as $3.6 billion (financeproject.org, 2007).

The rising popularity of after-school programs (ASPs) results primarily from new demands for accountability in education and the need for after-school care for children of working parents (Beckett, Hawken, & Jacknowitz, 2001; D. C. Gottfredson, Gertenblith, Soulé, Womer, & Lu, 2004; Kane, 2004; Lauer et al., 2006). Concerns about delinquency prevention are also linked to demand for ASPs, as the after-school hours present the highest risk of arrest for juveniles (D. C. Gottfredson, G. D. Gottfredson, & Weisman, 2001; Sickmund, Snyder, & Poe-Yamagata, 1997). The intuitive appeal of ASPs rests on the perception that unsupervised after-school time is either dangerous or simply wasted for adolescents. ASPs provide an opportunity to enhance academic learning and to introduce positive adult role models, and they may provide shelter from unsafe neighborhoods for low-income children in urban areas. ASPs are also a convenient platform on which to provide social and personal skills instruction that may be deemphasized in the classroom.

Addressing these objectives via ASPs presents the same hurdles faced by all community- and school-based intervention strategies, such as recruitment and retention of participants; determining the needs of the target group and setting reasonable goals for change; hiring,

training, and maintaining well-qualified staff; formulating and implementing a successful curriculum or tailoring an existing curriculum to suit the specific population and goals of the program; and gaining the support of community and governmental agencies. This process is more efficient when best practice recommendations are available, but best practices research on ASPs is still in its infancy. The picture painted by existing research on ASPs (summarized in the following section) is one of tremendous heterogeneity, in terms of both programming and outcomes.

## Prior Research

Recent reviews on the effectiveness of after-school programming generally agree that, "ASPs are *capable* of improving important youth outcomes" (emphasis added; Granger, Durlak, Yohalem, & Reisner, 2007, p. 3), but very little can confidently be said about how they can achieve success. Additionally, many programs have been shown to have no effect on youth outcomes, and in some cases, ASP participants experience negative outcomes (e.g., conduct problems, increased substance use, and negative peer influence) in comparison with nonparticipants (Dynarski et al., 2003; Mahoney, 2000; Weisman et al., 2002). The uncertainty about the direction and magnitude of the effects of ASP participation results from the generally poor quality of ASP research. Very few studies of ASP effectiveness meet contemporary standards for scientific rigor in program evaluation (Flay et al., 2005). Conclusions from reviews differ depending on where they have set the cutoff for scientific rigor in deciding which studies to include. Similarly, reviews differ in their inclusion criteria for ASP programs. Some reviews limit their data points to group-based activities that occur during the after-school hours and

combine recreation and youth development activities with academic support activities ("typical" ASPs) and others, including much more intensive and specialized programs that offer activities over the summer, on weekends, and during the school day in addition to activities that take place during the after-school hours. Finally, evaluations of ASPs differ greatly in terms of the populations studied. Some include only elementary-school-aged children, some only children at risk for academic failure, and some only socioeconomically disadvantaged children. Because the costs and benefits of participation in an ASP may differ for different populations, conclusions from reviews are expected to vary depending on the mixture of studies included in the review.

More than a dozen reviews of ASPs have been published in the past decade. Many of these reviews have focused broadly on "youth development" or "out-of-school" programs and, in so doing, have captured a broader set of programs than is of interest in this report. Eccles and Templeton (2002), for example, examined extracurricular activities, such as sports and leisure. Hollister (2003) included "out-of-school" programs that focused on youth development, many of which were not group-based programs but instead delivered tutoring or mentoring services to individual students. Fashola (1998) limited her review to group-based programs but included many programs that were not delivered during the after-school hours. Below we first summarize findings from five recent reviews that focused on programs delivered mostly during the after-school hours to groups of youths. These reviews also included only studies that used comparison groups and measured clearly defined outcomes. By summarizing results from a heterogeneous mixture of programs serving youths of different ages and different risk levels, these reviews focused on different slices of the ASP evaluation landscape and drew different conclusions about the effectiveness of ASPs. We follow this review of published reviews with a more focused summary of research on ASPs serving middle-school youths.

Among the five recent reviews selected for summary here, three (Kane, 2004; Lauer et al., 2006; Scott-Little et al., 2002) focused mainly on ASPs serving populations that were either socioeconomically disadvantaged or at risk for academic failure. Durlak and Weissberg (2007) included programs serving youths aged 5–18 years with no other limitations on youth characteristics, and Zief, Lauver and Maynard (2006) excluded ASP programs targeted specifically at youth with special needs such as learning disabilities, physical disabilities, emotional problems, or behavioral problems. The studies included in the Zief et al. (2006) review operated mostly in urban, school-based environments and served primarily low-income minority students in poor-performing schools. Most of the reviews included a mixture of studies of programs targeting elementary-school students and middle-school students, with a preponderance of the latter. Eighty percent of the studies included in Zief et al. (2006) served elementary-only populations.

The reviews also differed in the types of research designs included. As noted, all five limited their reviews to studies that included comparison groups and measured clearly defined outcomes. But they imposed different inclusion criteria related to the comparability of the treatment and comparison groups. Scott-Little et al. (2002), Lauer et al. (2006), and Durlak and Weissberg (2007) all required a comparison group but imposed no requirements regarding comparability of the comparison groups or the application of statistical controls. The studies reported in Kane (2004) used mostly nonequivalent comparison group designs, but all studies that did not use random assignment statistically controlled for observed pretreatment differences between the ASP participants and the nonparticipants. Zief et al. (2006) limited the studies in their review to "well implemented experimental design studies" (page 4). The probability is

much higher in these studies that observed outcome differences between ASP participants and nonparticipants are not a result of unmeasured characteristics (such as motivation to attend).

Not surprisingly, the conclusions reached about the effectiveness of ASPs in these reviews vary. Zief et al. (2006) reported that most of their positive findings were on measures of time expenditure in the after-school hours: Participation positively influenced youths' participation in athletics and art/music/dance/drama activities. Time spent in self-care was also lower for the ASP participants than for the controls. The review, however, found no significant effects of ASP participation on school attendance or behavioral outcomes. They noted "small but insignificant" (page 22) effects of ASP participation on school grades.

Kane's (2004) review of four large ASPs designed to address school performance found that participants' grades and test scores were improved slightly by academic programs delivered in an ASP setting. However, these improvements often failed to reach standard levels of statistical significance. He concluded that it may be unrealistic to expect a relatively small amount of after-school academic support to have a large impact on achievement. Furthermore, the improvements in school outcomes observed by Kane were conditional on program attendance. Across the programs Kane studied, typical students attended their programs only 1 or 2 days per week; students who attended more frequently experienced better outcomes.

The reviews that employed less stringent criteria on the research designs used in the included studies arrived at more positive conclusions about the effectiveness of ASPs for influencing academic and behavioral outcomes. Lauer at al.'s (2006) meta-analysis of 35 after-school and summer-school programs for predominantly elementary-level, high-risk youths showed that such programs do, on average, have a measurable impact on students' academic performance. These programs provided a mixture of individual and group instruction methods.

The average effect sizes for reading and math performance for after-school programs in Lauer et al.'s study were small ($d$ = .07 and .16, respectively) but statistically significant. Scott-Little et al. (2002) summarized results from 23 evaluations of ASPs that were not of a "drop-in" or "special activity" nature. They excluded from their review studies of tutoring and mentoring unless those activities were part of a broader program that was delivered in the after-school hours but included evaluations of programs that contained substantial in-school components. Scott-Little et al. (2002) concluded that the evaluations included in their review made "limited use of research designs that support causal conclusions and insufficient information to allow for meta-analysis of program effects" (p. 387). However, their summary of evidence from four studies reporting on academic outcomes revealed small, positive effects ($d$ = .21 and .16 for reading and math outcomes, respectively), and the authors tentatively concluded that ASPs can have positive impacts on participants.

Finally, Durlak and Weissberg's (2007) meta-analysis summarized effects for programs promoting personal and social skills. Their analysis included 66 studies of ASPs for youth aged 5 to 18 years that had stated goals of promoting personal development in the areas of leadership, decision making, self-control, and so on. The studies included in their review evaluated interventions that occurred "outside of normal school hours," but like Scott-Little et al. (2002) and Lauer at al. (2006), they also included summer programs and programs that contained a nontrivial in-school component. Durlak and Weissberg concluded on the basis of their meta-analysis that "after-school programs produced multiple benefits that pertain to youths' personal, social and academic life" (p. 7). Specifically, their results showed that, on average, ASPs have a positive impact on school bonding, attitudes about self-efficacy and self-esteem, behavioral adjustment indicators (e.g., prosocial and antisocial behaviors as well as drug use), and school

performance as measured by grades and achievement test scores (*d*'s ranging from .11 to .34, with an average of .22). Importantly, these positive outcomes were detected only for programs that used evidenced-based skill training approaches. Programs that failed to include evidence-based approaches were unsuccessful in improving any outcome.

What accounts for the more positive results reported in Durlak and Weissberg (2007)? As noted, there are several differences in terms of the populations and programs studied across the reviews that might lead to different conclusions. But a likely candidate is that the inclusion criteria for methodological rigor were different. This review included a larger number of studies than other reviews (66 compared with a range of 5 to 35 in the other reviews). Although the authors reported using a variety of strategies to rule out methodological factors as explanations for their findings, the fact remains that the mean effect sizes reported and most often repeated in summaries of the work contain several extremely high effect sizes from methodologically weak studies. We examined the studies contributing the seven highest effect sizes (all with overall effect sizes greater than .60) and found that five of the seven studies suffered either from severe attrition, differential attrition rates for the treatment and comparison groups, or obvious selection artifacts stemming from nonequivalent groups.

In summary, the existing reviews are inconsistent in their conclusions about the extent to which ASP participation influences important youth outcomes such as behavior and academic performance. Several reviews suggest that ASPs can produce small but measurable improvements in academic performance, and Durlak and Weissberg's (2007) review suggests that benefits extend beyond academic performance to other behavioral and attitudinal outcomes. But the conclusions seem to depend on the characteristics of the programs and on the methodological rigor of the studies included in the reviews. Most importantly, the

methodological rigor of almost all studies of ASPs is below par. More conclusive statements about ASP effectiveness must wait until more rigorous studies have been conducted.

It is also evident that the existing reviews of ASPs contain many programs for elementary-school-aged youths, who are likely to be more amenable to ASPs than middle-school-aged youths. As such, they are not directly relevant to our study of voluntary ASPs for middle-school youths. In an attempt to clarify what prior research tells us about the effects of ASPs for middle-school-aged youths, we summarized 12 studies that targeted this age group. Appendix Table A.1 provides detailed information about the methods used, type of students targeted, and results from these studies.[1] Like the studies included in the reviews summarized above, many studies of middle-school population fall short of contemporary standards for scientific rigor required to establish intervention effectiveness because they fail to randomize subjects to conditions and sometimes suffer from attrition problems. Only two studies (Lauver, 2002; Smith & Kennedy, 1991) used a randomized design, and one of those studies included only girls. For two studies (Girod, Martineau, & Zhao, 2004; Prenovost, 2001), preprogram differences between the treatment and comparison groups could not be ruled out as an explanation of posttreatment differences. For another study (Shelton, 2008), no difference findings could easily have resulted from lack of sufficient power. Of the remaining studies, many either did not address attrition at all or reported relatively high rates of attrition without examining possible differential attrition across study groups. Only 6 (50%) studies could be

[1] The studies included in Appendix Table A.1 were identified through a bibliographic search of published and unpublished evaluations of after-school programs that were similar to the one that is the subject of this evaluation in terms of population served and basic structure. We searched multiple online databases using the keywords "after school" and examined the reference lists of previous reviews of ASP research. Inclusion criteria were 1) the study reported analysis of student outcomes related to problem behavior or academic achievement; 2) the study included a comparison condition; 3) at least 50% of the study sample was middle-school aged (11–13 years old or in grades 6–8); 4) the program was delivered primarily during the after-school hours; and 5) the program included more than 10 sessions. This process yielded 12 qualifying studies.

considered reasonably rigorous in terms of their research designs, although none met the high standards described in Flay et al. (2005).

The stronger studies provided inconsistent findings regarding the benefits of ASP participation for middle-school youths. Two of these studies reported on ASP effects on school attendance: Lauver (2002) found no effects, and Dyanarski et al. (2004) found beneficial effects. Three of these studies reported on ASP effects on academic performance: Lauver (2002) found no effects, Dyanarski et al. (2004) found beneficial effects on social studies but not other grades, and Weisman et al. (2002) found negative effects on grade point average. Several studies reported on ASP effects on measures of student misconduct of school dropout. Some found positive effects (Huang, Sung Kim, Marshall, & Perez, 2005; Smith & Kennedy, 1991). But others found negative effects (Dyanarski et al., 2004; Weisman et al., 2002). Clearly, we do not yet have a clear picture of the effect of ASPs on the behaviors of middle-school youths.

Two studies of ASPs that share general population and structure characteristics with the ASP under current investigation attempted to identify characteristics of ASPs related to their effectiveness. In a study of 14 ASPs in Maryland, D. C. Gottfredson et al. (2004) concluded that an emphasis on social and character development distinguished ASPs that reduced problem behaviors from those that did not. A study of 35 ASPs in Maryland found that the delinquent behavior of participants decreased significantly in comparison with nonattenders (D. C. Gottfredson, Cross, & Soulé, 2007). Program characteristics associated with positive outcomes for youth included smaller program size, use of published curricula, and more educated staff.

Two studies that reported undesirable effects of ASP participation on conduct outcomes were for large, relatively unstructured programs. The evaluation of the nation's largest ASP, 21st Century Community Learning Centers (Dynarski et al., 2003, 2004), found that participating

students were no more likely to finish their homework or feel safe after school, despite these being stated goals of the program. In fact, middle-school participants were more likely to have had their property damaged, more likely to report they had used or sold drugs, and less likely to rate themselves positively at working out conflicts with others (Dynarski et al., 2003). The evaluated programs served an average of 60 youths per day and provided mostly homework assistance and recreational activities. Similar negative results were also uncovered by Weisman et al. (2002) in a quasi-experimental evaluation of 22 ASPs. Participants in the programs Weisman et al. studied reported more conduct problems, including a wider variety of drug use, and more drug-using peers than did nonparticipants. The authors noted that the evaluated programs, like the 21st Century Community Learning Centers programs, were too large and offered too much unstructured programming.

We conclude that the research on which the knowledge base about ASPs is based is relatively weak in terms of scientific rigor. Despite this, the research suggests that 1) the use of structured and tested program content is associated with better outcomes; 2) at-risk students likely benefit the most from participation in ASPs; 3) serving a large number of youth and 4) employing under-trained or under-educated staff is unadvisable; and 5) attendance is crucial because without consistent attendance, programs do not expose youth to a sufficient amount of treatment. But the inconclusive nature of the research to date, especially on middle-school populations, reduces confidence in these findings and demands that higher quality research be conducted.

## Purpose of Study

This study reports on a multisite ASP intervention in which an "enhanced" program model was provided to practitioners who routinely delivered ASPs in the state of Maryland. The plan[2] for the enhanced program model included three specific intervention strategies, or "research components," all of which had evidence of effectiveness in ASPs or other contexts. The research components were *Attendance Monitoring and Reinforcement* to increase school and ASP attendance (Brooks, 1975; Bry & George, 1980; D. C. Gottfredson, G. D. Gottfredson, & Hybl, 1990); *Structured Tutoring* to improve literacy and math skills (Cawelti, 1999; Cohen, Kulik, & Kulik, 1982; Lauer et al., 2006; Wasik, 1998); and the All Stars curriculum (http://www.allstarsprevention.com/) to reduce substance use and aggressive behavior and to increase social competency skills (Hansen, 1996; Hansen & Dusenbury, 2004; Harrington, Giles, Hoyle, Feeney, & Youngbluth, 2001; McNeal, Hansen, Harrington, & Giles, 2004).

The intent of the overall study was to assess the extent to which the routine practices of the implementing agency could be shifted in the direction of providing more research-based programming, and to measure the effects of doing so on a range of outcomes from middle-school youths, including personal conduct, classroom behavior, internalization of conventional beliefs, bonding to school, and learning outcomes such as standardized test scores and grade point average (GPA). The study is therefore an effectiveness trial of a combination of program components that have not heretofore been tested in an after-school setting. Although each component had been tested earlier, the combination has not been studied in an efficacy trial.

The results of this research should inform policy on ASPs by testing whether funds for ASPs could be better spent by introducing structured, tested program content into ASP models that are currently in use, such as the 21st Century model. As typically designed, ASPs allocate most time to unstructured or loosely structured activities with little likelihood of improving

[2] The program as delivered differed from the planned program, as will be described shortly.

academic outcomes. It is unknown how valuable ASPs can be in improving academic and social competency outcomes if they used research-based strategies with demonstrated effectiveness. This study was designed to provide a rigorous empirical evaluation of an enhanced ASP model that incorporates activities recommended in prior research implemented in existing ASPs.

## 3. Method

This study randomly assigned students within each of five participating schools to the enhanced ASP or to a "treatment as usual" control group. The ASP program operated for 9 hours per week for 30 weeks, and the plan for the enhanced ASP included attendance monitoring and reinforcement, structured tutoring, and the All Stars curriculum. The ASPs were located in public middle schools in Baltimore County, Maryland, that served high percentages of minority, socioeconomically disadvantaged youths. Youth surveys and school records were collected preprogram and postprogram to measure the key outcomes targeted as well as the key intermediate outcomes. Teacher ratings were collected at the end of the program year. A survey measuring exposure to after-school activities was administered midway through the program year. Program observations were conducted twice per month, and implementation data were collected daily. Regression models for each dependent variable compared treatment and control group means on each outcome and mediator. Standardized mean difference effects size statistics were calculated using the adjusted posttest means. Below we describe the study setting, sample, intervention, counterfactual condition, measures, and analysis strategy.

### Setting

The experimental after-school program was implemented in five low-performing middle schools in Baltimore County, Maryland, during the 2006–2007 school year via a partnership among four public agencies in Maryland.[3] The five school sites selected were the first among all

[3] The University of Maryland (UMD), Baltimore County Local Management Board (LMB), Baltimore County Department of Recreation and Parks (BCRP), and the Baltimore County Public Schools (BCPS) worked together to implement the experimental program in the five schools, to provide space, supplies, and employees, and to evaluate the success of the program. UMD provided all material and personnel support for the formal evaluation of the

low-performing middle schools in the county to express interest and agree to cooperate with the research procedures. The participating schools had high populations of minority youth, large numbers of students receiving subsidized meals, and high mobility. The principals at all five sites expressed the need for ASPs in their schools, stating that no comprehensive, school-based ASPs were available to their students. Demographic characteristics of the five participating school sites and all Baltimore County middle schools are presented in Table 3.1.

**Table 3.1. Demographic Characteristics of Schools and Number of Students Registered for the ASP at Each Participating School, by Site, 2006–2007**

| School | Total Enrollment | # ASP Registrations | % Minority | % Subsidized Meals | % Mobility[a] |
|---|---|---|---|---|---|
| A | 839 | 71 | 64.4 | 65.0 | 20.6 |
| B | 484 | 101 | 47.1 | 64.8 | 21.1 |
| C | 683 | 72 | 50.8 | 67.0 | 13.8 |
| D | 566 | 120 | 97.9 | 48.9 | 21.3 |
| E | 719 | 83 | 99.3 | 63.4 | 16.9 |
| All BC Middle Schools[b] | | | | | |
| Average | 879 | — | 51.5 | 37.2 | 12.3 |
| Minimum | 484 | — | 8.7 | 6.0 | 2.8 |
| Maximum | 1,490 | — | 99.3 | 67.0 | 22.0 |

[a]The percentage of students withdrawing for any reason during the school year.
[b]Alternative schools are omitted.

effects of program participation on youth outcomes. The LMB, a small county government agency charged with facilitating collaboration across public and private child-serving agencies, which funds, monitors, and evaluates a variety of after-school programs, facilitated involvement of BCPS and issued a request for proposals (RFP) to secure a vendor who would provide after-school services to youth. BCPS provided access to the student population, space for the program, and limited oversight to ensure compliance with school system regulations and coordination with the BCPS education program. BCRP was selected through a competitive process to run the programs. BCRP agreed to implement the enhanced program as described in the RFP, to conform to the requirements of the research design, and to assist with the process of data collection. A representative from the LMB managed and oversaw the contract performance of the vendor and provided continual coordination with the school system.

## Sample

All students who attended the five participating schools were eligible to register for the ASP. The study's recruitment goal was 100 students per school for a total of 500 students. Within each school, registered students had a 50% chance of being randomly assigned to the treatment group (i.e., invited to attend the after-school program) or to the control group.

Participant recruitment began in the spring of 2006 as a joint effort by UM and BCRP. Efforts included promoting the program at school and community events, including promotional fliers in every student's start-of-year-orientation packet, mailing multiple recruitment postcards to all students' homes, and placing an automated message on the home phones of eligible youth. Finally, school principals sent letters to academically or behaviorally "at-risk" students to encourage their enrollment in the ASP.

Recruitment goals were ultimately met or exceeded at two of the five school sites (see Table 3.1). Because recruitment lagged behind the ideal, principal referral and postcard recruitment efforts were continued into the fall 2006 semester after the programs had opened. When recruitment ended in January 2007, 447 students had registered and completed a pretest. Students who registered for the ASP were generally representative of the populations of their schools in terms of gender and socioeconomic status (SES, as determined by receipt of subsidized meals), but the ASP seemed to attract more minority youths. See Appendix Table B.1 for a comparison of study participants with their school populations.

The 447 students were randomized into treatment and control conditions by the principal investigator using a random number generator in SPSS version 11.0 (SPSS Inc., Chicago, IL). Students were randomized into conditions within their schools, such that each student had a 50%

chance of assignment to the treatment condition within his or her school. This method ensured that treatment and control groups would be of equivalent size at each school. Several rounds of randomization were conducted as new registrations were received throughout the recruitment period.

Randomization was successful in creating equivalent groups. Treatment and control students did not differ in terms of demographics (see Table 3.2) and differed significantly on only 1 of the 18 pretreatment measures shown in Table 3.3. The treatment youth scored higher in decision-making skills at pretest than did controls. This difference is also reflected in a significant difference on the social competence composite scale. One difference out of 18 tests conducted is approximately what would be expected by chance using a critical value of $p < .05$.

**Table 3.2. Demographic Characteristics of Sample, by Experimental Group**

| **Demographics** | **Total Sample *N* = 447** | | **Treatment *n* = 224** | | **Control *n* = 223** | |
|---|---|---|---|---|---|---|
| | **Mean or %** | ***N*** | **Mean or %** | ***N*** | **Mean or %** | ***N*** |
| Age[a] | 12.22 (.99) | 447 | 12.30 (1.03) | 224 | 12.15 (.94) | 223 |
| Family Income (Median) | $32,040 | 403 | $32,894 | 204 | $32,000 | 199 |
| % Male | 53.69 | 447 | 52.68 | 224 | 54.71 | 223 |
| % Black | 69.58 | 447 | 68.75 | 224 | 70.40 | 223 |
| % 6th Grade | 41.83 | 447 | 42.41 | 224 | 41.26 | 223 |
| % 7th Grade | 33.56 | 447 | 30.36 | 224 | 36.77 | 223 |
| % 8th Grade | 24.61 | 447 | 27.23 | 224 | 21.97 | 223 |
| % Living With Two Parents | 36.91 | 447 | 36.61 | 224 | 37.22 | 223 |
| % Subsidized Meals | 58.68 | 438 | 58.99 | 217 | 58.37 | 221 |
| % Mother is College Graduate | 12.56 | 438 | 13.57 | 221 | 11.52 | 217 |

[a]Standard deviation in parentheses.

**Table 3.3. Pretest Measures, by Experimental Group**

| Scale | Total Sample *N* = 447 | | | Treatment *n* = 224 | | | Control *n* = 223 | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Mean** | **SD** | ***N*** | **Mean** | **SD** | ***N*** | **Mean** | **SD** | ***N*** |
| **Composite Scales** | | | | | | | | | |
| Social competence | .00 | .73 | 447 | .09* | .70 | 224 | -.08 | .75 | 223 |
| Prosocial/antidrug attitudes | .74 | .22 | 447 | .74 | .22 | 224 | .75 | .22 | 223 |
| Academic performance | .02 | .84 | 400 | .01 | .86 | 203 | .02 | .82 | 197 |
| Conduct problems | .01 | .72 | 447 | .00 | .64 | 224 | .01 | .80 | 223 |
| **Unsupervised Socializing** | | | | | | | | | |
| Days with friends and no adults | 2.29 | 2.12 | 432 | 2.16 | 2.17 | 215 | 2.43 | 2.06 | 217 |
| **Positive Peer Influence** | | | | | | | | | |
| Positive peer influence | .84 | .19 | 443 | .85 | .18 | 222 | .83 | .19 | 221 |
| **School Bonding** | | | | | | | | | |
| Attachment to school | 2.46 | .89 | 443 | 2.46 | .89 | 223 | 2.46 | .89 | 220 |
| **Social Competence** | | | | | | | | | |
| Goal setting | 3.11 | .50 | 443 | 3.14 | .45 | 223 | 3.09 | .55 | 220 |
| Decision-making skills | 2.83 | .78 | 424 | 2.92* | .73 | 213 | 2.75 | .81 | 211 |
| Impulsiveness | .51 | .27 | 436 | .49 | .28 | 219 | .53 | .26 | 217 |
| **Prosocial/Antidrug Attitudes** | | | | | | | | | |
| Attitudes unfavorable to drug use | .78 | .25 | 446 | .77 | .25 | 223 | .78 | .25 | 223 |
| Belief in conventional rules | .71 | .24 | 447 | .70 | .24 | 224 | .72 | .24 | 223 |
| **School Attendance** | | | | | | | | | |
| % days absent (SR) | 4.33 | 4.31 | 405 | 4.33 | 4.13 | 205 | 4.34 | 4.45 | 200 |
| **Academic Performance** | | | | | | | | | |
| GPA[a] (SR) | 2.42 | .75 | 231 | 2.35 | .75 | 113 | 2.48 | .75 | 118 |
| MSA reading (SR) | 389.44 | 28.64 | 395 | 390.42 | 29.21 | 199 | 388.45 | 28.08 | 196 |
| MSA math (SR) | 385.13 | 39.20 | 395 | 385.36 | 39.98 | 199 | 384.90 | 38.50 | 196 |
| **Conduct Problems** | | | | | | | | | |
| Disruptive classroom behavior | 1.38 | .49 | 443 | 1.38 | .50 | 222 | 1.38 | .48 | 221 |
| Aggression | 1.58 | .59 | 441 | 1.60 | .60 | 220 | 1.57 | .58 | 221 |
| Delinquent behavior | .46 | 1.05 | 441 | .50 | 1.06 | 221 | .43 | 1.05 | 220 |
| Victimization | 1.31 | 1.64 | 437 | 1.30 | 1.69 | 219 | 1.33 | 1.59 | 218 |
| Last month drug use | .10 | .29 | 442 | .09 | .29 | 222 | .10 | .30 | 220 |
| Number suspensions (SR) | .38 | .84 | 404 | .38 | .82 | 205 | .38 | .87 | 199 |

*Notes.* Scales from youth survey unless otherwise noted.
*Abbreviations*. SD = standard deviation; SR = school records.
[a]Pretest GPA unavailable for 6th-grade students.
*$p < .05$.

## Power Analysis

Power analyses were carried out using software for "Power and Sample Size Calculations" (*PS,* V2.1.31), based on work by Dupont and Plummer (1990, 1998). Input to the program included the number of cases available for analysis, mean difference to be detected, the standard deviation of the outcome variable observed at pretest, and the Type I error rate desired. We conducted power analyses for two-tailed independent *t*-tests, fixing the Type I error rate (alpha) at 0.05. These power analyses estimated the power available, given the observed standard deviation of several outcomes, to detect an effect of 0.3 standard deviation units. We calculated power for several different outcomes, including delinquency-related risky behavior and aggression (measured continuously) and vulnerability to drug use and last-month frequency drug use (measured as binary outcomes). These analyses indicated that, using the entire sample, the power to detect differences of the specified size between the treatment and control groups is 0.99.

We anticipated that we might wish to examine treatment effects in subgroups of the population (e.g., by social class, latch-key status, or quality of implementation). We repeated the power analyses described above, but this time we estimated the number of cases necessary to detect an effect of 0.3 with 80% power. These analyses indicated that between 167 and 208 cases would be needed, depending on the dependent variable examined. Hence, we decided to refrain from examining subgroup differences in groups smaller than approximately 200 cases (or 100 cases per experimental group). This decision precludes examination of site-specific differences between treatment and control groups, as our largest site contained only 120 cases but allows for the examination of larger subgroups of interest.

Finally, we estimated the minimum effect that would be detected in our study. Following procedures outlined in Lipsey (1998), we used an estimate of the predictive power of the pretest covariate and school dummy variable as well as the intraclass correlations estimated from the data to determine that we will be able to detect effects of about 0.17 with 80% power. We conclude that that study has ample power to detect effects as small as 0.17, and that it has ample power to detect effects in the 0.3 range even in reasonably sized subgroups.

## Description of Study Conditions

### Basic Structure

The experimental ASP followed a traditional structure with research-based enhancements to program content. The program was offered on school grounds, 3 days per week (Tuesday, Wednesday, and Thursday), for 3 hours after the close of the regular school day. The number of youths served per day was restricted to a maximum of 50, but typically programs served far fewer than 50 students on a given day. Activities traditionally offered in ASPs (e.g., snack, sports, and crafts) occupied about two thirds of the program time. The program, free to participants, was intended to improve on traditional ASPs through three additional research components that were to occupy the remaining one third of program time: an attendance incentive system, tutoring/academic assistance, and the All Stars curriculum. Tuesdays and Wednesdays were designated “research days” when research-based components (All Stars and tutoring) would be delivered, occupying 1.5 hours on each of these days.

Figure 3.1 depicts the program model planned for the ASP and the intended outcomes of each of the main components. Simply attending the ASP was expected to reduce unsupervised socializing, increase positive peer influences, and promote bonding to school, all of which were expected subsequently to influence conduct problems. Increasing school bonding was also expected to increase academic performance. The three program enhancements were also expected to influence the three main outcomes. All Stars was intended to increase school bonding, social competence, and prosocial attitudes and beliefs, indirectly influencing academic performance and conduct problems via these mediators. The academic assistance component was expected to increase academic performance, and the attendance incentives component was expected to increase school attendance. Next we describe the three research components as they were developed. These components were implemented with varying degrees of integrity to the initial design. As discussed in the Results section, the attendance incentive system was partially implemented. Tutoring was not delivered, although academic assistance was. The All Stars curriculum was delivered more or less as planned.

### Attendance Incentives

The attendance incentives system was expected to award points for good attendance every Thursday in a brief ceremony. As planned, students would be awarded weekly points contingent on both absolute levels of and improvements in his or her *school* attendance (both absolute level of attendance and improvements in attendance earned points). Rewards for attendance at the ASP would be contingent on the attendance (both absolute and improved) of preestablished groups of

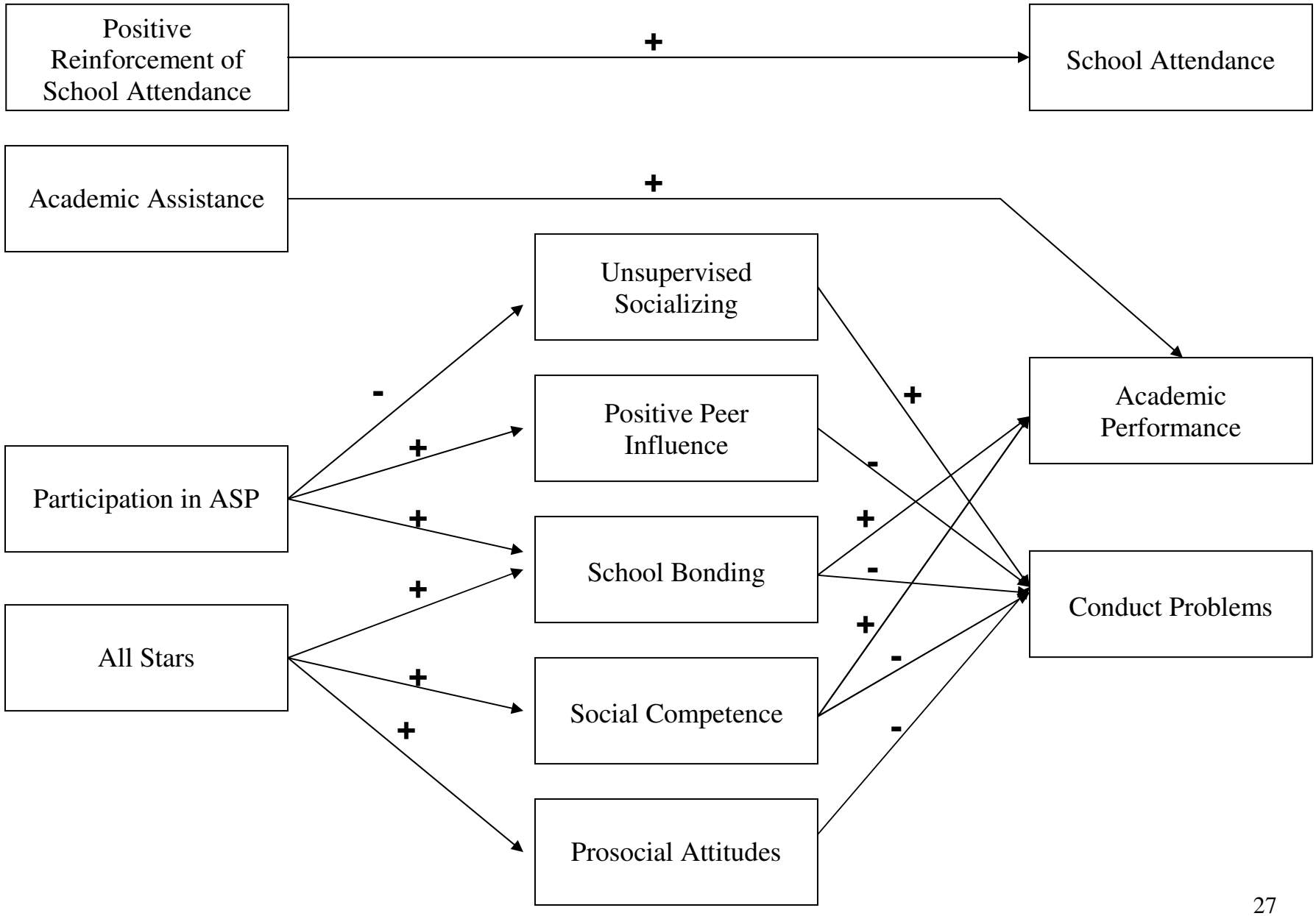
three or four students.[4] Group-based incentives were intended to create peer pressure to attend the program.

The Web-based management information system (MIS) developed for this project was equipped with a function that would automatically calculate and track each student's "point balance" based on attendance information entered by site staff. Attendance points could then be used to purchase a variety of prizes. Students were to receive points and praise every week and be given the opportunity to redeem points twice a month. A prize catalog was planned that would contain pictures and descriptions of a wide variety of prizes ranging from small, inexpensive items such as pens to more substantial rewards such as video games or graphing calculators. In this way, students could save points to spend on one or two large prizes or spend their points frequently on smaller prizes. The budget allowed for $80 per student in incentives. Prior to the beginning of the program, UM conducted focus groups with youth from participating schools to determine which rewards were most desirable.

Staff members received training in the purpose of the attendance incentive system and its planned execution. They were also trained in the use of the MIS to coordinate award and redemption of points. Monitoring of the attendance incentive system was achieved via staff recording into the MIS of attendance data, points awarded, and expenses as well as through UM observations of attendance incentive ceremonies.

[4] As will be described in the Results section, the group contingency portion of this component was not faithfully implemented.

**Figure 3.1. ASP Logic Model**



## Tutoring/Academic Activities

The original design for the academic component relied on the recruitment of volunteer tutors from the community, including high-school students and adult community members. The tutoring plan incorporated an assessment software tool that was to be administered to both tutors and students. The assessment would reveal areas in reading or math skills where each student needed the most improvement. This way the subject matter of the tutoring sessions could focus on areas of highest priority for individual students. Tutors were to be assessed for skill in math and reading so that they could be appropriately matched with students of lower skill. A math education expert created a customized math curriculum that was flexible and interactive for use in the tutoring program. Libraries containing books of varying difficulty were supplied to each site for use in reading tutoring. One-on-one tutoring sessions were scheduled for 45-minute periods on Tuesdays and Wednesdays. The staff at each site was expected to oversee the tutor's work. Staff received 6 hours of training in the tutoring model.

The tutoring program was not provided to students as planned.[5] An alternative academic activity consisting primarily of supervised homework assistance was substituted. This academic

[5] Implementation of the tutoring component encountered serious difficulties related to assessment software access and tutor recruitment. The school system requested that we use an assessment software package that was already in use in BCPS schools. The BCPS official assigned as a liaison to this project assured the research team that we could access the software for project purposes. We developed the content of the tutoring program to correspond with the specific diagnostic output provided by this software. However, when we requested the software in the beginning of the fall semester to allow volunteer tutors to use the assessment software, BCPS denied us access. BCPS informed us at that time that it was undertaking an update of their version of the software, and that no one could access it. However, they indicated that access would soon be granted. Prior to the beginning of winter break, when no progress had been made in gaining access to the software, we decided to substitute an alternative academic activity in the 45-minute time slot on Tuesdays and Wednesdays that had been set aside for tutoring. Also unclear to the research team is how many volunteers were actually recruited. A tutor coordinator, hired by BCRP as stipulated in its contract with the LMB, was responsible for tutor recruitment. When we obtained a list of volunteers committed to provide tutoring in fall 2007, it became clear that the number identified (27) was only a fraction of the number needed to attain a one-to-one student–tutor ratio. It is not clear to what extent tutor recruitment was slowed due to the problems with the software.

activity, designed by the vendor and Baltimore County Local Management Board as a replacement for the tutoring that had been designed by the researchers, was consistent with usual practice among ASP providers but, unlike the one-on-one tutoring it replaced, lacked an empirical foundation. As such, its substitution represented a major deviation from the plan for the project. Beginning in December, academic workbook activities in reading and math were provided to youths at all centers. Exercises from commercially available academic support workbooks were intended to augment and support classroom curriculum and were offered to students who did not have any homework on a given day. Independent reading, using age-appropriate books provided by the after-school program, was also offered to youths during this time. Staff members supervised academic activities and were available to answer student questions. However, these staff members were not always certified teachers, and the ratio of staff to students was the same as in the program overall—much higher than one to one, as initially planned.

No specialized training was provided for staff who supervised the alternative academic activity because this component was not a planned part of the program. Similarly, only limited monitoring of the quality of instruction during academic assistance was enacted because the model for the program had not anticipated the inclusion of homework assistance. During academic assistance sessions, UM observers rated student engagement and the structure of the activity. Program staff recorded the amount of time spent in academic assistance in the MIS.

### All Stars

The All Stars curriculum, a prevention curriculum focused on reinforcing commitment to abstain from substance use,[6] was delivered on Tuesdays and Wednesdays in 45-minute sessions. To reduce All Stars class sizes, sites divided students into at least two class groups and delivered separate sessions to each group. Twenty-seven separate All Stars lessons were available to site staff. All Stars instructors aimed to teach one lesson per week—half of the lesson on Tuesday and the other half on Wednesday.

One or more staff members from each site participated in a 3-day training conducted by the company that developed All Stars. Monitoring of All Stars implementation was conducted in two ways. All Stars instructors completed implementation fidelity checklists, created by the curriculum developer, and entered these data into the MIS. Each checklist asked which lesson and which specific activities within that lesson were taught, the instructor's impression of the quality of the lesson, if stated goals were achieved, and if students were engaged in the lesson. UM observers filled out similar checklists when they attended a site on a Tuesday or Wednesday and systematically assessed the level of student engagement in the All Stars lessons.

### Leisure Activities

The experimental ASP offered 9 hours of programming per week, 3 of which were dedicated to All Stars and academic activities. The remaining 6 hours of programming contained leisure activities planned by BCRP. These activities included a variety of activities typical of

[6] Gottfredson, Cross, Wilson, Rorie, and Connell (2010) describe the All Stars program and report in greater detail on the evaluation of this component of the experimental ASP.

child care environments such as snack, sports, crafts, board games, movies, field trips, and computer time. These activities were intended primarily to motivate students to attend. During recruitment, BCRP highlighted these activities as the main "pull" to encourage youths to register.

BCRP provided 19 hours of start-up training to ASP staff in areas related to leisure activities and general program operations. These trainings covered program orientation, first aid and CPR, cultural diversity, supervision, behavior management, reporting procedures for suspicions of abuse or neglect, and inclusion of people with disabilities, food handling, and transportation safety.

### Control Condition

The control condition was "treatment as usual" except that members of the control group were invited to attend one after-school activity per month. Sites usually planned a special event or party for the days that control students were invited to attend. Eight such control group days occurred at the sites during the program year. These events were not well attended by control students. Fifty-two percent never attended; 29% attended once, twice, or three times; and 17% attended between four and eight times. However, five control students (2%) attended more frequently than the 8 days planned, one of whom attended nearly every day the program was open. When days of attendance for these five students were only counted when they attended the appropriate days to which they were technically invited (up to eight days), the average attendance of the control group was 1.4 days.[7]

[7] Instrumental variables regression analyses that take into consideration the actual attendance patterns of study youths are reported in the Results section. These analyses show that the control group attendance at the program did not influence the results of the study.

Although the level of participation in the experimental ASP by control group students was trivial, they were free to participate in whatever other after-school activities were available to them. Virtually all (96%) members of the of the control group participated in some organized after-school activity. Nearly 60% participated in an after-school activity at their schools, and most also participated in community-based, after-school activities. These alternative leisure activities are described in greater detail in the Results section.

## Cost

The budget for running this program for one school year (including administrative costs, staff training and payroll, materials, snacks, attendance incentives, transportation, etc.) was approximately $450,000, or $90,000 per site. This budget, however, excluded certain costs that were contributed by the county, including rental costs for the space and the salary of the BCPR Program Coordinator. Adding estimates of the value of these contributions raises the total budget for the project to $574,720, or $114,944 per site. Using this figure and discounting costs related to the control group, the cost per treatment youth was approximately $2,566. The largest proportion of the operations budget was dedicated to staff payroll and the second largest to student transportation. It is not possible to estimate the cost of individual components because most of the budget was allocated to items that contributed to general program functions as opposed to specific components. It is also important to note that the “treatment as usual” control condition also involved provision of services, mostly of a recreational nature. Because we have no estimate of the costs of these services, we cannot estimate the marginal costs related to the enhancements that differentiate the experimental program from usual practice.

## Measures

Seven data sources contributed information used in this study. Table 3.4 provides an overview of the data sources and measures taken from each, and Appendix Table C.1 provides information on the number of cases and response rates for each source. The next section describes each of the data sources generally. Detailed information is then provided on measures from these sources used for 1) describing program implementation, 2) estimating program impacts on participation in after-school activities and programs, and 3) estimating program impacts on outcomes. Additional information on measures is provided in appendices D (outcome measures) and E (process and activity participation measures).

### Data Sources

*Registration form*. Before students were permitted to participate in the program, their parents had to complete a registration form. This form was used to obtain demographic information as well as tracking information for those students who withdrew from their registered school during the course of the evaluation. Demographic information reported on the registration form included age, race, gender, grade, family income, and parental education.

**Table 3.4. Overview of Data Sources and Measures**

| Data Source | Measures |
|---|---|
| Registration Form | Student demographics, family income, tracking information |
| Pretest and Posttest[a] Youth Surveys | Demographics, classroom behavior, aggression, delinquent behavior, victimization, substance use, attitudes unfavorable to drug use, drug resistance, impulsiveness, positive peer influence, self-reported grades, attachment to school, belief in conventional rules, commitment to education, reading for fun, goal setting, decision-making skills, parental monitoring, after-school time expenditure, ASP liking, friendship networks |
| Youth Experiences Survey[b] | After-school activity involvement |
| School Records | Attendance, GPA, achievement test scores, promotion, discipline records |
| Teacher Ratings[a] | Social competency, effectance motivation, expectancy of success, academic competence |
| Management Information System | Staff information and attendance, individual student school and program attendance, student withdrawal from the program, all stars fidelity information, all stars attendance and session recording, leisure activity recording and attendance, attendance incentive points awarding and redemption, weekly director checklist |
| Observations[b] | Program Observation: misbehavior, program content delivery, number of students, duration, supervision, social climate and structure |
| | Student Engagement Observation: number of students, duration, number of 5-minute intervals observed, structure, and student engagement rating |
| | All Stars Fidelity Observation: number of students, student engagement, overall quality of session |

[a]Form located in Appendix D.
[b]Form located in Appendix E.

*Youth surveys*. Participating youths completed pretest and posttest youth surveys measuring primarily the outcomes targeted by the enhanced ASP. These surveys consisted of 167 items. Pretest surveys were administered to all treatment and control youth ($N = 447$) after receipt of registration materials and signed consent forms from their parents or caregivers. Posttests were administered near the end of the program. Surveys were typically administered during one school period in a large room such as the cafeteria or library. Students received a $5

gift card for attending each survey administration. The response rate for posttest youth surveys was 96% ($N$ = 427), although surveys from 11 of these respondents were later discarded because more than 40% of the items in the surveys were left blank for a final response rate of 93% (Appendix Table C.1). Low study attrition is at least partially attributed to a $500 incentive offered to schools that achieved a 95% or higher response rate. All schools achieved this rate.

*Youth Experiences Survey 2.0 (YES).* Students completed the YES (Hansen & Larson, 2005) midway through the program year. This survey, which measured experiences during after-school activities, was administered in the same manner as the youth surveys except that no incentive was provided to the school. The YES response rate was 87% ($n$ = 389).

*School records*. School records were collected to measure student academic performance, attendance, and school suspensions for the year prior to the implementation of the program (2005–2006) as well as the year the program was implemented (2006–2007). At least one data element from these school records (both pre and post) was collected for all pretested youth. However, 2005–2006 GPA information was mostly unavailable for 6th graders, who were in elementary school during the 2005–2006 school year. GPA was not maintained electronically for these schools.

*Teacher ratings*. During the spring of 2007, science, math, social studies, and English teachers were asked to rate 427 study participants.[8] Teachers were offered $5 for each survey completed. These surveys measured student classroom behavior, social adjustment, and academic competence. A total of 1,696 surveys were distributed to 192 school teachers. At the close of data collection, 65% of teachers ($N$ = 125) had returned packets and 69% of student

[8] Twenty students had withdrawn from the BCPS system or had transferred to a new BCPS school too recently for the new teacher to rate the student reliably.

rating surveys ($n$ = 1,177) were returned. At least one survey was returned for 99% of students, and two or more were returned for 88% (Appendix Table C.1).

*Management information system (MIS).* The MIS was one of two methods used to collect implementation data. UM staff worked with a software developer to create this Web-based data entry system, which was used by program staff to record daily program procedures and events. See Table 3.4 for items captured by the MIS. Appendix Table E.1 provides additional detail on the contents of the MIS and the expected frequency of data entry. Training for ASP staff on the MIS occurred prior to the start of the ASP and upon hiring of new staff throughout the year. UM supplied each school with a laptop computer for daily MIS access. An additional work hour each day was allotted to staff for MIS data entry tasks.

*Program observations*. Data from program observations conducted by UM staff were also used to describe program implementation. Between October 2006 and April 2007, 80 site visits were conducted, usually by a pair of observers. On occasion (20% of the time), observations were conducted by one person. Observers rotated among sites to avoid observer-site bias. During site visits, observers filled out three[9] types of observation instruments: a general program observation form, a student engagement form,[10] and an All Stars fidelity form.[11] The content of each form is described in Table 3.4. The observation protocol directed observers to complete one program observation, one All Stars fidelity checklist (on days that All Stars was offered), and two engagement forms on each site visit. One engagement form was to be completed during All Stars or academics (if delivered), and the second was to be completed

[9] A deviancy training form was also completed. Data from this form are reported in a different report (Rorie, Gottfredson, Cross, Wilson, & Connell, in press.).

[10] At every 5-minute interval, the observer rated the level of structure in the observed activity from 1 to 5 (see Appendix E.3 for the operational definition of structure) and tallied the number of students who were engaged or not engaged in the activity during the interval. Engagement rates were recorded for intervals of activities that scored a 3 or higher on the structure scale.

[11] The All Stars checklists were provided by the developer of the curriculum (Tanglewood Research, 2008a) and were the same as those filled out by program staff in the MIS.

during a leisure activity. Five primary observers, all of whom attended the ASP training as well as approximately a week of training on observation procedures, conducted the observations.

## Outcome Measures

This section describes the specific measures derived from the data sources to measure program impacts on youth outcomes targeted in this study. Figure 3.1 shows the outcomes targeted by the ASP. Detail about the content of all measures discussed in this section along with reliability coefficients from pretest and posttests are shown in Appendix Table D.3. Higher scores on all measures indicate a higher level of the outcome. Scales were computed based on all valid items. Only when all items in the scale were missing was the scale score missing.

*Unsupervised socializing*. One item was used to measure supervision and time expenditure. This item is a count of the number of days spent hanging out with friends where no adult was present after school (range, 0–5).[12]

*Positive peer influence*. Positive peer influence (range 0–1) was created by averaging the items shown in Appendix Table D.3, after dichotomizing certain items as indicated. The items are from the *What About You* survey developed by G. D. Gottfredson and D. C. Gottfredson (1992) and the *Best Friend Influence* questionnaire (Poulin, 2003). The scale score was squared to reduce skew.

*School bonding*. School attachment (range 0–3) is a subset of items from Hansen's *Bonding* scale (Tanglewood Research, 2008b), which was constructed as a count of items endorsed.

[12] All ranges in the text reflect the possible range. Ranges in Appendix Table D.3 reflect the observed range.

*Social competence*. Three scales were created to measure social competence. Two of these scales, goal setting (range, 1–4) and decision-making skills (range, 1–4), are from Hansen's *Tanglewood Research Evaluation* (Tanglewood Research, 2008b) and were computed as averages of the items shown in Appendix Table D.3. Impulsiveness (range, 0–1) was adapted from Eysenck's $I_6$ *Impulsiveness* questionnaire (Eysenck, Easting, & Pearson, 1984) and was computed as an average of the items shown in Appendix Table D.3.

*Prosocial/antidrug attitudes*. Two scales were constructed to measure prosocial/antidrug attitudes. Attitudes unfavorable to drug use and belief in conventional rules (both range 0–1) are from the *What About You* survey and were created by averaging the items shown in Appendix Table D.3 (after dichotomizing certain items).

*School attendance*. BCPS provided the information used to compute the percentage of days absent (range, 0–100), which is the number of days absent divided by the number of days enrolled in school. The natural log was taken to reduce skew.

*Academic performance*. Four measures were used to assess academic performance. Academic competence (range, 1–5) was obtained from teacher ratings and was a one-item measure developed by UM researchers that measured the teacher's perception of the student's academic competence relative to his or her classmates. GPA (range, 0–4) and scores on the standardized Maryland State Assessment test in reading and math (range, 240–650) were obtained from BCPS records.

*Conduct problems*. Seven measures of problem behavior were drawn from the youth survey, school records, and teacher ratings. Disruptive classroom behavior (range, 1–3) was adapted from Hansen's *Problem Behaviors* scale (Tanglewood Research, 2008b), and aggression (range, 1–4) was adapted from Hansen's *All Star* questionnaire (Tanglewood Research, 2008b).

These scales were constructed by averaging the items shown in Appendix Table D.3. Delinquent behavior (range, 0–7) was measured from a subset of items contained in the *What About You* survey. Victimization (range, 0–7) was measured by a scale from the *Denver Youth Survey* (Huizinga & Esbensen, 1990). Scales for delinquent behavior and victimization were constructed by counting the number of items (shown in Appendix Table D.3) endorsed. Last month drug use contained three items that were dichotomized, indicating whether the youth had used any of the three substances in the previous month, also from the *What About You* survey. The scale was dichotomized so that any drug use in the last month was contrasted to none. The number of school suspensions for each student was a count of the number of suspensions during each school year, which was obtained from BCPS school records. Finally, a Social Competency Rating scale (G. D. Gottfredson, Jones, & Gore, 2002) was also computed as an average of the items shown in Appendix Table D.3.

*Composite scales*. Our study includes multiple measures for several outcomes targeted by the program. Conducting multiple hypothesis tests for impacts at a given α level of significance increases the chance of Type I errors to greater than α unless adjustments are made for multiple comparisons. To guard against such chance findings, we follow the advice offered by an expert panel recently convened by U.S. Department of Education Institute of Educational Sciences to explore ways of appropriately handling multiple comparisons (Schochet, 2007). This panel recommended that the data be structured and outcomes be prioritized to reflect the design of the intervention, and that confirmatory analyses be conducted to test global hypotheses within the main domains identified as central to the study's hypotheses. Accordingly, we developed scales to capture the eight outcomes identified in the program model (Figure 3.1).

Four outcomes (unsupervised socializing, positive peer influence, school bonding, and school attendance) were measured with a single indicator. The others (social competence, prosocial attitudes and beliefs, academic performance, and conduct problems) were measured with multiple indicators. These multiple indicators were combined to form four composite scales. Three of these scales: social competence, academic performance, and conduct problems (all ranging from –3 to +3) were computed by averaging the *z*-scores of the component items and scales. The social competency composite scale included goal setting, decision-making skills, and impulsiveness. Prior to computing the average, the impulsiveness *z*-score was reverse coded by multiplying the values by –1. The average correlation among these scales was .28 at both pretest and posttest. The academic performance composite scale included teacher reports of academic competence, GPA, MSA reading, and MSA math scores[13]. The average correlation among these scales was .45 at pretest and .51 at posttest. The conduct problems composite scale included disruptive classroom behavior, aggression, delinquent behavior, victimization,[14] last month drug use, number suspensions, and teacher reports of social competency.[15] Prior to computing the average, the teacher reports of social competency *z*-score was reverse coded by multiplying the values by –1. The average correlation among these scales was .30 at both pretest and posttest. Finally, the prosocial/antidrug attitudes composite scale (range .00–1.00) was computed by averaging the attitudes unfavorable to drug use and belief in conventional rules scales. The correlation among these scales was .62 at pretest and .68 at posttest.

[13] The pretest measures of the academic performance and the conduct problems composite scales exclude academic competence and social competency, respectively, because the scales from the teacher ratings survey were only measured at posttest.

[14] Although not a measure of conduct problems per se, victimization is included in this composite because it is highly correlated with the other problem behavior measures (for example, $r = .39$ and .48 for victimization and delinquent behavior at pretest and posttest, respectively). Evidence suggests that offenders and victims are often the same people.

[15] See footnote 13.

*Attrition and missing data*. Our primary source of outcome data, the posttest youth survey, was usable for 416 (93%) of the 447 registered students. The 31 students (13 treatment and 18 control) who were excluded from outcome analysis because of missing posttest data either refused to take the posttest ($n$ = 10), had transferred out of Maryland schools (n = 10), or left more than 40% of the survey items blank ($n$ = 11) . An attrition analysis (Appendix Tables C.2 and C.3) showed that registered youth who were excluded from the study ($n$ = 31) did not generally differ from those who were included ($N$ = 416), demographically or on a range of pretreatment measures. Exceptions were age and attitudes unfavorable to drug use. The excluded cases were older and had more favorable attitudes to drug use than those retained in the study.

Treatment by attrition interactions were conducted to test for differential attrition by treatment status that would bias the results of our study. Of 28 interactions, one (MSA math score) was statistically significant at the $p < .05$ level. This analysis suggested that higher achievers were more likely to attrit from the treatment than from the control group.

The amount of data missing at the item level from the surveys for the 416 usable cases is very low (see Appendix Table D.4), in part because of our decision to compute scales based on all valid items. Missing data only exceed 4% of the available cases for posttest unsupervised socializing, pre- and post-decision-making skills measures, and the pretest data from school records. Therefore, most analyses simply excluded the small number of cases for which outcome data were missing. However, for the measures listed above with more than 4% missing data, we employed maximum likelihood methods for imputing missing data (Allison, 2002). Outcome analyses involving these measures were conducted using both the imputed scores and listwise deletion. No substantive differences were observed across these two analyses, so we report only the results for the imputed scores.

Process Measures

This section describes the specific measures used to describe program implementation. Detail about these measures is available in Appendix E.

*ASP attendance measures*. The MIS described earlier provided information on days attended (see Appendix Table E.1). In addition, the youth survey contained three items measuring youth participation in after-school programs. The first item was a dichotomy asking youth whether they attended after-school activities with a group of youth and adults in a typical week this school year (the pretest referred to after-school activity participation in the previous year). Those students who answered "yes" to involvement in after-school programs were then asked to indicate how many days in a typical week they attended after-school programs at their school and not at their school (range, 0–5). These items were summed to count the number of days in a week students attended after-school programs (range, 1–10).[16]

*Program observation*. The program observation, developed by UM researchers, was used to measure program quality. The observation instrument (shown in Appendix E.2) contained 19 items covering the level of misbehavior, skillfulness of program content delivery, supervision, social climate, structure, and orderliness. The responses to the items were dichotomized and averaged to create one composite program quality scale that had an alpha reliability coefficient of 0.87.[17]

*Engagement*. The engagement rating form was developed by UM researchers (shown in Appendix E.3). Activity structure was rated in each 5-minute interval using the following 5-point

[16] This scale has a maximum of 10 days because youth often participate in multiple activities throughout the year; some of which are at school and others off campus.
[17] Cross, Gottfredson, Wilson, Rorie, and Connell (2010) describe the content of the scale in detail and show descriptive data for subscales that comprise the 19-item scale described here.

scale to describe how youths were expected to spend their time during the interval: no expectations (1); little expectation (2); some expectations but ambiguous/not communicated well (3); between clear expectations and some expectations (4); as well as clearly defined and all youth know what is expected at all times (5). The interval-level ratings were averaged to arrive at one structure rating per activity. The engagement rate was computed by dividing the sum of the total number of students engaged (as compared with unoccupied, socializing, or out of sight) by the sum of the total number of students observed across all intervals (range, 0–1).

*All Stars fidelity*. The MIS provided data on the fidelity of each All Stars session (see Appendix Table E.1). UM observers also measured All Stars fidelity using measures of engagement (1–4) and overall quality (1–5), both taken from the All Stars fidelity form provided by the developer (an example is shown in Appendix E.4) (Tanglewood Research, 2008a).

### After-School Activity Participation Measures

This section describes measures used primarily to compare the treatment and control conditions on after-school activity participation. In addition to the measures of ASP attendance from the youth surveys described earlier, data from the *Youth Experiences Survey* (YES; Hansen & Larson, 2005) were used to contrast treatment and control cases on after-school time expenditure. The YES measured the variety of activities in which youth engaged.[18] An activity checklist asked students to indicate in which of 47 activities they participated Monday to Friday between the hours of 3 p.m. and 6 p.m. Youth were then instructed to choose the activity they participated in the most. See Appendix E.5 for the activity checklist.

[18] The YES also measured positive and negative experiences in after-school activities. These experiences are discussed in a different report (Wilson, D. C. Gottfredson, Cross, Rorie, & Connell, in press). Only the variety of activities is discussed in this report.

## Statistical Methods Used to Compare Outcomes

Prior to comparing outcomes for the study groups, all outcome variables were examined to determine their best representation by identifying outliers and deviations from normality. Some variables were determined to be best represented as binary or count variables. When variables were transformed to reduce skew (for positive peer influence and school attendance), both the transformed and untransformed dependent variables were used in analyses testing for program effects. However, the results for the squared peer influence variable were similar to results using the untransformed variable, so the latter results are presented for the sake of simplification. In the case of school attendance, the logged variable is presented.

As mentioned in the Methods section, we used both "intent to treat" (ITT) and instrumental variables approaches to outcome analysis. We report ITT results first. In all outcome analyses, two-tailed tests of statistical significance were employed, with an alpha level of .05. First, regression models were run. The model for each dependent variable included a dummy variable measuring assignment to the treatment condition (1 = treatment; 0 = control), a measure of the dependent variable taken at pretest, gender (1 = male; 0 = female), race (1 = black; 0 = nonblack), age, and four dummy variables measuring school. We used regression models appropriate for each dependent variable. Specifically, ordinary least squares (OLS) models were used for normally distributed, continuous variables; negative binomial or Poisson regression was used for outcome variables that involved counts; and logistic regression was used for dichotomous outcome variables. Covariate-adjusted posttest means were calculated from these models. Standardized mean difference effects size statistics were calculated using the

difference between the treatment and the control group adjusted posttest mean in the numerator and the pooled standard deviation for the corresponding unadjusted posttest measures. In calculating the effect size for victimization, the control group posttest standard deviation was used instead of the pooled standard deviation because the variances of the treatment and control groups were statistically nonequivalent.

## Dosage Analysis

These ITT analyses answer important policy questions regarding the effect of adding an intervention to a "treatment as usual" control condition on developmental outcomes for adolescents. As will be described in a subsequent section, the treatment actually received did not overlap completely with the treatment assigned. Among the 416 cases included in the outcome analysis, 5 (2%) of the 205 control youths attended the ASP more frequently than the 8 days to which they were invited, whereas 53 (25%) of the 211 treatment youths failed to attend the program for more than 8 days. Hence, the ITT estimates may not reflect the effects of actual treatment participation. However, an analysis comparing outcomes for participants versus nonparticipants would produce biased estimates of the effect of participation to the extent that participation is related to unmeasured factors that also influence the outcomes (and were therefore included in the disturbance terms).

To obtain unbiased estimates of the effects of actual ASP participation, we conducted two sets of instrumental variable (IV) regressions using STATA (version 9.0; StataCorp LP, College Station, TX). The first IV regression estimated the effect of a variable that was coded "1" for

youths who attended the program for 9 or more days[19] and "0" for those who did not. The second IV regression estimated the effect of the actual days of attendance at the program (including days attended by control youths). For both sets of analyses, following Angrist (2006), the random assignment variable was used as an instrumental variable.[20] These IV estimates capture the effect of participation in the ASP only for individuals whose participation is influenced by random assignment to condition. For youths who would never participate or always participate regardless of condition assigned, the IV estimates are uninformative.

The IV regression models included the fitted value from a regression of the participation variable on the random assignment variable, four school dummies, and a small number of covariates that were significantly related to the outcome variable in each equation. These included the time-one measure of the outcome variable, and some equations also included additional time-one survey measures (time-one decision-making skills, drug resistance skills, or school attendance) that predicted the outcome variable. Although not required in the IV analyses, these covariates were included to reduce standard errors and increase efficiency in the estimates. The same set of covariates was included in the "first-stage" regressions predicting participation, as operationalized in the two manners discussed above.

[19] This cutoff was used because control youths were invited to attend for 8 days.
[20] IV estimates are generally accurate in models with a low number of instruments and in which the instrumental variable is a good predictor of the independent variable of interest—participation, in our case. Both of these conditions hold in our study, as only the random assignment variable is used as an instrument and this variable is highly correlated with actual days of attendance ($r$ = .63) and with binary participation ($r$ = .74).

## Exploration of Moderators

We also explored the possibility that the program might have been more effective for certain subgroups of the population. Prior research has hinted that latch-key youths (e.g., those left unattended during the after-school hours), lower socioeconomic status youths, and more at-risk youths might benefit the most from ASPs (Lauer et al., 2006). Other research has suggested that "moderate propensity" youths (e.g., those who have experimented with deviant behavior such as substance use but have not made a commitment to a deviant lifestyle) would be at greatest risk for being negatively influenced by other youths in an ASP (Dishion & Dodge, 2006). We tested these hypotheses in a series of interaction analyses. Additionally, we explored the relationship between student age and treatment effectiveness because previous research has found that age moderates effectiveness, but studies have found conflicting results about the direction of the relationship (D. C. Gottfredson et al., 2004; Vandell et al., 2005).

For the analyses involving interactions of treatment by student risk level, risk propensity was measured using a composite scale consisting of the following time-one measures: lifetime substance use, disruptive classroom behavior, aggression, delinquency, victimization, and impulsivity. It was both coded in its original form (i.e., with high scores indicating elevated propensity to engage in delinquent activities) and recoded so that moderate scores were coded as low; both high and low scores were coded as high. This was accomplished by centering the propensity measure at its mean and taking its absolute value before dichotomizing the variable based on the median split. This coding was used to test the possibility that youths who were committed to either a deviant or a nondeviant lifestyle would be less susceptible to negative peer influence as a result of ASP participation than would youths who had only experimented with

deviant behaviors (see Dishion & Dodge, 2006). These exploratory analyses were run using each of the eight main outcomes identified in the program model (Figure 3.1) as outcomes. For all of these analyses, regression models (OLS, negative binomial, Poisson, or logistic depending on the distribution of the dependent variable) were run, including the treatment variable, the potential moderator, a pretest measure of the dependent variable, a treatment by potential moderator interaction term, and four dummy variables for school.

Finally, additional analyses examined the extent to which the effect of assignment to treatment varied as a function of the quality of program implementation. These analyses were exploratory given the small number of programs and were run using only the eight main outcomes identified in the program model (Figure 3.1) as outcomes. Implementation quality was based on the 19 items from the program observation instrument described earlier measuring the quality of supervision, social climate, structure, behavior management, orderliness, and skillfulness of content delivery. As will be described in the Results section, two schools were consistently ranked higher and two consistently lower than the others across these quality measures. Consequently, we created dummy variables to contrast the high and low quality sites. Two sets of dummies were created, one placing the school that was consistently in the middle of the distribution in the higher and one on the lower quality set. For these analyses, regression models were run (OLS, negative binomial, Poisson, or logistic, depending on the distribution of the dependent variable), including the treatment variable, the quality dummy variable, a pretest measure of the dependent variable, and a treatment by quality dummy interaction term. The analyses were repeated for each quality dummy.

# 4. Results

## Program Implementation

As discussed, the intent of the study was to assess the extent to which the routine practices of the implementing agency could be shifted in the direction of providing more research-based programming, and to measure the effects of doing so on youth outcomes. The study is not an efficacy trial of the selected research components but a study of what happens when a typical youth-serving agency attempts to incorporate evidence-based programming into its routine. This section describes how the implementation of the program differed from the ideal.

### Program Staff

The program design called for a site director and three assistants at each of the five sites. This level of staffing was not achieved. Only 14 of 20 direct services positions were filled when the programs opened.[21] Thirteen individuals were hired after the beginning of the program to fill vacancies or replace lost staff. These new staff members did not receive the intensive startup training that the original staff received. Six of the original 14 staff members quit or were fired before the end of the year. Three staff members were relocated to new sites mid-year. Only six direct services staff worked at the site to which they were originally assigned for the entire program.

[21] This initial level of staffing was not regarded as problematic by the vendor because the student population was not yet at capacity at the start of the program.

On average, staff members worked at programs on 50.5 days, 53% of the 96 days the programs operated. Interesting differences in patterns of staffing were observed across sites (Table 4.1). Staff at sites B and E worked in their positions for more than 60 days on average, whereas staff at sites A and D worked far fewer average days, 35 and 48 days, respectively. Site C staff worked an average of 53 days.

Staff across sites received an average of 24.7 hours of job training, but this figure was far higher for original staff. The 14 original staff members received more than 40 hours of training on average, whereas the 13 replacement staff members received less than 6 hours. Consequently, sites where turnover was higher tended to employ fewer highly trained staff. For example, staff at Site A received 14 hours of training on average and staff at Site E received 33 hours. The average age of the 27 total site staff members was 33.0. They reported a median of 5.0 years of experience working with youth. Eight (30%) were males. Nineteen (70%) were African Americans. Nineteen (70%) had a bachelor's or more advanced degree. All had completed high school, and 9 (33%) were certified teachers. The large majority (75% or more) of staff members at all sites except Site D had earned a bachelor's degree. This was true for only 43% of Site D's staff.

**Table 4.1. Days Worked in ASP, Training, and Education, by Site**

| Site | Days Worked | Hours of Training | % BA or Higher |
|---|---|---|---|
| A | 35.13 | 13.69 | 75.00 |
| B | 61.00 | 27.40 | 80.00 |
| C | 53.40 | 22.50 | 80.00 |
| D | 47.71 | 31.07 | 42.86 |
| E | 65.40 | 32.60 | 80.00 |
| Overall | 50.47 | 24.65 | 70.00 |

In summary, staffing was particularly problematic at sites A and D, where staffing was unstable as indicated by fewer days worked, and staff quality was low in terms of either the level of training or education. Sites B and E had the most stable and the most highly qualified staffs.

## Program Management and Climate

Cross et al. (in press) describe the quality of implementation across the five sites in detail. Table 4.2 shows that the composite measure of program quality described earlier also varied significantly by site ($p < .001$). The average score on this scale was .59, and the range was .39 to .70. Sites B and E were rated fairly highly, with approximately 70% of the 19 items assessed favorably across observations. Sites A and D were rated less favorably, with half or fewer of the items assessed favorably. Student engagement rates, also measured by UM observers, followed a similar pattern, with youths at sites A and D being less engaged than youths in the other sites. The observation data make clear that, despite efforts to implement similar programs across the different sites, considerable heterogeneity was observed. In subsequent analyses of program quality, we distinguish sites B and E from the others.[22] These analyses test the effect of attending the highest quality programs versus programs of lesser quality.

[22] Alternative groupings, including one in which Site C was included with the higher implementation group, were tested. Results did not differ substantially.

**Table 4.2. Program Management and Climate and Engagement Rates, by Site**

| Site | Management and Climate | Engagement Rate |
|---|---|---|
| A | .50 | .71 |
| B | .70[a] | .81 |
| C | .68[a] | .79 |
| D | .39[a] | .73 |
| E | .70[a] | .80 |
| Overall | .59 | .77 |

[a]Significantly different from at least one other site.

Enhanced Services

*Attendance incentives*. The attendance incentive system was intended to increase the level of attendance. However, a slow start and deviations from the planned program likely contributed to the less than optimal effectiveness of this component. First, the system relied on staff entering attendance data and awarding points using the MIS each week. However, attendance was not recorded for approximately the first month of the program. When attendance was recorded, staff sometimes failed to award points every week. On average, points were awarded 23.2 times of the expected 32 (range, 18–26).

Once the staff adapted to the required data entry, individual incentive points for attending school were implemented fairly well. Group-based incentive points for program attendance, on the other hand, were not implemented as planned. Youth were to be placed unsystematically into groups containing three to four youth in an effort to encourage positive peer pressure for ASP attendance. Program staff thought this procedure was unfair and reorganized the groups so that high-attending youth were grouped together to ensure the attending students would receive the maximum points. In such groups, having members encourage their teammates to attend was unnecessary as these youth were already attending. This reorganization probably did not

encourage attendance among the lower attending youth because they were placed in groups with very low probabilities for receiving points.

It is also unclear whether the actual attendance incentives offered were valued by the participants. Researchers conducted focus groups with students in the participating schools to ascertain what types of incentives would be valued. Although this information was shared with the vendor, the attendance incentives offered did not always follow these recommendations. For example, rewards offered included carabineers and sports apparel that the youth in the focus group did not cite as desirable. Additionally, many rewards that focus group participants mentioned as especially desirable, such as tickets to movies or sports events, video games, jewelry, and mp3 players, were not offered. A catalogue containing all possible rewards was also not made available to students until midway through the program. It is unclear whether the incentive system might have been more effective in encouraging attendance if the rewards identified through the focus group were provided or if the full array of incentives were offered earlier.

Attendance award ceremonies, display of a visible attendance incentive chart for tracking points, and opportunities to spend points occurred far less frequently than planned. Staff reported holding award ceremonies on an average of 16.2 weeks (range, 9–23) and displaying an attendance chart for an average of 17.6 weeks (range, 4–32) of the 32 expected. Observers rarely witnessed award ceremonies or attendance incentive charts on display. Finally, sites provided the opportunity to spend points on an average of only 4.6 occasions (range, 2–7) of the expected 16 occasions (every 2 weeks). Students who were retained in the program ($n$ = 104) spent points an average of 1.8 times (range, 0–5). It is impossible to determine whether the relatively infrequent

point redemption resulted from the problems described above or because youth were saving their points for a larger reward at the end of the program.

*Academic activity*. The original design included a one-on-one paired tutoring model to be provided an hour and a half per week. As detailed in the Background and Methods sections, this component was abandoned midway through the year, despite extensive preprogram planning. This was primarily because of difficulties with assessment software access and tutor recruitment. Although an alternative academic activity was provided, it is possible that the failure to provide one-on-one tutoring contributed to withdrawal of youths whose parents viewed tutoring as the main draw of the program.

Once the alternative academic activity was substituted, it was implemented fairly regularly. Academic activities were offered an average of 30.4 days of the 34 expected (range, 22–46). These activities included workbooks, homework help, and independent reading. Observation data provide some insights into the quality of implementation of the component. Staff members always supervised these activities and were available to answer questions. The quality of interaction between staff and students varied across sites. In at least one site, staff remained fairly engaged with youth by checking their work, asking for clarification, and discussing ideas. Staff at other sites monitored the academic time in a more passive manner.

The academic component was offered regularly for half of the year, but exposure to the activity was limited by dropout and low attendance (to be discussed in a subsequent section of this report). Of the treatment youth who ever attended the ASP ($n$ = 205), 72.2% participated in an academic activity. The average youth who ever participated in an academic activity ($n$ = 148) received an average of 13.1 days of the 34 expected (range, 1–43).

*All Stars*. All Stars was well implemented by program staff, but similar to academic assistance, student exposure was less than anticipated because of the dropout and low attendance. The sites offered an average of 26 of the 27 available lessons (range 23–27) over the course of 102 sessions (range, 89–110). The number of sessions is large because students were split into two or three groups; thus, multiple sessions were held each day All Stars was offered. Also, lessons were most often delivered over two consecutive sessions. Almost all of the youth (91%) who ever attended the ASP ($n$ = 205) participated in an All Stars session. Even though All Stars was offered regularly and most students received some of the program, low attendance contributed to decreased student exposure. Of those youth who ever participated in All Stars ($n$ = 187), the average number of lessons received was 11.3 of the expected 27. The average hours received was 15.7 hours compared with 20.25 expected if All Stars were taught in 27, 45-minute sessions as recommended by the developer

The original intention was to have only those staff who participated in the extensive 3-day training prior to the start of the program deliver the All Stars curriculum. Staff turnover made it impossible to achieve this goal. One site, for example, had complete staff turnover (Site A), and it was, therefore, necessary for untrained replacement staff to deliver All Stars at that site. Despite high turnover, though, only 8% of the sessions ($n$ = 39) were led by untrained staff across all sites.

Staff who led All Stars sessions completed fidelity checklists in the MIS after each session (see Appendix Table E.1 for a description of this measure). Goal achievement was rated on a scale from 1 to 4, where 1 meant “poor” and 4 meant “very well.” Overall, staff reported a high level of session goal achievement ($M$ = 3.23). Each lesson comprised several activities with unique objectives. Staff were asked to report which activities they taught, the extent to which

they modified the activity, and the level of objective achievement. Staff reported a high percentage of activities taught per lesson (89%), a low level of modification ($M$ = 1.34) (range, 1–3, where 1 meant the teacher did not modify the activity at all and 3 meant the activity was modified a lot), and a high level of objective achievement ($M$ = 3.19; range, 1–4, where 1 meant not at all and 4 meant very well). In addition to the data reported solely by staff, we collected both staff and observer reports (see observation instrument, Appendix E.2) of quality of lesson as taught as well as a measure of student engagement in the activity. Both sources agreed that the lessons were well implemented ($M$ = 4.0 and 3.6, respectively; range 1–5) and that the students who were present were engaged ($M$ = 3.2 and 3.0, respectively, range, 1–4).[23] With the exception of higher quality ratings by observers at School B, there were no statistically significant differences by school on these measures.

### Leisure Activities and Overall Program Content

In addition to the research components that were expected to occupy about 3 hours per week, the program offered leisure activities during the remaining 6 hours of programming per week. On Tuesdays and Wednesdays, students engaged in leisure activities for 1.5 hours. On Thursdays, the entire 3 hours of programming was dedicated to leisure activities. The program coordinator created a weekly leisure activity schedule to ensure students would be exposed to a variety of activities. However, sites typically did not follow this schedule. Based on staff reports of activities provided, sports were the most commonly offered activity, consisting of 37% of all leisure activities. The second most commonly offered activities were board and card games as

[23] See also subsequent section on "Activity duration, structure, and student engagement" for engagement comparisons across all activities.

well as computers, each comprising 11% of all recorded leisure activities. Activities not listed on the schedule (classified as "other" in the MIS) made up nearly a quarter of recorded program content (23%). Despite intentions to provide a wide variety of activities to youth, it is clear that only a few activities were offered consistently during leisure activity time. Of course nonattendance and youth preference also influenced the activities actually experienced by individual youth.

Analyses were conducted using the data from the MIS to illustrate what activities youths participated in during a typical day. Activity-level attendance data were incomplete in the MIS. An estimate of the total hours spent in each activity[24] was obtained by imputing data for the missing activity-level data based on the pattern of activities recorded. The imputation occurred as follows. First, we calculated the percentage of time spent in each activity category using data from all recorded days. Next, total program time was calculated by multiplying 2.5 hours of program time by the number of days present.[25] Finally, the total program time was multiplied by the percentage of time spent in an activity category to obtain an imputed estimate of the percentage of total time each student spent in an activity category. These steps were taken to compute the average time spent in activities separately for each type of day (e.g., research or leisure) and overall. The imputation procedure assumes that the pattern of activities on unrecorded days was similar to the pattern on recorded days. In short, the following computations in Table 4.3 were based on imputations that estimated total time in each activity

[24] For ease of analysis, activities recorded in the MIS were grouped into five main categories. The academic category consisted mainly of the alternative academic activities, but it also included after-school council. The creative recreation category included arts and crafts, board games, computers, dance, drama, photography, and videography. The athletic recreation category consisted of sports. Life skills was dominated by All Stars but also included jump smart nutrition, service learning, and workforce skills. Finally, the content of activities classified as "other" in the MIS is largely unknown; however, observation data indicate that this category consisted of activities such as snack and program announcements, field trips, movies, and disciplinary quiet time.

[25] Although the programs ran for 3 hours per day, about 30 minutes per day were spent in snack, transition, and cleanup at the end of the day. These components were not recorded in the MIS by staff and thus were not factored into the percentage of time spent participating in activities.

based on each youth's total recorded hours in the program and percentage of that time spent in each activity category.

**Table 4.3. Time Spent in Activity Categories, by Type of Day**

| Activity Category | Research Days (Tuesdays and Wednesdays) | | Leisure Days (Thursdays) | | Overall (All Days) | |
|---|---|---|---|---|---|---|
| | Average Percentage of Time | Average Total Hours | Average Percentage of Time | Average Total Hours | Average Percentage of Time | Average Total Hours |
| Academic | 11.08% | 9.28 | 1.86% | .77 | 7.91% | 9.70 |
| Creative Recreation | 16.77% | 10.37 | 21.51% | 7.51 | 18.92% | 17.54 |
| Active Recreation | 21.25% | 13.65 | 35.15% | 13.50 | 27.30% | 27.16 |
| Life Skills | 36.78% | 22.72 | 4.66% | 1.48 | 25.76% | 23.08 |
| Other | 13.46% | 8.51 | 28.41% | 11.44 | 20.12% | 19.75 |

*Notes.* $n = 205$ (all students who ever attended). See footnote 25 for a description of each activity category.

Table 4.3 shows that the pattern of time expenditure differed for research and leisure days. Almost half of the day was spent in life skills and academic activities on research days, although as noted the academic activities delivered were not research-based. Alternatively, leisure days consisted of almost all non-research-based activities. Overall, a little more than a third of the time was spent in life skills and academic activities, whereas the remaining two thirds consisted of leisure-based activities.

*Activity duration, structure, and student engagement*. As discussed, observers rated duration, student engagement, and degree of structure in two or more distinct ASP activities each time they attended the ASP; one structured activity (e.g., All Stars or academics) and one unstructured activity (e.g., leisure activities). On Thursdays, when academics and All Stars were not offered, observers recorded engagement in only leisure activities. Engagement information

for 221 discrete activities was available for analysis. Some activities were observed by multiple observers. Observation data for these activities were aggregated to create one case per activity, resulting in 162 activity observations. Descriptive information on duration, structure, and engagement rates for each activity is provided in Table 4.4.

The average duration of activities (in number of minutes) overall was 42.99 minutes, and it varied by activity type ($p < .05$). Academic and All Stars activities, which were expected to be 45 minutes long, generally were close to this expectation.

Activities had an average structure rating of 3.15 (on a scale of 1–5). Structure peaked between 10 and 15 minutes into the activity, with the average structure of the third 5-minute interval being the highest at 3.48. The structure of the activity decreased after the first 15 minutes; those activities lasting over an hour (60–75 minutes) were rated 2.07 to 2.25 on the structure scale toward the conclusion of the activity. Structure ratings differed significantly by activity ($p < .01$). Post hoc tests indicated that All Stars activities had significantly higher structure than other activities.

The overall engagement rate across activities was 0.77, which means that on average 77% of youth were engaged in activities during the typical 5-minute interval. However, engagement was significantly different across activities ($p < .01$). Post hoc tests showed the engagement rate was significantly lower in academic activities than all other categories of activities.

Structured activities do not necessarily equate to engaging activities; active recreation activities had the lowest average structure rating but the highest engagement rates. However, All Stars sessions had high engagement rates *and* were more highly structured than all other activities. It seems that academic activities were the least engaging for youth. Academic

activities also had the second to lowest average structure rating, although not significantly different from creative recreation, active recreation, and other activities. The use of an evidence-based one-on-one tutoring program (as planned) might have resulted in more engagement and higher structure.

**Table 4.4. Activity Duration, Structure, and Student Engagement, by Activity Type**

| | Duration | | Structure | | Engagement Rate | |
|---|---|---|---|---|---|---|
| | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** |
| Overall | 42.99 | 14.05 | 3.15 | .82 | .77 | .20 |
| Academic | 40.92 | 9.80 | 2.89 | .67 | .52 | .24 |
| Creative Recreation | 48.07 | 16.26 | 2.96 | .61 | .82 | .16 |
| Active Recreation | 42.01 | 16.33 | 2.87 | .94 | .83 | .13 |
| All Stars | 44.66 | 11.04 | 3.63 | .71 | .81 | .15 |
| Other | 35.94 | 13.67 | 2.97 | .81 | .76 | .21 |
| *df*[a] | 156 | | 157 | | 140 | |
| *F* | 3.00* | | 7.82** | | 12.11** | |

*Notes.* The number of observations used in this analysis is 162. However, engagement rates were not calculated for activities in which the structure of each 5-minute interval was low (below 3). The *n* of cases for which engagement rate was calculated is 145.
[a]Within-groups. Between-groups degrees of freedom = 4.
*$p < .05$. **$p < .01$.

Program Attendance

The ASP sites were open for a total of 96 days beginning the third week in September 2006 and running through May 2007. One hundred and twenty youth, 54% of the treatment sample, withdrew from the ASP before the end of the year (see Table 4.5). When students withdrew, site staff noted the withdrawal date and the reason for withdrawal. The primary reason for dropout was voluntary withdrawal (63%). Another 20% were removed from enrollment by site staff because of very low or inconsistent attendance (usually after a month of unexplained

absence). Two students were asked to leave because of behavior problems, 1 student moved, and the remaining 17 students withdrew for unknown reasons. Site B had much less attrition than other sites with only 32% of students withdrawing before the end of the year. All other sites lost 50% or more of their participants to dropout. Site A lost 67%, whereas Site C lost 70%.

Comparisons on all demographic and pretest measures showed that retained ($n$ = 104) treatment students were more likely to be African American than withdrawn ($n$ = 120) treatment students (77.0% vs. 62.0%, $p < .05$), and they were absent from school about 3 days less during the previous school year (6.0 and 8.4, respectively, $p < .01$). Although withdrawal had a nonsignificant association with propensity for deviance overall, when propensity scores were recoded into low, medium, and high categories, 38% of withdrawn youth were coded as low propensity for deviance versus 27% of retained youth. These findings suggest a slight tendency for lower risk youths to leave the program.

Of the 96 possible days, the average days *enrolled* for treatment students was 54.2 days. Average days enrolled was (not surprisingly) higher for the retained students compared with withdrawn students (85.5 days and 27.0 days, respectively). The average days actually *attended* by all students was 35.6 days[26] (37.1% of possible days and 55.0% of days enrolled). Days attended was also higher for retained students (61.0 days: 63.5% of possible days and 71.0% of days enrolled) compared with withdrawn students (13.5 days: 14.1% of possible days and 31.0% of days enrolled). Table 4.5 shows that Site B students attended considerably more days (45.6) than students at other sites. Sites A and C had the lowest attendance, approximately 30 days. These differences approached statistical significance ($p$ = .06).

[26] The average rate of attendance was not different for "research" days (54.4%) versus leisure days (55.7%).

**Table 4.5. Days Attended and Percent Withdrawn, by Site, Treatment Only (*n* = 224)**

| | Days Attended | % Withdrawn |
|---|---|---|
| A | 29.83 | 66.67 |
| B | 45.58 | 32.00 |
| C | 29.24 | 70.27[a] |
| D | 35.90 | 53.33 |
| E | 33.76 | 53.66 |
| Overall | 35.59 | 53.57 |

[a]Significantly higher than at least one other site.

In summary, the level of withdrawal from the program was high and the rates of attendance were low, resulting in much lower levels of exposure to the research components than was intended. To place these figures in context, we note that they are similar to attendance levels reported in other evaluations of ASPs. The evaluation of the 21st Century Learning Center programs, for example, found that 50% of students dropped out of the program and students attended an average of 32.5 days during the school year (Dynarski et al., 2003). An evaluation of the Communities that Care (CTC) prevention system in which a variety of evidence-based program models were implemented in community settings (Fagan, Hanson, Hawkins, & Arthur, 2008) also reported that exposure to the evidence-based programming was lower in after-school than in school settings. In that study, only 77% of participants receiving any programming received at least 60% of the sessions (compared with 96% in the school-based CTC programs).

Control group contamination was low, but it did occur. As mentioned, five control students attended the ASP on more than 8 days. Program attendance for these students ranged between 11 and 89 days. Instrumental variables analysis presented later accounts for treatment exposure of these control students.

Estimated Impacts on Participation in ASPs and Activities

The enhanced ASP was intended to change the after-school activities of youth in the treatment group relative to youth in the control group who would have the "usual" after-school experience. Describing the activities of youth in the control group is important so that we can understand how their experiences differed from the treatment condition. The youth in the control condition were invited to the ASP on the last Thursday of every month, with the potential to attend the ASP eight times. On these days, youth participated only in recreational activities, as All Stars and academic activities were not offered. Of 223 control group youth, 48% ($n$ = 106) attended the monthly ASP activities at least once. On average, all control youth only attended 1.46 out of the possible 8 days (range, 0–8).

Of course, both treatment and control youths were free to participate in a variety of other after-school activities, both at school and elsewhere. Table 4.6 compares treatment and control group exposure to after-school activities using posttest and YES survey measures. It shows that the treatment group reported participating in after-school programs at school more than the control group both in the YES (75% treatment vs. 56% control) and at posttest (67% treatment vs. 55% control). However, the percentage of control youth reporting attending an ASP in their school exceeded the 48% that could have been expected based on their attendance in the enhanced ASP reported above. Therefore, at least some of the control youths must have attended an alternative ASP offered in their schools. Additional evidence that control youths participated in alternative programs is found in the youth's self-report at posttest on number of days spent in ASPs in the school. The treatment group reported spending about half a day more than control

youth in after-school programs at school (1.85 vs. 1.29). Although the intervention provided significantly more after-school programming to the treatment group, it seems that the control group was also able to find other activities at school in which to spend the after-school hours. What was the nature of these alternative activities?

**Table 4.6. Participation in ASPs, by Experimental Group**

| | Treatment | | | Control | | | |
|---|---|---|---|---|---|---|---|
| | ***n*** | **Mean** | **SD** | ***n*** | **Mean** | **SD** | ***p*-value** |
| Participation in ASP at School (YES; range, 0–1) | 196 | .75 | .43 | 193 | .56 | .50 | .00 |
| Participation in ASP at School (Posttest; range, 0–1) | 184 | .67 | .47 | 177 | .55 | .50 | .03 |
| Number Days per Week in ASP at School (Posttest; range, 1–5) | 180 | 1.85 | 1.67 | 173 | 1.29 | 1.65 | .00 |
| Number Days per Week in ASP not at School (Posttest; range, 1–5) | 175 | .92 | 1.54 | 168 | .93 | 1.54 | .96 |

*Sources*: YES survey and posttest youth survey.

Table 4.7 shows the proportion of youth reporting participation in all categories of activities as well as the number of activities reported, using responses to the YES survey described above. Overall, an equivalent proportion of youth in the treatment and control conditions reported participating in any activity during the after-school hours; 95% ($n$ = 187) of treatment youth and 96% ($n$ = 185) of control youth reported participating in at least one activity after school. Treatment youth reported participating in an average of 4.41 different activities during the after-school hours compared with 4.39 for the control group. Independent-sample $t$-tests showed no significant differences in treatment versus control reports of participation in *any* activity.

Youth in both groups reported participating in sports and performance/fine arts activities most often. Seventy-two percent of treatment youth and 73% of control youth reported participating in some form of sports after school (about 2.08 and 2.16 different sports on average, respectively). Forty-five percent of treatment youth reported participating in performance/fine arts activities compared with 48% of control youth (0.63 and 0.70 different performance activities, respectively). This analysis suggests that it is unlikely the enhanced ASP altered the variety of activities in which youths spent time during the after-school hours, although it did increase the amount of time spent in a school-based ASPs for treatment youths relative to controls.

**Table 4.7. Participation in After-School Activities, by Experimental Group**

| | Treatment *n* = 196 | | Control *n* = 193 | | |
|---|---|---|---|---|---|
| | **Mean** | **SD** | **Mean** | **SD** | ***p*-value** |
| **Proportion Reporting Participation** | | | | | |
| All Activities | .95 | .21 | .96 | .20 | .83 |
| Community/School-based Activity | .23 | .42 | .25 | .44 | .66 |
| Academic Activity | .24 | .43 | .25 | .43 | .93 |
| Performance/Fine Arts Activity | .45 | .50 | .48 | .50 | .52 |
| Faith-based/Service Activity | .34 | .48 | .33 | .47 | 1.00 |
| Sports Activity | .72 | .45 | .73 | .44 | .72 |
| **Number of Activities Reported** | | | | | |
| All Activities (range, 0–22) | 4.41 | 3.15 | 4.39 | 3.33 | .98 |
| Community/School-based Activity (range, 0–6) | .30 | .68 | .31 | .57 | .94 |
| Academic Activity (range, 0–3) | .26 | .48 | .30 | .57 | .45 |
| Performance/Fine Arts Activity (range, 0–5) | .63 | .89 | .70 | .91 | .37 |
| Faith-based/Service Activity (range, 0–4) | .40 | .63 | .36 | .54 | .68 |
| Sports Activity (range, 0–19) | 2.08 | 2.03 | 2.16 | 2.64 | .68 |

*Source*: YES survey.

*Comparison of enhanced ASP sample after-school activities with a national sample*. We explored the extent to which the surprisingly high availability of after-school alternatives to the experimental program available to study participants was unusual for middle-school youths. Appendix Table F.1 compares the percentage of 8th graders in our sample reporting participation in various after-school activities with 8th graders in two national samples. The BCPS sample in our study actually reported lower levels of participation in the selected activities than those youth included in Monitoring the Future and similar levels of activity participation to parent reports in the National Household Education Surveys program. It is, therefore, unlikely that the schools selected for participation in our study were highly unusual in the availability of alternative after-school activities. The control group, even with its relatively high level of participation in ASP activities, does seem to provide a reasonable "treatment as usual" comparison. ASPs are intended to provide activities for youth who would otherwise be unoccupied, but the data from this study suggest that even without access to a comprehensive ASP such as the one designed as part of this study, students readily found access to after-school activities.

## Estimated Impacts of Assignment to ASP on Student Outcomes

Table 4.8 presents the adjusted posttest means, significance level, and effect sizes comparing treatment and control for the five mediators and three outcome measures shown in the ASP program model (Figure 3.1).[27] Posttest means were adjusted for the pretest measure of each variable, race, age, gender, and school site. The main effect for treatment reaches the $p < .05$ level of statistical significance for only one measured outcome: unsupervised socializing. No significant differences between treatment and control youths were found on measures of conduct

[27] Results of exploratory analyses for all outcomes measured is provided in Appendix Table G.1.

problems, academic performance, school attendance, prosocial/antidrug attitudes, social competence, school bonding, or positive peer influence. Effect sizes range from a high of –.26 for unsupervised socializing (indicating that the experimental ASP participants scored approximately one quarter of 1 standard deviation lower, which is in the desired direction for this measure) to a low of –.05 for positive peer influence (indicating that the experimental ASP participants scored in the more negative or undesirable direction on this measure). The magnitude of the effect for the one difference that was statistically significant (unsupervised socializing) was small relative to our expectation: Youths attending the ASP reported being with their friends with no adults present for approximately one-half day less per week than did control youths. We anticipated a larger difference on this outcome, given that the program ran 3 days per week. No treatment by school interactions were found for these eight outcomes.

**Table 4.8. Adjusted Posttest Means and Effect Sizes**

| | Adjusted Posttest Mean | | | | | |
|---|---|---|---|---|---|---|
| **Scale** | **Control (*n*)** | | **Treatment (*n*)** | | ***p*-level** | ***d*** |
| Intermediate Outcomes | | | | | | |
| Unsupervised Socializing (days per week) | 2.40 | (205) | 1.89 | (211) | .01* | – .26 |
| Positive Peer Influence | .79 | (202) | .78 | (210) | .54 | – .05 |
| School Bonding | 1.99 | (202) | 2.00 | (210) | .85 | .01 |
| Social Competence | – .01 | (194) | – .01 | (202) | .92 | .01 |
| Prosocial/Antidrug Attitudes | .66 | (193) | .66 | (201) | .74 | .03 |
| Distal Outcomes | | | | | | |
| School Attendance | .05 | (203) | .05 | (207) | .57 | – .05 |
| Academic Performance | .01 | (205) | – .02 | (211) | .63 | – .03 |
| Conduct Problems | – .01 | (205) | .00 | (211) | .84 | .01 |

*Notes*. Means are adjusted for pretest level of the dependent variable, race, age, gender, and school site.

*$p < .05$.

## Dosage Analysis

Treatment youths' attendance ranged from 0 to 94 days (of 96 possible days), and five control participants received nontrivial exposure to the program. Although appropriate for answering policy questions about the overall effectiveness of adding an ASP to the activity options available to youth after school, the results of the ITT analysis reported above do not address the effects of actual attendance in the ASP. We, therefore, conducted a series of instrumental variable regressions, described earlier, to estimate the effect of actual days of attendance on the outcomes. Of course, given the overall null effects reported from the ITT analysis, the positive effects resulting from more days attended would imply that youths who attended infrequently were harmed. Although unlikely, this pattern of effects is possible, especially given findings from earlier evaluations of relatively unstructured programs of a drop-in nature that have shown negative outcomes for ASP participants in comparison with nonparticipants (Dynarski et al., 2003; Mahoney, 2000; Weisman et al., 2002).

The results of the IV regressions are shown in Table 4.9. Two models were run, one testing a dummy variable measuring whether the youth attended the program for 9 or more days (control youth had eight opportunities to attend the program) and another model using a continuous measure of days attended. These models show, first, that the ITT results indicating a positive effect of treatment on days with friends and no adult is apparently not from a linear relationship between actual days attended and time spent unsupervised with peers. The significant zero-order association with each participation variable is no longer significant in the model that includes controls for school and a pretest measure of unsupervised time usage. More days of actual attendance does not translate into significantly fewer days reported with friends

unsupervised. Continued investigations showed that program dropouts, who had fewer actual days of ASP attendance, engaged in alternative activities during the after-school hours that reduced their unsupervised time with friends.

The IV regressions yielded no other significant associations with program attendance.

**Table 4.9. Regression of Study Outcomes on Number of Days Attended (Instrumental Variables)**

| | Attended Nine or More Days | | Number of Days Attended[a] | |
|---|---|---|---|---|
| **Scale** | **Model 1** | **Model 2** | **Model 1** | **Model 2** |
| Intermediate Outcomes | | | | |
| Unsupervised Socializing (days per week) | –.58* | –.45 | –1.22* | –.96 |
| Positive Peer Influence | –.04 | –.07 | –.09 | –.10 |
| School Bonding | –.02 | –.06 | –.04 | .01 |
| Social Competence | –.00 | –.10 | –.00 | –.02 |
| Prosocial/Antidrug Attitudes | –.04 | –.02 | –.00 | .00 |
| Distal Outcomes | | | | |
| School Attendance | –.00 | –.00 | –.00 | –.00 |
| Academic Performance | –.05 | –.02 | –.01 | –.00 |
| Conduct Problems | .02 | .01 | .00 | .00 |

*Notes.* The coefficient is the unstandardized coefficient for attendance. Random assignment to treatment versus control is the instrument for attendance. Model 1 includes only the participation variable. Model 2 includes the participation variable as well as four dummy variables for school, a time 1 measure of the dependent variable, and in some equations time 1 decision-making skills, drug resistance skills, or school attendance. *N* of cases ranges from to 410 to 416.
[a]The coefficient for number of days attended is multiplied by 10.
*$p < .05$. **$p < .01$.

Moderator Analysis

As noted in the Methods section, we explored the possibility that the program might have been more effective for certain subgroups of the population. We tested hypotheses that the ASP would be more beneficial for these subgroups of youths: latchkey, lower SES, more at risk, and moderately at risk. We also investigated whether youth age interacted with program

effectiveness, as well as for conditional effects by level of program implementation measured in two different ways.

**Table 4.10. Adjusted Posttest Means, by Moderator and Experimental Group**

| | Adjusted Posttest Mean | | | |
|---|---|---|---|---|
| | **Control (*n*)** | **Treatment (*n*)** | ***p*-level** | ***d*** |
| **Social Competency** | | | .04 | |
| Latchkey | .09 (133) | –.08 (142) | | –.25 |
| Nonlatchkey | –.06 (59) | .03 (59) | | .13 |

*Notes*. Means are adjusted for pretest level of the dependent variable, school site, age, gender, and race.

In total, we conducted 56 (eight outcomes by seven potential moderator variables) tests for moderator effects. Only one of these produced a significant interaction, fewer than the number that would be expected by chance. Program effects did not differ for students who were lower versus higher SES, less versus more at risk, or who displayed moderate propensity for deviance versus very high or very low propensity. However, latchkey status interacted significantly with treatment to produce differential effects on social competency. Youth in the treatment group who were never supervised during the after-school hours at the time of the pretest (who represented 30% of the study population) declined in social competency relative to their controls, whereas treatment youth who were supervised at least some of the time at pretest increased on this construct as their controls declined. The direction of this interaction is opposite what was expected. See Table 4.10 for adjusted means for these subgroups, *p*-level for the interaction term, and effect sizes.

Tests for interaction by program implementation quality yielded no significant differences on the eight outcome variables examined. Therefore, we conclude that although we observed variability across the sites in implementation quality, this variability was not related to program effectiveness.

# 5. Discussion

## Conclusions about Intervention's Effectiveness

In this study, we have described what happened when an agency specializing in providing after-school care to students attempted to shift its practices to incorporate more research-based programming. We also sought to measure the effects of the "enhanced" after-school program on a range of youth outcomes. We were only modestly successful in accomplishing the desired shift in practices. In the end, the study provided a rigorous test of an ASP that more closely resembled the traditional after-school program implemented by BCRP than it did the planned intervention. That is, the tested program was a reasonably well-structured, but largely typical, school-based ASP similar to the routine programs delivered by BCRP except for the addition of a research-based prevention curriculum. As such, the study adds to the existing research on ASPs (summarized earlier) by providing a much-needed, well-implemented randomized trial of a typical ASP for middle-school students. The main conclusion from this study is that the program as implemented did not produce the desired effect on the participating youths.

The program as implemented fell short of expectation in all areas. Although the planned incentives for attending school were implemented, the group-based incentives for attending the ASP were not. The anticipated outcomes of this intervention, high program and school attendance, were not realized. The planned one-on-one tutoring program was replaced with a group-based academic assistance program that resembled what BCRP was accustomed to providing. Academic performance did not improve as a result of participation in the program. The only real shift in practices was the addition of the All Stars prevention curriculum. This

component, although delivered as planned, was crippled because high dropout and sporadic attendance limited student exposure. Neither the mediators—school bonding, social competence, and prosocial/antidrug attitudes—nor the more distal outcomes—academic performance and conduct problems—targeted by All Stars were influenced by program participation. Two additional outcomes were expected to improve as a result of simply attending the program. Students were expected to reduce their time spent in unsupervised socializing with peers, and peer influence was expected to become more positive. No effect for peer influence was observed. A significant reduction in time spent in unsupervised socializing was observed, but the magnitude of that effect was small relative to our expectation: Youths attending the ASP reported being with their friends with no adults present for approximately one-half day less per week than control youths. We anticipated a larger difference on this outcome, given that the program ran 3 days per week.

Variability in student exposure to the program was observed both within and between sites, and the quality of implementation also varied across the five implementing sites. We found no evidence that either more exposure to the ASP or higher quality implementation resulted in more beneficial outcomes for participating youths. At least within the range of program quality observed in this study, the program, as delivered, did not produce the expected effects.

The results for this experiment are similar to results from other experimental trials of ASPs. Zief et al. (2006), for example, limiting their meta-analysis to "well-implemented experimental design studies," found that of the 97 impacts measured by the five studies included in their review, 84% showed no significant differences between the program and control youth. As was the case in our study, Zief et al.'s positive findings were on measures of time expenditure in the after-school hours. Our only significant program effect was on unsupervised time spent

with peers in the after-school hours. Zief et al. found no significant effects on academic or behavioral outcomes.

Our results are not consistent with prior research that has suggested that ASPs are effective under certain conditions. The research summarized earlier suggested that more positive effects would be observed for more at-risk students and students who attended more, and in programs that were more structured, smaller, and were staffed by highly trained and educated staffs. We did not observe stronger effects for students exposed to more of the program or for more at-risk students, and we did not observe more positive effects in sites characterized by higher quality implementation. Within the range of implementation quality observed in this study, none of the programs could be regarded as effective.

Recognizing that the results from our study do not generalize beyond the five participating sites, our conclusion is that programs like this are not strong enough to increase academic performance, reduce problem behavior or school nonattendance, or influence any of the targeted intermediate behaviors and attitudes other than time expenditure. The program as delivered was not attractive enough to middle-school youths to keep them attending regularly. But the fact that the regular attenders did not benefit suggests that the problem extends beyond low exposure. We believe that the shortcomings of the program can be traced in large part to problems with the staffing of the programs. As detailed, BCRP was slow to fully staff the ASP sites, which resulted in many staff missing the major training in the research-based components that was provided prior to the opening of the programs. Staff turnover was also high, with only 6 of the 20 direct service staff staying in their positions for the entire program. We observed that coping with this staff instability occupied a great deal of the program coordinator's time and effort. Had a more highly qualified and stable staff been available, more efforts could have been

devoted to ensuring that the program components (especially the one-on-one tutoring and attendance incentives) were implemented more faithfully.

Unfortunately, staff instability of this nature is not unusual for after-school programs or for child-care programs more generally. A National Study on Child Care Staffing (Whitebook, Howes, & Phillips, 1998) found high turnover to be a problem in general and linked the high turnover to extremely low wages and poor benefits for these workers. Also, our Baltimore County partners who had been operating ASPs for years in Baltimore County confirmed that the staffing quality and turnover experiences in our program were common. The part-time nature of the work and the mid-afternoon hours makes the ASP staff position unattractive to potential staff, with the possible exception of teachers. However, for teachers, the rate of pay (direct service worker hourly pay ranged from $11.20 to $22.30, with an average of $17.13) was low relative to the teacher pay scale. What we regard as the most challenging obstacle to achieving high-quality programming in after-school programs is one feature that is likely to be extremely difficult to change.

Working within the constraint of staff who are likely to leave as soon as a better opportunity arises, it seems that much more attention to quality control will be needed. One possible strategy is a structure in which a central bank of workers is constantly being hired and trained, so that trained workers are always available to replace those who leave and to substitute for absent staff. In our study, the availability of such a pool of workers would have reduced the need for the program coordinator to scurry from site to site trying to cover absent workers and would have allowed the emphasis to shift toward ensuring quality programming on a day-to-day basis. If the problem of short staffing were solved, the central staff could then focus on providing

feedback to workers and retraining as necessary. This model would add considerably to the overhead costs of the program but would likely result in higher quality implementation.

Another mechanism for increasing quality would be to add a pilot year. Our first year was a planning year during which we secured our partners, developed memoranda of understanding, and worked with the participating school system to coordinate the programming in the ASP with school system objectives. Because the vendor had to be located through a time-consuming bidding process as required by the Baltimore County government, it was necessary to have a 1-year lead time before implementation. In retrospect, 1 year was insufficient. We encountered numerous challenges to implementation in addition to the staffing issues discussed above. Despite elaborate and detailed planning, the academic assistance component was thrown off course at the beginning of the school year by a school system decision to switch its diagnostic testing software (after we had developed the entire component around the "old" software system at the request of the school system). This obstacle could have been overcome during a pilot year. We experienced similar (although not as devastating) challenges with the attendance incentive system and with the leisure activities offered. All of these problems either were resolved or it was clear how they could be resolved by the end of the first year. With another year of implementation, we would likely have achieved higher quality implementation at more sites.

## Significance of Results

### For Practitioners

This study cannot offer guidance about the extent to which ASPs in general are effective for achieving their desired aims. Our results instead suggest that programs such as those implemented in the context of our study are not effective. The program evaluated in this study

resembled a typical ASP model routinely implemented by an organization such as BCRP except for the addition of a structured, evidence-based prevention component. This program was not sufficiently attractive to youths to achieve their consistent attendance. We conclude that the standard comprehensive ASP model may not be effective for middle-school youths. Even in the better implemented sites in this study, dropout and inconsistent student attendance were the rule rather than the exception; the best site had a 32% dropout rate and youths attended only an average of 47 of the 96 possible days. The examination of after-school time expenditure among the control group made clear that the program did not substantially alter after-school time usage for middle-school students. The ASP was only one of several options for after-school activity available to youths, and even the wide array of opportunities available to youths in the after-school hours is not sufficient to reduce substantially the amount of time youths spend unsupervised with their peers. Although at some point in the past opportunities for after-school activities may have been scarce, this is certainly no longer the case. Offering more voluntary programs in the after-school hours will not substantially alter the way middle-school youths choose to spend their time unless the programs can effectively compete with the variety of other available activities.

We therefore believe that it will be beneficial to explore alternative "high-interest" models for middle-school ASPs. Qualitative impressions of the programs summarized in Cross et al. (in press) suggest that the most consistent attendance was achieved in the sites in which staff were more effective at creating emotional bonds with the youth participants. These observations are consistent not only with criminological theory that links social bonding with several prosocial outcomes (e.g., Hirschi, 1969) but also with prior reviews that have found that the most effective programs are those in which staff have more positive relationships with youth (Beckett et al.,

2001; McComb & Scott-Little, 2003 ). This suggests that middle-school youths may respond better to after-school activities that focus on developing bonds with prosocial adults. Such models might be organized more like mentoring activities such that a small number of youths might be connected to an adult who would help youths develop a particular skill or ability. The structure for these programs might be more fluid and flexible than the typical comprehensive ASP model, allowing youths to participate in a variety of competing activities as well. This model could be organized around much more focused activities (e.g., photography, acting club, and math club) keyed to specific youth interests and could incorporate content shown to produce desired outcomes in prior research. But they would involve much closer relationships between the youths and the adults than is typical in a comprehensive ASP. Any one experience might last for a shorter duration, and youths might opt to participate in more than one throughout the school year.

The idea behind our research—that incorporating more evidence-based programming into existing comprehensive ASPs for middle-school youths will improve their effectiveness—continues to make sense but only if the programs can be delivered in such a way as to hold youths' interest. In this study, the All Stars prevention curriculum was implemented in a reasonably high-quality fashion at all five sites. Staff at all five sites were trained to implement All Stars, and a high proportion of lessons at all sites were implemented by trained staff (92%). But the typical student received only slightly more than half of the program because of nonattendance, and the outcomes most directly targeted by the All Stars curriculum were no different for students who were and were not exposed to the program. This of course does not imply that All Stars and similar evidence-based prevention programs are not effective but rather that voluntary ASPs are not ideal settings in which to attempt to deliver such a program. Unless

more stable staffing and more regular attendance can be achieved, practitioners are advised to incorporate prevention curricula into the regular school day where greater exposure and implementation quality are more likely (Fagan et al., 2008).

### For Researchers

As noted, we recommend that future studies of ASPs include sufficient time to pilot the model to resolve implementation difficulties prior to the beginning of the outcome study period. We also recommend that future research be undertaken to test alternative models for middle-school ASPs that focus on developing specific youth interests in the context of smaller, more tailored programs that encourage the development of meaningful bonds between youths and adults in addition to research-based content.

Furthermore, it is essential that ASPs continue to be studied using rigorous methods such as were used in this study. Experimental studies of ASPs are rare, and the results from these studies tend not to be as positive as those from studies using nonexperimental methods (see Zief et al., 2006). Because ASPs are voluntary and allow for substantial self-selection, experimental studies can be expected to more precisely identify program effects.

Finally, prior research (most recently summarized in Durlak & Weissberg, 2007) has suggested that the use of structured, evidence-based practices is important to the success of ASPs. Yet, the studies on which these findings are based have not generally assessed multiple dimensions of program quality. It is possible, and a question for future research, that the use of structured, evidence-based content is correlated with some or all of the other dimensions of implementation discussed in this report, including staff quality and general program management. To the extent this is true and these other dimensions are not controlled, the

correlation between program content and outcomes may be at least in part spurious. To determine the characteristics of truly effective ASPs, research assessing multiple dimensions of implementation quality is required.

## 6. References


Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: Sage.

Angrist, J. D. (2006). Instrumental variables methods in experimental criminological research: What, why and how. *Journal of Experimental Criminology, 2*, 23–44.

Beckett, M., Hawken, A., & Jacknowitz, A. (2001). *Accountability for after-school care: Devising standards and measuring adherence to them*. Santa Monica, CA: Rand Corporation.

Brooks, B. D. (1975). Contingency management as a means of reducing school truancy. *Education, 95*, 206–211.

Bry, B. H., & George, H. E. (1980). The preventive effects of early intervention on the attendance and grades of urban adolescents. *Professional Psychology: Research & Practice, 11*, 252–260.

Cawelti, G. (Ed.). (1999). *Handbook of research on improving student achievement*. Arlington, VA: Educational Research Service.

Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19*, 237–248.

Cross, A. B., Gottfredson, D. C., Wilson, D. M., Rorie, M. R., & Connell, N. M. (2010) Implementation quality and positive experiences in after-school programs, *American Journal of Community Psychology.* (Published online March, 2010: DOI: 10.1007/s10464-010-9295-z.)

Dishion, T. J., & Dodge, K. A. (2006). Deviant peer contagion in intervention programs. In K. A. Dodge, T. J. Dishion, & J. E. Lansford (Eds.), *Deviant peer influences in programs for youth problems and solutions*. New York: Guilford Press.

Dupont, W. D., & Plummer, W. D. (1990). Power and sample size calculations: A review and computer program. *Controlled Clinical Trials, 11*, 116–128.

Dupont, W. D., & Plummer, W. D. (1998). Power and sample size calculations for studies involving linear regression. *Controlled Clinical Trials, 19*, 589–601.

Durlak, J. A., & Weissberg, R. P. (2007). *The impact of after-school programs that promote personal and social skills*. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning.

Dynarski, M., James-Burdumy, S., Moore, M., Rosenberg, L., Deke, J., & Mansfield, W. (2003). *When schools stay open late: The national evaluation of the 21st Century Community Learning Centers Program: Final Report*. U.S. Department of Education, National Center for Education Evaluation and Regional Assistance. Washington, DC: U.S. Government Printing Office.

Dynarski, M., James-Burdumy, S., Moore, M., Rosenberg, L., Deke, J., & Mansfield, W. (2004). *When schools stay open late: The national evaluation of the 21st Century Community Learning Centers Program: New findings*. U.S. Department of Education, National Center for Education Evaluation and Regional Assistance. Washington, DC: U.S. Government Printing Office.

Eccles, J. S., & Templeton, J. (2002). Extracurricular and other after-school activities for youth. *Review of Research in Education, 26*, 113–180.

Eysenck, S. B., Easting, G., & Pearson, P. R. (1984). Age norms for impulsiveness, venturesomeness and empathy in children. *Personality and Individual Differences, 5*, 315–321.

Fabiano, L., Pearson, L. M., & Williams, I. J. (2005). *Putting students on a pathway to academic and social success: Phase III findings of the Citizen Schools evaluation*. Washington, DC: Policy Studies Associates.

Fagan, A. A., Hanson, K., Hawkins, J. D., & Arthur, M. W. (2008). Bridging science to practice: Achieving prevention program implementation fidelity in the Community Youth Development study. *American Journal of Community Psychology, 41*, 235–249.

Fashola, O. S. (1998). *Review of extended-day and after-school programs and their effectiveness*. Baltimore, MD: Center for Research on the Education of Students Placed at Risk.

Financeproject.org. (2007). *Estimated federal investment in out-of-school time*. Retrieved April 2008, from http://www.financeproject.org/publications/estimatedfederalOST_TFPflyer.pdf

Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D. C., Kellam, S., et al. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6*, 151–176.

Girod, M., Martineau, J., & Zhao, Y. (2004). After-school computer clubhouses and at-risk teens. *American Secondary Education, 32*, 63–76.

Gottfredson, D. C., Cross, A. B., & Soulé, D. A. (2007). Distinguishing characteristics of effective and ineffective after-school programs to prevent delinquency and victimization. *Criminology & Public Policy, 6,* 601–631.

Gottfredson, D. C., Cross, A. B., Wilson, D. M., Rorie, M., & Connell, N. (2010). An experimental evaluation of the All Stars prevention curriculum. *Prevention Science, 11*, 2, 142-154. (DOI: 10.1007/s11121-009-0156-7).

Gottfredson, D. C., Gerstenblith, S. A., Soulé, D. A., Womer, S. C., & Lu, S. (2004). Do ASPs reduce delinquency? *Prevention Science, 5*, 253–266.

Gottfredson, D. C., Gottfredson, G. D., & Hybl, L. G. (1990). Managing adolescent behavior: A multi-year, multi-school experiment. *American Educational Research Journal, 30*(1), 179–216.

Gottfredson, D. C., Gottfredson, G. D., & Weisman, S. A. (2001). The timing of delinquent behavior and its implications for ASPs. *Criminology & Public Policy, 1*(1), 61–80.

Gottfredson, G. D., & Gottfredson, D. C. (1992). *What about you*. Ellicott City, MD: Gottfredson Associates, Inc.

Gottfredson, G. D., Jones, E. M., & Gore, T. W. (2002). Implementation and evaluation of a cognitive-behavioral intervention to prevent problem behavior in a disorganized school. *Prevention Science, 3*, 43–56.

Granger, R. C., Durlak, J. A., Yohalem, N., & Reisner E. (2007). *Improving after-school program quality*. The William T. Grant Foundation Working Paper.

Hansen, D. M., & Larson, R. (2005). The Youth Experience Survey 2.0: Instrument revisions and validity testing. Retrieved January 2008, from http://web.aces.uiuc.edu/youthdev/UnpublishedManuscriptonYES2111%20(2).doc

Hansen, W. B. (1996). Pilot test results comparing the All Stars program with seventh grade DARE: Program integrity and mediating variable analysis. *Substance Use and Misuse, 31*, 1359–1377.

Hansen, W. B., & Dusenbury, L. (2004). All Stars Plus: A competence and motivation enhancement approach to prevention. *Health Education, 104*, 371–381.

Harrington, N. G., Giles, S. M., Hoyle, R. H., Feeney, G. J., & Youngbluth, S. C. (2001). Evaluation of the All Stars™ character education and problem behavior prevention program: Effects on mediator and outcome variables for middle school students. *Health Education Research, 28*, 533–546.

Hirschi, T. (1969). *Causes of delinquency*. Berkeley: University of California Press.

Hollister, R. G. (2003). *The growth in after-school programs and their impact*. Washington, DC: Brookings Institution.

Huang, D., Sung Kim, K., Marshall, A., & Perez, P. (2005). Keeping kids in school: An LA's BEST example. National Center for Research on Evaluation, Standards, and Student Testing (CRESST) Center, Center for the Study of Evaluation (CSE), Graduate School of Education and Information Studies, University of California, Los Angeles.

Huizinga, D., & Esbensen, F. (1990). *Scales and measures of the Denver Youth Survey*. Boulder, CO: Institute of Behavioral Science.

Kane, T. J. (2004). *The impact of after school programs: Interpreting the results of four recent evaluations*. The William T. Grant Foundation Working Paper.

Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research, 76*, 275–313.

Lauver, S. C. (2002). Assessing the benefits of an after-school program for urban youth: An impact and process evaluation (doctoral dissertation, University of Pennsylvania, 2002). *Dissertation Abstracts International, 63*(02), 553A.

Lipsey, M. W. (1998). Design sensitivity: Statistical power for applied experimental research. In Bickman, L. and Rog, D.J. (Eds.), *Handbook of applied social science research methods*. Thousand Oaks, CA: Sage Publications.

Mahoney, J. L. (2000). School extracurricular activity participation as a moderator in the development of antisocial patterns. *Child Development, 71*, 502–516.

McComb, E. M., & Scott-Little, C. (2003). *After-school programs: Evaluations and outcomes*. Greensboro, NC: The Regional Educational Laboratory at SERVE.

McNeal, R. B., Hansen, W. B., Harrington, N. G., & Giles, S. M. (2004). How All Stars™ works: An examination of program effects on mediating variables. *Health Education and Behavior, 31*, 165–178.

Poulin, F. S. (2003). *A self-report measure of deviancy training and susceptibility to peer pressure in young adolescents' friendships*. Paper presented at the European Conference on Developmental Psychology, Milan, Italy.

Prenovost, J. K. E. (2001). A first-year evaluation of after school learning programs in four urban middle schools in the Santa Ana Unified school district (doctoral dissertation, University of California, Irvine, 2001). *Dissertation Abstracts International, 62*, 03A.

Rorie, M., Gottfredson, D. C., Cross, A., Wilson, D. M., & Connell, N. (in press). *S*tructure and deviancy training in after-school programs*. Journal of Adolescence.* (DOI*:* 10.1016/j.adolescence.2010.01.007).

Schochet, P. Z. (2007). *Guidelines for multiple testing in experimental evaluations of educational interventions*. Princeton, NJ: Mathematica Policy Research, Inc.

Scott-Little, C., Hamann, M.S., & Jurs, S.G. (2002). Evaluations of after-school programs: A meta-evaluation of methodologies and narrative synthesis of findings. *American Journal of Education, 23,* 387-419.

Shelton, D. (2008). Translating theory into practice: Results of a 2-year trial for the LEAD programme. *Journal of Psychiatric and Mental Health Nursing, 15*, 313–321.

Sickmund, M., Snyder, H. N., & Poe-Yamagata, E. (1997). *Juvenile offenders and victims: 1997 update on violence*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Smith, C., & Kennedy S. D. (1991). *Final impact evaluation of the Friendly PEERsuasion program for Girls Incorporated*. New York: Girls Incorporated.

St. Pierre, T. L., Mark, M. M., Kaltreider, D. L., & Aikin, K. J. (1997). Involving parents of high risk youth in drug prevention: A three-year longitudinal study in Boys & Girls clubs. *Journal of Early Adolescence, 17*, 21–50.

Tanglewood Research. (2008a). *All Stars Quality Assessments*. Retrieved July 2006, from http://www.allstarsprevention.com/programs/programsEvaluationQA.asp

Tanglewood Research. (2008b). *All Stars Student Surveys*. Retrieved July 2006, from http://www.allstarsprevention.com/programs/programsEvaluationOA.asp

Vandell, D. L., Reisner, E. R., Brown, B. B., Dadisman, K., Pierce, K. M., Lee, D., et al. (2005). *The study of promising after-school programs: Examination of intermediate outcomes in year 2*. Report to the Charles Stewart Mott Foundation.

Wasik, B. A. (1998). Using volunteers as reading tutors: Guidelines for successful practices. *The Reading Teacher, 51*, 562–570.

Weisman, S. A., Womer, S. C., Kellstrom, M. A., Bryner, S. L., Kahler, A., Slocum, L., & Gottfredson, D. C. (2002). *Maryland after school community grant program, report on the 2001-2002 school year evaluation of the phase 3 after school programs*. Technical report available from authors.

Whitebook, M., Howes, C., & Phillips, D. (1998). *Worthy work, unlivable wages: The National Child Care Staffing Study, 1988-97*. Washington, DC: National Center for the Early Childhood Work Force.

Wilson, D. M., Gottfredson, D. C., Cross, A. B., Rorie, M., & Connell, N. (in press). Youth development in after-school leisure activities. *Journal of Early Adolescence*.

Zief, S. G., Lauver, S. & Maynard, R. (2006). *The impacts of after-school programs on student outcomes: A systematic review for the Campbell Collaboration*. Retrieved April 2008, from http://www.campbellcollaboration.org/doc-pdf/zief_afterschool_review.pdf

Appendix A: Previous Research on ASPs Serving Middle-School Youth

Appendix Table A.1 Summary of Previous Research on ASPs Serving Middle-School Youth

| Author | Intervention and Duration | Design | Sample | % Attrition/ Potential Differential Attrition Bias | Outcomes Reported | Results | Favorable? |
|---|---|---|---|---|---|---|---|
| Dynarski et al. (2004) | 21st Century Community Learning Centers, 2 years | Pre-Post, NECG, adequate pretest controls | 1782 TX<br>2482 C | 91% TX<br>90% C<br>/ no | Self-care, location after school, days staying after school for activities, activity participation, teacher reports of effort, school discipline records, homework habits, educational aspirations, social and emotional outcomes, feelings of safety, negative behaviors, victimization | Treatment effects for days after school for activities, participation in lessons and clubs, less likely to be with sibling after school, better school attendance, social studies grade. Also increase in negative behavior composite scale for the treatment group. | D, N, A |
| Fabiano, Pearson, & Williams (2005) | Citizen Schools Program, 3 years | Pre-Post, NECG, adequate pretest controls | 855 TX<br>855 C | no data | Attendance, suspension, promotion to next grade, math and English grades, standardized test score in English and math | **6th & 7th graders:** Treatment effect attendance, suspension, promotion to next grade, test score in English<br>**8th graders:** Treatment effect attendance, promotion to next grade | D, N |
| Girod, Martineau, & Zhao (2004) | KLICK! After-school technology club, 2 years | Pre-Post, NECG, lacking adequate pretest controls | 231 total at pre-test; at posttest 81 TX, 78 C | 31% / yes | Student–teacher relations, readiness for classroom instruction, perceived parental involvement in school, overall school value, and experience using computer technologies | Treatment effect on valuing school and experience with computers | N |
| Gottfredson et al. (2004) | Maryland After School Community Grant Program, programs in operation for various durations | Pre-Post, NECG, adequate pretest controls | Older Sample—239 TX 201 C | Overall older and younger samples—11% TX 13% C / no | Delinquent behavior, rebellious behavior, last-year drug use, intentions not to use drugs, hours/week in self-care, involvement in constructive activities, social skills, positive peers, peer drug models | Treatment effects for constructive activities, drug-using peers, and last year drug use—Structural equations model finds treatment effect for latent "delinquent behavior" | D, N, A |

| Author | Intervention and Duration | Design | Sample | % Attrition/ Potential Differential Attrition Bias | Outcomes Reported | Results | Favorable? |
|---|---|---|---|---|---|---|---|
| Gottfredson et al. (2007) | Maryland After School Opportunities Fund Program, programs in operation for various durations | Pre-Post, NECG, adequate pretest controls | 389 TX<br>108 C | 41% TX<br>31% C / yes | Delinquency, victimization, and substance use | Treatment effect for delinquency | D, N |
| Huang et al. (2005) | LA's Best Program, three years | Pre-Post, NECG, adequate pretest controls | 5827 TX<br>5816 C | 0% | School dropout | Treatment effect for dropout | D, N, A |
| Lauver (2002) | After-school recreation program, three years | Pre-post, RCT | 126 TX<br>101 C | 3%/ yes | Constructive activities, self-care, time spent on homework, educational aspirations, attendance, grades, standardized test scores | Treatment effects for participation in fitness activities, time spent on homework, educational aspirations | D, N, A |
| Prenovost (2001) | After-school learning program, first year of program operation | Pre-Post, NECG, lacking adequate pretest controls | 300 High-Dose<br>304 Low-Dose<br>828 C | No data | Standardized reading and math test scores, study effort, school attendance, feelings of safety at school | High-dose treatment group improved more in attendance | N |
| Shelton (2008) | LEAD expressive art program, 14 weeks | Pre-Post, NECG, adequate pretest controls | 46 TX<br>43 C | 0% | Self-esteem, resilience, behavioral self-control, and protective factors | No treatment effects | D, A |
| Smith & Kennedy (1991) | Friendly PEERsuasion program, 14 weeks | RCT, Pre-post | All girls.<br>152 TX<br>202 C | 17% TX<br>19% C / no | Avoiding substance use, leaving situations in which substances are being used | Study used critical value of $p < .10$. Treatment effects on drinking, drinking initiation, and leaving situations where drinking was occurring | D, N, A |

| Author | Intervention and Duration | Design | Sample | % Attrition/ Potential Differential Attrition Bias | Outcomes Reported | Results | Favorable? |
|---|---|---|---|---|---|---|---|
| St. Pierre, Mark, Kaltreider, & Aikin (1997) | Boys and Girls Clubs implementing a drug prevention program, 3 years | Pre-Post, NECG, adequate pretest controls | 411 TX<br>105 C | 39% TX<br>46% C/ yes | Basic social skills, drug knowledge, attitudes about drugs, drug use, and drug refusal skills | Treatment effects for drug refusal skills, drug knowledge and drug attitudes.<br>No effects on drug use outcomes | D, N |
| Weisman et al. (2002) | Maryland After School Community Grant Program, programs in operation for various durations | Pre-Post, NECG, adequate pretest controls | 594 TX<br>476 C | 21% TX<br>23% C / no | Social skills, GPA, rebelliousness, commitment, intentions not to use drugs | Iatrogenic effects for social skills, GPA, rebelliousness, commitment, but positive treatment effect for Intentions not to use drugs | D, N, A |

*Notes*. Includes only studies of programs delivered primarily during after-school hours and that included more than 10 sessions.
*Abbreviations*. RCT = randomized, controlled trial; NECG = nonequivalent comparison group; TX = treatment group; C = comparison group; D = meets criteria for sound research design having either a randomized design or using sufficient controls on identified pretest differences between groups; N = meets criteria for sufficient sample size having more than 100 participants per experimental group; A = no unaddressed problems with attrition, coded when attrition was < 20%, or evidence was presented demonstrating that differential attrition was not introducing bias.

Appendix B: Demographic Characteristics of ASP Sample and Participating Schools

Appendix Table B.1. Demographic Characteristics of ASP Sample and Participating Schools

| School | % Eligible Registered | % Male | | % Minority | | % Subsidized Meals | |
|---|---|---|---|---|---|---|---|
| | | School | ASP | School | ASP | School | ASP |
| A | 8.46 | 52.08 | 50.70 | 64.36 | 76.06 | 64.96 | 66.67 |
| B | 20.87 | 48.76 | 45.54 | 47.11 | 67.33 | 64.75 | 60.20 |
| C | 10.54 | 51.68 | 47.22 | 50.80 | 65.28 | 66.98 | 65.28 |
| D | 21.20 | 55.48 | 58.33 | 97.88 | 98.33 | 48.89 | 48.74 |
| E | 11.54 | 52.85 | 65.06 | 99.30 | 100.00 | 63.36 | 58.75 |

# Appendix C: Response Rates and Attrition Analyses

Appendix Table C.1. Youth Survey, YES, Teacher Ratings, and School Records Response Rates, by Experimental Group

| | **Total Sample** | ***N*/Response Rate** | **Total Treatment** | ***n*/Response Rate** | **Total Control** | ***n*/Response Rate** |
|---|---|---|---|---|---|---|
| Pretest Youth Survey | 447 | 447/100.00 % | 224 | 224/100.00 % | 223 | 223/100.00 % |
| Posttest Youth Survey | 447 | 416/93.06% | 224 | 211/94.20% | 223 | 205/91.93% |
| YES | 447 | 389/87.02% | 224 | 196/87.50% | 223 | 193/86.55% |
| Teacher Ratings[a,b] | 427 | 423/99.06% | 215 | 213/99.07% | 212 | 210/99.06% |
| Pretest School Records [b] | 447 | 447/100.00 % | 224 | 224/100.00 % | 223 | 223/100.00 % |
| Posttest School Records [b] | 447 | 447/100.00 % | 224 | 224/100.00 % | 223 | 223/100.00 % |

[a]Teacher ratings were not sought for 20 students because they had withdrawn from the BCPS system or had transferred to a new school too recently for the teacher to rate reliably the student.
[b]Percentage reflects student with at least one teacher rating or at least one data element from school records.

Appendix Table C.2. Demographic Characteristics of Sample, by Attrition Status

| **Demographics** | **Included Cases *n* = 416** | | **Excluded Cases[a] *n* = 31** | |
|---|---|---|---|---|
| | **Mean or %** | ***N*** | **Mean or %** | ***N*** |
| Age[b] | 12.20* (.99) | 416 | 12.60 (.95) | 31 |
| Family Income (Median) | $32,454 | 376 | $30,210 | 27 |
| % Male | 52.88 | 416 | 64.52 | 31 |
| % Black | 69.47 | 416 | 70.97 | 31 |
| % 6th Grade | 42.31 | 416 | 35.48 | 31 |
| % 7th Grade | 33.17 | 416 | 38.71 | 31 |
| % 8th Grade | 24.52 | 416 | 25.81 | 31 |
| % Living With Two Parents | 37.50 | 416 | 29.03 | 31 |
| % Subsidized Meals | 58.05 | 410 | 67.86 | 28 |
| % Mother is College Graduate | 12.71 | 409 | 10.35 | 29 |

*Notes*. No experimental group/attrition status interactions significant.
[a]Excluded cases did not take the posttest or had more than 40% of the items missing.
[b]Standard deviation in parenthesis.
*$p < .05$.

Appendix Table C.3. Pretest by Attrition Status

| Scale | Included Cases *n* = 416 | | | Excluded Cases[b] *n* = 31 | | |
|---|---|---|---|---|---|---|
| | **Mean** | **SD** | ***N*** | **Mean** | **SD** | ***N*** |
| **Composite Scales** | | | | | | |
| Social competence | –.00 | .73 | 416 | .08 | .76 | 31 |
| Prosocial/antidrug attitudes | .75* | .22 | 416 | .66 | .26 | 31 |
| Academic performance | .02 | .83 | 374 | .00 | 1.10 | 26 |
| Conduct problems | –.02 | .63 | 416 | .36 | 1.48 | 31 |
| **Unsupervised Socializing** | | | | | | |
| Days with friends and no adults | 2.27 | 2.12 | 406 | 2.73 | 1.97 | 26 |
| **Positive Peer Influence** | | | | | | |
| Positive peer influence | .84 | .19 | 413 | .85 | .21 | 30 |
| **School Bonding** | | | | | | |
| Attachment to school | 2.49 | .86 | 414 | 2.07 | 1.19 | 29 |
| **Social Competence** | | | | | | |
| Goal setting | 3.11 | .51 | 414 | 3.14 | .37 | 29 |
| Decision-making skills | 2.83 | .76 | 398 | 2.90 | 1.02 | 26 |
| Impulsiveness[c] | .51 | .27 | 409 | .52 | .28 | 27 |
| **Prosocial/Antidrug Attitudes** | | | | | | |
| Attitudes unfavorable to drug use | .78* | .25 | 416 | .68 | .26 | 30 |
| Belief in conventional rules | .71 | .24 | 416 | .62 | .30 | 31 |
| **School Attendance** | | | | | | |
| % days absent (SR) | 4.19 | 3.90 | 376 | 6.20 | 7.74 | 29 |
| **Academic Performance** | | | | | | |
| GPA[a] (SR) | 2.42 | .72 | 213 | 2.34 | 1.11 | 18 |
| MSA reading (SR) | 389.40 | 28.42 | 371 | 390.04 | 32.45 | 24 |
| MSA math (SR) | 385.60 | 38.44 | 371 | 377.79 | 49.99 | 24 |
| **Conduct Problems** | | | | | | |
| Disruptive classroom behavior | 1.38 | .49 | 413 | 1.44 | .54 | 30 |
| Aggression | 1.58 | .59 | 412 | 1.60 | .59 | 29 |
| Delinquent behavior | .43 | .99 | 413 | .93 | 1.74 | 28 |
| Victimization | 1.31 | 1.64 | 411 | 1.35 | 1.65 | 26 |
| Last month drug use | .09 | .29 | 412 | .13 | .34 | 30 |
| Number suspensions (SR) | .35 | .74 | 376 | 0.82 | 1.68 | 28 |

*Notes*. Scales from youth survey unless otherwise noted; SR = school records. Significant treatment by attrition interaction for MSA math scores.

[a]Pretest GPA unavailable for 6th-grade students.

[b]Excluded cases either did not take the posttest or had more than 40% of the items missing.

*$p < .05$.

Appendix D: Outcome Measures

Appendix D.1. Youth Survey Posttest

# After School Program Youth Survey

The purpose of this research is to learn more about after school programs and the students who attend these programs. These pages ask questions about you, what you think, and what you do at school and away from school. In order for this survey to be helpful, it is important that you answer each question as thoughtfully and honestly as possible.

Your help with this survey is up to you. You have the right to not answer any or all the questions. But we want you to know that your answers are very important. All of your answers will be kept private and will never be seen by your parents, friends, teachers, or anyone else at your school or in your community.

Be sure to read the instructions below before you begin to answer. Thank your very much for being an important part of this project.

**Instructions**

1. This is **NOT** a test, so there are no right or wrong answers.
2. Please do **NOT** write your name anywhere on this survey.
3. All of the questions should be answered by marking one of the answer spaces. If you don't find an answer that fits exactly, use the one that comes closest.
4. Your answers will be read automatically by a computer. Please follow these instructions carefully.
   - Make heavy marks inside the circle.
   - Draw an **X** over any answer you wish to change.
   - Make no other markings or comments on the survey, since they may interfere with the automatic reading.

**This kind of mark will work:**
**Correct Mark**

**These kinds of marks will <u>NOT</u> work:**
**Incorrect Marks**

## Some Questions About You

*Please answer the following questions about yourself…*

1. Are you?

O Male
O Female

2. How old are you?

O 9 years
O 10 years
O 11 years
O 12 years
O 13 years
O 14 years

3. What grade are you in?

O 6th grade
O 7th grade
O 8th grade

4. How do you describe yourself? (check all that apply)

O White
O Black
O Native American/Alaskan Native
O Asian/Pacific Islander
O Latino/a
O Other

5. Which of the following adults do you live with most of the time? (check all that apply)

O Mother
O Father
O Stepmother
O Stepfather
O Grandmother
O Grandfather
O Other adult relative
O Other adult who is not a relative

6. Do you get a free or reduced lunch at school?

O Yes
O No

7. How many people live in your household?

O 1
O 2
O 3
O 4
O 5
O 6
O 7
O 8 or more

## School and Your Education

*Think about your school, would you say the following statements are mostly true or mostly false?*

| | | Mostly True | Mostly False |
|---|---|---|---|
| 8. | I feel like I belong at this school. | O | O |
| 9. | I wish I did not attend this school. | O | O |
| 10. | This school is a pretty good school to go to. | O | O |

---

11. What grades do you earn in school?

O Mostly A's
O About half A's and half B's
O Mostly B's
O About half B's and half C's
O Mostly C's
O About half C's and half D's
O Mostly below D's
O Not applicable

12. Do you think you will get a college degree?

O Yes
O No
O Not Sure

13. Do you expect to complete high school?

O I am certain I will finish high school.
O I probably will finish high school.
O I probably will not finish high school.

14. Some students think it is important to work hard in school and others do not. How important do you think it is to work hard in school?

O Very important
O Important
O Not important
O Not at all important

15. Compared to other students how hard do you work in school?

O Much harder
O Harder
O Not as hard
O Much less hard

16. How many hours did you spend reading for fun (not for school work) yesterday?

O I did not read for fun.
O Less than 15 minutes
O 15 to 30 minutes
O More than 30 minutes

## How Do You Spend Your Time After School?

17. In a typical **week** during the school year how many **days** (Monday – Friday) did you spend doing something where **NO ADULTS were present** after school?

O None → Skip to question # 20

O 1 Day
O 2 Days
O 3 Days
O 4 Days
O 5 Days

18. Of these days, how many did you spend **hanging out with your friends** with no adults present?

O None → Skip to question # 20

O 1 Day
O 2 Days
O 3 Days
O 4 Days
O 5 Days

19. On these days, how many hours each day (between 3 and 6 pm) did you spend with your friends and without adults?

O 1 Hour
O 2 Hours
O 3 Hours

20. In a typical **week** during the school year how many **days** (Monday – Friday) did you spend doing something where **AN ADULT WAS present** after school?

O None
O 1 Day
O 2 Days
O 3 Days
O 4 Days
O 5 Days

## Your Goals

| *Think about goals you have set for yourself…* | | Never | Sometimes | Often | All the time |
|---|---|---|---|---|---|
| 21. | How often do you set goals? | O | O | O | O |
| 22. | How often do you work on goals that you have set for yourself? | O | O | O | O |
| 23. | I think about what I would like to be when I become an adult. | O | O | O | O |
| 24. | When I set a goal, I think about what I need to do to achieve that goal. | O | O | O | O |

| | | Strongly Agree | Agree | Disagree | Strongly Disagree |
|---|---|---|---|---|---|
| 25. | Once I set a goal, I don't give up until I achieve it. | O | O | O | O |
| 26. | Whenever I do something I always give it my best. | O | O | O | O |

## Relationship With Your Parents

| *Think about your relationship with your parents…* | | Strongly agree | Agree somewhat | Disagree somewhat | Strongly disagree |
|---|---|---|---|---|---|
| 27. | When I go someplace, I leave a note for my parents. | O | O | O | O |
| 28. | My parents know where I am when I am not at home or at school. | O | O | O | O |
| 29. | I know how to get in touch with my parents if they are not at home. | O | O | O | O |
| 30. | My parents know who I am with if I am not at home. | O | O | O | O |
| 31. | My parents keep close track of how well I am doing in school. | O | O | O | O |
| 32. | My parents let me stay away from the house when I want. | O | O | O | O |
| 33. | My parents usually know if I do something wrong. | O | O | O | O |
| 34. | My parents usually let me go wherever I want after school. | O | O | O | O |

## Your Behavior

| *Have you **ever**…* | | Yes | No |
|---|---|---|---|
| 35. | Smoked cigarettes? | O | O |
| 36. | Drunk beer, wine, or "hard" liquor? | O | O |
| 37. | Smoked marijuana (weed, pot)? | O | O |

| | | Yes | No |
|---|---|---|---|
| 38. | Used inhalants (sniffed or huffed glue, gas, sprays, marking pens)? | O | O |
| 39. | Used another illegal drug? | O | O |

| *In the last **month** how often have you…* | | Not at all | Once or twice | A few times a week | Every day |
|---|---|---|---|---|---|
| 40. | Smoked cigarettes? | O | O | O | O |
| 41. | Drunk beer, wine, or “hard” liquor? | O | O | O | O |
| 42. | Smoked marijuana (weed, pot)? | O | O | O | O |

| *Since the **beginning of this academic school year ,** how often have you…* | | Never | Once | More Than Once |
|---|---|---|---|---|
| 43. | Been sent out of a classroom by a teacher for bad behavior? | O | O | O |
| 44. | Been suspended from school? | O | O | O |
| 45. | Gone to school, but skipped (cut) a class? | O | O | O |

| *Are the following statements mostly true or mostly false?* | | Mostly True | Mostly False |
|---|---|---|---|
| 46. | I will never smoke cigarettes. | O | O |
| 47. | I will never try marijuana or other drugs. | O | O |
| 48. | People my age who smoke are show-offs. | O | O |
| 49. | I will never drink beer, wine, or “hard” liquor. | O | O |
| 50. | Being honest is more important than being popular. | O | O |
| 51. | I admit it when I have done something wrong. | O | O |
| 52. | Sometimes you have to be a bully to get respect. | O | O |
| 53 | You have to be willing to break some rules if you want to be popular with your friends. | O | O |
| 54. | Sometimes a lie helps to stay out of trouble with the teacher. | O | O |
| 55. | The grades I get in school are important to me. | O | O |
| 56. | I turn my homework in on time. | O | O |
| 57. | If a teacher gives a lot of homework, I try to finish all of it. | O | O |
| 58. | I am satisfied with the way I am doing in school. | O | O |
| 59. | My grades at school are good. | O | O |
| 60. | I am proud of my school work. | O | O |
| 61. | I won’t let anything get in the way of my school work. | O | O |
| 62. | I usually quit when my school work is too hard. | O | O |
| 63. | I try to do my best at school work. | O | O |
| 64. | It is important to me to complete assignments given by teachers. | O | O |

| *In the **last 30 days**, how often…* | | **Not at all** | **Once or twice** | **A few times a week** | **Every day** |
|---|---|---|---|---|---|
| 65. | Did you tease someone else your age? | O | O | O | O |
| 66. | Did you encourage other people your age to fight? | O | O | O | O |
| 67. | Were you angry most of the day? | O | O | O | O |
| 68. | Did you push, shove, hit, or kick someone? | O | O | O | O |
| 69. | Did you call someone your age a bad name to their face? | O | O | O | O |
| 70. | Did you threaten to hurt or hit someone? | O | O | O | O |

| *Since the **beginning of this academic school year** have you…* | | **Yes** | **No** |
|---|---|---|---|
| 71. | Purposely damaged or destroyed property belonging to a school? | O | O |
| 72. | Purposely damaged or destroyed other property that did not belong to you, not counting family or school property? | O | O |
| 73. | Been involved in gang fights? | O | O |
| 74. | Used force or threat to get money or things from a person? | O | O |
| 75. | Stolen or tried to steal things worth less than $50? | O | O |
| 76. | Stolen or tried to steal something at school, such as someone's coat from a classroom, locker, or cafeteria, or a book from the library? | O | O |
| 77. | Belonged to a gang that has a name and engages in fighting, stealing, or selling drugs? | O | O |
| 78. | Had someone use a weapon, force, or threat to get money or things from you? | O | O |
| 79. | Been hit by someone trying to hurt you? | O | O |
| 80. | Had your wallet or purse stolen, or an attempt made to do so? | O | O |
| 81. | Had some of your things other than a wallet or purse, stolen from you? (Including books, clothes, money taken from a car, locker, home or cafeteria, etc.) | O | O |
| 82. | Been attacked by someone with a weapon or someone trying to seriously hurt you? | O | O |
| 83. | Had someone threaten to beat you up? | O | O |
| 84. | Had some of your things damaged on purpose (such as your bike tires slashed, or your books and clothing ripped)? | O | O |

| *Do you…* | | Yes | No |
|---|---|---|---|
| 85. | Often get in trouble because you do things without thinking? | O | O |
| 86. | Usually work quickly without checking your answers? | O | O |
| 87. | Usually think carefully before doing anything? | O | O |
| 88. | Sometimes break the rules without thinking about it? | O | O |
| 89. | Mostly speak without thinking things out? | O | O |
| 90. | Often get involved in things you later wish you could get out of? | O | O |
| 91. | Get bored more easily than most people doing the same old things? | O | O |
| 92. | Need to use a lot of self control to keep yourself out of trouble? | O | O |
| 93. | Get very annoyed if someone keeps you waiting? | O | O |
| 94. | Get very restless if you have to stay around home for any length of time? | O | O |

## Behaviors of People Your Age

*How wrong is it for someone your age to do each of the following things?*

| | | Not wrong at all | A little bit wrong | Wrong | Very wrong |
|---|---|---|---|---|---|
| 95. | Use marijuana | O | O | O | O |
| 96. | Use alcohol | O | O | O | O |
| 97. | Get drunk once in awhile | O | O | O | O |
| 98. | Use prescription drugs (ex: speed, downers, Valium, Ritalin, Prozac) without a prescription | O | O | O | O |
| 99. | Give or sell alcohol to a person under 21 | O | O | O | O |
| 100. | Cheat on school tests | O | O | O | O |
| 101. | Purposely damage or destroy property that does not belong to them | O | O | O | O |
| 102. | Steal something worth less than $5 | O | O | O | O |
| 103. | Hit or threaten to hit someone without any reason | O | O | O | O |
| 104. | Break into a vehicle or building to steal something | O | O | O | O |
| 105. | Steal something worth more than $50 | O | O | O | O |

## Your Friends

*Think about **your friends** when answering the following questions. Would you say these statements are mostly true or mostly false?*

| | | Mostly True | Mostly False |
|---|---|---|---|
| 106. | My friends often try to get me to do things the teacher doesn't like. | O | O |
| 107. | Most of my friends think getting good grades is important. | O | O |

*Now think about your **best friend.** Would you say these statements are mostly true or mostly false about him or her? My best friend…*

| | | Mostly True | Mostly False |
|---|---|---|---|
| 108. | Is interested in school. | O | O |
| 109. | Always attends classes. | O | O |
| 110. | Gets into trouble at school. | O | O |

*Think about your **best friend** and respond to the following situation…*

| | | Not hard at all | Not very hard | Pretty hard | Very hard | I would not want to say "no" |
|---|---|---|---|---|---|---|
| 111. | Pretend your best friend offered you a cigarette and you did not want it. How hard would it be to say "no"? | O | O | O | O | O |
| 112. | Pretend your best friend offered you a drink of beer or wine and you did not want it. How hard would it be to say "no"? | O | O | O | O | O |
| 113. | Pretend your best friend offered you some marijuana and you did not want it. How hard would it be to say "no"? | O | O | O | O | O |

*How often do you and your **best friend** talk about these topics?*

| | | Never | Infrequently | Sometimes | Often | Very Often |
|---|---|---|---|---|---|---|
| 114. | How we could get cigarettes | O | O | O | O | O |
| 115. | How to make trouble in the neighborhood | O | O | O | O | O |
| 116. | How we could get alcohol or drugs | O | O | O | O | O |

| *How much do you agree with the following statements?* | **Strongly agree** | **Agree** | **Disagree** | **Strongly disagree** |
|---|---|---|---|---|
| 117. My friends think fighting is an OK way to settle differences. | O | O | O | O |
| 118. Most people my age stay away from getting into fights. | O | O | O | O |
| 119. My friends think people who pick fights are really dumb. | O | O | O | O |

*During the past **three months**, how many of your friends would you say have…*

| | **None** | **Some** | **Most** |
|---|---|---|---|
| 120. Used marijuana? | O | O | O |
| 121. Gotten drunk once in a while? | O | O | O |
| 122. Sold or given beer or wine to another student? | O | O | O |

| *Is the following statement true or false?* | **True** | **False** |
|---|---|---|
| 123. I have been at a party where someone brought beer, wine or wine coolers to drink. | O | O |

| *Think about your friends when responding to the following questions…* | **Yes** | **No** |
|---|---|---|
| 124. If your friends got in trouble with the police, would you lie to protect them? | O | O |
| 125. If you found that your group of friends was leading you into trouble, would you still spend time with them? | O | O |
| 126. If a friend asked to copy your homework, would you let the friend copy it even though it might get you in trouble with the teacher? | O | O |
| 127. If one of your friends was smoking marijuana and offered you some, would you smoke it? | O | O |

## After School Activities

128. In a typical week this school year, did you attend after school activities with a group of youths where **adults were present**?

O Yes → Answer questions 129 to 133
O No → Skip to question **134**

129. How many days in a typical week this school year did you attend after school activities **at your school**?

O 0 Days
O 1 Day
O 2 Days
O 3 Days
O 4 Days
O 5 Days

130. How many days in a typical week this school year did you attend other after school activities **not at your school**?

O 0 Days
O 1 Day
O 2 Days
O 3 Days
O 4 Days
O 5 Days

*How much do you agree with the following statements?*

| | | **Strongly agree** | **Agree somewhat** | **Disagree somewhat** | **Strongly disagree** |
|---|---|---|---|---|---|
| 131. | I like the after school activities I go to very much. | O | O | O | O |
| 132. | The after school activities I go to are very important to me. | O | O | O | O |
| 133. | The things I learn at my after school activities are helpful in my life. | O | O | O | O |

## Your Closest Friends

*Write the first and last name of your **best friend** (the friend you believe to be closest to you) and think of him/her when answering the following 5 questions.*

**First Name** **Last Name**

| | | **Yes** | **No** |
|---|---|---|---|
| 134. | _____ | | |
| 135. | Is this friend a boy? | O | O |
| 136. | Does this friend go to your school? | O | O |
| 137. | Did you meet this friend to hang out or go somewhere during the last week? | O | O |
| 138. | Did this friend call you on the phone during the last week? | O | O |
| 139. | Does this friend get in trouble at school? | O | O |

*Write the first and last name of your **second closest friend** and answer the following 5 questions about him/her.*

| | First Name | Last Name | Yes | No |
|---|---|---|---|---|
| 140. | _____ | | | |
| 141. | Is this friend a boy? | | O | O |
| 142. | Does this friend go to your school? | | O | O |
| 143. | Did you meet this friend to hang out or go somewhere during the last week? | | O | O |
| 144. | Did this friend call you on the phone during the last week? | | O | O |
| 145. | Does this friend get in trouble at school? | | O | O |

*Write the first and last name of your **third closest friend** and answer the following 5 questions about him/her.*

| | First Name | Last Name | Yes | No |
|---|---|---|---|---|
| 146. | _____ | | | |
| 147. | Is this friend a boy? | | O | O |
| 148. | Does this friend go to your school? | | O | O |
| 149. | Did you meet this friend to hang out or go somewhere during the last week? | | O | O |
| 150. | Did this friend call you on the phone during the last week? | | O | O |
| 151. | Does this friend get in trouble at school? | | O | O |

*Write the first and last name of your **fourth closest friend** and answer the following 5 questions about him/her.*

| | First Name | Last Name | Yes | No |
|---|---|---|---|---|
| 152. | _____ | | | |
| 153. | Is this friend a boy? | | O | O |
| 154. | Does this friend go to your school? | | O | O |
| 155. | Did you meet this friend to hang out or go somewhere during the last week? | | O | O |
| 156. | Did this friend call you on the phone during the last week? | | O | O |
| 157. | Does this friend get in trouble at school? | | O | O |

*Write the first and last name of your **fifth closest friend** and answer the following 5 questions about him/her.*

| | First Name | Last Name | Yes | No |
|---|---|---|---|---|
| 158. | _____ | | | |
| 159. | Is this friend a boy? | | O | O |
| 160. | Does this friend go to your school? | | O | O |
| 161. | Did you meet this friend to hang out or go somewhere during the last week? | | O | O |
| 162. | Did this friend call you on the phone during the last week? | | O | O |
| 163. | Does this friend get in trouble at school? | | O | O |

## Your Decisions

| *Think about the decisions you make every day…* | | Never | Sometimes but not often | Often | All the time |
|---|---|---|---|---|---|
| 164. | How often do you stop to think about your options before you make a decision? | O | O | O | O |
| 165. | How often do you stop to think about how your decisions may affect others' feelings? | O | O | O | O |
| 166. | How often do you stop and think about all of the things that may happen as a result of your decisions? | O | O | O | O |
| 167. | I make good decisions. | O | O | O | O |

# Thank you for your help!

Appendix D.2. Teacher Rating Form

Dear Teacher:

Attached are rating forms for your students who are participating in a research study being conducted by the University of Maryland about after-school programs. The form asks you to rate each student's conduct, academic competence, and social skills. It will take up to ten minutes to rate each student and assuming you complete this work outside of school hours, you will be paid $5 for every complete rating form. You are being asked to rate only those students whose parents have given their consent for a teacher to provide these ratings. Please read and sign the enclosed consent form which explains the study and procedures.

To protect the identity of each student, remove and discard the label attached to each rating form containing the child's name after you have completed the form. Once you have completed all of your ratings, please fill out the attached invoice for your services by indicating the number of completed ratings and the address to which your payment should be sent. Remember to sign the invoice.

Place the completed rating forms, the signed informed consent form and the invoice in the envelope provided and mail it to Dr. Gordon Bonham.

Thank you for participating in this study. If you have questions about the study, please call me at 301-405-4717.

Sincerely,

Denise C. Gottfredson
Professor

**For each answer, using pen please fill in marks like this: ● not like this: ⊗ ⊘ ✓**

**Please Affix Label**

Please indicate how much each statement describes the usual behavior of this student in the last month:

| | | Never/Almost Never | Sometimes | Often | Very Often |
|---|---|---|---|---|---|
| 1. | Acts without thinking | O | O | O | O |
| 2. | Acts in ways that annoy or bother others | O | O | O | O |
| 3. | Articulates different ways to solve a problem | O | O | O | O |
| 4. | Asks an adult for help or advice about ways to resolve difficult situations | O | O | O | O |
| 5. | Expresses concern for others | O | O | O | O |
| 6. | Gossips or spreads rumors | O | O | O | O |
| 7. | Helps others | O | O | O | O |
| 8. | Hits, kicks at, or jumps on other children | O | O | O | O |
| 9. | If provoked by peers, shows self-control | O | O | O | O |
| 10. | If upset, responds with verbal aggression (swearing, calling names) | O | O | O | O |
| 11. | If angered, expresses anger without being aggressive or destructive | O | O | O | O |
| 12. | Is impulsive in interacting with peers | O | O | O | O |
| 13. | Is able to see things from other children's perspectives | O | O | O | O |
| 14. | Is teased, hit, or bullied by other kids | O | O | O | O |
| 15. | Lets others know how she/he feels about situations | O | O | O | O |
| 16. | Removes him or herself from potential problem situations | O | O | O | O |
| 17. | Resists peer pressure when appropriate | O | O | O | O |

| | | Never/Almost Never | Sometimes | Often | Very Often |

| | | | | | |
|---|---|---|---|---|---|
| 18. | Responds with physical aggression to problems with peers | O | O | O | O |
| 19. | Shows defiance in interactions with parents or other adults | O | O | O | O |
| 20. | Shows respect for others | O | O | O | O |
| 21. | Solves problems with peers through compromise or discussion | O | O | O | O |
| 22. | Takes time to calm down when dealing with problem situations | O | O | O | O |
| 23. | Takes other people's feelings into account before acting | O | O | O | O |
| 24. | Takes or steals things that belong to others | O | O | O | O |
| 25. | Takes responsibility for own actions (for example, apologizes) | O | O | O | O |
| 26. | Teases, insults, provokes or threatens others | O | O | O | O |
| 27. | Tells lies or cheats | O | O | O | O |
| 28. | Tries a new approach to a problem when first approach is not working | O | O | O | O |
| 29. | Understands the likely consequences of his or her own actions | O | O | O | O |

How true is each statement of the student in the last month?

| | | **Mostly False** | **Somewhat False** | **Somewhat True** | **Mostly True** |
|---|---|---|---|---|---|
| 30. | Child works earnestly, doesn't take it lightly. | O | O | O | O |
| 31. | Child works for the pleasure it gives him or her. | O | O | O | O |
| 32. | Child carries out requests responsibly. | O | O | O | O |
| 33. | Child works hard even when no reward is available. | O | O | O | O |
| 34. | Child is a self-starter. | O | O | O | O |
| 35. | Child sticks with a goal or task until it is complete. | O | O | O | O |

**Mostly False** **Somewhat False** **Somewhat True** **Mostly True**

| | | | | | |
|---|---|---|---|---|---|
| 36. | Child finishes school work quickly. | O | O | O | O |
| 37. | Child is confident. | O | O | O | O |
| 38. | Child is easily discouraged. | O | O | O | O |
| 39. | Child is sure things will work out well when she/he has a problem at school. | O | O | O | O |
| 40. | Child is sure things will work out well when she/he has new work to do at school. | O | O | O | O |
| 41. | Child expects to succeed at most things. | O | O | O | O |

42. Compared with other children in my classroom, the **overall academic performance** of this child is:

O Very Poor (Lowest 10%)

O Poor (Next Lowest 20%)

O Average (Middle 40%)

O Good (Next Highest 20%)

O Excellent (Highest 10%)

O Inadequate Opportunity to Observe

**Thank you for your time!**

Appendix Table D.3. Survey Scale Content and Reliabilities

| Scale | *N* Items | Items Included | α pretest | α posttest |
|---|---|---|---|---|
| **Unsupervised Socializing** | | | | |
| Days with friends and no adults | 1 | YS18 | -- | -- |
| **Positive Peer Influence** | | | | |
| Positive peer influence | 12 | YS106R, YS107, YS108, YS109, YS110R, Y114RD, YS115RD, YS116RD, YS120D, YS121D, YS122D, YS123R | .75 | .77 |
| **School Bonding** | | | | |
| Attachment to school | 3 | YS8, YS9R, YS10 | .68 | .73 |
| **Social Competence** | | | | |
| Goal setting | 6 | YS21, YS22, YS23, YS24, YS25, YS26 | .70 | .76 |
| Decision-making skills | 4 | YS164, YS165, YS166, YS167 | .83 | .86 |
| Impulsiveness | 10 | YS85, YS86, YS87R, YS88, YS89, YS90, YS91, YS92, YS93, YS94 | .75 | .77 |
| **Prosocial/Antidrug Attitudes** | | | | |
| Attitudes unfavorable to drug use | 10 | YS46, YS47, YS48, YS49, YS95D, YS96S, YS97D, YS98D, YS99D, YS127R | .81 | .86 |
| Belief in conventional rules | 14 | YS50, YS51, YS52R, YS53R, YS54R YS100D, YS101D, YS102D, YS103D, YS104D, YS105D, YS124R, YS125R, YS126R | .80 | .81 |
| **School Attendance** | | | | |
| % days absent | 1 | SR | -- | -- |
| **Academics** | | | | |
| Academic competence | 1 | TR42 | -- | -- |
| GPA | 1 | SR | -- | -- |
| MSA reading | 1 | SR | -- | -- |
| MSA math | 1 | SR | -- | -- |
| **Conduct Problems** | | | | |
| Disruptive classroom behavior | 3 | YS43, YS44, YS45 | .64 | .64 |
| Aggression | 6 | YS65, YS66, YS67, YS68, YS69, YS70 | .83 | .86 |
| Delinquent behavior | 7 | YS71, YS72, YS73, YS74, YS75, YS76, YS77 | .72 | .81 |
| Victimization | 7 | YS78, YS79, YS80, YS81, YS82, YS83, YS84 | .74 | .74 |
| Last month drug use | 3 | YS41D, YS42D, YS43D | .69 | .74 |
| Number suspensions | 1 | SR | -- | -- |
| Social competency (TR) | 29 | TR1R, TR2R, TR3, TR4, TR5, TR6R, TR7, TR8R, TR9, TR10R, TR11, TR12R, TR13, TR14R, TR15, TR16, TR17, TR18R, TR19R, TR20, TR21, TR22, TR23, TR24R, TR25, TR26R, TR27R, TR28, TR29 | -- | .96 |

*Notes. N*s for scales range from 407 to 416.

*Abbreviations*. YS = youth survey posttest; SR = school records; TR = teacher ratings; R = reverse code; D = dichotomized.

Appendix Table D.4. Percent Missing, Pretest and Posttest Outcome Scales

| Scale | Pretest % Missing | Posttest % Missing |
|---|---|---|
| **Composite Scales** | | |
| Social competence | .00 | .00 |
| Prosocial/antidrug attitudes | .00 | .00 |
| Academic performance[a] | 10.10[c] | .00 |
| Conduct problems[b] | .00 | .00 |
| **Unsupervised Socializing** | | |
| Days with friends and no adults | 2.40 | 13.46[c] |
| **Positive Peer Influence** | | |
| Positive peer influence | .72 | .24 |
| **School Bonding** | | |
| Attachment to school | .48 | .48 |
| **Social Competence** | | |
| Goal setting | .48 | .24 |
| Decision-making skills | 4.33[c] | 8.41[c] |
| Impulsiveness | 1.68 | .96 |
| **Prosocial/Antidrug Attitudes** | | |
| Attitudes unfavorable to drug use | .00 | .00 |
| Belief in conventional rules | .00 | .00 |
| **School Attendance** | | |
| % days absent (SR) | 9.62[c] | 1.44 |
| **Academic Performance** | | |
| Academic competence (TR) | — | 2.88 |
| GPA (SR) | 48.80[c] | 1.92 |
| MSA reading (SR) | 10.82[c] | 1.44 |
| MSA math (SR) | 10.82[c] | 1.44 |
| **Conduct Problems** | | |
| Disruptive classroom behavior | .72 | 1.68 |
| Aggression | .96 | .48 |
| Delinquent behavior | .72 | .96 |
| Victimization | 1.20 | .96 |
| Last month drug use | .96 | .96 |
| Number suspensions (SR) | 9.62[c] | .48 |
| Social competency (TR) | — | 2.64 |

*Notes*. Includes only cases included in outcome analysis ($N = 416$). Scales from youth survey unless otherwise noted.

*Abbreviations*. SR = school records, TR = teacher ratings.

[a]Pretest academic performance scale excludes academic competence because this scales was only measured at posttest.

[b]Pretest conduct problems scale excludes social competency because this scale was only measured at posttest.

[c]Missing data imputed for this scale.

Appendix Table D.5. Scale Descriptives

| Scale | Pretest | | | | Posttest | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | *N* | Range | Mean | SD | *N* | Range |
| **Composite Scales** | | | | | | | | |
| Social competence | .00 | .73 | 416 | –2.35–1.73 | .01 | .72 | 416 | –1.89–1.77 |
| Prosocial/antidrug attitudes | .75 | .22 | 416 | .12–1.00 | .66 | .25 | 416 | .04–1.00 |
| Academic performance[a] | .02 | .83 | 374 | –3.17–-1.95 | –.00 | .78 | 416 | –2.20–1.99 |
| Conduct problems[b] | –.02 | .63 | 416 | –.66–2.87 | –.00 | .63 | 416 | –.91–1.90 |
| **Unsupervised Socializing** | | | | | | | | |
| Days with friends and no adults | 2.27 | 2.12 | 406 | 1.00–5.00 | 2.14 | 1.98 | 360 | 1.00–5.00 |
| **Positive Peer Influence** | | | | | | | | |
| Positive peer influence | .84 | .19 | 413 | .00–1.00 | .79 | .22 | 415 | 0.00–1.00 |
| **School Bonding** | | | | | | | | |
| Attachment to school | 2.49 | .86 | 414 | .00–3.00 | 1.99 | 1.13 | 414 | .00–3.00 |
| **Social Competence** | | | | | | | | |
| Goal setting | 3.11 | .51 | 414 | 1.17–4.00 | 3.02 | .58 | 415 | 1.17–4.00 |
| Decision-making skills | 2.83 | .76 | 398 | 1.00–4.00 | 2.72 | .82 | 381 | 1.00–4.00 |
| Impulsiveness | .51 | .27 | 409 | .00–1.00 | .56 | .28 | 412 | .00–1.00 |
| **Prosocial/Antidrug Attitudes** | | | | | | | | |
| Attitudes unfavorable to drug use | .78 | .25 | 416 | .00–1.00 | .70 | .30 | 416 | .00–1.00 |
| Belief in conventional rules | .71 | .23 | 416 | .07–1.00 | .62 | .26 | 416 | .00–1.00 |
| **School Attendance** | | | | | | | | |
| % days absent (SR) | 4.20 | 3.90 | 376 | 0–34.71 | 5.14 | 4.83 | 410 | 0–29.51 |
| **Academics** | | | | | | | | |
| Academic competence (TR) | — | — | — | — | 3.20 | .94 | 404 | 1.00–5.00 |
| GPA (SR) | 2.42 | .72 | 213 | .75–4.00 | 2.51 | .73 | 408 | .50–4.00 |
| MSA reading (SR) | 389.40 | 28.42 | 371 | 314.00–465.00 | 386.68 | 27.23 | 410 | 312.00–455.00 |
| MSA math (SR) | 385.60 | 38.44 | 371 | 240.00–472.00 | 389.75 | 25.53 | 410 | 329.00–478.00 |
| **Conduct Problems** | | | | | | | | |
| Disruptive classroom behavior | 1.38 | .49 | 413 | 1.00–3.00 | 1.63 | .58 | 409 | 1.00–3.00 |
| Aggression | 1.58 | .59 | 412 | 1.00–4.00 | 1.87 | .74 | 414 | 1.00–4.00 |
| Delinquent behavior | .43 | .99 | 413 | .00–6.00 | .86 | 1.55 | 412 | .00–7.00 |
| Victimization | 1.31 | 1.64 | 411 | .00–7.00 | 1.49 | 1.75 | 412 | .00–7.00 |
| Last month drug use | .09 | .29 | 412 | .00–1.00 | .18 | .39 | 412 | .00–1.00 |
| Number suspensions (SR) | .35 | .72 | 416 | .00–4.00 | .52 | 1.01 | 414 | .00–6.00 |
| Social competency (TR) | — | — | — | — | 2.71 | .52 | 405 | 1.34–3.88 |

*Notes*. Includes only cases included in outcome analysis (*N* = 416). Scales from youth survey unless otherwise noted.
*Abbreviations*. SR = school records; TR = teacher ratings.
[a]Pretest academic performance scale excludes academic competence because this scales was only measured at posttest.
[b]Pretest conduct problems scale excludes social competency because this scale was only measured at posttest.

Appendix E: Process and ASP Activity Measures

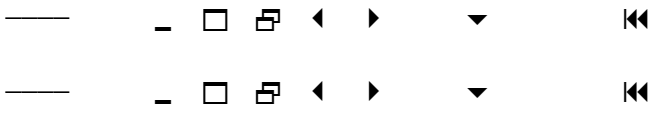
Appendix Table E.1. Management Information System Components

| Component | Description | Frequency of Data Entry |
|---|---|---|
| Staff information and credentials | Database for staff information, including log-in ID and password, name, date employed, date terminated, date of birth, gender, race, job title, highest level of education, teaching certification, years of experience providing direct services to youth and whether employed elsewhere | Data entered prior to employee's start date and modified as necessary |
| Staff attendance | Time sheet for staff attendance entry | Daily |
| Staff training | Database for recording training, including date, length of training, type of training, additional notes, and staff attendance at trainings | Entry required for each training held |
| Student information | Database for student information, including date of registration, date started program, name, date of birth, gender, 2005–2006 grade, 2006–2007 grade, race, name of registered siblings, primary caretaker, address, phone number, alternative contact address and phone number, medical information, and parent employment information | Data entered after registration and was modified as necessary |
| School attendance | Attendance sheet for marking students present, absent excused, or absent unexcused from school | Daily |
| Program attendance | Attendance sheet for marking students present, absent excused, absent unexcused, tardy excused, or tardy unexcused from the ASP | Daily |
| Withdrawal | Included date of withdrawal, reason for leaving the program, whether the parent was notified by the program of withdrawal, and whether the parent notified the ASP of student withdrawal | Required each time a student withdrew from the program or upon a long period of no attendance |
| Activity fidelity | Included group leader supervising the activity, date, the activity type, start and end time, the number of staff present, and student attendance sheet | Daily; one checklist required for each leisure or alternative academic activity held |
| All Stars fidelity | Included the lesson number and name, group leader supervising the | Daily; one checklist required for each All Stars |

| | | |
|---|---|---|
| | activity, date, start and end time, session goal achievement rating, student engagement rating, overall quality rating of the session as written and as taught, whether each activity within the lesson was taught, rating for how well each activity objective within the lesson was achieved, whether the activity was modified and in what way, and student attendance sheet | Session held |
| Award attendance incentive points | Button for awarding attendance incentives that were automatically calculated based on entered school and program attendance | Once per week |
| Redeem attendance incentive points | Points spent were recorded and an optional field for incentive description was provided | Bi-weekly following each attendance incentive ceremony |
| Director's checklist | Included questions about program implementation, including data entry of the All Stars lessons presented in that week, whether the ASP is on the one All Star lesson per week schedule and reasons for not following the schedule, whether new students were enrolled and if so whether an attendance card was signed, whether all students had signed an attendance card, whether an award ceremony was held, and whether an attendance incentive chart was clearly viable for all students to see | Weekly |

Appendix E.2. Program Observation Instrument

**Total Number of Youth Present:**



**Start Time: _____** **End Time: _____**

## Overall Program Atmosphere

**Level of Supervision:**

| | | All | Most | Some | Few or None |
|---|---|---|---|---|---|
| 1. | How many of the youth are under the direct supervision of an adult (the adult can at least see the youths) for most or all of the time? | O | O | O | O |

| | | Always | Often | Sometimes | Never |
|---|---|---|---|---|---|
| 2. | How often are staff members vigilantly attending to the behavior of youth (they can both see and hear the youths' activities and appear to be attending closely to them)? | O | O | O | O |
| 3. | How often are the youth allowed to come and go between activities without explicit permission (excluding bathroom trips)? | O | O | O | O |

| | | Yes | No |
|---|---|---|---|
| 4.[a] | Are there ever opportunities for youths to leave the program activities and go to an unsupervised area (e.g. outside, empty classroom, etc.)? | O | O |

| | | Regularly | Sometimes | Infrequently |
|---|---|---|---|---|
| 5.[a] | If yes, how often? | O | O | O |

**Social Climate:**

| | | None | 1 or 2 | More than 2 |
|---|---|---|---|---|
| 6. | Do you see any youth that appear disconnected from the program staff? | O | O | O |

| | | Yes | No |
|---|---|---|---|
| 7. | Do you see any evidence of friction between youth and program staff? | O | O |
| 8.[b] | Do you see any evidence of friction between program staff? | O | O |

| | | Always | Often | Sometimes | Never |
|---|---|---|---|---|---|
| 9. | How often are program staff members offering support to youth? | O | O | O | O |
| 10. | How often does the program staff behave in an unprofessional manner (e.g. talking on the phone or with a friend for a personal matter, using inappropriate language, losing temper, etc.)? | O | O | O | O |

| | | All | Most | Some | None |
|---|---|---|---|---|---|
| 11. | How many of the youth do the program staff seem to know as individuals (e.g. know names and specific things about each youth)? | O | O | O | O |

**Behavior Management:**

| | | No | Infrequently | Sometimes | Frequently |
|---|---|---|---|---|---|
| 12. | Did you observe misbehavior at this program? | O | O | O | O |

| | | Yes | No | Staff not aware of Misbehavior |
|---|---|---|---|---|
| 13.[b] | If you did observe misbehavior, was the program staff response to it appropriate? | O | O | O |

Example of misbehavior – anything requiring a disciplinary action by the teacher i.e. cursing, name calling, bullying, breaking or misusing program supplies, hitting, directly defying teachers

Examples of appropriate responses to misbehavior – Time-out, reminder of rules, apology to others involved, phone call home, suspension

**Level of Structure:**

| | | Mostly True | Mostly False |
|---|---|---|---|
| 14. | Activities seem to be planned well in advance, with very little improvisation. | O | O |
| 15. | Transitions between activities are efficient and orderly. | O | O |
| 16. | Opening procedures are efficient and orderly. | O | O |
| 17. | Snack time is efficient and orderly. | O | O |
| 18. | Closing procedures are efficient and orderly. | O | O |
| 19. | I observe very little dead time. | O | O |

| | | Skillfully | Adequate | Not Skillfully |
|---|---|---|---|---|
| 20. | Program content is delivered skillfully. | O | O | O |

| | | Mostly True | Mostly False | Unclear | No Defined Schedule |
|---|---|---|---|---|---|
| 21. | Program activities begin and end at scheduled times. | O | O | O | O |

| | | Yes | No |
|---|---|---|---|
| 22. | The program began at the scheduled time. | O | O |

[a] Items combined into a single item for scale construction.
[b] Items excluded from scale due to lack of variability or irrelevance.

Appendix E.3. Student Engagement Observation

Observer Name:_____ School Name:_____ Group Leader's Name: _____

Date:_____ Number of Students in Room: _____ Activity: _____

Start Time: _____

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Observation Time (5 minute Intervals)** | | | | | | | | | | | | | | | | |
| **Structure Rating (1-5)** | | | | | | | | | | | | | | | | **Total** |
| **Engaged** | | | | | | | | | | | | | | | | (1) |
| **Non-Engaged Time:** | | | | | | | | | | | | | | | | |
| **Socializing** | | | | | | | | | | | | | | | | (2) |
| **Unoccupied** | | | | | | | | | | | | | | | | (3) |
| **Cannot Tell** | | | | | | | | | | | | | | | | (4) |

Number of Students Engaged: Enter the total of #1:----------→ (5) _____

Number of Student Observations: Enter the total of #1 through #4: ----------→ (6) _____

Total Engagement Rate (from Total Column) = Box# 5 _____ divided by Box #6 _____ = _____ (7)

Comments: _____

_____

_____

Level of Structure

1 – There are no expectations for how youth spend their time in the activity

3 – There are some expectations for how youth spend their time in the activity but these expectations are ambiguous or not communicated well

5 – Expectations for how youth should be spending their time in the activity are clearly defined and all youth know what is expected of them at all times

Appendix E.4. All Stars Fidelity Observation

# Core Lesson 1
# Program Orientation

1. How many students participated in the All Stars class? _____

2. Group Leader's Name: _____

| | Not at All | Not Very | Mostly | Very |
|---|---|---|---|---|
| 3. How engaged were students during this session? | O | O | O | O |

Rate the following on a scale from 1 – 5 (1 indicating the lowest quality; 5 indicating the highest quality)

4. Rate the overall quality of this session as taught. ☐ ☐ ◂ ▸

5. Comments _____

_____

_____

_____

# Activity 1
# Setting the Stage

1. Was the activity taught?

O Yes
O No

| | Not at All | Not Very | Mostly | Very |
|---|---|---|---|---|
| 2. Objective: Welcome students and set a high standard for participation and interaction | | | | |
| How well was this objective achieved? | O | O | O | O |
| 3. Objective: Learn all students' names | | | | |
| How well was this objective achieved? | O | O | O | O |

## Activity 2
## Program Description

1. Was the activity taught?

O Yes
O No

| | Not at All | Not Very | Mostly | Very |
|---|---|---|---|---|
| 2. Objective: Encourage students to attend and actively participate in All Stars' sessions | | | | |
| How well was this objective achieved? | O | O | O | O |

## Activity 3
## Standards for Getting Along

1. Was the activity taught?

O Yes
O No

| | Not at All | Not Very | Mostly | Very |
|---|---|---|---|---|
| 2. Objective: Establish standards for getting along that encourage participation as well as respect | | | | |
| How well was this objective achieved? | O | O | O | O |

## Activity 4
## Surveys

1. Was the activity taught?

O Yes
O No

| | | Not at All | Not Very | Mostly | Very |
|---|---|---|---|---|---|
| 2. | Objective: Identify peer opinion leaders and social isolates and collect opinions about risky behaviors | | | | |
| | How well was this objective achieved? | O | O | O | O |

## Activity 5
## Closure

1. Was the activity taught?

O Yes
O No

| | | Not at All | Not Very | Mostly | Very |
|---|---|---|---|---|---|
| 2. | Objective: Answer students' questions about All Stars | | | | |
| | How well was this objective achieved? | O | O | O | O |

*Note*. Form is an example, observations are specific to the lesson taught.

Appendix E.5. Activity Checklist from the Youth Experiences Survey

**Directions:** Look at the following lists of activities. Fill in the circle for any activity that you participate in during the after school hours in a **typical week** during **this school year**. The after-school hours are **Monday to Friday** between **3:00 pm and 6:00 pm**. Then, on the line next to that activity, write down the number of days per week that you participate in the activity. When you are finished, circle the activity that you spend the most time participating in during the week.

**Community or School Organizations and Clubs:**

| Activity | Number of Days Per Week |
|---|---|
| O After-School Program At Your School | _____ |
| O Boys/Girls Club | _____ |
| O YMCA/YWCA | _____ |
| O Scouts | _____ |
| O 4-H Club | _____ |
| O Teen Court | _____ |
| O Other Clubs or Other After-School Programs | _____ |

**Academic and Leadership Activities:**

| Activity: | Number of Days Per Week |
|---|---|
| O Student Government | _____ |
| O Yearbook | _____ |
| O Newspaper | _____ |
| O Honor Societies | _____ |
| O Chess Club | _____ |
| O Language Club | _____ |
| O Computer Club | _____ |
| O History Club | _____ |
| O Other Academic Club | _____ |

**Performance and Fine Arts Activities:**

| Activity: | Number of Days Per Week |
|---|---|
| O Dance | _____ |
| O Band/Music Lessons | _____ |
| O Art Club | _____ |
| O Chorus | _____ |
| O Drama | _____ |
| O Other Clubs | _____ |

**Faith-Based and Service Activities:**

| Activity: | Number of Days Per Week |
|---|---|
| O Youth Groups | _____ |
| O Community Service | _____ |
| O Tutoring | _____ |
| O Key Club | _____ |
| O SADD | _____ |
| O Other Clubs | _____ |

**Sports:**

| Activity: | Number of Days Per Week |
|---|---|
| O Basketball | _____ |
| O Football | _____ |
| O Baseball | _____ |
| O Swimming | _____ |
| O Soccer | _____ |
| O Cheerleading | _____ |
| O Softball | _____ |
| O Track | _____ |
| O Wrestling | _____ |
| O Volleyball | _____ |
| O Gymnastics | _____ |
| O Golf | _____ |
| O Exercise | _____ |
| O Hockey | _____ |
| O Weight Lifting | _____ |
| O Bike Riding | _____ |
| O Aerobics | _____ |
| O Horse-back Riding | _____ |
| O Other Sports | _____ |

If you are not involved in any of these activities please raise your hand.

Appendix F: Comparison with National Samples

Appendix Table F.1. Percent of Students Reporting After-School Activity Participation, BCPS and National Samples of 8th Graders

| **Type of Activity** | **BCPS** | **MTF** | **NHES** |
|---|---|---|---|
| Community/School-based | 25.9 | 70.0 | 16.5 |
| Academic | 28.2 | 51.2 | 21.0 |
| Performance/Fine Arts | 41.2 | 49.6 | 42.0 |
| Faith-based/Service | 33.7 | — | 41.5 |
| Sports | 75.3 | 86.2 | 74.0 |

*Abbreviations.* BCPS = Baltimore County Public Schools; MTF = Monitoring the Future; NHES = National Household Education Surveys program.

# Appendix G: Detailed Outcome Analysis

Appendix Table G.1. Adjusted Posttest Means and Effect Sizes, All Outcomes

| Scale | Adjusted Posttest Mean: Control (*n*) | | Adjusted Posttest Mean: Treatment (*n*) | | *p*-level | *d* |
|---|---|---|---|---|---|---|
| **Composite Scales** | | | | | | |
| social competence | – .01 | (194) | – .01 | (202) | .92 | .01 |
| prosocial/antidrug attitudes | .66 | (193) | .66 | (201) | .74 | .03 |
| academic performance | .01 | (205) | – .02 | (211) | .63 | – .03 |
| conduct problems | – .01 | (205) | .00 | (211) | .84 | .01 |
| **Unsupervised Socializing** | | | | | | |
| days with friends and no adults | 2.40 | (205) | 1.89 | (211) | .01* | – .26 |
| **Positive Peer Influence** | | | | | | |
| positive peer influence | .79 | (202) | .78 | (210) | .54 | – .05 |
| **School Bonding** | | | | | | |
| attachment to school | 1.99 | (202) | 2.00 | (210) | .85 | .01 |
| **Social Competence** | | | | | | |
| goal setting | 3.06 | (203) | 2.98 | (210) | .12 | – .14 |
| decision-making skills[b] | 2.72 | (205) | 2.70 | (211) | .77 | – .03 |
| impulsiveness | .56 | (199) | .55 | (206) | .92 | – .01 |
| **Prosocial/Antidrug Attitudes** | | | | | | |
| attitudes unfavorable to drug use | .72 | (205) | .69 | (211) | .32 | – .09 |
| belief in conventional rules | .62 | (205) | .61 | (211) | .81 | – .02 |
| **School Attendance** | | | | | | |
| % days absent from school (SR) | .05 | (203) | .05 | (207) | .45 | – .05 |
| **Academic Performance** | | | | | | |
| academic competence[a] (TR) | 3.24 | (198) | 3.17 | (206) | .42 | – .08 |
| GPA[b] (SR) | 2.51 | (202) | 2.51 | (206) | .93 | .00 |
| MSA reading (SR) | 387.17 | (203) | 386.19 | (207) | .62 | – .04 |
| MSA math (SR) | 389.77 | (203) | 389.73 | (207) | .98 | .00 |
| **Conduct Problems** | | | | | | |
| disruptive classroom behavior | 1.66 | (200) | 1.59 | (206) | .14 | – .12 |
| aggression | 1.88 | (202) | 1.86 | (208) | .70 | – .03 |
| delinquent behavior | .82 | (202) | .90 | (210) | .61 | .05 |
| victimization | 1.42 | (199) | 1.54 | (208) | .86 | .07 |
| last month drug use | .18 | (201) | .18 | (207) | .91 | .01 |
| number suspensions (SR) | .56 | (205) | .49 | (209) | .59 | – .07 |
| social competency[a] (TR) | 2.74 | (198) | 2.68 | (207) | .18 | – .13 |

*Notes*. Scales from youth survey unless otherwise noted; SR = school records; TR = teacher ratings. Means are adjusted for pretest level of the dependent variable, school site, age, gender, and race.

[a]Pretest information not available.

[b]Treatment by site interaction significant, $p < .05$.

*$p < .05$.