

**The Study of Promising After-School Programs:
Examination of Intermediate Outcomes in Year 2**

March 2005

Deborah Lowe Vandell
Elizabeth R. Reisner
B. Bradford Brown
Kimberly Dadisman
Kim M. Pierce
Dale Lee
Ellen M. Pechman

Chapter 1: Introduction	1
Goals of the Study of Promising After-School Programs	2
Theory of Change	4
Examination of Intermediate Outcomes	6
Chapter 2: Sample Description	9
After-School Programs	9
Sample Recruitment	9
Sample Characteristics	11
Program and Comparison Groups	12
Program Dosage	15
Other After-School Experiences	15
Chapter 3: Measures of Child and Youth Functioning	18
Student Measures	18
Teacher and Program Staff Measures	21
Chapter 4: Findings: After-School “Clusters”	26
Demographic Characteristics Associated with the After-School Clusters	28
Relations Between Cluster Membership and Child Functioning at Baseline	33
Relations Between Cluster Membership and Child Functioning at the First-Year Follow Up	36
Chapter 5: Conclusions and Next Steps	45
Key Findings and Their Implications	45
Next Steps	49
References	51
Appendix A: Program vs. Comparison Analyses	55

Baseline Scores on Outcome Measures	55
Assessment of Intermediate Outcomes	55
Appendix B: Program Dosage Analyses	60

List of Tables

Table 2.1	Observer Ratings of Key Process Features in the Programs	10
Table 2.2	Demographic Characteristics of the Study Sample and Participating Schools	11
Table 2.3	Fall 2003 and Spring 2004 Program Participation	13
Table 2.4	Percentage Distribution of Program and Comparison Samples on Background Variables	14
Table 2.5	Program Dosage during the 2003-04 School Year	16
Table 2.6	Participation in After-School Activities Outside of the Selected Programs	17
Table 3.1	Student Measures of Child and Youth Functioning	19
Table 3.2	Teacher Measures of Child and Youth Functioning	22
Table 3.3	Program Staff Measures of Child and Youth Functioning	23
Table 4.1	Means (Standard Deviations) and Percentages on Clustering Variables, Elementary Clusters	29
Table 4.2	Means (Standard Deviations) and Percentages on Clustering Variables, Middle School Clusters	30
Table 4.3	Percentage Distribution of Elementary School Clusters on Background Variables	31
Table 4.4	Percentage Distribution of Middle School Clusters on Background Variables	32
Table 4.5	Adjusted Means of Student Performance at Baseline, Elementary School Clusters	34
Table 4.6	Adjusted Means of Student Performance at Baseline, Middle School Clusters	35
Table 4.7	HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Elementary School Sample	38
Table 4.8	HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Middle School Sample	39
Table 4.9	HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Program Staff Reports	40

Table 4.10	Effect Sizes for Significant Effects in HLM Analyses	43
Table A-1	Baseline Scores for Program and Comparison Students	57
Table A-2	HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Program versus Comparison: Elementary	58
Table A-3	HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Program versus Comparison: Middle School	59
Table B-1	HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Program Dosage: Elementary	62
Table B-2	HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Program Dosage: Middle School	63
Table B-3	HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Program Dosage: Program Staff Reports	64

List of Figures

Figure 1.1	Theory of Change for the Study of Promising After-School Programs	5
------------	---	---

Chapter 1

Introduction

Substantial numbers of youth in the United States grow up in poverty. Studies suggest that education remains a critical pathway out of poverty for these young people, but activities within the school classroom alone cannot provide youth with the educational, social, and personal resources they need to overcome economically disadvantaged backgrounds (Eccles et al., 1993). Children and youth also need varied opportunities, experiences, and supports after school. Programs that provide constructive, supervised activities during the after-school hours can help to meet this need.

Some investigations have indicated that participation in after-school programs can improve academic and behavioral outcomes (Baker & Witt, 1996; Dishion, McCord, & Poulin, 1999; Hamilton & Klein, 1999; Huang, Gribbons, Kim, Lee, & Baker, 2000; Posner & Vandell, 1994; Schinke, Cole, & Poulin, 2000; Tierney, Grossman, & Resch, 1995; University of California at Irvine & California Department of Education, 2002; Reisner, White, Russell, & Birmingham, 2004). Other researchers have reported no effects or, in some cases, negative consequences of program participation (Catalano, Berglund, Ryan, Lonczak, & Hawkins, 1998; Grossman et al., 2002; Vandell & Corasaniti, 1988; Vandell & Pierce, 1995; Welsh, Russell, Williams, Reisner, & White, 2002). Critiques of the extant research (Vandell & Posner, 1999; Vandell, Pierce, & Dadisman, in press) have identified several factors that may contribute to these differences: the quality of the programs, how frequently young people attended the programs, the family and community context in which the programs were embedded, availability of other types of after-school activities, and family and child selectivity bias.

Numerous studies have identified the program features and practices that are associated with positive youth development outcomes (American Youth Policy Forum, 1997, 1999;

Beckett, Hawken, & Jacknowitz, 2001; Catalano et al., 1998; Eccles & Gootman, 2002; MacDonald & Valdivieso, 2000; McLaughlin, 2000; Pierce, Hamm, & Vandell, 1999; Pittman, Irby, & Ferber, 2000; Rosenthal & Vandell, 1996; Scales & Leffert, 1999). In general, key features of high-quality programs for elementary children include positive relationships with staff, positive relationships with peers, diverse activities, and opportunities to exercise choice and autonomy. Key features of positive, high-quality developmental settings for adolescents include physical and psychological safety; appropriate structure; supportive relationships; opportunities to belong; positive social norms; support for efficacy and mattering; opportunities for skill building; and integration of family, school, and community efforts.

Goals of the Study of Promising After-School Programs

In much of the prior research, methodological issues such as investigation of only one or two programs, failure to address selection bias, assessment of a narrow set of outcomes, and/or failure to use longitudinal designs to track outcomes over time limit confidence about the impact of the programs, and scholars have called for more rigorous and creative research designs (American Youth Policy Forum, 1997, 1999; Catalano et al., 1998; Eccles & Gootman, 2002; Fashola, 2002; MacDonald & Valdivieso, 2000). One notable large-sample study that addressed some of these methodological issues is the national evaluation of the 21st Century Community Learning Centers (CLC) Program conducted by Mathematica Policy Research (Dynarski et al., 2003). That study featured a random assignment design to assess the impact of participation in after-school programs in the elementary grades and a quasi-experimental design to assess the impact of programs in the middle school grades.

The Study of Promising After-School Programs differs from the 21st Century CLC evaluation, which selected typical programs irrespective of quality, by concentrating on high-quality programs. We will examine longitudinally the effects of participation in high-quality

after-school programs on various outcomes among economically disadvantaged youth in both the elementary school and middle school years. The study has two major goals: (1) to identify elementary school and middle school programs that feature promising practices for enhancing students' academic and social development and their emotional and physical well-being, and (2) to test the hypothesis that economically disadvantaged youth between the ages of 8 and 14 who participate in high-quality after-school programs achieve significantly greater developmental and learning gains over a two-year period than do disadvantaged youth who do not participate in similar opportunities. We focus on **promising** after-school programs, rather than a random or representative sample of programs, in order to assess the potential for programs to exert positive effects on youth and to identify common elements that might account for their ability to foster positive youth development. Promising programs are those that offer high-quality after-school environments for youth, manifest sustainability, and exhibit characteristics believed to promote positive youth outcomes.

The need for effective programs may be especially acute for economically and socially disadvantaged youth who have fewer options for constructive activities after the regular school day ends (Kleiner, Nolin, & Chapman, 2004) and for whom unsupervised time may be particularly risky (Pettit, Laird, Bates, & Dodge, 1997). The after-school hours represent a particular challenge for low-income families because their incomes are needed for basic necessities and they cannot afford the educational enrichment and extracurricular activities that are used routinely by middle-class families to supplement the school day. For many low-income families, especially single-parent households, having children return home to a parent after school is not an option because the incomes of all available adults in the household are needed. Consequently, in the Study of Promising After-School Programs we concentrate on low-income children, families, and communities that are among those in greatest need for external supports

of all types, including after-school programming. We target school-based or school-linked after-school programs that are consistent with the program model of the 21st Century Community Learning Centers. These foci reflect the priorities of the Charles Stewart Mott Foundation and the Foundation's special interest in seeking "pathways out of poverty" for the vulnerable populations who have the greatest claim on public and philanthropic resources.

Theory of Change

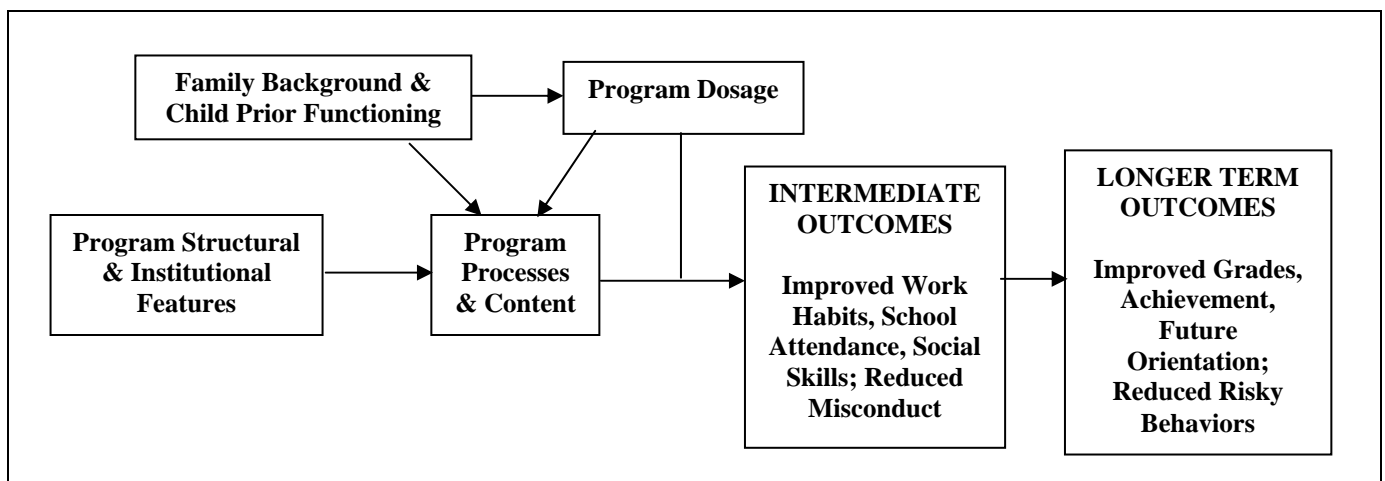
The identified features of positive after-school experiences point to both structural and organizational aspects of programs as well as the type of activities they offer to children and youth. Mindful of these findings, we derived a theory of change (see Figure 1.1) that includes the following propositions:

- To achieve positive life outcomes, children and youth require opportunities and supports in multiple developmental domains, including academic, social, psychological, and behavioral areas.
- High-quality after-school programs can stimulate positive experiences and outcomes for children and youth in these core developmental areas, employing varied program content foci (e.g., arts, academics, sports).
- Certain structural and institutional features support the implementation of high-quality programs.
- Children and youth who participate in high-quality programs more often, thus receiving a higher "dosage" of programming, will derive greater benefits than children and youth who participate less often.
- Because pre-existing family and child differences may affect the participation of children and youth in voluntary after-school programs, the examination of program effects must take these differences into account.

- Effects of high-quality programs are expected to “build” over time, with “intermediate” outcomes providing skills that contribute to “longer term” outcomes.

As shown in Figure 1.1, the study’s theory of change cites both structural and institutional features as well as process and content features that characterize effective programs for children and youth. **Structural and institutional** features are those elements of a program that establish the setting and context for positive relationships and high-quality activities. These features include staff qualifications and support, program size and group configuration, financial and physical resources, external affiliations, and sustainability efforts. **Process and content** features are those practices that participating children and youth experience directly. They include adults’ interactions and relationships with participants, relationships among participants, program content and activities, and content delivery strategies. These features of after-school programs are discussed in greater detail in our report on the first year of the study (Vandell et al., 2004).

Figure 1.1
Theory of Change for the Study of Promising After-School Programs



The elements of our theoretical framework are reflected in the study's design and implementation, including:

- The criteria for screening nominated programs for potential selection into the study.
- The sampling plan that guides recruitment of study participants into the treatment and comparison groups.
- The domains to be measured and the types of instruments used to gauge the implementation and outcomes of targeted after-school programs.
- The analysis plan, which will use research data to trace the unfolding operation of the elements that make up the change theory.

The Study of Promising After-School Programs has been designed with the understanding that effective after-school programs must incorporate certain process and content features and also certain structural and institutional features in order to achieve positive effects for at-risk children and youth. The study is *not* designed to examine the effects of variations in program quality, nor to ascertain the effects of programs that are strong in some areas and not others. Rather, the study is examining the impact of sustained participation in high-quality programs on the academic, social, and behavioral development of participants.

Examination of Intermediate Outcomes

This report presents information from the second year of the Study of Promising After-School Programs, during which we collected data from students, parents, teacher, and program staff and examined the effects of participation in the selected after-school programs on intermediate child and youth outcomes. According to our theory of change, these effects should be stepwise and cumulative. They first should be manifest in improved work habits and social skills, reduced misconduct, and improved school attendance, outcomes selected because they contribute to long-term school success. We expected that the effects on the intermediate

outcomes of interest would be stronger for those whose participation in the programs was regular and extensive. We also expected that effects would be evident first at the after-school programs (as reported by program staff) and then at school and at home. To test the model properly, we also conducted preliminary analyses of long-term effects including improved academic performance (grades) and reduced risky behaviors. We expected that these effects would not be as strong as the intermediate effects after the first year of data collection.

Children's after-school options are not randomly assigned to them but are influenced by varying child, family, and neighborhood factors. Moreover, these contexts exert multiple influences on child and youth development that are confounded with participation in after-school programs. Our theory of change recognizes that multiple factors influence (a) whether children and youth participate in after-school programs, and (b) whether those who participate benefit from the experience. To address the fact that initial differences between participants and nonparticipants are likely to contribute to differential decisions regarding whether to attend after-school programs, we measured and controlled for child and family characteristics suggested by prior research as potential sources of selection bias, including:

- Child prior behavioral adjustment
- Child gender
- Child ethnicity
- Parent/guardian employment
- Parent/guardian education
- Family structure
- Family income

In Chapter 2 of this report, we describe the study sample. Following a description of our research methods and measures (Chapter 3), we present our findings on the effects of

participation in the selected after-school programs (Chapter 4). Chapter 5 presents conclusions drawn thus far from this phase of our work and outlines the next year of the study.

Chapter 2

Sample Description

After-School Programs

As detailed in our report of the study's first year (Vandell et al., 2004), during Spring 2003 we identified 20 high-quality after-school programs serving elementary school students and 18 serving middle school students. Four elementary programs and five middle school programs subsequently were withdrawn from the study due to loss of funding, changes in administrative personnel, or school district concerns. During early Fall 2003, we screened and selected replacement programs, resulting in a final program sample of 19 elementary and 16 middle school programs located in 13 cities in 9 states across the United States.

During the program selection process, we observed each program for two afternoons in Fall 2002 using the *Promising Practices Rating System* (PPRS) to quantify key program processes on a 4-point scale (1 = highly uncharacteristic, 4 = highly characteristic): supportive relationships with staff and with peers, student engagement in activities, appropriate program structure, opportunities for cognitive growth, opportunities for autonomy, and mastery orientation (see www.wcer.wisc.edu/childcare/des3.html for a copy of the PPRS). We observed the programs again in Fall 2003 to confirm that they continued to meet our quality criteria, using a modified version of the PPRS that omits the rating of opportunities for autonomy and adds ratings of chaos and over-control. Table 2.1 shows the mean scores on the observed process features in the programs during Fall 2002 and Fall 2003. As shown in the table, the programs maintained their ratings on nearly all of the quality indicators from Year 1 to Year 2.

Sample Recruitment

During early Fall 2003, we recruited elementary school students in the third and fourth grades, and middle school students in the sixth and seventh grades, to participate in the study.

Table 2.1
Observer Ratings of Key Process Features in the Programs

	Elementary programs <i>M (SD)</i>			Middle school programs <i>M (SD)</i>		
	Fall 2002 <i>N</i> = 15	Fall 2003 <i>N</i> = 19	<i>t</i> ₍₃₂₎	Fall 2002 <i>N</i> = 11	Fall 2003 <i>N</i> = 16	<i>t</i> ₍₂₅₎
Supportive relations with staff	3.69 (.41)	3.63 (.50)	0.39	3.77 (.33)	3.69 (.48)	0.51
Supportive relations with peers	3.79 (.30)	3.68 (.58)	0.71	3.95 (.15)	3.63 (.62)	2.04
Student engagement	3.78 (.40)	3.68 (.48)	0.64	3.92 (.13)	3.69 (.48)	1.88
Opportunities for cognitive growth	3.16 (.46)	2.89 (.81)	1.18	3.17 (.60)	2.69 (.79)	1.70
Appropriate structure	3.59 (.60)	3.74 (.45)	0.79	3.61 (.37)	3.63 (.62)	0.05
Mastery orientation	3.20 (.68)	2.84 (.90)	1.28	3.52 (.54)	2.75 (.86)	2.64*
Opportunities for autonomy	3.13 (.82)	--	--	3.51 (.60)	--	--
Chaos	--	1.26 (.56)	--	--	1.25 (.45)	--
Over-control	--	1.26 (.56)	--	--	1.19 (.40)	--

Note. Fall 2002 means are based on ratings made in those programs that remained in the study at the start of the 2003-04 school year. Ratings in Fall 2003 include several programs that were added to the study after the Fall 2002 ratings had been completed.

* $p < .05$

Recruitment was conducted in the schools affiliated with the after-school programs selected for the study. In order to recruit both program participants and students who did not attend the programs, we employed a “blanket” recruitment strategy. We explained the study and distributed parent consent letters to all students enrolled in the targeted grade levels. In schools with 500 or more students in the targeted grades, we randomly selected five classrooms for blanket recruitment. The students returned the signed parental consent forms to their classroom teachers.

We approached 4,002 students in third and fourth grades across the 19 partner elementary schools. Signed parental consent was obtained for 1,820 students, or 45.5% of the recruited

group. Of the 2,926 students in sixth and seventh grades whom we approached in the 16 partner middle schools, we obtained parental consent (and student assent) from 1,119, or 38.2% of the possible respondents.

Sample Characteristics

In general, both our sample and the host schools can be described as low income and from a variety of ethnic backgrounds, but predominantly ethnic minorities (see Table 2.2). Compared to the host schools, both the elementary and middle school samples contain smaller proportions of boys (47% vs. 51%) and students receiving free or reduced-price lunch (90% vs. 94% elementary, 76% vs. 82% middle). The middle school sample contains a higher proportion of White (31%) and Black students (13%) than the host schools (21% and 10%, respectively), and a smaller proportion of Hispanic students (49% vs. 62%). The elementary sample and schools are similar in terms of ethnic minority populations, with both comprised of close to 90% ethnic minority children, predominantly Hispanic.

Table 2.2
Demographic Characteristics of the Study Sample and Participating Schools

	Elementary			Middle		
	Sample N = 1,820	Schools	$\chi^2_{(1)}$	Sample N = 1,119	Schools	$\chi^2_{(1)}$
Male	47%	51%	7.43**	47%	51%	7.76**
Free/reduced-price lunch	90%	94%	23.63***	76%	82%	23.00***
Asian/other	3%	4%	2.08	7%	8%	2.93
Black	8%	8%	0.08	13%	10%	15.16***
Hispanic	77%	77%	0.01	49%	62%	70.54***
White	12%	11%	1.82	31%	21%	62.09***

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

Program and Comparison Groups

A critical issue in the study of after-school programs is the conceptualization and measurement of program participation. We initially planned to define “program” and “comparison” students by program attendance at the time of recruitment, an approach taken in the 21st Century Community Learning Centers evaluation (Dynarski et al., 2003). In that study, if a student was registered for the targeted after-school program at recruitment, the student was placed in the program group. Comparison students were those who attended the same school as the program students but were not enrolled in the targeted program at recruitment. In the Study of Promising After-School Programs, however, attendance data revealed substantial mobility in and out of the selected programs, especially between fall and spring of the academic year.

Table 2.3 shows the numbers of students who did and did not attend the selected programs for 5 or more days during Fall 2003 and Spring 2004. As seen in the table, 134 of the 907 elementary students (15%) who attended the programs during Fall 2003 did not attend in Spring 2004, and 110 of the 913 elementary students (12%) who did not attend the programs during Fall 2003 subsequently attended in Spring 2004. Similar discrepancies appeared in the middle school sample: 75 of 453 students (17%) who attended the programs in Fall 2003 did not attend in Spring 2004, and 87 of 666 students (13%) who did not attend in Fall 2003 subsequently attended in Spring 2004. In some cases, across the academic year as a whole, a participant initially placed in the comparison group actually attended the targeted after-school program more frequently than several members of the program group. In view of this, we altered the criteria for group assignment. Participants who attended the targeted program 5 days or more during at least one semester (fall or spring) were placed in the program group; all others were placed in the comparison group. Using this strategy, 1,017 elementary school students and 540 middle school students were designated as program participants.

Table 2.3
Fall 2003 and Spring 2004 Program Participation

		Elementary school students	
		Spring 2004	
		Attended (<i>n</i> = 883)	Did not attend (<i>n</i> = 937)
Fall 2003	Attended (<i>n</i> = 907)	773 (43%)	134 (7%)
	Did not attend (<i>n</i> = 913)	110 (6%)	803 (44%)
		Middle school students	
		Spring 2004	
		Attended (<i>n</i> = 465)	Did not attend (<i>n</i> = 654)
Fall 2003	Attended (<i>n</i> = 453)	378 (34%)	75 (6%)
	Did not attend (<i>n</i> = 666)	87 (8%)	579 (52%)

Note. Students were included in the Attended group if they participated in the programs 5 or more days during the applicable semester.

Table 2.4 shows the demographic characteristics of the program and comparison groups at the elementary and middle school levels. For **elementary** students, the groups did not differ on gender, father's employment, parental education, and family income. Over half of each group had annual family incomes under \$20,000, and more than a third of the youths' mothers did not graduate from high school. Most of the elementary school participants are Hispanic, but there are modest ethnic differences: The program group was more likely to be Black or White and less likely to be Hispanic than the comparison group. The program group also was more likely to reside in single-parent homes (35% vs. 26%), with mothers who worked full time (49% vs. 37%). Nearly all elementary respondents qualified for free or reduced-price lunch in the 2003-04 school year, but the proportion is significantly higher among comparison (94%) than program youth (87%).

Table 2.4
Percentage Distribution of Program and Comparison Samples on Background Variables

	Elementary		Middle	
	Program N = 1017	Comparison N = 803	Program N = 540	Comparison N = 579
Gender	$\chi^2_{(1)} = 2.73$		$\chi^2_{(1)} = 11.68^{***}$	
Male	45	49	52	42
Ethnicity	$\chi^2_{(3)} = 43.52^{***}$		$\chi^2_{(3)} = 7.88^*$	
Asian/other	3	4	5	7
Black	10	6	15	12
Hispanic	72	83	52	46
White	15	7	28	34
Free/reduced-price lunch	$\chi^2_{(1)} = 25.78^{***}$		$\chi^2_{(1)} = 0.18$	
Yes	87	94	76	75
Family structure	$\chi^2_{(2)} = 13.97^{***}$		$\chi^2_{(2)} = 14.78^{***}$	
Two parents	59	69	60	71
Single parent	35	26	32	25
Alternate caregiver	6	5	8	4
Maternal employment	$\chi^2_{(2)} = 15.12^{***}$		$\chi^2_{(2)} = 7.30^*$	
Full time	49	37	53	45
Part time	20	20	18	16
Not employed	31	43	29	38
Paternal employment	$\chi^2_{(2)} = 1.63$		$\chi^2_{(2)} = 0.15$	
Full time	72	70	75	75
Part time	12	15	11	11
Not employed	16	15	14	14
Maternal education	$\chi^2_{(3)} = 2.82$		$\chi^2_{(3)} = 7.45$	
Did not graduate high school	34	37	27	35
High school diploma/GED	26	27	25	24
Some college	30	28	33	30
4-year college degree	10	8	15	11
Paternal education	$\chi^2_{(3)} = 5.56$		$\chi^2_{(3)} = 4.61$	
Did not graduate high school	39	41	27	33
High school diploma/GED	30	27	26	27
Some college	20	24	29	25
4-year college degree	11	8	18	15
Family income	$\chi^2_{(3)} = 2.82$		$\chi^2_{(3)} = 2.72$	
Less than \$11,000	29	32	23	21
\$11,000-19,999	25	24	25	22
\$20,000-29,999	22	21	17	21
\$30,000 or more	24	23	35	36

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

Within the **middle school** sample, the program and comparison students did not differ in terms of receipt of free or reduced-price lunch, father's employment, parental education, and family income. However, the program group was relatively more likely to be male (52% vs. 42%), to live in single-parent homes (32% vs. 25%), and to have mothers who were employed full time (53% vs. 45%). Hispanic youth comprise only about half the middle school sample (compared to nearly 80% of the elementary sample). The program group contains a slightly smaller percentage of White students than found in the comparison group (28% vs. 34%).

Program Dosage

A key element of our theory of change is that the degree to which after-school programs affect children and adolescents should be related to the frequency with which they attend the programs. Therefore, we also measured program dosage by collecting attendance information from staff at the selected programs on a monthly basis. In most cases, the programs provided raw daily attendance data and we extracted monthly totals for the study participants. From these data, we created a program dosage variable based on the number of days respondents attended the programs across the academic year. We stipulated five dosage levels: minimal (0-4 days of attendance), low (5-29 days), moderate (30-59 days), substantial (60-89 days), and high (90 or more days). Students originally identified as comparison students as well as youth originally identified as program participants were included in the analysis. As shown in Table 2.5, the sample distribution on the program dosage variable was skewed among both elementary and middle school students. Most program respondents experienced minimal to moderate program dosage (65% elementary and 82% middle school students); few experienced high levels of program dosage (28% and 12%, respectively).

Other After-School Experiences

In addition to collecting program attendance data, we obtained information from the

Table 2.5
Program Dosage during the 2003-04 School Year

	Elementary <i>N</i>	Middle <i>N</i>
Minimal, 0-4 days	800 (44%)	577 (52%)
Low, 5-29 days	213 (12%)	201 (18%)
Moderate, 30-59 days	172 (9%)	129 (11%)
Substantial, 60-89 days	120 (7%)	82 (7%)
High, 90+ days	515 (28%)	130 (12%)

study participants about their involvement in other types of experiences after school. In both Fall 2003 and Spring 2004, students reported how many days each week they participated in each of 10 types of structured activities, as well as the number of days they were without adult supervision in the after-school hours—at home alone or with siblings, or “hanging out” with peers. We calculated annual mean scores from the fall and spring reports of time spent in several types of structured activities and two unsupervised situations (home alone or with sibling, hanging out with peers), utilizing a 4-point scale (1 = not at all/once or twice, 2 = about once a week, 3 = 2-3 days a week, 4 = 4 or more days a week).

Table 2.6 shows the average amount of time across the school year that the program and comparison groups reported they participated in each activity or circumstance. Among both elementary and middle school respondents, with the exception of coached sports among the younger participants, program youth participated in other types of structured, supervised activities to a greater extent than comparison youth. There were no significant differences in the time that these two groups reported in unsupervised settings (home alone/with siblings or hanging out with peers). The rates of participation in supervised enrichment activities in the program group are higher than those reported in national surveys of low-income children and

youth (Tout, Scarpa, & Zaslow, 2002; Kleiner et al., 2004). These differences may suggest that the students in this study reported activities that occurred at their after-school programs, as additional enrichment activities not associated with the promising programs.

Table 2.6
Participation in After-School Activities Outside of the Selected Programs

	Elementary			Middle		
	Program <i>N</i> = 1017 <i>M</i> (<i>SE</i>)	Comparison <i>N</i> = 803 <i>M</i> (<i>SE</i>)	<i>F</i>	Program <i>N</i> = 540 <i>M</i> (<i>SE</i>)	Comparison <i>N</i> = 579 <i>M</i> (<i>SE</i>)	<i>F</i>
Organized activities ¹	1.85 (.02)	1.59 (.03)	$F_{(1,1281)} = 50.13^{***}$	1.89 (.03)	1.59 (.03)	$F_{(1,830)} = 48.77^{***}$
Coached sports	1.74 (.04)	1.69 (.04)	$F_{(1,1281)} = 0.79$	1.95 (.05)	1.71 (.05)	$F_{(1,829)} = 11.25^{***}$
School-based activities	1.64 (.03)	1.28 (.03)	$F_{(1,1280)} = 59.82^{***}$	1.89 (.04)	1.46 (.04)	$F_{(1,828)} = 49.25^{***}$
Lessons	2.19 (.04)	1.78 (.04)	$F_{(1,1278)} = 48.05^{***}$	2.15 (.05)	1.76 (.05)	$F_{(1,825)} = 31.97^{***}$
Other programs/clubs	1.83 (.04)	1.61 (.04)	$F_{(1,1281)} = 14.75^{***}$	1.58 (.04)	1.44 (.04)	$F_{(1,830)} = 4.77^*$
Home unsupervised ²	1.35 (.02)	1.39 (.03)	$F_{(1,1281)} = 1.13$	1.64 (.04)	1.66 (.04)	$F_{(1,830)} = 0.13$
Home alone	1.40 (.03)	1.44 (.03)	$F_{(1,1279)} = 0.49$	1.80 (.05)	1.83 (.05)	$F_{(1,830)} = 0.21$
Home with siblings	1.30 (.02)	1.33 (.03)	$F_{(1,1280)} = 0.92$	1.47 (.04)	1.48 (.04)	$F_{(1,829)} = 0.03$
With peers unsupervised	1.46 (.02)	1.46 (.03)	$F_{(1,1281)} = 0.01$	1.81 (.05)	1.85 (.05)	$F_{(1,830)} = 0.32$

Note. Scores are mean annual participation rates (1 = not at all/once or twice, 2 = about once a week, 3 = 2-3 days a week, 4 = 4 or more days a week) and are adjusted for child gender and ethnicity, family structure, maternal education and employment, and family income.

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

¹ Mean of coached sports, school-based activities, lessons, other programs or clubs.

² Mean of home alone and home with siblings.

Chapter 3

Measures of Child and Youth Functioning

Baseline data on child and youth functioning were collected during Fall 2003 from students, parents, teachers, and program staff. During late Spring 2004, first-year follow-up data were collected from students, teachers, and program staff. The surveys we administered can be obtained at www.wcer.wisc.edu/childcare/des3.html.

During the 2003-04 school year, we measured both intermediate and longer term outcomes that we hypothesized would be affected by participation in high-quality after-school programs, as outlined in the study's theory of change. We also collected school attendance data from students' school records. However, due to delays at some schools in providing this information, we were not able to examine this intermediate outcome in this report. We will report school attendance as both an intermediate and longer term outcome in the next year of the study. Parent reports of child and youth functioning were collected at baseline and will be collected again at the second-year follow up. Because the parent measures cannot be analyzed in terms of intermediate outcomes, we do not provide information about them in this report.

Student Measures

Students completed baseline and first-year follow-up measures of several intermediate and longer term outcomes addressed in this report, including work habits, self-efficacy, misconduct, and substance use (see Table 3.1 for psychometric information, including mean item scores, standard deviations, score ranges, and Cronbach's coefficient alphas). The surveys were administered to small groups of students, either during the school day or at the after-school program. The administration was conducted in Spanish for those students who were not able to complete the surveys in English. As is common in low-income samples, some students moved and transferred out of the participating schools between the time of recruitment and survey

Table 3.1
Student Measures of Child and Youth Functioning

Outcome	Response scale	Elementary			Middle		
		<i>M (SD)</i>	Range	Alpha	<i>M (SD)</i>	Range	Alpha
Work habits	1-4						
Fall 2003		3.42 (0.56)	1.17-4	.72	3.21 (0.58)	1-4	.78
Spring 2004		3.30 (0.61)	1-4	.76	3.13 (0.58)	1-4	.79
Self-efficacy	1-4						
Fall 2003		--	--	--	3.16 (0.51)	1.43-4	.63
Spring 2004		--	--	--	3.17 (0.49)	1.57-4	.66
Misconduct	0-4						
Fall 2003		0.47 (0.60)	0-4	.82	0.53 (0.56)	0-4	.82
Spring 2004		0.52 (0.62)	0-4	.84	0.62 (0.59)	0-3.73	.84
Substance use	0-4						
Fall 2003		--	--	--	0.08 (0.34)	0-4	.83
Spring 2004		--	--	--	0.09 (0.33)	0-3.50	.77

Note. Elementary $N = 1,713$ in Fall 2003, 1,611 in Spring 2004; middle school $N = 1,024$ in Fall 2003, 973 in Spring 2004.

administration. Nonetheless, survey completion rates were high at both baseline and the first-year follow up. Of the 1,820 elementary students in the recruited sample, 1,713 (94%) completed the baseline survey and 1,611 (89%) completed the first-year follow-up survey. Of the 1,119 recruited middle school students, 1,024 (92%) completed the baseline survey and 973 (87%) completed the follow-up survey.

Work habits. Elementary and middle school students completed an adaptation of the 6-

item Work Habits scale from the *Mock Report Card* (Pierce et al., 1999), a measure of children's classroom work habits initially intended for teachers to report on student behavior. The items were reworded for the student self-report and the response scale was modified (1 = not at all true, 4 = really true). Sample items include "I follow the rules in my classroom" and "I finish my work on time." On average, the elementary students judged their work habits to be very good, whereas the middle school students' judgments were more modest.

Self-efficacy. Middle school students completed a 7-item modification of the *Self-Efficacy* scale developed by Walker & Arbreton (2001). Three items were reworded for this study, one item was omitted, and the response scale was modified (1 = not at all true, 4 = really true). Sample items include "I give up on things before finishing them" and "I am not sure how good I am at things." Several negatively worded items were reverse coded prior to scoring. On average, the students had positive but modest feelings of self-efficacy.

Misconduct. Elementary and middle school students completed the 11-item *Misconduct Scale*, an adaptation of the *Self-Reported Behavior Index* (Brown, Clasen, & Eicher, 1986), about their behavior since the start of the school year (Fall 2003 administration) or since January (Spring 2004 administration). Sample items include "Gotten into a fight at school," "Taken something from a store without paying for it," and "Done something your parents told you not to do." The measure utilizes a 5-point response scale (0 = never, 4 = 4 or more times a week). On average, misconduct was infrequent in both the elementary and middle school samples; however, some students reported that they engaged in numerous negative behaviors almost daily.

Substance use. Middle school students completed four items about their use of tobacco, alcohol, marijuana, and other drugs. The items are from the *Substance Use and Risk Behaviors* measure used in the 2002-03 Youth Survey for the Baltimore Youth Places Evaluation, conducted by Policy Studies Associates, Inc. We modified the response scale (0 = never, 4 = 4 or

more times a week) and asked students to indicate the frequency with which they used substances since the beginning of the school year (Fall 2003 administration) or since January (Spring 2004 administration). On average, the middle school students used substances rarely, although some students reported that they used a variety of substances on a frequent basis.

Teacher and Program Staff Measures

Teachers and after-school program staff completed baseline and first-year follow-up measures of a number of students' intermediate outcomes, including work habits, task persistence, social skills, and aggressive and prosocial behavior with peers. Teachers also provided information about students' academic performance, a longer term outcome in our theory of change. At the elementary level, we asked classroom teachers to complete surveys about the study participants in their class. For middle school participants, language arts teachers were asked to complete surveys. We consulted with after-school program directors to identify the staff member who was most familiar with each study participant in the program group, and asked that staff member to complete a survey. Psychometric information for these measures appears in Tables 3.2 (teacher reports) and 3.3 (program staff reports).

Teacher and program staff survey completion rates were high at both baseline and the first-year follow up. For the elementary students in the recruited sample, we received completed teacher surveys for 81% at baseline and 84% at the first-year follow up. For the elementary students in the program group, we received completed program staff surveys for 78% at baseline and 80% in Spring 2004. Completion rates were similar for the middle school sample. We received completed teacher surveys for 88% and 85% of the recruited middle school students at baseline and the first-year follow up, respectively. For the students in the program group, we received completed program staff surveys for 83% and 79% at baseline and the first-year follow up, respectively. In general, program staff surveys were collected only for those students in the

Table 3.2
Teacher Measures of Child and Youth Functioning

Outcome	Response scale	Elementary			Middle		
		<i>M</i> (<i>SD</i>)	Range	Alpha	<i>M</i> (<i>SD</i>)	Range	Alpha
Work habits	1-5						
Fall 2003		3.30 (1.01)	1-5	.97	3.46 (1.13)	1-5	.98
Spring 2004		3.43 (1.10)	1-5	.97	3.50 (1.12)	1-5	.98
Task persistence	1-4						
Fall 2003		2.89 (0.73)	1-4	.93	2.96 (0.71)	1-4	.93
Spring 2004		2.93 (0.72)	1-4	.93	2.99 (0.70)	1-4	.93
Social skills	1-5						
Fall 2003		3.33 (0.97)	1-5	.96	3.47 (0.96)	1-5	.96
Spring 2004		3.46 (0.98)	1-5	.96	3.49 (1.00)	1-5	.96
Aggressive w/peers	0-2						
Fall 2003		0.31 (0.45)	0-2	.93	0.27 (0.43)	0-2	.94
Spring 2004		0.35 (0.48)	0-2	.93	0.32 (0.46)	0-2	.93
Prosocial w/peers	0-2						
Fall 2003		1.50 (0.49)	0-2	.93	1.48 (0.48)	0-2	.93
Spring 2004		1.52 (0.47)	0-2	.93	1.47 (0.48)	0-2	.93
Academic performance	1-5						
Fall 2003		2.96 (0.96)	1-5	.95	2.87 (1.18)	1-5	.95
Spring 2004		2.94 (1.03)	1-5	.96	2.99 (1.13)	1-5	.95

Note. Elementary $N = 1,473$ in Fall 2003, 1,522 in Spring 2004; middle school $N = 980$ in Fall 2003, 947 in Spring 2004.

Table 3.3
Program Staff Measures of Child and Youth Functioning

Outcome	Response scale	Elementary			Middle		
		<i>M (SD)</i>	Range	Alpha	<i>M (SD)</i>	Range	Alpha
Work habits	1-5						
Fall 2003		3.71 (0.97)	1-5	.96	3.75 (1.00)	1-5	.97
Spring 2004		3.57 (0.97)	1-5	.96	3.59 (0.97)	1.17-5	.97
Task persistence	1-4						
Fall 2003		3.03 (0.61)	1-4	.89	3.09 (0.61)	1-4	.91
Spring 2004		2.96 (0.60)	1-4	.89	3.01 (0.63)	1-4	.92
Social skills	1-5						
Fall 2003		3.59 (0.82)	1-5	.94	3.64 (0.85)	1.14-5	.95
Spring 2004		3.46 (0.84)	1-5	.94	3.52 (0.83)	1.29-5	.95
Aggressive w/peers	0-2						
Fall 2003		0.31 (0.43)	0-2	.92	0.24 (0.35)	0-1.56	.90
Spring 2004		0.34 (0.45)	0-2	.93	0.28 (0.41)	0-2	.93
Prosocial w/peers	0-2						
Fall 2003		1.38 (0.44)	0-2	.88	1.40 (0.45)	0.13-2	.90
Spring 2004		1.36 (0.45)	0-2	.89	1.40 (0.50)	0-2	.93

Note. Elementary $N = 798$ in Fall 2003, 815 in Spring 2004; middle school $N = 446$ in Fall 2003, 426 in Spring 2004.

program group who were attending the after-school programs at the time the surveys were administered.

Work habits. Teachers and program staff completed the 6-item Work Habits scale from the *Mock Report Card* (Pierce et al., 1999), a measure of children's classroom work habits. Sample items include "Follows classroom procedures" and "Completes work promptly." Four additional items were included in the teacher measure of work habits: "Completes assignments to my satisfaction," "Is attentive in class," "Participates in class," and "Turns in homework promptly." On average, both elementary and middle school teachers judged their students' work habits to be slightly above average; program staff ratings were a bit higher.

Task persistence. Teachers and program staff completed an adaptation of Walker & Arbretson's (2001) *Self-Efficacy*, a child self-report measure containing seven items. The items were reworded for administration to teachers and program staff and rated on a 4-point scale (1 = not at all true, 4 = really true). Sample items include "This student gives up on things before finishing them" and "This student is unsure about his/her ability to do things." Several negatively worded items were reverse coded prior to scoring. On average, both elementary and middle school teachers, and program staff, rated their students as having moderate task persistence.

Social skills. Teachers and program staff completed the Prosocial Behavior scale from the *Teacher Checklist of Peer Relations* (Coie & Dodge, 1988), a measure of children's social skills with peers. The seven items are rated on a 5-point scale (1 = very poor, 5 = very good). Sample items include "Accurately interprets what peers are trying to do" and "Is aware of the effects of his/her behavior on others." On average, teachers and program staff judged their students to have average social skills.

Behavior toward peers. Teachers and program staff completed 17 items from the *Child Behavior Scale* (Ladd & Profilet, 1996), a measure of children's aggressive, withdrawn, and

prosocial behaviors, utilizing a 3-point response scale (0 = not true, 1 = sometimes true, 2 = often true). The selected items form two scales, **Aggressive with Peers** and **Prosocial with Peers**. Sample items measuring aggressive behavior are “Is an aggressive child” and “Annoys or irritates classmates;” sample items measuring prosocial behavior are “Compromises in conflicts with classmates” and “Offers help or comfort when classmates are upset.” On average, teachers and program staff rated their students as exhibiting low amounts of aggressive behavior and moderately high levels of prosocial behavior with their peers.

Academic performance. Teachers completed the *Academic Performance* scale from the *Mock Report Card* (Pierce et al., 1999), a measure that was developed in order to obtain standardized information about students’ academic performance across school districts that utilize different grading systems. Performance in several subject areas was rated on a 5-point scale (1 = child is performing below grade level, 5 = child is performing beyond grade level). Middle school language arts teachers and elementary teachers rated their students’ performance in reading, oral language, and written language. Elementary teachers also rated performance in math, science, and social studies. An overall academic performance score was computed as the mean of the item scores. On average, teachers rated their students as performing near grade level.

Chapter 4

Findings: After-School “Clusters”

The reports of program participation, program attendance, and after-school activities we obtained (see Chapter 2) indicate that the study participants’ after-school hours encompass several different types of experiences—attending the selected promising programs in varying amounts; participating in coached sports, school-based extracurricular activities, lessons, and other clubs or programs; and spending time unsupervised at home or hanging out with peers. Researchers typically have studied these experiences as independent events and have assessed the unique variance associated with a particular type of experience. Another possibility, however, is that after-school programs and other types of after-school contexts are experienced in combination, and it is these *sets of experiences* (not any single setting in isolation) that are associated with student outcomes. In order to test this hypothesis, we used cluster analyses to identify subgroups of students who differed in the ways that they spent their time after school.¹

Cluster analysis is a statistical technique that places individuals in groups, or “clusters,” based on their scores on a set of variables. We included eight variables in our cluster analyses: program status (whether respondents belonged to the program or comparison group); program dosage, or amount of participation in one of the selected promising after-school programs; and the amount of time respondents spent in six other after-school contexts—school-based extracurricular activities, coached sports, lessons, home alone, home with siblings (but no adults present), and hanging out with peers (unsupervised by adults). We used student reports of their involvement in activities and time spent unsupervised, rather than parent reports, for several

¹ Data analysis procedures were revised from our initial analysis plan due to student participation patterns. Initially, we planned to compare program and comparison respondents and assess differences among program respondents by dosage levels. Many students were involved in structured activities beyond the selected programs, so we used cluster analysis procedures, which are more informative. Findings related to the initial analysis plan appear in Appendices A and B.

reasons. First, parents reported their children's activities in Fall 2003 only, whereas student reports were obtained in both Fall 2003 and Spring 2004, allowing us to construct an annual report of time spent in each of the six contexts. Second, student reports of the amount of time they spent at the participating programs were highly correlated with attendance reports provided by the programs, suggesting that the students reported their activities accurately. Finally, student and parent reports of time spent in the six contexts were significantly correlated.

We performed hierarchical agglomerative cluster analyses, separately on data from elementary and middle school participants, using SLEIPNER version 2.1 and squared Euclidean distance based on Ward's method (Aldenderfer & Blashfield, 1984). The most sensible solution was a four-cluster solution that yielded remarkably similar clusters for the two age groups. The first cluster, *high program / high activity*, accounted for 16% ($n = 278$) of elementary and 18% ($n = 195$) of middle school respondents and was comprised of youth who had high levels of involvement in both the selected after-school programs and other enrichment activities. In the second cluster, *high program / low activity*, respondents were characterized by high rates of participation in the after-school programs and low scores on most of the other clustering variables. About a third of elementary school youth ($n = 582$) and 28% of the middle schoolers ($n = 312$) fell into the second cluster. A third cluster, *low supervision*, contained respondents who had low levels of participation in the targeted programs, moderate to high amounts of time in activities, and high amounts of time in unsupervised settings (especially hanging out with peers). It encompassed 16% ($n = 284$) of elementary and 15% ($n = 162$) of middle school respondents. The final cluster, *supervised at home*, was the largest, comprising over a third of both samples (601 elementary and 409 middle school youth). This group manifested relatively low scores on all of the clustering variables, suggesting that they spent much of their time after school at home under adult supervision.

Tables 4.1 and 4.2 present mean scores (and standard deviations) on each of the clustering variables for each cluster in the elementary and middle school samples, respectively. The significant differences between the clusters on these variables are to be expected, because the variables were used to construct the clusters. The differences observed across the four clusters are consistent with the cluster names and serve to validate our four-cluster solution.

Three issues were examined with respect to the after-school clusters: (1) their demographic characteristics, (2) differences in their performance at baseline, and (3) relative changes in their functioning from baseline to the first year follow-up.

Demographic Characteristics Associated with the After-School Clusters

Our theory of change recognizes that multiple factors influence whether children and youth attend after-school programs. In our analyses, we considered six demographic factors in relation to cluster membership: child gender, child ethnicity, family structure (two-parent vs. all others), maternal employment status (full time vs. all others), maternal education, and family income. Tables 4.3 and 4.4 indicate the demographic characteristics of the elementary and middle school clusters, respectively.

In the **elementary** sample (see Table 4.3), the *high program / high activity* group had a higher proportion of girls than the other clusters, whereas males predominated in the *low supervision* cluster. The *high program / low activity* group had somewhat higher family incomes than the other clusters, although over half of all groups had an annual family income of less than \$20,000. The *low supervision* group had the greatest proportion of mothers who did not graduate from high school. The *supervised at home* group contained relatively more youth from two-parent households and youth whose mothers did not work full time. Consistent with the study sample, all of the clusters had high proportions of Hispanic students, but the proportion was especially high in the *high program / high activity* and *supervised at home* groups.

Table 4.1
Means (Standard Deviations) and Percentages on Clustering Variables, Elementary Clusters

	High program / high activity <i>N</i> = 278	High program / low activity <i>N</i> = 582	Low supervision <i>N</i> = 284	Supervised at home <i>N</i> = 601	
Group assignment					
Comparison	15 (5%)	0	131 (46%)	601 (100%)	$\chi^2_{(3)} = 1874.09^{***}$
Program	263 (95%)	582 (100%)	153 (54%)	0	
Program dosage ¹	3.27 ^a (1.22)	2.79 ^b (1.25)	1.39 ^c (1.58)	.005 ^d (.07)	$F_{(3,1741)} = 894.31^{***}$
Coached sports ²	2.07 ^b (1.12)	1.51 ^c (.87)	2.35 ^a (1.11)	1.54 ^c (.84)	$F_{(3,1741)} = 72.01^{***}$
School-based activities ²	2.89 ^a (.74)	1.11 ^c (.30)	1.63 ^b (.83)	1.20 ^c (.47)	$F_{(3,1741)} = 746.22^{***}$
Lessons ²	2.92 ^a (.97)	1.86 ^c (.96)	2.49 ^b (1.08)	1.62 ^d (.86)	$F_{(3,1741)} = 148.09^{***}$
Home alone ²	1.19 ^b (.49)	1.25 ^b (.57)	2.49 ^a (1.01)	1.21 ^b (.52)	$F_{(3,1741)} = 318.20^{***}$
Home with siblings ²	1.18 ^b (.49)	1.12 ^b (.39)	2.31 ^a (.95)	1.16 ^b (.46)	$F_{(3,1741)} = 344.52^{***}$
Hang out with peers ²	1.39 ^b (.69)	1.25 ^c (.57)	2.61 ^a (1.03)	1.24 ^c (.55)	$F_{(3,1741)} = 310.53^{***}$

Note. Means with different superscripts are significantly different at $p < .05$.

*** $p \leq .001$

¹ Program dosage is a categorical variable: 0 = 0-4 days, 1 = 5-29 days, 2 = 30-59 days, 3 = 60-89 days, 4 = 90 or more days across the school year.

² Time spent in other after-school contexts is a categorical variable: 1 = not at all/once or twice, 2 = about once a week, 3 = 2-3 days a week, 4 = 4 or more days a week.

Table 4.2
Means (Standard Deviations) and Percentages on Clustering Variables, Middle School Clusters

	High program / high activity N = 195	High program / low activity N = 312	Low supervision N = 162	Supervised at home N = 409	
Group assignment					
Comparison	45 (23%)	0	94 (58%)	409 (100%)	$\chi^2_{(3)} = 1063.06^{***}$
Program	150 (77%)	312 (100%)	68 (42%)	0	
Program dosage ¹	1.71 ^b (1.37)	2.38 ^a (1.23)	0.77 ^c (1.13)	.002 ^d (.05)	$F_{(3,1074)} = 375.85^{***}$
Coached sports ²	2.54 ^a (1.09)	1.68 ^c (.87)	2.11 ^b (1.11)	1.49 ^d (.80)	$F_{(3,1074)} = 63.98^{***}$
School-based activities ²	2.92 ^a (.80)	1.35 ^c (.58)	1.69 ^b (.84)	1.31 ^c (.63)	$F_{(3,1074)} = 275.41^{***}$
Lessons ²	3.09 ^a (.84)	1.71 ^c (.81)	2.10 ^b (1.02)	1.49 ^d (.74)	$F_{(3,1074)} = 176.12^{***}$
Home alone ²	1.75 ^b (.90)	1.45 ^c (.75)	3.12 ^a (.85)	1.56 ^c (.85)	$F_{(3,1074)} = 166.19^{***}$
Home with siblings ²	1.29 ^b (.55)	1.21 ^{b,c} (.45)	2.99 ^a (.83)	1.17 ^c (.42)	$F_{(3,1074)} = 519.24^{***}$
Hang out with peers ²	1.71 ^b (.92)	1.64 ^b (.90)	2.86 ^a (1.03)	1.67 ^b (.95)	$F_{(3,1074)} = 73.20^{***}$

Note. Means with different superscripts are significantly different at $p < .05$.

*** $p \leq .001$

¹ Program dosage is a categorical variable: 0 = 0-4 days, 1 = 5-29 days, 2 = 30-59 days, 3 = 60-89 days, 4 = 90 or more days across the school year.

² Time spent in other after-school contexts is a categorical variable: 1 = not at all/once or twice, 2 = about once a week, 3 = 2-3 days a week, 4 = 4 or more days a week.

Table 4.3
Percentage Distribution of Elementary School Clusters on Background Variables

	High program / high activity N = 278	High program / low activity N = 582	Low supervision N = 284	Supervised at home N = 601
Child gender, $\chi^2_{(3)} = 65.50^{***}$				
Male	34	46	67	46
Female	66	54	33	54
Child ethnicity, $\chi^2_{(9)} = 74.95^{***}$				
Asian/other	1	4	1	4
Black	7	10	12	5
Hispanic	84	67	74	84
White	8	19	13	7
Family structure, $\chi^2_{(3)} = 19.09^{***}$				
Two parents	56	60	62	71
Other	44	40	38	29
Maternal employment, $\chi^2_{(3)} = 20.90^{***}$				
Full time	44	50	48	36
Other	56	50	52	64
Maternal education, $\chi^2_{(9)} = 24.68^{**}$				
Did not graduate high school	41	28	44	36
High school diploma/GED	20	31	24	27
Some college	29	32	23	29
4-year college degree	10	9	9	8
Family income, $\chi^2_{(9)} = 19.54^*$				
Less than \$11,000	29	27	38	31
\$11,000-19,999	32	24	22	23
\$20,000-29,999	23	21	19	22
\$30,000 or more	16	28	21	24

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

Table 4.4
Percentage Distribution of Middle School Clusters on Background Variables

	High program / high activity N = 195	High program / low activity N = 312	Low supervision N = 162	Supervised at home N = 409
Child gender, $\chi^2_{(3)} = 24.74^{***}$				
Male	39	58	48	42
Female	61	42	52	58
Child ethnicity, $\chi^2_{(9)} = 28.72^{***}$				
Asian/other	7	6	6	8
Black	14	13	17	11
Hispanic	44	58	37	46
White	35	23	40	35
Family structure, $\chi^2_{(3)} = 11.89^{**}$				
Two parents	65	61	61	72
Other	35	39	39	28
Maternal employment, $\chi^2_{(3)} = 9.67^*$				
Full time	56	48	57	44
Other	44	52	43	56
Maternal education, $\chi^2_{(9)} = 19.27^*$				
Did not graduate high school	28	28	24	37
High school diploma/GED	22	24	30	25
Some college	32	32	37	28
4-year college degree	18	16	9	10
Family income, $\chi^2_{(9)} = 18.79^*$				
Less than \$11,000	17	24	17	23
\$11,000-19,999	25	25	18	22
\$20,000-29,999	15	18	19	22
\$30,000 or more	43	33	46	33

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

In the **middle school** sample (see Table 4.4), the *high program / high activity* cluster had a comparatively high proportion of girls and, along with the *low supervision* group, a greater percentage of mothers who worked full time. These clusters also had somewhat higher incomes, although at least one-third of each cluster had incomes of less than \$20,000 a year. The *high program / low activity* cluster was predominantly male and Hispanic, with a lower proportion of White members than the other clusters. The *supervised at home* group had the highest proportion of two-parent households, the lowest percentage of mothers working full time, and somewhat lower levels of maternal education.

Relations Between Cluster Membership and Child Functioning at Baseline

To determine whether the clusters differed on measures of child functioning at baseline, we conducted a series of analyses of covariance (ANCOVA) on the Fall 2003 baseline scores for the four clusters, controlling for background characteristics. As seen in Table 4.5, the most consistent finding from these analyses at the **elementary** level is that the *low supervision* group tended to have the most worrisome profile, with higher ratings of self-reported misconduct and consistently poorer ratings from teachers. Program staff perceived that children in the *high program / high activity* cluster evinced better work habits and better social skills at baseline than students in the *high program / low activity* and *low supervision* clusters. These differences remind us that this study was unable to assess students prior to any program experience. Indeed, in these established programs, some students may have attended the program for some time, so these baseline measurements cannot be viewed as pre-program (or pretest) scores. Rather, they simply reflect student performance early in the 2003-04 school year.

Differences at baseline also were evident for the **middle school** youth (Table 4.6). The *low supervision* cluster had the highest self-reported rates of misconduct and substance use, and they were rated by classroom teachers as more aggressive than the *high program / low activity*

Table 4.5
Adjusted Means of Student Performance at Baseline, Elementary School Clusters

	High program / high activity <i>N</i> = 278 <i>M</i> (<i>SE</i>)	High program / low activity <i>N</i> = 582 <i>M</i> (<i>SE</i>)	Low supervision <i>N</i> = 284 <i>M</i> (<i>SE</i>)	Supervised at home <i>N</i> = 601 <i>M</i> (<i>SE</i>)	
Student report					
Work habits	3.53 ^a (.04)	3.38 ^b (.03)	3.41 ^{a, b} (.04)	3.45 ^{a, b} (.03)	$F_{(3,1243)} = 3.84^{**}$
Misconduct	0.38 ^b (.04)	0.41 ^b (.03)	0.80 ^a (.04)	0.37 ^b (.03)	$F_{(3,1245)} = 29.63^{***}$
Teacher report					
Work habits	3.42 ^a (.08)	3.36 ^{a, b} (.05)	3.11 ^b (.08)	3.54 ^a (.05)	$F_{(3,1084)} = 6.29^{***}$
Task persistence	3.00 ^a (.06)	2.91 ^{a, b} (.04)	2.76 ^b (.06)	3.02 ^a (.04)	$F_{(3,1082)} = 5.49^{***}$
Social skills	3.37 ^{a, b} (.08)	3.34 ^b (.05)	3.16 ^b (.07)	3.53 ^a (.05)	$F_{(3,1081)} = 6.32^{***}$
Aggressive w/peers	0.34 ^{a, b} (.03)	0.25 ^{b, c} (.02)	0.40 ^a (.03)	0.23 ^c (.02)	$F_{(3,1077)} = 7.96^{***}$
Prosocial w/peers	1.52 ^a (.04)	1.53 ^a (.02)	1.40 ^b (.04)	1.58 ^a (.03)	$F_{(3,1077)} = 5.45^{***}$
Academic performance	2.80 ^a (.07)	2.81 ^a (.05)	2.56 ^b (.08)	2.82 ^a (.05)	$F_{(3,1079)} = 3.08^*$
Program staff report					
Work habits	3.96 ^a (.07)	3.68 ^b (.05)	3.61 ^b (.11)	--	$F_{(2,589)} = 5.57^{**}$
Task persistence	3.12 (.05)	3.02 (.03)	3.04 (.07)	--	$F_{(2,585)} = 1.72$
Social skills w/peers	3.77 ^a (.04)	3.56 ^b (.04)	3.43 ^b (.09)	--	$F_{(2,590)} = 5.97^{**}$
Aggressive w/peers	0.33 (.03)	0.29 (.02)	0.34 (.05)	--	$F_{(2,588)} = 1.06$
Prosocial w/peers	1.42 (.03)	1.39 (.02)	1.34 (.05)	--	$F_{(2,589)} = 0.96$

Note. Means are adjusted for child gender and ethnicity, family structure, maternal education and employment, and family income. Means with different superscripts are significantly different at $p < .05$.

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

Table 4.6
Adjusted Means of Student Performance at Baseline, Middle School Clusters

	High program / high activity <i>N</i> = 195 <i>M</i> (<i>SE</i>)	High program / low activity <i>N</i> = 312 <i>M</i> (<i>SE</i>)	Low supervision <i>N</i> = 162 <i>M</i> (<i>SE</i>)	Supervised at home <i>N</i> = 409 <i>M</i> (<i>SE</i>)	
Student report					
Work habits	3.33 (.05)	3.18 (.04)	3.18 (.05)	3.21 (.03)	$F_{(3,781)} = 2.40$
Self-efficacy	3.23 (.04)	3.14 (.03)	3.12 (.05)	3.16 (.03)	$F_{(3,779)} = 1.32$
Misconduct	0.51 ^b (.04)	0.49 ^b (.04)	0.88 ^a (.05)	0.44 ^b (.03)	$F_{(3,785)} = 20.61^{***}$
Substance use	0.08 ^{a, b} (.03)	0.05 ^b (.02)	0.18 ^a (.03)	0.06 ^b (.02)	$F_{(3,786)} = 5.45^{***}$
Teacher report					
Work habits	3.60 (.09)	3.59 (.08)	3.35 (.10)	3.59 (.06)	$F_{(3,735)} = 1.73$
Task persistence	3.06 (.06)	2.98 (.05)	2.87 (.06)	3.04 (.04)	$F_{(3,733)} = 2.26$
Social skills	3.59 (.08)	3.44 (.07)	3.37 (.09)	3.56 (.05)	$F_{(3,707)} = 1.83$
Aggressive w/peers	0.28 ^{a, b} (.03)	0.20 ^b (.03)	0.34 ^a (.04)	0.23 ^{a, b} (.02)	$F_{(3,731)} = 2.85^*$
Prosocial w/peers	1.52 (.04)	1.52 (.03)	1.44 (.04)	1.50 (.03)	$F_{(3,729)} = 0.93$
Academic performance	3.10 (.09)	2.94 (.08)	2.86 (.10)	3.00 (.06)	$F_{(3,733)} = 1.17$
Program staff report					
Work habits	3.82 (.10)	3.91 (.07)	3.60 (.15)	--	$F_{(2,311)} = 1.53$
Task persistence	3.12 ^{a, b} (.06)	3.16 ^a (.04)	2.90 ^b (.09)	--	$F_{(2,311)} = 3.36^*$
Social skills w/peers	3.70 (.08)	3.75 (.06)	3.52 (.13)	--	$F_{(2,320)} = 1.31$
Aggressive w/peers	0.24 (.03)	0.19 (.02)	0.28 (.05)	--	$F_{(2,331)} = 1.93$
Prosocial w/peers	1.45 (.04)	1.41 (.03)	1.39 (.07)	--	$F_{(2,323)} = 0.40$

Note. Means are adjusted for child gender and ethnicity, family structure, maternal education and employment, and family income. Means with different superscripts are significantly different at $p < .05$.

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

cluster. Program staff rated youth in the *high program / low activity* cluster as exhibiting greater task persistence than youth in the *low supervision* cluster.

Relations Between Cluster Membership and Child Functioning at the First-Year Follow Up

For our primary analyses, we performed a set of hierarchical linear modeling (HLM) analyses in which we contrasted the *high program / high activity* cluster, the *high program / low activity* cluster, and the *supervised at home* cluster with the *low supervision* cluster. We tested these contrasts because the *low supervision* group personifies the situation of strongest concern to parents, educators, and policymakers. These contrasts allowed us to examine whether the selected after-school programs and other enrichment activities would be protective for children and youth who are at risk for social and academic problems.

Prior to conducting the HLM analyses, we used a multiple imputation procedure (Rubin, 1987) to address missing observations in the data set. In this procedure, missing data were replaced by observations drawn randomly from a multivariate distribution fit to the variable and covariates. Multiple datasets are created in which different samples are selected for missing observations. Analyses then are performed on each data set as though all data had actually been obtained. The HLM software is able to accommodate the multiple imputation procedure, and was used to pool results across the analyses. In the current analyses, five imputed datasets were used. Missing observations were sampled using a Markov chain Monte Carlo procedure implemented in LISREL 8.71 (Jöreskog & Sörbom, 2004). The appeal of the multiple imputation approach is that the level of uncertainty associated with missing observations is addressed, and thus the standard errors of model parameters are computed correctly.

A two-level HLM model was fit in which students (Level 1) were nested within schools (Level 2) for each outcome variable. This allowed us to separate variability in scores that can be traced to site-level factors (the schools the study participants attended) from variability related to

the individual's cluster membership. In comparison to a standard single-level regression model that ignores school effects, HLM controls for differences between schools in terms of average changes in student functioning that are unrelated to cluster membership. Using schools rather than programs as the site-level factor allowed us to include all students, regardless of program attendance, in the analyses.

In the HLM analyses, we controlled for a number of demographic characteristics, including child gender and ethnicity, and family background (family income, family structure, maternal education, and maternal work status). We also controlled for student functioning at baseline on each outcome variable. This control of performance at baseline allowed us to assess residualized change scores in which relative increases and decreases from fall to spring of the school year are assessed. In the case of the after-school clusters, we contrasted performances of three clusters (*high program / high activity*, *high program / low activity*, *supervised at home*) with the *low supervision* cluster.

Results of the HLM analyses of the student- and teacher-reported outcomes are summarized in Table 4.7 for the elementary sample and Table 4.8 for the middle school sample. Analyses of after-school program staff reports of both elementary and middle school student outcomes are summarized in Table 4.9. The coefficients shown in the tables quantify the residualized changes in the outcomes. For example, for the elementary student reports of misconduct, the coefficient of $-.284$ indicates that students in the *high program / low activity* cluster decreased $.284$ units relative to students in the *low supervision* group from fall to spring. Similarly, for the elementary teacher report of academic performance, the coefficient of $.162$ for the *high program / low activity* cluster indicates that these students exhibited a relative gain of $.162$ units in comparison to the *low supervision* cluster when performance at baseline was controlled.

Table 4.7
HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Elementary School Sample

	Student report		Teacher report					
	Work habits	Misconduct	Work habits	Task persistence	Social skills	Aggressive w/peers	Prosocial w/peers	Academic performance
FIXED EFFECT								
Intercept	2.085***	.678***	.934***	.863***	1.389***	.209*	.769***	.719***
Gender (1 = female)	.110***	-.187***	.176***	.061*	.155***	-.081***	.096***	.067
White	.033	-.004	.094	.016	-.052	.024	.012	-.047
Black	.018	-.058	-.055	-.021	-.181	.106	-.067	-.061
Hispanic	.018	-.040	.104	.078	.000	.035	-.009	.011
Family income	.002	-.001	.003	.001	.009	.003	-.003	.017
Family structure (1 = two parents)	.031	-.074*	.049	.056	.019	-.043	.027	.031
Maternal education	-.011	.017	-.004	.011	.004	.005	-.008	.020
Maternal work (1 = full time)	.000	-.020	.011	.027	-.036	-.003	-.017	-.001
Fall 2003 baseline	.303***	.418***	.661***	.622***	.561***	.612***	.471***	.720***
High program / high activity cluster	.197***	-.213***	.062	.079	.072	-.021	.044	.074
High program / low activity cluster	.091*	-.284***	.142*	.122**	.135*	-.061*	.065*	.162**
Supervised at home cluster	.105*	-.240***	.126*	.101*	.166*	-.049	.042	.131*
RANDOM EFFECT								
Site mean, variance component ¹	.007***	.005***	.012***	.001	.013***	.001**	.003***	.026***
Level-1 effect, variance component ²	.301	.241	.563	.273	.546	.132	.144	.473

Note. Clusters are compared against the *low supervision* cluster. * $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

¹ Indicates whether there are differences between sites (schools) in terms of degree of change.

² Indicates the amount of residual (unexplained) variance within clusters.

Table 4.8
HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Middle School Sample

	Student report				Teacher report					
	Work habits	Self-efficacy	Misconduct	Substance use	Work habits	Task persistence	Social skills	Aggressive w/peers	Prosocial w/peers	Academic performance
FIXED EFFECT										
Intercept	1.522***	1.868***	.633***	.150*	.941***	1.063***	1.041***	.326***	.548***	.781***
Gender (1 = female)	.064*	-.025	-.077*	.002	.222***	.096**	.136**	-.093***	.100***	.043
White	.035	.027	-.022	.047	.056	.011	-.086	.013	-.019	.118
Black	-.075	-.023	.067	.062	.008	-.004	-.023	.033	.031	.004
Hispanic	-.012	.010	.003	.041	.076	-.013	.032	-.016	.022	.041
Family income	.002	.001	-.004	.001	.023*	.006	.024*	-.013**	.020***	.024
Family structure (1 = two parents)	-.026	.006	-.054	-.014	.014	.005	.028	-.016	-.018	.012
Maternal education	.011	.018	-.010	-.010	-.004	.019	.020	-.008	.007	.022
Maternal work (1 = full time)	-.018	-.024	-.028	-.002	-.011	-.020	.024	.016	-.034	-.040
Fall 2003 baseline	.483***	.409***	.484***	.286***	.621***	.601***	.597***	.622***	.491***	.631***
High program / high activity cluster	.066	.018	-.148**	-.109***	.172*	.017	.041	-.001	.002	.106
High program / low activity cluster	-.040	-.087	-.152**	-.092**	.095	-.029	.027	.003	.032	.044
Supervised at home cluster	.017	-.049	-.162**	-.111***	.108	.005	.105	-.023	.031	.146*
RANDOM EFFECT										
Site mean, variance component ¹	.004*	.001	.000	.001	.024***	.005**	.043***	.002**	.006***	.018***
Level-1 effect, variance component ²	.225	.180	.232	.086	.577	.255	.490	.114	.141	.548

Note. Clusters are compared against the *low supervision* cluster. * $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

¹ Indicates whether there are differences between sites (schools) in terms of degree of change.

² Indicates the amount of residual (unexplained) variance within clusters.

Table 4.9
HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Program Staff Reports

	Elementary					Middle				
	Work habits	Task persistence	Social skills	Aggressive w/peers	Prosocial w/peers	Work habits	Task persistence	Social skills	Aggressive w/peers	Prosocial w/peers
FIXED EFFECT										
Intercept	1.248*	1.059*	1.109*	.263**	.371	2.138***	1.659***	2.133***	.288**	.821***
Gender (1 = female)	.205***	.090*	.156**	-.097***	.117***	.230***	.100*	.207***	-.090**	.117***
White	.021	.104	.074	-.052	.055	-.017	.026	.004	.003	.015
Black	-.002	.049	.032	-.051	.062	-.207	.017	-.028	-.027	.023
Hispanic	-.001	.070	.091	-.034	.047	-.121	.065	.043	-.010	.040
Family income	.008	.018	.012	.003	.007	.016	.011	.031**	-.000	.009
Family structure (1 = two parents)	.049	-.044	.005	-.015	-.004	.025	.079	-.005	-.041	.013
Maternal education	-.010	-.006	-.007	.011	.004	.011	.022	.003	-.010	.001
Maternal work (1 = full time)	.070	.054	.055	.010	.031	-.105	-.075	-.096	.011	-.018
Fall 2003 baseline	.445***	.447***	.446***	.403***	.428***	.332***	.335***	.282***	.500***	.275***
High program / high activity cluster	.442*	.307*	.488**	.004	.228***	.179	.128	.138	-.037	.031
High program / low activity cluster	.417*	.292*	.447**	-.008	.189**	.068	.046	.035	-.004	-.001
RANDOM EFFECT										
Site mean, variance component ¹	.034***	.009***	.053***	.003***	.010***	.024***	.022***	.046***	.003**	.036***
Level-1 effect, variance component ²	.778	.357	.647	.135	.150	.566	.230	.398	.101	.128

Note. Clusters are compared against the *low supervision* cluster. * $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

¹ Indicates whether there are differences between sites (schools) in terms of degree of change.

² Indicates the amount of residual (unexplained) variance within clusters.

As shown in Table 4.7, changes from baseline to follow up were found for the **elementary** students' reports of work habits and misconduct, and for teacher reports of work habits, task persistence, social skills, aggressive behavior with peers, prosocial behavior with peers, and academic performance. Students in the *high program / low activity* cluster posted the most consistent differences relative to the *low supervision* group. The *high program / low activity* cluster had relative declines in teacher-reported aggressive behavior with peers and relative gains in teacher-reported prosocial behavior compared to the *low supervision* group. The *high program / low activity* and *supervised at home* clusters posted relative gains in teacher reports of academic performance, work habits, social skills, and task persistence in comparison to the *low supervision* cluster. When contrasted with the *low supervision* cluster, the two *program* clusters and the *supervised at home* cluster displayed relative gains in self-reported work habits and relative declines in self-reported misconduct.

Program staff reports of elementary student outcomes (see Table 4.9) exhibited similar patterns of findings. The *high program / high activity* and *high program / low activity* clusters exhibited relative improvements in work habits, social skills, task persistence, and prosocial behavior with peers. There were no significant differences between the clusters on change in staff reports of students' aggressive behavior with peers.

In the **middle school** sample (see Table 4.8), changes from baseline to follow up in student functioning were not as evident as in the elementary sample. The middle school students in the *high program / high activity* cluster, the *high program / low activity* cluster, and the *supervised at home* cluster reported relative reductions in misconduct and substance use when contrasted with the *low supervision* cluster. For teacher-reported outcomes, two significant differences were observed. The *high program / high activity* cluster exhibited relative gains in work habits, and the *supervised at home* cluster had relative gains in academic performance, in

contrast to the *low supervision* cluster. There were no significant differences between the *low supervision* cluster and the *high program / high activity* and *high program / low activity* clusters in terms of program staff reports of student outcomes (see Table 4.9).

Effect sizes. In order to quantify the effectiveness of the selected after-school programs and supervision at home on student outcomes relative to being unsupervised during the after-school hours, we calculated the effect size (using Cohen's *d*) for each of the significant differences between the *low supervision* cluster and the other clusters. These effect sizes provide information about how well the different after-school contexts (*high program / high activity*, *high program / low activity*, and *supervised at home*) protected the students from adverse outcomes.

The effect sizes for both the elementary and middle school samples are shown in Table 4.10. The effect sizes ranged from .17 to .61 in the elementary sample, and from .20 to .38 in the middle school sample. The table also shows approximate percentiles for the effect sizes. Where effect sizes are positive, the percentiles represent the percentage of students in the *low supervision* cluster whose gain scores were smaller than those of the average student in the comparison cluster. Where effect sizes are negative, these represent the percentage of students in the *low supervision* cluster whose score changes are larger than those of the average student in the comparison cluster. For example, the effect size of -.58 for the elementary *high program / low activity* group's self-reported misconduct is at approximately the 72nd percentile, meaning that 72% of the students reported larger decreases in misconduct from baseline to the first-year follow up than the average student in the *low supervision* cluster. Similarly, the effect size of .24 for teacher reports of academic performance in this comparison means that approximately 60% of the students in the *high program / low activity* cluster had larger gains in academic performance than the average student in the *low supervision* cluster. Effect sizes also can be

Table 4.10
Effect Sizes for Significant Effects in HLM Analyses

	High program / high activity vs. low supervision				High program / low activity vs. low supervision				Supervised at home vs. low supervision			
	Elementary		Middle		Elementary		Middle		Elementary		Middle	
	Effect size	%	Effect size	%	Effect size	%	Effect size	%	Effect size	%	Effect size	%
Student report												
Work habits	.36	64 th			.17	56 th			.19	58 th		
Misconduct	-.43	67 th	-.31	62 nd	-.58	72 nd	-.32	62 nd	-.49	69 th	-.34	64 th
Substance use	NA	NA	-.37	64 th			-.31	62 nd			-.38	65 th
Teacher report												
Work habits			.23	59 th	.19	58 th			.17	56 th		
Task persistence					.23	59 th			.19	58 th		
Social skills					.18	56 th			.23	59 th		
Aggressive w/peers					-.17	56 th						
Prosocial w/peers					.17	56 th						
Academic performance					.24	60 th			.19	58 th	.20	58 th
Program staff report												
Work habits	.50	69 th			.47	68 th			NA	NA	NA	NA
Task persistence	.51	69 th			.49	69 th			NA	NA	NA	NA
Social skills w/peers	.61	73 rd			.56	71 st			NA	NA	NA	NA
Prosocial w/peers	.59	73 rd			.49	69 th			NA	NA	NA	NA

Note. % refers to percentile, or the percentage of students in the *low supervision* cluster who scored below the average student (or above the average student, in the case of negative effect sizes) in the comparison cluster.

interpreted as the proportion of a standard deviation between the two groups. For example, the difference between the *high program / low activity* group and the *low supervision* group in misconduct is about .6 of a standard deviation, which is considered a large effect.

The effect sizes we obtained are comparable to those obtained in educational research. For example, effect sizes of .21-.30 and .15-.31 for elementary students' (kindergarten through Grade 3) achievement test performance in reading and math, respectively, were observed in a study of the effects of reducing class size from 23 students to 15 students (Finn, Gerber, Achilles, & Boyd-Zaharias, 2001). These effect sizes increased with each additional year that students were enrolled in the smaller classes, and were strongest for the youngest students. Additionally, effect sizes of .08-.20 were observed for the influence of K-3 enrollment in small classes on reading and math achievement test performance in Grades 4-8. Larger effect sizes are evident in studies of more intensive interventions with at-risk students. For example, the Success for All (SFA) program provides resources and support to children starting with literacy activities in the preschool years and continuing into elementary school with a reading program, individual tutoring, and family support to encourage parental involvement with the school and provide assistance with health or home problems. Effect sizes of .43 to .68 for reading achievement test scores have been observed across Grades 1-5 in comparisons of SFA students and matched control groups (Ross, Smith, Slavin, & Madden, 1997). Effects were strongest where SFA implementation quality was high; where implementation was poor, effects were weaker or nonexistent.

Chapter 5

Conclusions and Next Steps

Key Findings and Their Implications

There is growing consensus among scholars and practitioners about features of after-school environments that foster healthy social, emotional, and intellectual development for children and youth. In designing programs to address these features, it is common to assume that youth will choose a particular program or setting in which to spend their after-school hours. Our findings indicate, however, that among elementary and middle school students who spend time in structured after-school activities, few are devoted exclusively or even primarily to a single program. Instead, they construct an after-school schedule that involves *sets of experiences*. They may spend part of the week in a comprehensive after-school program such as the promising programs on which our study focuses, but they are also likely to take advantage of other opportunities at their school or in their community: special interest clubs, sports teams, music or dance lessons, religious activities, and so on. For many of these youth, the particular combination of activities and the relative amount of time devoted to each shifts over the course of a school year.

The tendency to create sets of after-school experiences has profound implications for those who design and manage programs for youth. It means, first of all, that one must allow for irregular attendance—youth who come a couple of days a week, or who attend regularly for several weeks and then disappear for several weeks—in designing specific programs or activities. More importantly, it encourages more collaboration among programs available to youth in a given school or community. Rather than being “all things to all students,” each after-school program may need to be more attentive to how its strengths are coordinated with other after-school environments in which young people in their area may be involved.

Our discovery that young people tend to have sets of experiences in the after-school hours, rather than enrollment in a specific program (or lack of participation), also prompted us to change our approach to data analyses. Rather than relying on comparisons of “program” and “comparison” respondents, we used cluster analysis to identify groups of respondents with common patterns of after-school activities. Four clusters emerged from these analyses, and they were remarkably similar among elementary and middle school youth.

Two of the clusters reported low or virtually no involvement in our target programs. Attendance was especially rare in one cluster, which accounted for over a third of elementary and nearly 40% of middle school respondents. This group was also unlikely to be involved in any other structured after-school experiences. Rather, they spent most of their time after school at home, presumably under the care of parents or other adults. This group had a distinctive demographic profile: They were more likely to come from two-parent households, with mothers who did not work full time. Mothers had modestly lower educational levels in this group as well. In other words, the *supervised at home* cluster seemed to be dominated by youth whose parents were more available to look after them in the hours after school. The distinctive demographic characteristics of this group made it inappropriate as a comparison point for clusters that had high involvement in our targeted programs.

Another cluster of respondents, about 15% of the elementary and middle school samples, also had low participation rates in structured activities. Unlike the *supervised at home* cluster, however, these youth spent much of their time after school in unsupervised settings (at home without adults, or hanging out with peers). The dangers of this arrangement, which have been emphasized by other investigators (Smolansky & Gootman, 2003), were reaffirmed in our study. It would be inaccurate, however, to portray the *low supervision* cluster as *never* supervised. In fact, this group had higher rates of participation in coached sports, school-based activities, and

lessons than the other clusters. Their distinguishing feature was spending time in unsupervised environments such as being home alone or hanging out with peers, in addition to reporting spending time in structured activities. Again, this points to the importance of *sets of experiences*. Youth who move among a variety of activities, all of which feature structure and adult supervision, may be as well protected against problem behavior as those who devote almost all of their after-school hours to one supervised environment. Youth who vacillate between supervised and unsupervised settings appear to be more vulnerable to negative outcomes. There may even be a “critical proportion” of time that young people from disadvantaged backgrounds such as those of our respondents need to spend in well-structured and supervised environments to reap the general benefits of these settings. Assessing such a critical proportion, however, lies beyond the scope of our study.

The *low supervision* cluster was somewhat distinctive demographically, but not consistently so across age groups. Among elementary school youth, it was dominated by males and had comparatively low levels of maternal education. Among middle school respondents, more of its members were White and fewer were from Hispanic backgrounds than the other clusters. The lack of consistent demographic differences and the fact that some of the *low supervision* cluster members did attend our promising after-school programs made it a good criterion or reference group for our major data analyses.

The two remaining clusters had high rates of participation in our targeted after-school program, but were differentiated in terms of their level of involvement in other structured activities. A natural question is whether it is better for youth to devote most of their time to one, comprehensive after-school program (enriched by modest involvement in other supervised activities) or to divide their time more evenly among promising after-school programs and other community activities with more specific foci. We will address this question more confidently

next year, when data from the long-term follow up are available. Based on our assessment of intermediate outcomes, however, it appears as if neither group is consistently better.

Among elementary students, both program groups (*high program / high activity* and *high program / low activity*) were advantaged in comparison to the *low supervision* cluster, in child as well as staff reports of outcomes. Teachers, however, perceived more difference between the *low supervision* cluster and the *high program / low activity* cluster than the *high program / high activity* group. In fact, youth who focused their after-school hours on our target programs displayed more favorable intermediate outcomes than the *low supervision* cluster on all measures rated by teachers, whereas youth with levels of participation in an array of programs did not score significantly stronger on any of the measures rated by teachers.

Among middle school students, the comparative advantage of these two groups (relative to *low supervision* youth) was confined almost exclusively to self-reported outcomes. Both the *high program / high activity* and *high program / low activity* groups reported more favorable (and roughly comparable) changes in patterns of misconduct and substance use than the *low supervision* cluster. There was only one instance, however, in which one of the groups was distinctive from the *low supervision* group on staff or teacher ratings (teachers reported greater gains for the *high program / high activity* youth in work habits).

Collectively, these findings suggest that consistent attendance at a promising after-school program has significant short-term benefits for elementary school youth; these benefits are observable in student behavior in program as well as classroom contexts. The absence of evidence of these same benefits among middle school youth is puzzling. Do our findings suggest that most advantages of attendance in promising after-school programs accrue during the early years of schooling, or might they be attributable to the different organization of programs among youth of different age groups—for example, the more sporadic attendance patterns of older

youth? We hope to pursue this question in the next phase of our work. Likewise, we hope to pursue the issue of why focused involvement is more beneficial than diversified involvement among elementary students. Is it possible that, during the initial years of schooling, students benefit from the consistency and security of a single program structure and a single set of caring adults?

In examining findings from our intermediate analyses, it is important to keep three caveats in mind. First, our sample participation rate was relatively low, and the resulting sample was significantly different in several ways from the school population. We need to be careful not to generalize findings too readily to the populations of participating schools. Second, in most cases, effect sizes were modest. This is consistent with previous, related research, but it indicates that many of the differences we observed among clusters were relatively modest in size or scope. Coupled with this is the fact that our respondents may have had prior experience in the promising after-school programs that we studied or in other school or community activities. Students may have accrued much of the benefit of participation in these programs in the years prior to the initiation of our study. For many of our respondents, our study assesses the *additional* benefit of another year of participation, rather than the total benefit that students realize from these programs. Without a more rigid and artificial design (randomly assigning youth to programs in circumstances in which they have had no opportunity in the past to participate in organized after-school activities), it is not possible to specify the full impact of such programs on young people's development.

Next Steps

Phase 3 of the Study of Promising After-School Programs began in Fall 2004. Program directors reported on program features and provided information on the program staff. We continue to collect after-school program attendance data. In Spring 2005, we will begin to collect

longer term outcome data from students, parents, teachers, and after-school program staff.

We expect longer term outcomes to be evident in home, after-school program, and school settings, and that they will derive from skills developed during two program years. The longer term outcomes we are targeting include:

- Academic, as measured by school attendance, grades, achievement test scores, and on-time promotion
- Social, as measured by social skills at home, in the after-school program and in the classroom and positive relationships with peers
- Psychological, as measured by positive future orientation and positive emotions at home
- Behavioral, as measured by reduced engagement in misconduct or delinquent behaviors.

As we did with the cluster analyses and the intermediate effects described in this report, we will continue to investigate the relationship between *sets of experiences* and long-term outcomes. In addition, we will test the programs' effects on long-term outcomes by comparing results for program participants and comparison youth, as well as among participant youth who participate at varying levels of intensity.

After-school program quality verification visits will also be conducted during Spring 2005. We will conduct 2 days of observation at each promising program. The purpose of these activities is to examine the structural components of each program, participants' relationships with peers and program staff, program content, and content delivery strategies. These data will be used to verify that the programs continue to offer high-quality programming for participants.

References

- Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster analysis* (Sage university papers series No. 07-044). Beverly Hills, CA: Sage Publications.
- American Youth Policy Forum (1997, 1999). *Some things do make difference for youth: A compendium of evaluations of youth programs and practices*. Washington, DC: Author.
- Baker, D. & Witt, P. (1996). Evaluation of the impact of two after-school recreation programs. *Journal of Park and Recreation Administration, 14*, 23-44.
- Beckett, M., Hawken, A., & Jackowitz, A. (2001). *Accountability for after-school care: Devising standards and measuring adherence to them*. Los Angeles, CA: RAND Corporation.
- Brown, B. B., Clasen, D. R., & Eicher, S. A. (1986). Perceptions of peer pressure, peer conformity, dispositions, and self-reported behavior among adolescents. *Developmental Psychology, 22*, 521-530.
- Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (1998). *Positive youth development in the United States: Research findings on evaluation of positive development programs*. Seattle, WA: University of Washington, Social Development Research Group.
- Coie, J. D., & Dodge, K. A. (1988). Multiple sources of data on social behavior and social status in the school: A cross-age comparison. *Child Development, 59*, 815-829.
- Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*, 755-764.
- Dynarski, M., Moore, M., Mullens, J., Gleason, P., James-Burdumy, S., Rosenberg, L., et al. (2003). *When schools stay open late: The national evaluation of the 21st Century Community Learning Centers program*. Princeton, NJ: Mathematica Policy Research.
- Eccles, J. S. & Gootman, J. A. (Eds.). (2002). *Community programs to promote youth development*. Washington, DC: National Academies Press.
- Eccles, J. S., Midgley, C., Wigfeld, A., Buchanan, C. M., Reuman, D., Flanagan, C., & MacIver, D. (1993). Development during adolescence: The impact of stage-environment fit on adolescents' experiences in schools and families. *American Psychologist, 48*, 90-101.
- Fashola, O. S. (2002). *Building effective after-school programs*. Thousand Oaks, CA: Corwin Press.
- Finn, J. D., Gerber, S. B., Achilles, C. M., & Boyd-Zaharias, J. (2001). The enduring effects of small classes. *Teachers College Record, 103*(2), 145-183.
- Grossman, J. B., Price, M. L., Fellerath, V., Juvocy, L., Kotloff, L., Raley, R., & Walker, K.

- (2002). *Multiple choices after school: Findings from the extended-services schools initiatives*. Philadelphia, PA: Public/Private Ventures & Manpower Demonstration Research Corporation.
- Hamilton, V. L. & Klein, S. P. (1999). *Foundations school-age enrichment program: Evaluation of student achievement*. Washington, DC: RAND Corporation.
- Huang, D., Gribbons, B., Kim, K. S., Lee, C., & Baker, E. L. (2000). *A decade of results: The impact of the LA's Best After School Enrichment Program on subsequent student achievement and performance*. Los Angeles, CA: UCLA Center for the Study of Evaluation.
- Jöreskog, K., & Sörbom, D. (2004). LISREL 8.71 [computer program]. Chicago, IL: Scientific Software International.
- Kleiner, B., Nolin, M. J., & Chapman, C. (2004). *Before- and after-school care, programs, and activities of children in kindergarten through eighth grade: 2001* (NCES 2004-008). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Ladd, G. W., & Profilet, S. M. (1996). The Child Behavior Scale: A teacher-report measure of young children's aggressive, withdrawn, and prosocial behaviors. *Developmental Psychology*, 32(6), 1008-1024.
- MacDonald, G. B. & Valdivieso, R. (2000). Measuring deficits and assets: How we track youth development now, and how we should track it. In *Youth development: Issues, challenges and directions* (pp.149-184). Philadelphia, PA: Public/Private Ventures.
- McLaughlin, M. W. (2000). *Community counts: How youth organizations matter for youth development*. Washington, DC: Public Education Network.
- Pettit, G. S., Laird, R. D., Bates, J. E., & Dodge, K. A. (1997). Patterns of after-school care in middle childhood: Risk factors and developmental outcomes. *Merrill-Palmer Quarterly*, 43, 515-538.
- Pierce, K. M., Hamm, J. V., & Vandell, D. L. (1999). Experiences in after-school programs and children's adjustment in first-grade class classrooms. *Child Development*, 70, 756-767.
- Pittman, K., Irby, M., & Ferber, T. (2001). Unfinished business: Further reflections on a decade of promoting youth development. In *Youth development: Issues, challenges, and directions*. Philadelphia, PA: Public/Private Ventures.
- Posner, J. & Vandell, D. L. (1994). Low-income children's after-school care: Are there beneficial effects of after-school programs? *Child Development*, 65, 440-456.
- Reisner, E. R., White, R. N., Russell, C., & Birmingham, J. (2004, November 3). *Building quality, scale, and effectiveness in after-school programs*. Washington, DC: Policy Studies Associates.

- Rosenthal, R., & Vandell, D. L. (1996). Quality of care at school-aged child care programs: Regulatable features, observed experiences, child perspectives, and parent perspectives. *Child Development, 67*, 2434-2445.
- Ross, S. M., Smith, L. J., Slavin, R. E., & Madden, N. A. (1997). Improving the academic success of disadvantaged children: An examination of Success for All. *Psychology in the Schools, 34*(2), 171-180.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.
- Scales, P. C., & Leffert, N. (1999). *Developmental assets: A synthesis of the scientific research on adolescent development*. Minneapolis, MN: Search Institute.
- Schinke, S. P., Cole, K. C., & Poulin, S. R. (2000). Enhancing the educational achievement of at-risk youth. *Prevention Science, 1*, 51-59.
- Smolansky, E., & Gootman, J. A. (Eds.) (2003). *Working families and growing kids: Caring for children and adolescents*. Washington, DC: National Academies Press.
- Tierney, J., Grossman, J. B., & Resch, N. (1995). *Making a difference: An impact study of Big Brothers/Big Sisters*. Philadelphia, PA: Public/Private Ventures.
- Tout, K., Scarpa, J., & Zaslow, M. J. (2002, March). *Children of current and former welfare recipients: Similarly at risk*. Washington, DC: Child Trends.
- University of California at Irvine & California Department of Education (2002). *Evaluation of California's After School Learning and Safe Neighborhoods Partnerships Program: 1999-2001. Preliminary report*. Sacramento, CA: California Department of Education.
- Vandell, D. L. & Corasaniti, M. A. (1998). The relation between third graders' after-school care and social, academic, and emotional functioning. *Child Development, 59*, 868-875.
- Vandell, D. L. & Pierce, K. M. (1995). *Safe Haven program evaluation (1994-1995)*. A report prepared for the city of Madison, WI and the Madison Metropolitan School District.
- Vandell, D. L., Pierce, K. M., & Dadisman, K. (in press). Out-of-school settings as a developmental context for children and youth. In R. Kail (Ed.), *Advances in child development and behavior* (Vol. 33). New York: Academic.
- Vandell, D. L., & Posner, J. K. (1999). Conceptualization and measurement of children's after-school environments. In S. L. Friedman & T. D. Wachs, (Eds.) *Measuring environment across the lifespan: Emerging methods and concepts* (pp. 167-196). Washington, DC: American Psychological Association Press.
- Vandell, D. L., Reisner, E. R., Brown, B. B., Pierce, K. M., Dadisman, K., & Pechman, E. M. (2004, February). *The Study of Promising After-School Programs: Descriptive report of the promising programs*. Madison, WI: Wisconsin Center for Education Research, University of Wisconsin-Madison. Available at www.wcer.wisc.edu/childcare

Walker, K. E., & Arbreton, A. J. A. (2001). *Working together to build Beacon Centers in San Francisco*. Philadelphia: Public/Private Ventures.

Welsh, M., Russell, C., Williams, I., Reisner, E., and White, R. N. (2002, October 31). *Promoting learning and school attendance through after-school programs*. Washington DC: Policy Studies Associates.

Appendix A

Program vs. Comparison Analyses

Baseline Scores on Outcome Measures

To examine whether the program and comparison students differed at baseline, we compared their mean scores on the outcome measures reported by students and teachers, controlling for six selection factors (child gender and ethnicity, family structure and income, maternal education and employment). Results of these analyses are presented in Table A-1 for both elementary and middle school students. There was only one significant difference between the program and comparison groups: At the elementary level, teachers rated the program students as having poorer social skills at baseline than the comparison students.

Assessment of Intermediate Outcomes

We employed HLM procedures to examine whether students in the program and comparison groups differed on the first-year outcomes. In these analyses, we controlled for background variables (child gender and ethnicity, family income and structure, maternal education and employment), Fall 2003 baseline scores, and time spent in other after-school contexts (organized activities, home alone or with siblings, hanging out with peers). There were no significant differences between the program and comparison groups in either the elementary or middle school samples in the Spring 2004 outcomes. We did find significant associations between the outcomes and time spent in after-school contexts other than the selected after-school programs, however. In the **elementary** sample (see Table A-2), students who spent more after-school time hanging out with peers reported poorer work habits and more misconduct compared to students who spent less time with peers, and according to their teachers, they exhibited poorer work habits, less task persistence, and less prosocial behavior with peers. Elementary students who spent more unsupervised time at home alone or with siblings after school, compared to

students who spent less time with siblings, had greater self-reported misconduct and poorer teacher-reported work habits, social skills, and academic performance at the first-year follow up. More time spent in organized after-school activities was associated with better self-reported work habits and more misconduct.

In the **middle school** sample (see Table A-3), students who spent more time hanging out with peers during the after-school hours, compared to students who spent less time with peers, reported that they had poorer work habits and self-efficacy, and more misconduct and substance use, and their teachers reported that they had poorer social skills. Middle school students who spent more time home alone or with siblings reported that they used substances more often than students who spent less time at home unsupervised by adults. More time spent in organized after-school activities was associated with better self-reports of work habits and self-efficacy, but poorer teacher reports of social skills.

Table A-1
Baseline Scores for Program and Comparison Students

	Elementary			Middle		
	Program <i>N</i> = 1017	Comparison <i>N</i> = 803		Program <i>N</i> = 540	Comparison <i>N</i> = 579	
	<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)		<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)	
Student report						
Work habits	3.42 (.02)	3.44 (.02)	$F_{(1,1249)} =$ 0.26	3.22 (.03)	3.21 (.03)	$F_{(1,787)} =$ 0.01
Self-efficacy	--	--		3.17 (.03)	3.15 (.02)	$F_{(1,785)} =$ 0.19
Misconduct	0.45 (.02)	0.45 (.02)	$F_{(1,1251)} =$ 0.00	0.57 (.03)	0.50 (.03)	$F_{(1,791)} =$ 3.24
Substance use	--	--		0.09 (.02)	0.07 (.02)	$F_{(1,792)} =$ 0.73
Teacher report						
Work habits	3.35 (.04)	3.45 (.05)	$F_{(1,1093)} =$ 2.50	3.53 (.06)	3.55 (.05)	$F_{(1,750)} =$ 0.05
Task persistence	2.92 (.03)	2.96 (.03)	$F_{(1,1091)} =$ 0.48	2.97 (.04)	3.02 (.03)	$F_{(1,748)} =$ 1.26
Social skills	3.33 (.04)	3.46 (.04)	$F_{(1,1090)} =$ 5.52*	3.48 (.05)	3.52 (.05)	$F_{(1,722)} =$ 0.29
Aggressive w/peers	0.29 (.02)	0.26 (.02)	$F_{(1,1086)} =$ 1.34	0.24 (.02)	0.26 (.02)	$F_{(1,746)} =$ 0.29
Prosocial w/peers	1.51 (.02)	1.54 (.02)	$F_{(1,1086)} =$ 0.70	1.51 (.02)	1.48 (.02)	$F_{(1,744)} =$ 0.40
Academic performance	2.79 (.04)	2.76 (.04)	$F_{(1,1088)} =$ 0.23	2.93 (.06)	3.03 (.06)	$F_{(1,748)} =$ 1.60

Note. Scores are adjusted for child gender and ethnicity, family structure, maternal education and employment, and family income. The asterisk indicates a significant difference between the program and comparison groups within school level at $p < .05$.

Table A-2
HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Program versus Comparison: Elementary

	Student report		Teacher report					
	Work habits	Misconduct	Work habits	Task persistence	Social skills	Aggressive w/peers	Prosocial w/peers	Academic performance
FIXED EFFECT								
Intercept	2.297***	.065	1.287***	1.100***	1.769***	.085	.871***	1.030***
Gender (1 = female)	.103***	-.164***	.172***	.054	.154***	-.070***	.089***	.055
White	.000	.016	.064	.008	-.069	.025	.009	-.038
Black	-.028	-.045	-.078	-.019	-.166	.093	-.064	-.057
Hispanic	-.022	.002	.084	.075	.007	.033	-.012	.014
Family income	.001	-.001	.006	.005	.014	.001	-.000	.016
Family structure (1 = two parents)	.020	-.065*	.054	.050	.025	-.044	.022	.042
Maternal education	-.014	.021	-.005	.009	.000	.006	-.011	.021
Maternal work (1 = full time)	-.001	-.032	.023	.037	-.023	-.013	-.013	.021
Fall 2003 baseline	.300***	.368***	.648***	.616***	.551***	.611***	.473***	.711***
Program (1 = program)	.003	-.014	-.008	.006	-.033	-.003	.013	.002
Other organized activities	.060**	.047*	-.010	-.009	-.045	.007	.005	-.018
Home alone/with siblings	-.022	.066**	-.071*	-.039	-.091**	.031	-.025	-.077*
Hang out with peers	-.077***	.135***	-.056*	-.049**	-.027	.023	-.027*	-.039
Program x gender	.019	-.062	.120	.074	.197**	-.033	.058	.007
Program x family structure	-.078	.051	-.008	-.001	-.032	.022	-.022	-.038
Program x maternal work	.014	-.023	.046	-.004	-.012	.032	-.036	.037
RANDOM EFFECT								
Site mean, variance component ¹	.005**	.004***	.011***	.001	.016***	.001**	.003***	.027***
Level-1 effect, variance component ²	.300	.234	.562	.269	.541	.130	.142	.478

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

¹ Indicates whether there are differences between sites (schools) in terms of degree of change.

² Indicates the amount of residual (unexplained) variance within groups.

Table A-3
HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Program versus Comparison: Middle School

	Student report				Teacher report					
	Work habits	Self-efficacy	Misconduct	Substance use	Work habits	Task persistence	Social skills	Aggressive w/peers	Prosocial w/peers	Academic performance
FIXED EFFECT										
Intercept	1.655**	1.759***	.222*	-.078	1.165***	1.094***	1.377***	.236**	.674***	1.067***
Gender (1 = female)	.064*	-.018	-.065*	.010	.235***	.106**	.151**	-.099***	.100***	.056
White	.038	.028	-.012	.053	.040	-.003	-.105	.006	-.034	.121
Black	-.085	-.030	.049	.041	-.018	-.026	-.034	.027	.019	-.021
Hispanic	-.029	.007	.018	.047	.057	-.028	.014	-.019	.005	.024
Family income	-.000	.002	-.004	-.000	.023*	.007	.026*	-.012*	.021***	.028*
Family structure (1 = two parents)	-.035	-.001	-.037	-.009	.011	.005	.015	-.013	-.020	.011
Maternal education	.004	.018	-.002	-.009	-.006	.016	.017	-.007	.006	.008
Maternal work (1 = full time)	-.013	-.027	-.032	-.007	-.002	-.023	.019	.009	-.030	-.042
Fall 2003 baseline	.463***	.406***	.434***	.268***	.622***	.604***	.589***	.617***	.493***	.627***
Program (1 = program)	-.019	-.031	-.017	.018	.014	.001	-.039	.018	-.014	-.031
Other organized activities	.054*	.059**	.008	.005	-.014	-.011	-.092*	.035	-.035	-.039
Home alone/with siblings	.022	.026	.006	.028*	-.043	.008	.032	-.014	.006	-.022
Hang out with peers	-.069***	-.030*	.130***	.038***	-.013	-.008	-.056*	.020	-.024	-.020
Program x gender	.038	.049	-.064	-.004	.038	.118	.034	-.033	-.043	.137
Program x family structure	.040	.034	-.041	-.019	-.077	.039	-.036	-.055	.007	-.070
Program x maternal work	-.058	-.026	.031	-.027	-.188	-.105	-.211*	.038	-.113*	-.176
RANDOM EFFECT										
Site mean, variance component ¹	.004**	.002	.000	.000	.025***	.005**	.042***	.001*	.005***	.015***
Level-1 effect, variance component ²	.220	.176	.216	.086	.554	.248	.475	.111	.136	.539

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

¹ Indicates whether there are differences between sites (schools) in terms of degree of change.

² Indicates the amount of residual (unexplained) variance within groups.

Appendix B

Program Dosage Analyses

We examined the effects of varying amounts of program attendance, or dosage, across the 2003-04 school year on the first-year outcomes, utilizing HLM procedures. As detailed in Chapter 2, we stipulated five dosage levels: minimal (0-4 days of attendance), low (5-29 days), moderate (30-59 days), substantial (60-89 days), and high (90 or more days). The minimal group was omitted from the analyses because our goal was to examine the effects of dosage for the program participants only. In the analyses, we controlled for background variables (child gender and ethnicity, family income and structure, maternal education and employment), Fall 2003 baseline scores, and time spent in other after-school contexts (organized activities, home alone or with siblings, hanging out with peers). Results of these analyses are shown in Tables B-1 (elementary student and teacher reports), B-2 (middle school student and teacher reports), and B-3 (both elementary and middle school after-school program staff).

Program dosage was not associated with student and teacher reports of outcomes in either the elementary or the middle school sample (see Tables B-1 and B-2). According to after-school program staff (see Table B-3), however, both elementary and middle school students exhibited increased aggressive behaviors with peers when they attended the programs more often during the school year.

There were some significant associations between the Spring 2004 outcomes and time spent in after-school contexts other than the selected after-school programs. In the **elementary** sample (see Table B-1), students who spent more after-school time hanging out with peers, compared to students who spent less time with peers, reported poorer work habits and more misconduct. According to their teachers, they exhibited less task persistence; program staff reported them to have poorer work habits, task persistence, and social skills, and to exhibit fewer

prosocial behaviors and more aggressive behavior with peers. Elementary students who spent more unsupervised time at home alone or with siblings after school, compared to students who spent less time in this context, reported more misconduct; their teachers reported that they had poorer work habits and social skills, and program staff reported poorer work habits. More time spent in other organized after-school activities was associated with better self-reported work habits and more misconduct, but was not associated with teacher and program staff reports.

Associations between time spent in other after-school contexts and outcomes in the **middle school** sample were less evident than in the elementary sample (see Table B-2). Middle school students who spent more unsupervised time with peers after school, compared to students who spent less time with peers, reported that their work habits were poorer at the end of the school year, and they engaged in more misconduct and substance use. Time spent in other organized activities was associated with teacher reports of poorer social skills and less prosocial behavior with peers. There were no significant associations between time in other contexts and program staff reports of middle school student outcomes.

Table B-1
HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Program Dosage: Elementary

	Student report		Teacher report					
	Work habits	Misconduct	Work habits	Task persistence	Social skills	Aggressive w/peers	Prosocial w/peers	Academic performance
FIXED EFFECT								
Intercept	2.225***	.065	1.575***	1.256***	2.102***	-.078	.998***	1.065***
Gender (1 = female)	.102**	-.187***	.227***	.084*	.241***	-.089***	.114***	.064
White	-.108	.036	-.141	-.060	-.205	.097	-.033	-.190
Black	-.124	-.060	-.271	-.066	-.297	.162	-.088	-.106
Hispanic	-.133	.017	-.144	-.022	-.200	.130	-.092	-.085
Family income	-.007	.001	-.002	.001	.001	.007	-.004	.014
Family structure (1 = two parents)	-.008	-.041	.045	.050	.013	-.034	.012	.020
Maternal education	.005	-.001	-.010	.011	.000	.005	-.013	.029
Maternal work (1 = full time)	.015	-.039	.066	.043	.004	-.011	-.017	.038
Fall 2003 baseline	.333***	.385***	.636***	.597***	.532***	.598***	.466***	.713***
Program dosage	-.007	.003	-.022	-.008	-.035	.019	-.010	.003
Other organized activities	.121***	.054*	.014	.003	-.023	.004	.023	-.016
Home alone/with siblings	-.039	.076*	-.099*	-.045	-.119**	.042	-.044	-.076
Hang out with peers	-.083***	.127***	-.041	-.050*	-.033	.018	-.030	-.030
Dosage x gender	.006	-.016	.084	.038	.084*	-.037	.025	.022
Dosage x family structure	-.022	-.036	-.001	.018	.024	-.009	.002	-.045
Dosage x maternal work	.001	.002	-.042	-.047	-.026	-.005	-.001	-.014
RANDOM EFFECT								
Site mean, variance component ¹	.004*	.004*	.017***	.002	.016***	.001	.002	.025***
Level-1 effect, variance component ²	.303	.230	.570	.276	.552	.131	.142	.479

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

¹ Indicates whether there are differences between sites (schools) in terms of degree of change.

² Indicates the amount of residual (unexplained) variance within groups.

Table B-2
HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Program Dosage: Middle School

	Student report				Teacher report					
	Work habits	Self-efficacy	Misconduct	Substance use	Work habits	Task persistence	Social skills	Aggressive w/peers	Prosocial w/peers	Academic performance
FIXED EFFECT										
Intercept	1.678***	1.932***	.147	-.117	1.486***	1.215***	1.659***	.232	.810***	1.186***
Gender (1 = female)	.083	.002	-.092*	.010	.254***	.170***	.158*	-.112**	.073	.132*
White	.116	.004	.077	.079	-.052	.012	-.283	.017	-.067	.075
Black	-.046	-.137	.141	.025	-.165	-.072	-.338*	.021	.012	-.199
Hispanic	.043	-.090	.118	.046	-.041	-.041	-.167	-.037	.002	-.113
Family income	.007	.003	-.006	.004	.012	.002	.025	-.016*	.020*	.023
Family structure (1 = two parents)	-.028	.010	-.055	-.025	-.002	.037	-.018	-.041	-.016	-.015
Maternal education	-.014	-.009	.010	-.012	-.012	.019	.015	-.004	.001	.001
Maternal work (1 = full time)	-.048	-.043	-.011	-.029	-.083	-.068	-.083	.033	-.081	-.126
Fall 2003 baseline	.453***	.416***	.430***	.246***	.618***	.586***	.597***	.611***	.514***	.619***
Program dosage	-.031	-.018	.017	.013	-.042	-.026	-.049	.012	-.017	-.010
Other organized activities	.057	.040	-.036	-.000	-.022	-.044	-.112*	.043	-.060*	-.063
Home alone/with siblings	.027	.014	.006	.027	-.066	.005	.042	-.003	-.003	.001
Hang out with peers	-.069**	-.024	.133***	.061***	.026	.022	-.034	.019	-.035	.025
Dosage x gender	-.075*	-.028	.040	-.004	.106	.083*	.041	-.028	.036	.072
Dosage x family structure	.041	-.017	.026	.021	.022	-.020	-.002	-.023	-.028	.001
Dosage x maternal work	.053	.047	-.026	-.046	-.038	-.012	-.055	.014	-.040	-.037
RANDOM EFFECT										
Site mean, variance component ¹	.006	.003	.001	.000	.024**	.002	.019**	.006***	.008***	.015*
Level-1 effect, variance component ²	.242	.188	.238	.119	.569	.258	.492	.112	.141	.505

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

¹ Indicates whether there are differences between sites (schools) in terms of degree of change.

² Indicates the amount of residual (unexplained) variance within groups.

Table B-3
HLM Analyses of Relative Change from Baseline to First-Year Follow Up, Program Dosage: Program Staff Reports

	Elementary					Middle				
	Work habits	Task persistence	Social skills	Aggressive w/peers	Prosocial w/peers	Work habits	Task persistence	Social skills	Aggressive w/peers	Prosocial w/peers
FIXED EFFECT										
Intercept	2.356***	1.933***	2.323***	.107	.770***	2.422***	1.705***	2.200***	.179	.771***
Gender (1 = female)	.253***	.115***	.203***	-.101***	.139***	.291***	.139**	.262***	-.119***	.144***
White	.070	.057	.086	-.036	.100	-.064	.037	-.008	-.008	.063
Black	-.017	.006	.002	-.040	.073	-.392	-.032	-.159	.002	.020
Hispanic	.050	.048	.101	-.043	.067	-.227	.019	-.004	-.000	.058
Family income	-.005	.008	.002	.004	.002	.009	.003	.029*	-.003	.007
Family structure (1 = two parents)	.078	-.016	.044	-.018	.017	.068	.103	.026	-.046	.027
Maternal education	-.020	-.018	-.017	.011	.001	.021	.030	.009	-.010	.003
Maternal work (1 = full time)	.070	.015	.008	.006	.012	-.089	-.069	-.091	.007	-.007
Fall 2003 baseline	.363***	.373***	.325***	.390***	.356***	.300***	.355***	.265***	.507***	.276***
Program dosage	-.008	-.004	.023	.029*	.015	.002	-.026	.014	.043**	.005
Other organized activities	-.004	-.019	-.025	.018	-.000	.053	.041	.023	-.005	.013
Home alone or w/siblings	-.102*	-.049	-.074	-.017	-.029	-.026	-.010	.002	.018	.016
Hang out with peers	-.079*	-.047*	-.070*	.035*	-.048**	-.054	-.019	-.013	.002	-.017
Dosage x gender	.116**	.049	.095*	-.053**	.049*	.018	-.040	.001	-.016	.004
Dosage x family structure	.044	.010	.041	-.017	.019	-.046	-.058	-.021	.020	-.030
Dosage x maternal work	-.021	-.012	.021	.009	-.003	.020	-.025	.002	.012	.010
RANDOM EFFECT										
Site mean, variance component ¹	.017**	.005*	.043***	.002*	.007***	.031***	.019***	.053***	.000	.042***
Level-1 effect, variance component ²	.621	.256	.468	.140	.133	.611	.245	.427	.112	.130

* $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

¹ Indicates whether there are differences between sites (schools) in terms of degree of change.

² Indicates the amount of residual (unexplained) variance within groups.