When Schools Stay Open Late: The National Evaluation of the 21st-Century Community Learning Centers Program

First Year Findings

2003

U.S. DEPARTMENT OF EDUCATION OFFICE OF THE UNDER SECRETARY

When Schools Stay Open Late: The National Evaluation of the 21st-Century Community Learning Centers Program

First Year Findings

2003

Prepared by:

Mark Dynarski
Mary Moore
John Mullens
Philip Gleason
Susanne James-Burdumy
Linda Rosenberg
Carol Pistorino
Tim Silva
John Deke
Wendy Mansfield
Sheila Heaviside
Daniel Levy

Mathematica Policy Research, Inc.
Decision Information Resources, Inc.

U.S. DEPARTMENT OF EDUCATION **OFFICE OF THE UNDER SECRETARY**

This report was prepared for the U.S. Department of Education under Contract No. ED-99-CO-0134. The project monitor was Elizabeth Warner in the Office of the Under Secretary. The Charles Stewart Mott Foundation also contributed to the evaluation under Grant No. 20-205. The views expressed herein are those of the contractor. No official endorsement by the U.S. Department of Education is intended or should be inferred.

U.S. Department of Education
Rod Paige
Secretary

Office of the Under Secretary
Eugene Hickok
Under Secretary

January 2003

To order copies of this report, write to:

ED Pubs
Education Publications Center
U.S. Department of Education
P.O. Box 1398
Jessup, MD 20794-1244;

or **fax** your request to: (301) 470-1244;

or **e-mail,** send your request to edpubs@inet.ed.gov;

or **call** in your request toll-free: 1-877-433-7828 (1-877-4-ED-PUBS). If 877 service is not yet available in your area, call 1-800-872-5327 (1-800-USA-LEARN); Those who use a telecommunications device for the deaf (TDD) or a teletypewriter (TTY), should call 1-800-437-0833.

Or **order online,** at: www.ed.gov/about/ordering.jsp.

This report is also available on the Department's Web site at:
http://www.ed.gov/pubs/21cent/firstyear.

On request, this publication is available in alternate formats, such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at (202) 260-9895 or (202) 205-8113.

Contents

Figures

Figure **Page**

Tables

Table **Page**

Table		Page

Acknowledgments

This report resulted from the combined efforts of researchers, data collection experts, and school staff who are too numerous to all be thanked by name. The authors want to recognize Elizabeth Warner for her encouragement and support and for her incisive review as the report went from draft to final. Also, we wish to acknowledge An-Me Chung and the C.S. Mott Foundation for support that enabled us to broaden the focus and refine the conceptualization of the evaluation. We also thank Alan Ginsburg, David Goodwin, Ricky Takai, Audrey Pendleton, Robert Stonehill, and Adriana de Kanter of the Department of Education for helpful comments on the design of the evaluation and drafts of this report, and staff of the 21st-Century program office at the U.S. Department of Education for their assistance and support in launching the evaluation.

The evaluation's Technical Working Group has offered constructive advice and comments that have helped shape the evaluation's design and the report. The group includes Joan Bissell, Christopher Cross, Jennifer Davis, Aaron Pallas, Michael Puma, Elizabeth Reisner, Deborah Vandell, Constancia Warren, and Richard Weissbourd.

Keith Appleby, Heather Berry, Jan Fertig, Annette Luyegu, Jessica Wilkins, Valerie Williams, Claire Wilson, and Julie Young assisted in managing the data collection. Richard Heman-Ackah, Larry Snell, and Marianne Stevenson, and their interviewing and data collection staff at phone centers in Columbia, Md., and Princeton, N.J., put in many hours of effort to collect the data. We also thank Anne Bloomenthal, Arianna Freeman, Holly Gerhart, Josh Hart, Leonard Hart, Ece Kalay, Nazmul Khan, Barbara Kolln, Sophia Kuan, Marcel Paret, Sedhou Ranganathan, and Angela Richardson for information systems and programming support, and Lara Hulsey for her steady managerial assistance.

Many school district and after-school program staff assisted the evaluation by providing data and by assisting in site visits. We appreciate their willingness to respond to many requests and to contribute their perspectives and insights during interviews.

The report was produced by Jill Miller with assistance from Cindy McClure and was edited by Roy Grisham.

When Schools Stay Open Late:
The National Evaluation of the 21st-Century Community Learning Centers Program

Summary of First-Year Findings

In an era when most parents work, many Americans want their children to have access to safe and supervised after-school activities that can help develop academic, personal, and social skills. In 1994, Congress authorized the 21st-Century Community Learning Centers (21st-Century) program to open up schools for broader use by their communities. In 1998, the program was refocused on supporting schools to provide school-based academic and recreational activities after school and during other times when schools were not in regular session, such as on weekends, holidays, and during summers. As an after-school program, 21st-Century grew quickly from an appropriation of $40 million in fiscal year 1998 to $1 billion in fiscal year 2002. It now supports after-school programs in about 7,500 rural and inner-city public schools in more than 1,400 communities. Programs operate in public school buildings and offer academic, recreational, and cultural activities during after-school hours. A distinguishing characteristic of 21st-Century programs is the inclusion of academic activities. Grants made after April 1998 included a requirement that programs include academic activities.

This study, conducted for the U.S. Department of Education with support for additional data collection and analysis from the Charles Stewart Mott Foundation, presents the first-year findings of the largest and most rigorous examination to date of school-based after-school programs.[1] The study was designed to examine the characteristics and outcomes of typical programs and did not attempt to define or identify the characteristics of the best programs. Programs selected to be in the study operated in elementary and middle schools. Some were in their second year of funding when the study began collecting data and others were in their third year of funding. Most grantees that were part of the study had operated some type of after-school program before receiving a 21st-Century grant and were using their grant funds to expand or modify their services and activities. About 65 percent of middle school grantees and about 57 percent of elementary school grantees in the study had operated after-school programs in one or more schools that were part of the 21st-Century grant.

The study currently is collecting another year of follow-up data and has expanded to include more programs serving elementary school students. The additional data from the second follow-up year and from the newly included programs will be the basis for two future reports. The first will update the findings for middle school students using another year of follow-up data and will present first year findings for elementary school students using a larger number of elementary school programs. The second will update the findings for elementary school students using another year of follow-up data.

[1]This study focuses on school-based programs that are part of the 21st-Century program. Results do not extrapolate to all after-school programs in general.

Key Impact Findings

The first-year findings reveal that while 21st-Century after-school centers changed where and with whom students spent some of their after-school time and increased parental involvement, they had limited influence on academic performance, no influence on feelings of safety or on the number of "latchkey" children and some negative influences on behavior.[2] In brief, the key findings are:

- ***Limited Academic Impact.*** At the elementary school level, reading test scores and grades in most subjects were not higher for program participants than for similar students not attending the program. In addition, on average, programs had no impact on whether students completed their homework or completed assignments to their teacher's satisfaction.

  For middle school students, grades in most subjects were not different than for similar students not attending the 21st-Century program. Grades for math were higher for 21st-Century participants, but the overall difference was small. A subgroup analysis found larger grade point improvements for black and Hispanic middle school students and their teachers also reported less absenteeism and tardiness compared with nonparticipants. Teachers for middle school students were more likely to say assignments were completed to their satisfaction, although program participants were not more likely to do or complete the homework assigned. Another subgroup analysis found that students who attended programs more frequently, both at the middle school and elementary school levels, did not have higher academic outcomes compared with students that attended less frequently. Other analyses did not find statistically significant relationships between program characteristics, including program maturity, and academic impacts.

- ***Adult Care Increased but Self-Care Unaffected.*** The findings indicate that programs reduced the proportion of students being cared for by parents and by older siblings, and increased the proportion of students being cared for by non-parent adults. The net effect was to increase the proportion of students being cared for by an adult (either a parent or a non-parent adult), by reducing the proportion being cared for by an older sibling.

  Programs did not reduce the percentage of students in self-care (who are commonly referred to as "latchkey" children). Students were defined to be in self-care if they (or their parents, for elementary school students in grades K-2) indicated that they were not in the presence of adults or older siblings after school (they were by themselves, with others their age, or with younger siblings after school). Other definitions of self-care, such as whether students ever said they were by themselves after school, were analyzed with similar results. The most common care arrangement for nonparticipants was for students to go home after school and be cared for by a parent, which was true for about 53 percent of middle school students in the comparison group and 67 percent of elementary school students in the control group.

[2] A "center" refers to after-school services operated in one school, and a "program" refers to one or more centers operated in one school district. The study measured impacts at the program level but not at the center level.

- ***No Improvements in Safety and Behavior.*** Programs did not increase students' feeling of safety after school. At the middle school level, participants were more likely to report that they had sold drugs "some" or "a lot" and somewhat more likely to report that they smoked marijuana "some" or "a lot" (although the incidence was low). Participants also were more likely to have had their property damaged. (Data on these items were not collected for elementary school students.) No impacts were found on other measures of behavior.

- ***Increased Parental Involvement.*** At the middle school level, programs were associated with increased parent involvement at their child's school. Parents of program participants were more likely to volunteer at their child's school and attend open houses or parent-teacher organization meetings. Parents of elementary school level program participants were more likely to help their child with homework or ask about things they were doing in class.

- ***Negligible Impact on Developmental Outcomes.*** Programs had no impacts on developmental outcomes, such as whether students felt they were better able to plan, set goals, or work with a team. At the middle school level, program participants were less likely to rate themselves as "good" or "excellent" at working out conflicts with others.

Key Implementation Findings

The first-year findings indicate that grantees generally had succeeded in implementing their planned programs and in gaining support from and creating working relationships with school principals and teachers. Most programs provided academic, enrichment, and recreation activities, with homework help being the most common academic activity. The mix across the three activity areas varied according to locally determined needs and preferences. A few programs focused only on providing academic activities, but none focused only on providing recreational activities. The federal grant and other funding sources enabled programs to spend about $1,000 for each student enrolled during the school year, equivalent to about a 16 percent increase in education spending. Other implementation findings include:

- ***Low Levels of Student Participation.*** Attendance in the programs was low, averaging less than two days a week, despite the fact that programs typically were available to participants four to five days a week.

- ***Programs Staffed Predominantly by School-Day Teachers.*** A third of the program coordinators and three out of five program staff members were school-day teachers. To accommodate the varying schedules and requirements of teachers, staff members often worked only a few days a week and for short periods.

- ***Limited Efforts to Form Partnerships and Plan for Sustainability.*** Programs did not collaborate much with other community organizations. In general, centers contracted with community agencies to provide specific after-school sessions rather than as partners with shared governance or combined operations. Programs also were slow to begin planning to sustain themselves after the 21st-Century grant ends. Even among those grantees within months of their grant's end, sustainability planning was almost nonexistent.

Overall, the findings suggest that policymakers and program developers need to consider ways to address low student participation and low academic content. Considering program structures that would facilitate more frequent attendance, such as focusing on serving students having difficulty in reading or math and asking them to participate a minimum number of days each week, may be worth considering. Efforts to increase the academic content and quality of activities also may be fruitful. Especially for middle school students, the challenge will be how to both attract students and help students improve their academic performance.

Methodology

While research has evaluated other after-school programs, this study—conducted by Mathematica Policy Research, Inc. (MPR) and its partner, Decision Information Resources, Inc.—is one of the few that is consistent with the principles of scientifically based research set out in the recent No Child Left Behind Act. The study is unique in the large number of after-school programs that were included and in its application of rigorous techniques for measuring impacts.

The evaluation's design includes a middle school study and an elementary school study. The middle school study is based on a nationally representative sample of after-school programs and participants and a matched comparison group of students which is similar to the program participant group. Similar students were identified in host schools or in other schools in the participating districts. Thirty-four school districts and 62 centers in the districts are included in the study.

The elementary school study uses random assignment of students to treatment and control groups. The study involved 14 school districts and 34 centers. Results presented here are from seven school districts selected in the first year of the study; another seven school districts were added in the second year of the study and data currently are being collected in these districts. The elementary school programs that were part of the study appear to be typical of elementary school 21st-Century programs along most dimensions (although they tended to be more urban and served a larger percentage of minority students than the average elementary program). However, caution should be exercised in applying the findings to all elementary school programs. Programs in the study had more applicants for their slots than they could serve, which facilitated the use of an experimental design, but the programs were not statistically sampled.

The findings presented in this report are based on one year of data collected in school year 2000-2001 from students, parents, teachers, principals, program staff members, and school records. Evaluators collected baseline and follow-up data for 4,400 middle school students and 1,000 elementary school students, and conducted site visits, lasting between two and four days, to all grantees at least once. MPR is continuing to study the programs and will prepare two additional reports based on another year of follow-up data and another round of visits to each program.

General Information about 21st-Century Programs

Annual performance reports submitted by grantees to the U.S. Department of Education indicate general characteristics and context of 21st-Century programs. The reports also are

informative about centers in the study. Nationwide, the average grantee ran three or four centers that together reported enrollment of almost 700 students over the course of the school year. Attendance varied by day, with some students attending regularly and others more occasionally, and with students enrolling and exiting from the program at different points during the year. Fifty-seven percent were minority students, compared with 37 percent of students nationwide. Most centers (95 percent) were located in elementary or middle schools or located in schools that included some combination of K-8. Typically, centers were open 10 or more hours a week, after school, and a third were open 20 hours or more a week. Some were open on Saturdays, and many offered summer programs. Sixty-six percent of host schools were considered high-poverty (at least half their students were eligible for free or reduced-price lunches). Nationally, 17 percent of schools are high-poverty. Center budgets averaged about $196,000 a center, or about $1,000 per enrolled student, with the 21st-Century grant accounting for about 70 percent of budgets. Programs typically were free both for students and parents.

The rest of this summary looks at findings for middle school programs, then at findings for elementary school programs. These findings are based on the various samples that were drawn by this study. We present the findings for middle and elementary schools separately because of differences in how the programs were selected for the evaluation and how impacts were measured.

Findings for Middle School Programs

Middle school centers in the study usually offered the following activities:

- ***Academic help,*** primarily supervised daily homework sessions. Nearly 9 of 10 middle school centers (89 percent) provided homework help. Slightly more than half (54 percent) provided homework help and other academic support, such as tutoring, state test preparation, and help sessions in reading, writing, and math skills. Help sessions usually were scheduled between one and three days a week, staffed by certified teachers, and targeted to particular students, such as those referred by a classroom teacher or those performing poorly on state standardized tests.

  In spite of the focus on homework support, fewer than two in five students (38 percent) said that the centers were a good place to get homework done. Consistent with this finding, site visitors observed that homework sessions usually were organized with students in large groups proctored by teachers or other staff members, with students talking to each other and staff members not checking the homework for quality or completeness.

- ***Recreation activities,*** such as using the gym, playing board games, or using computers. These often were part of the daily student fare although content varied according to the day.

- ***Cultural and interpersonal enrichment,*** including crafts, drama, music, mentoring, role modeling and conflict resolution, and issue forums. These activities were offered most days of the week but not necessarily every day. Specific activities might occur just once or twice during the week.

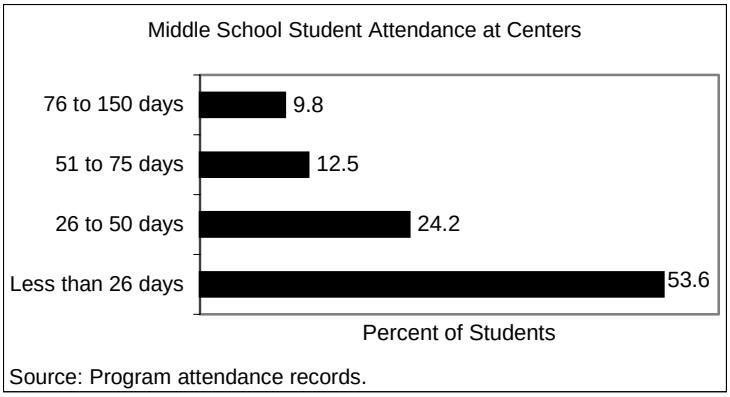
Characteristics of Staff in Middle School Centers

- ***Average Student-Staff Ratio across Centers:*** About 11-to-1. Academic activities had much lower ratios than recreational activities.
- ***Average Work Week:*** For coordinators, four to five days a week, five hours a day. For other staff members, three days a week, three hours a day, often in cycles and not continuously throughout the school year. About a third of coordinators and three-fifths of other staff members were teachers.
- ***Compensation:*** Fifty-five percent of middle school coordinators were paid by the hour, with an average hourly wage of about $17. Most other staff members also were paid by the hour, with an average hourly wage of about $16.

SOURCE: Survey of program staff for grantees in the national evaluation. Staff in the elementary school centers that were part of the national evaluation had similar characteristics.

Management and Staffing

Officials from the host school or district oversaw most middle school programs. Program directors usually had supervisory and administrative roles, while program coordinators handled day-to-day details of the centers, such as recruitment, scheduling, staffing, parent and community outreach, and attendance monitoring. Nearly all other staff members were directly involved in student activities or instruction and spent most of their time working with students. Survey data showed that middle school teachers believed that, as a result of working with students at the centers, they improved their teaching skills and had better relationships with some students.



Source: Program attendance records.

Student Participation

Middle school students in the study attended centers for 32 days—about one day a week—during the 2000-2001 school year. More than half attended for fewer than 25 days, a quarter attended for more than 50 days, and almost 10 percent attended for more than 75 days (see box). Program staff attributed the low attendance to the lack of interesting or appealing activities and to competition from other organized activities, especially sports. Center policies also made it

easy not to attend—many allowed students to participate on a drop-in basis, choosing each day whether or not to participate.

Not all students chose to participate in 21st-Century programs. Students who had chosen not to participate (surveyed in six selected programs) said that they would rather "hang out" after school, were involved in other organized activities after school, or were not interested in the activities. Almost half of the students thought the centers were "mostly a place kids go when their parents are at work," and a quarter considered them "just for kids who need help in school." Participants who had stopped attending echoed these sentiments.

A Typical Middle School Center

The center is open four days a week for two and a half hours a day. About 60 students participate on a given day. Activities begin with a homework session at 2:30 p.m., when the regular school day ends. Homework sessions are held in regular classrooms in one wing of the school. To participate in other recreational and enrichment activities, students must attend the homework sessions. In these sessions, students eat a snack provided by the program and work on their assignments. Each session has about 15 students and a teacher. Homework time ends at 3:45 p.m., and students then participate in a mix of recreational and enrichment activities. The center's activities include table tennis, Pep Club, tennis, golf, and board games. Enrichment activities include classes in martial arts, cooking, and choral music. Some activities, such as martial arts classes, are popular and are scheduled throughout the year. Others, such as cooking, change every 12 weeks to reflect changing student interest. The center's activities end at 5 p.m. and students go home on school buses.

Learning Outcomes

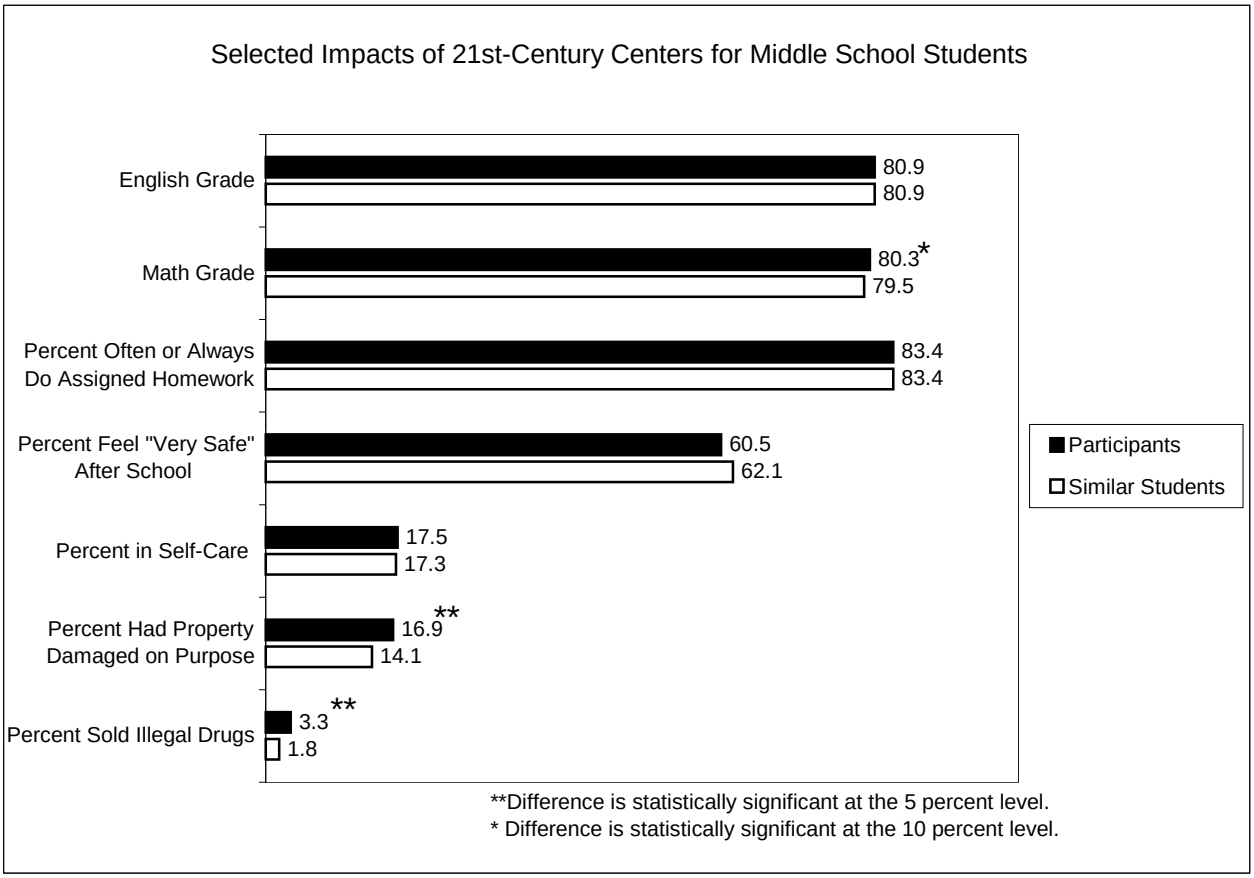
The objective of improving learning outcomes distinguished 21st-Century after-school programs, and more than 75 percent of parents of participants said they believed participation would help their child do better in school. However, participants were just as likely as comparison group students to complete homework, although they were more likely to do so to their teachers' satisfaction, and participants had about the same English, science, and social studies or history grades as similar students. Participants had slightly higher math grades (see box on next page), and slightly higher school attendance.

Additional Analyses and Other Outcomes

The evidence on the effect of programs on student effort in school is mixed. According to teachers, program students were more likely than similar students to try hard in reading or English class, be attentive in class, and participate and volunteer in class. However, teachers also report similar rates of frequent homework completion for program participant and nonparticipants. In addition, program participants report spending a similar number of hours watching TV.

Another program objective was to reduce students' exposure to unsafe settings. However, programs did not increase the extent to which students felt safer after school, and, although rates were not high, participants were more likely to report that they sold drugs, smoked marijuana, and, especially for girls, had their personal property damaged or were "picked on." Other measures of behavior—such as suspensions, absences, and teacher reports of discipline problems—were the same in both groups.

In general, program participation did not change students' interpersonal skills. Program students were no more likely to report getting along with others their age, feeling included, being good at working with others in a team, or setting a goal and working to achieve it. In fact, middle school participants were less likely their nonparticipant peers to rate themselves as good or excellent at working out conflicts with others.



NOTE: Reported impacts were estimated using regression models to adjust for baseline differences between program participants and the similar students. The adjustment variables in the regression models included student demographic characteristics, household socioeconomic status, and students' baseline test scores, attendance, disciplinary problems, and self-reported grades.

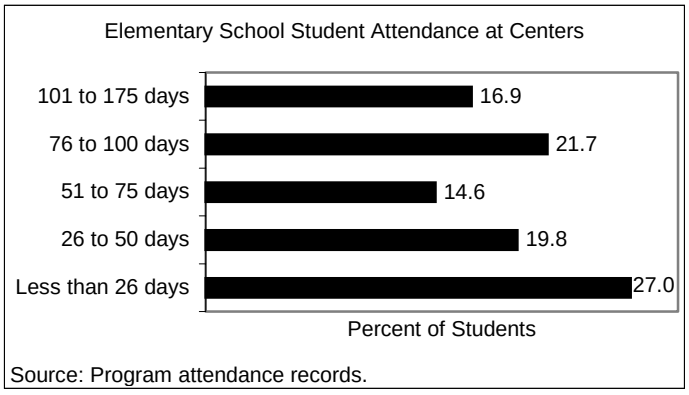
Impacts by program characteristics were also estimated. These analyses focused on two types of program characteristics: (1) program emphasis on academics and (2) levels of participant attendance. Interestingly, programs that emphasized academic activities over

recreation and other activities were not more likely to increase test scores or grades. Similarly, no relationship is evident between average attendance of a program and impacts by program.

Additional analysis looked at the impacts for frequent participants compared to infrequent participants. The analysis suggests that frequent participants were more likely to be from disadvantaged households and to want to improve in school, as their better behavior in school and their more frequent attendance itself indicate. However, the analysis did not reveal that more frequent participation led to better outcomes.

Findings for Elementary School Programs

Researchers selected elementary school centers that had more applicants than they could accept, because these centers could implement experimental designs. Elementary school programs in the study were more likely to be in urban areas and to serve more disadvantaged students than other elementary school programs, but most characteristics were similar to other elementary school programs.

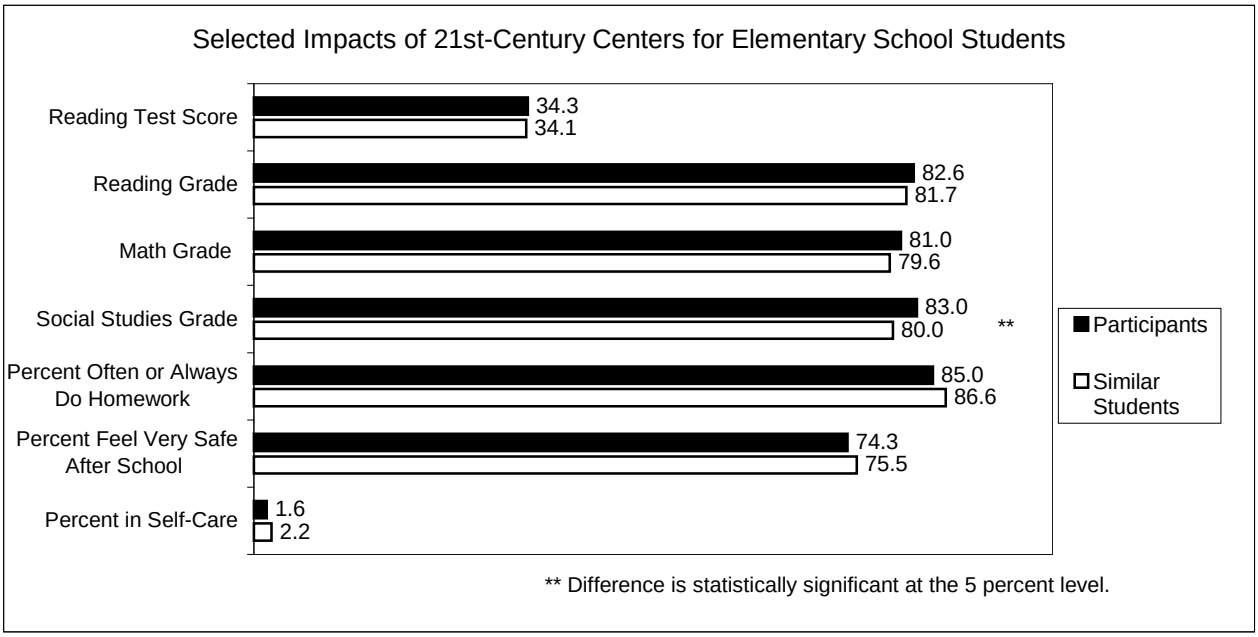


Elementary school students attended for 58 days, on average, during the school year, and more than one-third of students attended for more than 75 days. These attendance levels may not be typical of attendance levels of elementary school programs in general because the evaluation looked only at oversubscribed programs.

The elementary school programs in the study increased the time students spent at school or outside the home and reduced the time spent at home after school cared for by a parent or sibling. Programs did not reduce self-care, the incidence of which was low (about two percent of students).

The programs had no effects on reading or math grades or reading test scores. For example, in spring 2001, program students had an average percentile reading score of 34.3, compared with a score of 34.1 for similar students. Social studies grades were higher by a statistically significant margin (83, compared with 80), but grades in other subjects were not.

Programs did not appear to improve student effort in school. Parents and teachers had different views about whether effort improved. According to teachers, program students were more likely than similar students to try hard in reading or English class. According to parents, however, program students were less likely than similar students to work hard in school. However, students reported no differences in homework completion, time spent watching television, or time spent reading for fun.



NOTE: Impacts were estimated using regression models to adjust for differences between treatment group and control groups in fall 2000. The adjustment variables in the regression included indicators of students' demographic characteristics, household socioeconomic status, and students' fall test scores, as well as previous year attendance, disciplinary problems, and self-reported grades.

Programs did not affect whether students felt safe or unsafe after school and did not affect student behavior in school. Suspensions, absences, and teacher reports of discipline problems were the same for both groups.

Program participation did not change students' interpersonal skills. Program students were no more likely to report getting along with others their age, feeling included, being good at working with others in a team, or setting a goal and working to achieve it.

A Typical Elementary School Center

The center is open five days a week for two and a half hours a day. About 80 students participate every day, with most participating three or four times a week. After the school day ends, students have a snack provided by the program and play outside for 30 minutes. At 2:30 p.m., third-, fourth-, and fifth-grade students participate in a homework session. Kindergarten, first-, and second-grade students have "story time." To participate in other recreational and enrichment activities, students must attend the homework session (or story time). In the homework session, students work on assignments or read a book if they have completed their homework. Each homework classroom has about 20 students, two at a table, and a college student or paraprofessional. At 3:30 p.m., homework and story time end, and recreational and enrichment activities begin. All students participate in two 45-minute electives. Recreational activities include arts and crafts, games, computers, and team sports. Enrichment activities include music, drama, and dance. Homework assistance and access to computers are provided throughout the year. Other electives change quarterly based on student interest. At 5 p.m., the second elective ends, and students gather in the school library to be picked up by school buses. If they have parental permission, some older students walk home after signing out.

Directions for the Future

These findings reflect the challenges school-based after-school programs face to improve student outcomes. Even for after-school programs oriented toward providing academic support as well as recreational and social activities, there were few improvements in homework completion, grades, and test scores. The lack of academic improvement may be due to the low attendance rates and the length of the follow-up period. However, analyses of those who participate more frequently found that more attendance alone may not make measurable differences in outcomes. In addition, too few participants may have received sustained, substantive academic support. Both participation rates and the content of program academic offerings may need more attention.

The No Child Left Behind Act restructures the 21st-Century program and focuses more attention on the program's potential for improving academic outcomes, especially for disadvantaged students. An additional year of follow-up and the expansion of the number of elementary school programs in the study will provide another opportunity to assess whether the programs (as they are currently implemented) are likely to meet these objectives.

I. Introduction

Most parents work, but most schools dismiss their students hours before the workday ends. During the intervening hours—"out-of-school time"—many parents want their children to be able to develop academic, personal, and social skills in safe, supervised settings. In 1994, Congress created the 21st-Century Community Learning Centers program to support efforts by communities to make greater use of school buildings when schools were not in session. The program, operated by the U.S. Department of Education (ED), later refocused its efforts to provide after-school opportunities and made its first grants supporting after-school activities in 1998. By 2002, more than 1,400 school districts and communities were participating.

In September 1999, ED selected Mathematica Policy Research, Inc. (MPR) and its partner, Decision Information Resources, Inc., to conduct a national evaluation of the program. The Charles Stewart Mott Foundation also contributed to the evaluation through a grant to MPR. This report provides findings from the evaluation's first year of qualitative and quantitative data collection from students, parents, teachers, principals, and program staff members, as well as from visits to the programs and observations of their activities.

Underlying the interest in after-school programs is their potential to improve a wide range of outcomes.[3] Programs could improve academic outcomes by helping students become more capable in the classroom, learn more subject matter, and have higher grades and test scores. They could improve developmental outcomes by helping children and youths learn social skills, appreciate their own and other cultures, and become more sure of themselves and their own values. And programs could keep children and youths safe. The evaluation examined how

[3]See, for example, *Safe and Smart* (1998) and *Working for Children and Families (*2000), both prepared by ED and the U.S. Department of Justice.

21st-Century programs were implemented, how they were structured, whom they served, and the issues they faced in meeting their objectives. It also examined whether and how programs improved academic and developmental outcomes.

A. The 21st-Century Community Learning Centers Program

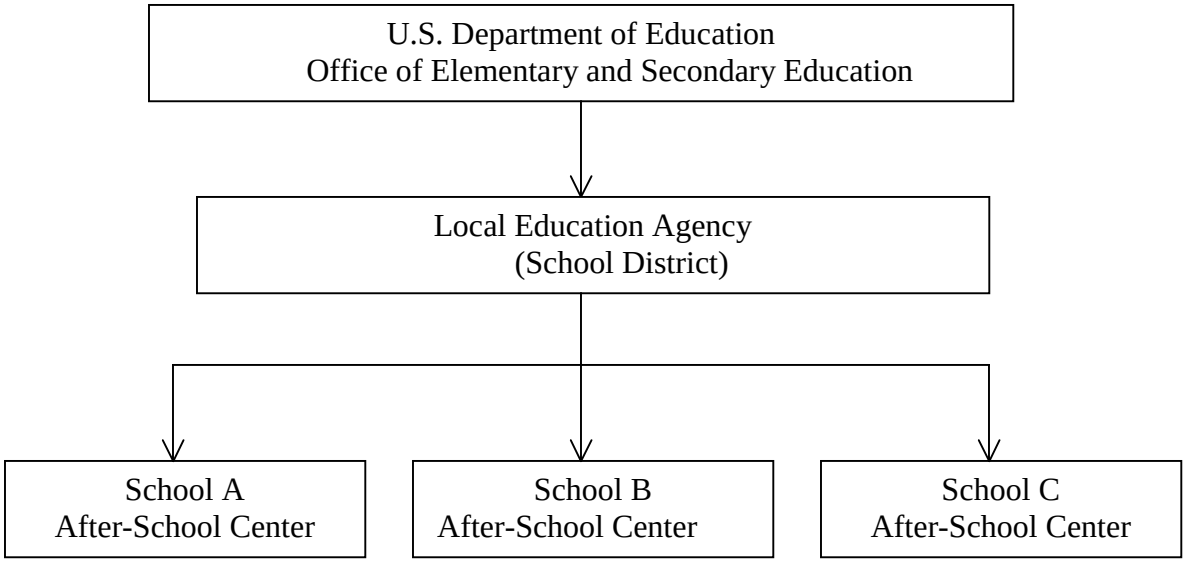
The Improving America's Schools Act of 1994 (P.L. 103-382) created the 21st-Century program. In fiscal year 1998, Congress appropriated $40 million for it, and ED awarded grants to 99 school districts. Subsequently, the appropriation and the program's scale increased substantially, with the appropriation rising to $1 billion in fiscal year 2002 and the number of grantees to nearly 1,600. In all, ED has funded seven cohorts of grantees. The national evaluation focuses mostly on the first three rounds (or cohorts) of grantees receiving 21st-Century grants.

The legislation stipulated grants of three years' duration, awarded only to local education agencies (which usually are school districts). The average grant award for the first three cohorts of grantees was slightly under $400,000. Figure I.1 diagrams the general structure of a 21st-Century grant as it moves from ED to the level of local schools. The school districts receiving grants must use their funds to operate school-based programs. Often, the grants support after-school centers in public school buildings, although grant funds can be used to operate summer and before-school programs.[4] The programs the grantees offer at the after-school centers must incorporate at least 4 of 13 activities listed in the authorizing legislation. These activities include integrated education, health, social service, recreational, or cultural programs, literacy education programs, children's day care services, and telecommunications and technology education programs.

[4]The national evaluation did not investigate grantees' summer and before-school program offerings.

Figure I.1

Structure of 21st-Century Community Learning Center Grants



Although only districts and schools are eligible to receive grants, the federal statute strongly encouraged grantees to collaborate with other public agencies, nonprofit organizations, and businesses in their communities. The statute required local education agencies to describe "the collaborative efforts to be undertaken" with such organizations. It also defined community learning centers as places to be operated by local education agencies "in conjunction with" organizations external to schools.

Data from annual performance reports that grantees submitted to ED in April and October 2001 give the size of the centers that grantees operated, the most common services the centers

provided, and the types of schools in which centers were operating.[5] The following information from the reports gives highlights of centers' features:

- The average grantee operated between three and four centers and reported enrolling nearly 700 students and 250 adults (parents or other adults from the community) over the course of the year.
- Nearly all centers were open 10 or more hours a week, usually after school, and a third were open 20 hours or more. Some were also open on Saturdays, and many offered summer programs.
- Most centers were in elementary and middle schools or in schools that included a combination of K-8 grade levels. Five percent of centers operated in high schools.
- Nearly all centers reported providing reading, math, and science activities. Enrichment activities, such as art and music, and technology activities also were common.
- Schools that centers served had more minority students than the average school. Fifty-seven percent of students in schools that centers served were minority students, compared with 37 percent of students nationwide.
- Schools that centers served were more likely to be high-poverty schools. Sixty-six percent were high-poverty (meaning that more than half their students were eligible for free or reduced-price lunches), while 17 percent of schools nationwide fall into this category.
- Most centers reported that they collaborated with local organizations to provide services, set goals and objectives, and share techniques.
- Nearly all centers reported communicating with their host schools to recruit and refer students, provide feedback on students, set goals and objectives, communicate curricula, and share instructional practices. Nearly all centers reported that at least one of their staff members worked in the host school.

The highlights do not convey the substantial variation in centers' schedules, staffing, and emphases, which we discuss in subsequent chapters.

Other after-school programs around the country, such as the Beacons, LA's BEST, programs supported by the After-School Corporation in New York, and Boys and Girls Clubs, are similar

[5]Annual performance report data are for grantees in the first through fifth cohorts that submitted reports to ED in April and October 2001. Grantees are responsible for gathering and reporting the data.

to the 21st-Century program in many respects. An important difference is the requirement that 21st-Century grantees offer academic activities.[6] Furthermore, until recently, only local education agencies could receive a 21st-Century grant, and the centers supported by the grant had to be located in a public education facility.

The No Child Left Behind Act, which became law in January 2002 (P.L. 107-110), changed the 21st-Century program in major ways. As the program operated before this legislation, ED received funds, carried out grant competitions, and made awards under set criteria to those submitting the highest-rated grant applications. Grants were for three years, and grantees were not required to match federal funds with state or local funds. For the new program, each state will be allotted funds and will carry out its own grant competition and make awards. Local education agencies, as well as community and nonprofit organizations, will be eligible for awards. States may specify up to a dollar-for-dollar match in making awards, and the grant period can be from three to five years, at the discretion of the state.

B. A Conceptual Framework for the National Evaluation

A previous report (Dynarski et al. 2001) describes the evaluation's design. An accompanying concept paper (Moore et al. 2000) examines design aspects enhanced by the grant from the Mott Foundation and the integration of the evaluation components. The report and the concept paper together laid out the evaluation's conceptual framework, discussed statistical aspects regarding how grantees were selected for the evaluation, and presented the instruments and protocols used to gather data. The key highlights of the design include an emphasis on rigorous estimation of effects and multifaceted data collection that allowed the evaluation to

[6]Many of these other programs, however, include improving educational performance among their goals and offer time for homework, reading, and, sometimes, tutoring.

explore many questions about program operation and implementation in addition to its impacts on children and youths.

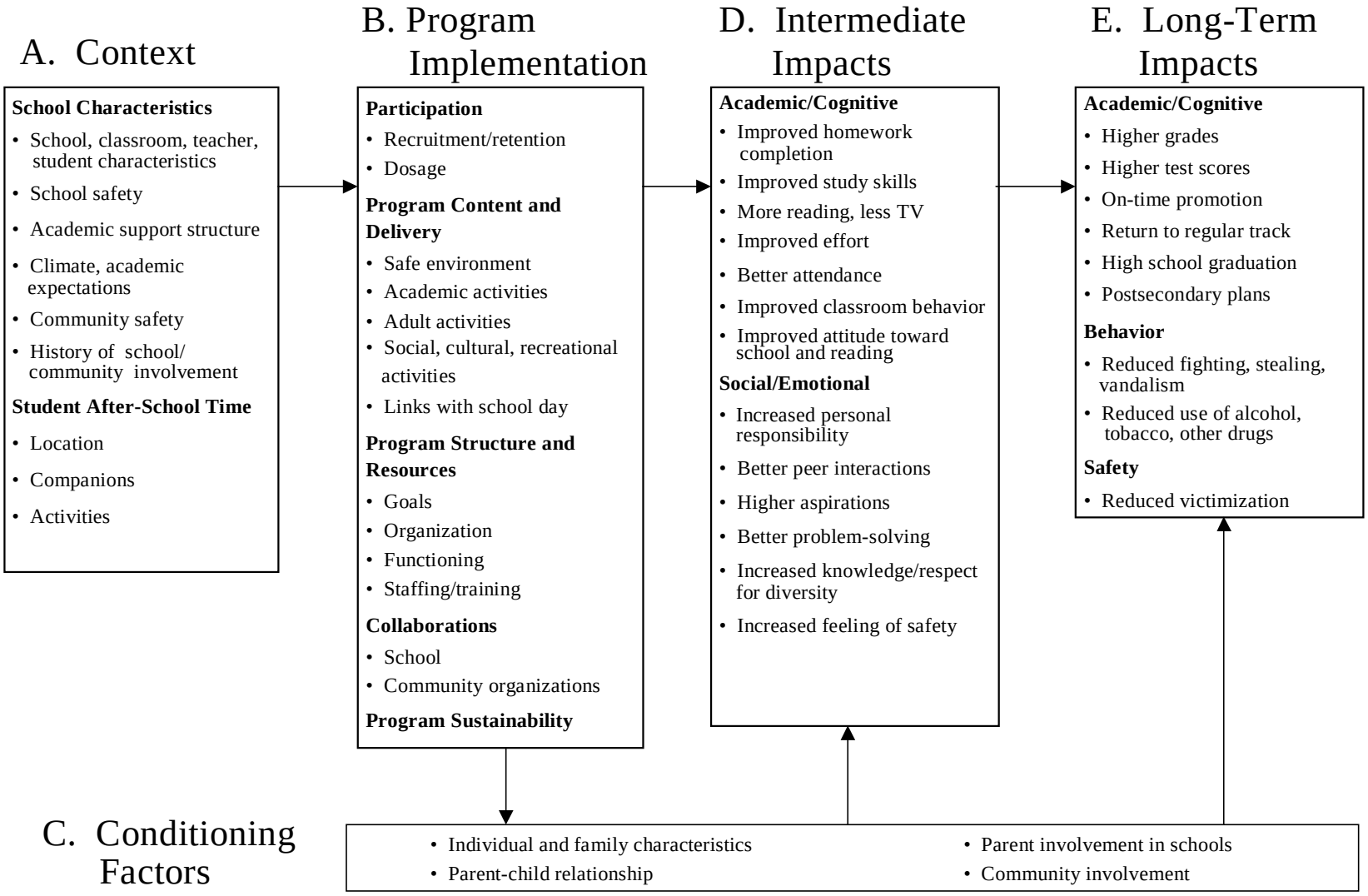
The logic model shown in Figure I.2, which presents the five topical areas central to the national evaluation, guided the outcomes measured and the issues considered in studying program implementation. These areas are (A) the context in which an after-school program operates; (B) the implementation of the after-school program itself; (C) family, individual, and community conditioning factors that influence after-school programs and that, in turn, affect (D and E) student intermediate and long-term outcomes. The figure highlights how after-school programs are embedded in the larger constellation of school, community, and family influences that contribute to student outcomes in and out of school.

Context. The national evaluation set out to identify the circumstances under which after-school programs are implemented. These circumstances include the educational and policy climate, perceptions about safety, community relationships, and demographic characteristics of the school, district, and community.

Program Implementation. To help us learn which practices and approaches are effective in different settings and for different student groups, it is necessary to know the details of how programs operate. The following measures of program implementation help us understand how programs were implemented and structured: student participation, program content and structure, collaboration with host schools and community organizations, and efforts toward sustainability.

Intermediate and Long-Term Impacts. Because 21st-Century programs provide many services and activities, the programs could have many impacts, including improved safety, better

Figure I.2
Logic Model for Understanding the Impacts of 21st-Century Programs



academic performance, positive behavioral changes, and increased personal competence. The conceptual framework separates effects into intermediate effects and the longer-term effects that are presumed to follow. For example, if students attend school more often and try harder in the classroom (intermediate effects), they are more likely to improve their grades and test scores (long-term effects). Similarly, if students exhibit greater personal responsibility and associate with peers who share positive values, risky behaviors are more likely to decline.

Conditioning Factors. External factors and relationships may influence the effects of after-school programs on students. For example, specific features of after-school programs may affect older students differently from the way they affect younger students. Students with learning deficits may benefit more from after-school programs than students at less risk. Students with behavioral difficulties may benefit differently from students without them.

C. Key Features of the Design

The design report presents more detail about the aspects of the evaluation's implementation and impact data collection and analysis design. Here, we discuss the key features of the design.

1. Different Designs Used to Evaluate Centers Serving Middle School Students and Those Serving Elementary School Students

The national evaluation was designed to use rigorous techniques to measure impacts of after-school programs on students in middle and elementary schools. Comparison-student designs (with matching to identify comparison students) were used to measure the impacts of middle school programs, and experimental designs were used to measure the impacts of elementary school programs.[7] We chose a comparison-student design for middle school

[7]In general, students selected for the middle school sample attended grades 6 through 8, while students sampled at the elementary school level attended kindergarten through grade 5. Some districts had middle schools

programs because of the paucity of oversubscribed middle school centers. Because centers serving elementary school students were more likely than those serving middle school students to be oversubscribed, the evaluation was able to identify a set of elementary school centers that could implement rigorous experimental designs, with random assignment of students to treatment groups and control groups.

Middle School Sample. To evaluate middle school centers, we selected as a probability sample a set of grantees that operated such centers. The evaluation team used students attending 21st-Century centers during a one-month window in the fall of 2000 to form the treatment groups, and identified similar students not attending these centers to form the comparison groups. Thirty-five grantees from cohorts one through three that operated centers serving middle school students were selected at random from 16 strata. (Grantees were not excluded if they operated elementary school centers, as long as they operated at least one center that served middle school students.) The stratification ensured representation of grantees' geographic region and urban and rural areas. Findings for the middle school centers in the evaluation generalize to first- through third-cohort grantees serving middle school students, because the grantees were a random sample of all middle school grantees in those cohorts. At the start of data collection activities, first-cohort grantees were beginning their third year of funding and second and third cohort grantees were beginning their second full year of funding. Second-cohort grants were awarded in November 1998 and some grantees may not have begun serving students until the fall of 1999.

Ultimately, 34 middle school grantees agreed to participate in the evaluation. Annual performance report data provide a sense of how the sample of middle school grantees

(continued)
that included fifth grade or elementary schools that included sixth grade, in which case the national evaluation used those definitions.

represented the full set of middle school grantees. Table I.1 presents characteristics of all centers and of those in the evaluation from which performance report data were gathered (25 of the 34 grantees). The table shows that characteristics of middle school centers in the national evaluation, as expected from the random sampling procedure used, are similar to those of middle school centers in general. For example, the average center had 23 staff members, and the average sampled center had 21. Middle school centers in the evaluation served fewer students (238 compared with 243) and had fewer attending for 30 days or more (79, compared with 101). The racial and ethnic composition of enrolled students was similar.

For each grantee, the evaluation team used propensity score matching techniques (Rosenbaum and Rubin 1983) to select a group of comparison students for program participants. (Appendix B describes technical aspects of how the matching was done.) Ultimately, we obtained parental consents and follow-up data for 4,264 students. The evaluation was able to collect the first wave of follow-up data for 32 grantees. Delays in getting baseline data from two grantees impinged on the first follow-up effort in those districts, so data from the second follow-up effort in those two districts will be included in the next report.

The middle school comparison design offers a rigorous assessment of the impacts of after-school programs on middle school students. The design used for the assessment, however, was dictated by the lack of oversubscription for most middle school programs. The findings lack the same high degree of internal validity of random-assignment designs. We used analytic techniques to try to minimize this shortcoming, but, ultimately, the shortcomings temper our ability to attribute measured effects to the 21st-Century programs alone. Nevertheless, the

Table I.1

Characteristics of Centers in the National Evaluation

	Middle School Centers	National Evaluation Sample of Middle School Centers	Elementary School Centers	National Evaluation Sample of Elementary School Centers
Characteristics of Staff Members				
Average Number	23.0	21.0	22.0	24.0
School Day Teachers (%)	39.6	40.6	34.1	35.2
College Students (%)	11.5	13.6	14.8	27.2
High School Students (%)	10.0	5.4	11.8	1.4
Parents (%)	9.1	5.7	10.7	4.5
Youth Development Workers (%)	7.6	4.6	6.3	12.8
Other Community Members (%)	12.1	9.4	10.5	13.4
Student Attendance				
Average Number Attending Fewer than 30 Days	142.0	159.0	88.0	51.0
Average Number Attending 30 Days or More	101.0	79.0	97.0	103.0
Percent Attending Fewer than 30 Days	58.4	66.8	47.6	33.1
Percent Attending 30 Days or More	41.6	33.2	52.4	66.9

Table I.1 (*Continued*)

	Middle School Centers	National Evaluation Sample of Middle School Centers	Elementary School Centers	National Evaluation Sample of Elementary School Centers
Race/Ethnicity of Enrolled Students (Percent)				
White	40.0	45.3	39.7	28.2
Black or African American	25.3	23.8	22.8	66.8
Hispanic	22.1	22.9	27.6	1.8
Asian	2.4	1.5	2.0	1.9
Native Hawaiian or Other Pacific Islander	1.0	0.1	0.7	0.3
American Indian or Alaska Native	5.5	2.3	5.4	1.0
Percent Eligible for Free or Reduced-Price Lunches				
Less than 25 percent	12.3	19.0	11.1	0.0
25 to 49 percent	24.1	14.3	19.6	28.6
50 to 74 percent	25.9	33.3	24.3	0.0
75 to 100 percent	37.7	33.3	45.1	71.4
Sample Size	**768**	**46[a]**	**792**	**11[a]**

SOURCE: Annual performance reports.

[a]Annual performance report data were not available for 16 middle school centers and 7 elementary school centers that were included in the national evaluation.

information in this report provides the best available estimate to date of the impacts nationwide of 21st-Century programs on students in this age group.

Elementary School Sample. To implement the elementary school evaluation design, we randomly assigned about 1,000 students at seven grantees during fall 2000 and collected data for them at baseline and in spring 2001. These seven grantees and their student samples generate the elementary school findings presented in this report. To augment the size of the elementary school sample, we randomly assigned another 1,600 students in seven additional elementary grantees in fall 2001 and collected baseline data. We will include results for these students in the next report, scheduled for winter 2003. Because the first-year findings may change when the full set of elementary school grantees is included in the analysis, the elementary school findings in this report should be viewed as preliminary.

Findings for the elementary school centers in the evaluation do not generalize to all elementary school centers, because the ones in the evaluation were chosen for their ability to carry out the experimental design. Table I.1 shows characteristics of all elementary school centers that submitted annual performance report data, as well as characteristics of 11 of the 18 centers in the first cohort of the national study for which performance report data were available. In general, the elementary centers for which we report results at this time serve a larger percentage of minority students, especially African Americans, than elementary centers in the same grantee cohorts, and are in schools with higher poverty levels. However, although not a representative sample, the elementary school findings have strong internal validity for attributing student outcome differences to the 21st-Century program. Consequently, they have important implications for understanding how after-school programs serving younger students can affect outcomes.

2. Impacts Estimated for a Range of Academic and Nonacademic Outcomes, Using Data from Students, Parents, and Teachers

We gathered data for a wide range of outcomes from student questionnaires, school records, parent questionnaires, and teacher questionnaires. (Appendix A describes the evaluation's data collection procedures and response rates for the instruments.) Outcomes included academic performance and homework completion, behavior, feelings of safety, and personal and social development. Teachers who received questionnaires were English teachers of the middle school students and regular classroom teachers of the elementary school students the evaluation sampled.

Two major considerations in evaluating after-school programs are the different levels of participation among enrolled students and the activities that were available on the days they did attend. The national evaluation addressed the first by collecting attendance information from each center for students in the sample. However, because of the lack of routinely kept records of attendance at each activity in the centers, as well as the burden associated with imposing such a system on center staff, the evaluation could not obtain a detailed breakdown of the degree to which each student in the sample participated in specific activities.

3. Program Implementation Assessed Based on Visits to Grantees and Other Data

Research staff conducted site visits to all grantees in the evaluation at least once during the 2000-2001 school year. Most visits lasted two to four days and included interviews with staff members associated with centers, host schools, districts, and collaborating community organizations. Six grantees were visited twice as part of the enhancement study supported by a grant from the Mott Foundation. Site visit reports were coded using qualitative analysis software, and site visitors completed several assessment forms that allowed researchers to categorize center programs (for example, to distinguish the degree of emphasis placed on

academics and developmental activities).[8] In addition, questionnaires were administered to principals, to all center staff members, and, in the six grantees that were part of the enhancement study, to a sample of students in the host schools who did not attend the centers during the school year. The data on nonparticipants provided insights into why the students did not participate and factors that would have encouraged them to participate.

D. Organization of Report and Presentation of Findings

The chapters that follow describe the implementation of the middle school centers, the impacts of middle school centers, and the implementation and impacts of elementary school centers. We separate centers by the grade levels served because of the differences in how we measured impacts for the two types. We studied implementation using the same methods for collecting and analyzing the data, but the design differences give a somewhat different meaning to the middle school and elementary school implementation findings. The middle school findings can be generalized to first- to third-cohort grantees serving middle school students, whereas the elementary school findings cannot be generalized to all grantees serving elementary school students.

Throughout the rest of this report, we present findings from perspectives (for example, from the grantee, the center staff members in the schools, principals and teachers at host schools, and participating students and comparison students) appropriate to the topic being discussed. Furthermore, for subsequent chapters that relate impact results, we analyzed student impacts at

[8]Volume 2 of the design report contains examples of the assessment forms.

the grantee level, not the individual center level.[9] A range of implementation findings are presented at the individual center level.

Throughout the report, we use the terms "grantee" or "project" to apply to activities and operations at the level of the school districts sampled as part of the national evaluation. We reserve the term "programs" for the activities and offerings of the 21st-Century learning centers in the evaluation. "Project directors" are those charged with oversight of the grant by the school district, and "center coordinators" are those who directly oversee after-school programs in the school buildings.

[9]We found less variation for centers within grantees, for example, than across them, which supports our use of the grantee as a unit of analysis. Furthermore, student matches at the middle school level were based on the treatment sample for the grantee as a whole, which precluded analyses by individual centers within a given grantee.

II. A Comprehensive Look at 21st-Century After-School Programs in Middle Schools

This chapter focuses on program implementation at centers that serve middle school students. We discuss program offerings, student participation levels, staffing, linkages to schools and community organizations, and funding and sustainability. Several themes emerge from the results presented in this chapter. To their credit, all the middle school grantees in the evaluation's nationally representative sample had organizational structures in place and were providing a range of after-school activities to students. Academic offerings—especially homework assistance—were common components of programs. In addition, many parents, students, principals, and school faculty expressed support for the programs.

Some problems also came to light, however. Students in the middle school centers often were less engaged in academic classes than in recreation or cultural activities. Centers had difficulty recruiting and retaining students, and students did not attend centers frequently or consistently. Center coordinators spent a lot of time finding appropriate, available staff members for activities. Many staff worked at the centers for only a few days each week or only some months in of the school year, which led to less daily consistency of staff. Programs rarely had strong links to community organizations; instead, they generally used such organizations as sources for someone to lead a recreational or cultural activity of particular interest. Sustaining these programs beyond the federal grant was proving to be a substantial challenge.

Five main sources of information shed light on the middle school centers:

1. Site visit observations and interviews with staff members at district offices, centers, and host schools that were part of the national evaluation.
2. Surveys of project directors, principals, center coordinators, center staff members, and students.

3. Center attendance records for participants.

4. A survey of participants in six sites about their reasons for attending, activities in which they participated, and perceived outcomes. In addition, a survey of a sample of students who had not participated in centers about their reasons for not participating and their perceptions of centers.

5. Annual performance report data that grantees submitted to the U.S. Department of Education (ED).

We often used different sources of information to confirm patterns, but to simplify our presentation, we cite the most direct source.

A. After-School Programs in Middle Schools Were Designed for Broad Student Appeal

In addition to the broad list of activities the federal statute encouraged, three prominent considerations shaped 21st-Century programs in middle schools. Program designers sought to:

1. Create offerings that had broad student appeal and were responsive to rapidly changing student interests, which prompted them to give students choices about the activities in which they participated and to vary the offerings.

2. Find staff members who could lead activities and work well with students and who could work after school, which led them to segment program schedules to suit staff members, especially teachers, who had limited availability after school.

3. Accommodate staff members', parents', and teachers' views of what students needed to improve and develop, which resulted in providing a range of activities that spanned academic, physical, social, and cultural dimensions.

Not surprisingly, site visitors found that most centers focused on multiple objectives that extended beyond academic improvement. While most centers had academic improvement as a major objective, noteworthy percentages also placed major emphasis on recreation, safety, and cultural opportunities (Table II.1).

Table II.1

Objectives of 21st-Century Middle School Centers

	Percentage		
	Major Objective	Minor Objective	Not an Objective
Help Children Improve Academic Performance	69	31	0
Provide Recreational Opportunities for Children	56	30	15
Provide a Safe Environment for Children After School	56	38	7
Provide Cultural Opportunities for Children	41	39	20
Help Children to Develop Socially	31	56	13
Help Parents or Other Adults Develop Skills	10	31	59

SOURCE: Site visitor assessments based on visits to 61 centers.

Middle school centers typically viewed recreation sessions and, to a lesser extent, enrichment activities as the focal points of the program and the components that attracted students. Centers encouraged or required students to attend the academic sessions before they engaged in other activities that provided more choice, variety, and potential for fun. "Fun" activities were the reward for doing homework or engaging in other academic activities.

Choice was a frequent method that centers used to appeal to middle school students (see box). Choice was least common for academic assistance, particularly homework or test preparation sessions, and most common for activities emphasizing recreation, culture, and interpersonal skills. When centers restricted students' choice of activities, they did so to address the needs of particular students or to achieve a balance of academic and other activities. Centers also restricted student choice for some activities to maintain desired student-teacher ratios or to obtain an appropriate mix of students.

Policies on Student Choice of Activities

Full Choice 41%

Some Choice 41%

No Choice 18%

Middle school centers paid significant attention to staffing their programs adequately. A center typically had 12 or 13 paid staff members working with students. The average student-staff ratio across the centers was about 11-to-1, ranging from 3-to-1 to as high as 50-to-1.[10] Staff expressed to site visitors that to ensure a positive and productive experience for both groups, they wanted to prevent student-staff ratios from rising too high, but "too high" varied by place and type of activity. Some grantees had general goals for centers to keep the student-staff ratio in virtually all activities below a certain level, 10-1 to 15-1. Other grantees had set maximum student-staff ratios only for specific activities such as tutoring and, generally, academic activities has lower student-staff ratios than recreational activities. Also, in practice the ratios are likely to have been lower due to participant absences from the program (just as regular-school class sizes in practice may be lower if some students are absent from school).

1. Offerings of the Typical Middle School Center

Three types of activities were common to most centers: (1) academic assistance, (2) recreation, and (3) cultural enrichment and interpersonal skill development.[11] Generally, the first time slot was devoted to academic assistance, followed by one or two slots for recreation and enrichment and interpersonal activities. Academic assistance sessions usually lasted 45 minutes to an hour, during which students often had a snack, which almost every center provided. Centers also sponsored occasional field trips or presentations for the school community that highlighted accomplishments from after-school activities.

[10]These estimates are based on the total number of students enrolled in a center and the total number of paid staff working with students; more precise estimates are difficult because of variations in the number of students and staff at a center on a given day.

[11]Only five centers provided academic activities with no cultural or recreational activities.

a. Academic Assistance Offerings

Middle school centers used two main strategies to deliver academic assistance: (1) homework sessions and, more commonly, (2) homework sessions combined with "other academic assistance" (see box). This second category varied across centers and included tutoring, preparation for state assessment tests, or sessions to improve reading, writing, or math skills.

Middle School Center Academic Assistance Activities

Homework Plus Other Academic Assistance	54%
Homework Only	33%
Only Other Academic Assistance	10%
None of the Above	3%

Homework sessions were the most prevalent type of after-school academic assistance that centers offered. Site visitors observed that most homework sessions resembled study halls in which students were expected to know their assignments, bring their materials, and work independently. These sessions typically consisted of about 20 students monitored by two staff members (usually certified teachers or a certified teacher and a paraprofessional). Although having teachers from the host school oversee homework sessions offered a potentially fruitful path for helping students after school, the caliber of homework assistance was low. This weak assistance may help explain why only 38 percent of students thought the centers were a good place to get homework done. The sessions focused on providing students with the opportunity to complete homework, not on ensuring they completed it. Site visitors rarely observed staff members checking homework for completeness and accuracy. Help was nearly always available, but students had to ask for it. Although the teachers' role was to maintain order and a quiet atmosphere, sessions often were noisy.

A few centers did develop strategies to strengthen homework sessions. About one-fifth of centers used written documents to monitor students' homework assignments or academic needs. Some centers in host schools required some or all students to record homework assignments in journals. Centers used these journals to find out what students were supposed to be working on,

then made a note for the regular teacher that the student had completed the homework. In other centers, the coordinator used a list of failing students provided by the host school to identify those who needed extra attention. A few center coordinators generated lists of after-school participants and shared them with the school day teachers, who then identified students who needed extra help.

Centers offered students other types of academic assistance, which included tutoring, classes in practicing concepts and skills for state assessment tests, and computer-based instruction to improve skills. By far the most common forms of academic assistance other than homework were sessions to build reading, writing, or math skills: 51 percent of middle school centers provided help with reading and writing, and 46 percent provided help with math. Because centers often targeted particular students, such as those referred by a classroom teacher for extra help or those performing poorly on state tests, this academic help reached only a portion of students. Tutoring sessions, for example, typically grouped five to seven students with a teacher to work on specific skills, often using materials similar to those used in class. Practice sessions for state tests often had 7 to 10 students working with a teacher on specific reading or math skills that were to be tested. Computer classes of up to 15 students featured software practice on academic concepts learned during the day. These other forms of academic assistance were less frequent than homework sessions, typically occurring between one and three days a week.

Illustrative Recreation Activities

Swimming
Weight Training
Bowling
Dance
Rock Climbing
Fishing
Kickboxing
Cheerleading
Basketball
Breakdancing
Martial Arts

b. Recreation, Cultural Enrichment, and Interpersonal Skills Offerings

In general, recreation was the most prevalent activity other than academic assistance that centers offered (provided more than

once a week by 84 percent of middle school centers). Recreation activities often incorporated structured opportunities for students to learn a skill or develop specialized skills (see box). Other, less-structured, recreational activities, such as open gym, free play, board games, or general computer use, provided some supervised relaxation and physical outlets for students.

Most centers (77 percent) offered cultural activities more than once a week, but less frequently than they offered recreation and academic assistance. Interpersonal skills activities were the least frequently offered, although most centers (54 percent) did offer them more than once a week. In addition, students could develop interpersonal skills in activities that had other stated objectives. For example, recreation activities involving teamwork could reinforce leadership and conflict resolution skills.

Illustrative Activities

Cultural Enrichment	**Interpersonal Development**
Japanese	Leadership
Manners Training	Conflict Resolution
Crafts (Sewing, Rug Making)	Positive Peer Modeling
Photography	Mentoring
Drama	Teen Issue Forums
Broadcasting	Peer Risk Prevention
Sandstone Sculpture	
Choir, Band, Orchestra	

Center coordinators favored changes in the mix of activities to attract new students as well as to keep already enrolled students attending the center. Supply and demand governed many decisions about activities. When too few students enrolled in an activity, centers introduced new activities. When too many students wanted to participate in an activity, centers tried to add sections if instructors were available. Accommodating instructors' work schedules was also an important consideration. Cycling the activities helped coordinators tap the expertise and interests of teachers and outside staff members, many of whom wanted to limit their time commitment.

c. How Students Spent Their Time at Middle School Centers

Additional questions asked of participants at six middle school programs provide a bit more detail on the range of activities that students participated in at middle school centers. The

students' perspective is consistent with that of other sources. Homework, sports, and computers were the dominant activities participants cited, followed by reading, writing, or science activities; tutoring; lessons in art, music, and dance; and volunteering or community service (Table II.2).[12] Students also reported special activities they did at centers. The most common were being in a special tournament of some kind (cited by 29 percent of participants), performing in a play or show (cited by 26 percent), and a range of other activities, such as giving a speech or creating artwork, each cited by less than 20 percent.

d. Offerings for Adults

Although the federal statute specifies several services to adults in its list of allowed grantee activities, centers' offerings for adults were minimal. Nearly two-thirds of middle school centers offered no services or activities for adults, and other centers offered sporadic activities that varied widely.[13] Dismayed by low attendance at early attempts to serve adults, most center directors had concluded that parents (the main group they felt they could tap for activities) did not want additional commitments on their schedules, which were already filled with jobs and child care obligations.

B. Participants Did Not Attend Centers Often

Student attendance is a critical element of 21st-Century middle school centers. Centers could improve student outcomes only if students attended. Attendance typically was voluntary,

[12]The high rate at which participants cited tutoring as a center activity suggests that they were considering homework help as tutoring. A student meeting one-on-one with a tutor at centers was rare, according to site visitors.

[13]Among the offerings that site visitors noted were family involvement (family art, holiday celebrations), adult literacy (GED classes, English as a Second Language classes, and job skills), adult enrichment (courses in computers, Spanish, gardening), and adult basic needs (medical services and parenting skills classes).

Table II.2

Participant Activities
(Middle School Centers in Six Sites)

Students Who Participated in the Following Center Activities "Some" or "A Lot":	Percentage
Homework	62.8
Organized sports	56.5
Surfing the Internet or other things on a computer	51.3
Non-homework reading, writing, or science activities	38.5
Tutoring	33.9
Lessons (music, dance, art, others)	34.0
Volunteering or doing community service	21.3
Percentage of Participants Who Participated in the Following Activities in a Center	
Participated in a special tournament	28.6
Performed in a play or show	25.7
Gave a dance performance	18.1
Performed a piece of music	16.0
Other accomplishments	13.1
Gave a speech	12.5
Produced a piece of art that was displayed	12.5
Participated in a debate	10.4
Produced a newspaper or newsletter	5.8
None of the above	30.3

SOURCE: Participant survey module in six middle school sites. The sample size is 263 participants.

however, so centers believed they had to attract students through center offerings and relationships with staff and schools.

We explored how centers recruited students, what the frequency and patterns of student attendance were, what attendance policies had been established, and how participants' and nonparticipants' perceptions of centers may have affected attendance. The picture that emerged suggests that limited participation is likely to be the norm for middle school programs. Participants came because they wanted to come and they perceived positive outcomes from participating, but the average participant did not participate much. This suggests that most students consider the 21st-Century programs to be acceptable places to go after school, but they do not find the activities so compelling that they want to attend every day or often. Students who

had not participated in the centers thought centers were less attractive than other after-school opportunities or faced barriers such as household obligations that made participation difficult.

The primary recruiting approach used by programs was to appeal to students at the host school (used by 85 percent of middle school centers). Common recruiting techniques included letters to parents, teacher and parent referrals, presentations at school events (registration, open houses, and parent-teacher conferences), and announcements over the school's public address system. Some centers also used newspaper articles, announcements on the school bulletin board or outdoor sign, and radio ads. Some centers relied on referrals or recommendations from school staff members. Others targeted their efforts to students with particular needs or characteristics (such as those with low grades or test scores), sometimes contacting their parents to encourage enrollment. The data show a consistent pattern—participants usually heard about centers from teachers or other school staff members or found out about them through posters or school announcements (Table II.3).[14]

Table II.3

How Students Heard about Centers
(Middle School Centers in Six Sites)

How Students Heard about the Center	Percent of Students
A teacher, counselor, or other adult at school	61.2
Posters or announcements at school	38.9
Friends	34.2
Their parents	13.9
Some other way	4.6

SOURCE: Participant survey module in six middle-school sites. The sample size is 263 participants. Some items had smaller sample sizes due to nonresponse. Respondents could cite more than one way they had heard about centers so percentages do not add to 100.

[14]Data on how students had heard about centers were collected from students in six middle school programs in the evaluation. The additional questions were part of the enhancement supported by a Mott Foundation grant.

Students considered attendance at centers to be voluntary on their part (Table II.4). Parent wishes were influential for less than half of participants and reports that their school required them to attend were rare. Parents of middle school participants corroborated that their son or daughter wanted to go to the center. They also indicated that they believed centers would help their child do better in school (78 percent), and more than half said they thought the center was a safe place that would keep their son or daughter safe and out of trouble. More than a third also said they needed child care.

Participating students perceived a range of positive outcomes from their center experiences. For example, many (64 percent) reported that they learned to help others "some" or "a lot" at centers (Table II.5). Participants also said they were more confident in their schoolwork, felt more comfortable with students who were different, and felt more able to work out problems by talking about them.[15]

The voluntary nature of attendance and the positive views of both students and parents did not mean that students attended the centers often, however. Records show that attendance was low during the 2000-2001 school year (Table II.6). Participants who were part of the national evaluation sample ultimately attended centers for 32 days on average during the school year.[16] More than half (54 percent) attended fewer than 25 days. A small percentage of students did

[15]Evaluations of other after-school programs also have found that participants report positive outcomes. See, for example, evaluations of the Beacons program (Warren et al. 1999), the Extended Services Schools program (Grossman et al. 2002), the After-School Corporation program (Reisner et al. 2001), and the California After-School Learning and Safe Neighborhoods Partnership Program (Bissell and Malloy 2002).

[16]In general, participants in the sample were students who had attended centers for three days or more during the first month that centers were open (the basis for selecting students into the sample as participants) and for whom parental consent to be in the study had been given. Appendix A provides more details on how participants were identified.

Table II.4
Reasons Middle School Students Attend Centers
(Middle School Centers in Six Sites)

Why Students Participated	Percent of Students
They wanted to go	66.2
Their parents wanted them to attend	40.0
A teacher, counselor, or other adult at school wanted them to attend	26.2
Someone else wanted them to go	14.5
Why Students Participated according to Their Parents	**Percent of Parents**
My child wanted to go	87.0
It will help my child do better in school	77.5
It is a safe place for my child after school	58.8
It will help my child stay out of trouble	56.7
School staff suggested that my child enroll	43.0
It provides dependable after-school care	39.2
It provides affordable after-school care	33.8
I work and need after-school care for my child	26.4

SOURCE: Participant survey module in six middle school sites and parent survey in all sites. Sample sizes are 1,494 parents and 263 participants. Some items had smaller sample sizes due to nonresponse. Respondents could cite more than one reason for participating so percentages do not add to 100.

Table II.5
Perceived Outcomes of Participation in Middle School Centers
(Middle School Centers in Six Sites)

Percentage of Students Who Report Having Learned to Do the Following "Some" or "A Lot" in the Center:	
Help others	63.9
Feel more confident about my school work	62.1
Feel more comfortable with kids who are different from me	61.1
Feel more confident solving math problems	57.6
Work out problems by talking about them	52.0
Deal with peer pressure	44.6
Speak and understand English better	42.8
Enjoy reading more	42.6

SOURCE: Participant survey module in six middle school sites. Sample size is 514.

Table II.6

21st-Century Middle School Center Attendance
(2000-2001 School Year)

Average Days Attended in the 2000-2001 School Year	32.5
Number of Days Attended (Percent of Participants)	
25 Days or less	53.6
26 to 50 Days	24.2
51 to 75 Days	12.5
76 to 150 Days	9.8
Attendance Rate[a] (Percent of Participants)	
10 Percent or less	18.2
11 to 25 Percent	36.1
26 to 50 Percent	22.4
51 to 70 Percent	9.0
71 to 85 Percent	6.4
86 to 100 Percent	8.0

SOURCE: Center Attendance Records. The sample size is 1,869.

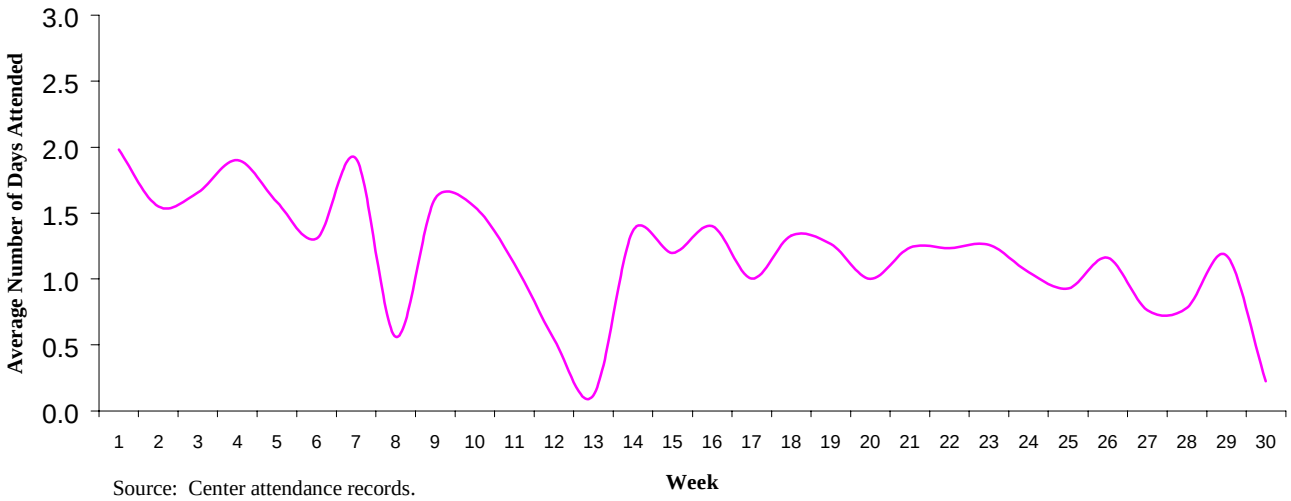
[a]The attendance rate is the number of days participants attended as a proportion of the number of days centers were open, which they provided in their annual performance reports. Totals may not add to 100 percent because of rounding.

attend regularly, however. For example, 22 percent of participants attended centers for more than 50 days, and nearly 10 percent for more than 75 days. Consistent with the low number of days attended, more than half of students had attendance rates of less than 25 percent, meaning they attended centers less than 25 percent of the days that centers were open. The attendance rates reported here are similar to rates reported for the Extended Services Initiative (Grossman et al. 2002).

The pattern of attendance during the school year reveals a slow but steady decline in attendance, with sharper (and temporary) declines around major holidays. Figure II.1 plots average days attended each week during the 2000-2001 school year. The downward trend of

Figure II.1

Average Days Attended Per Week
(Middle School Centers, 2000-2001 School Year)



Source: Center attendance records.

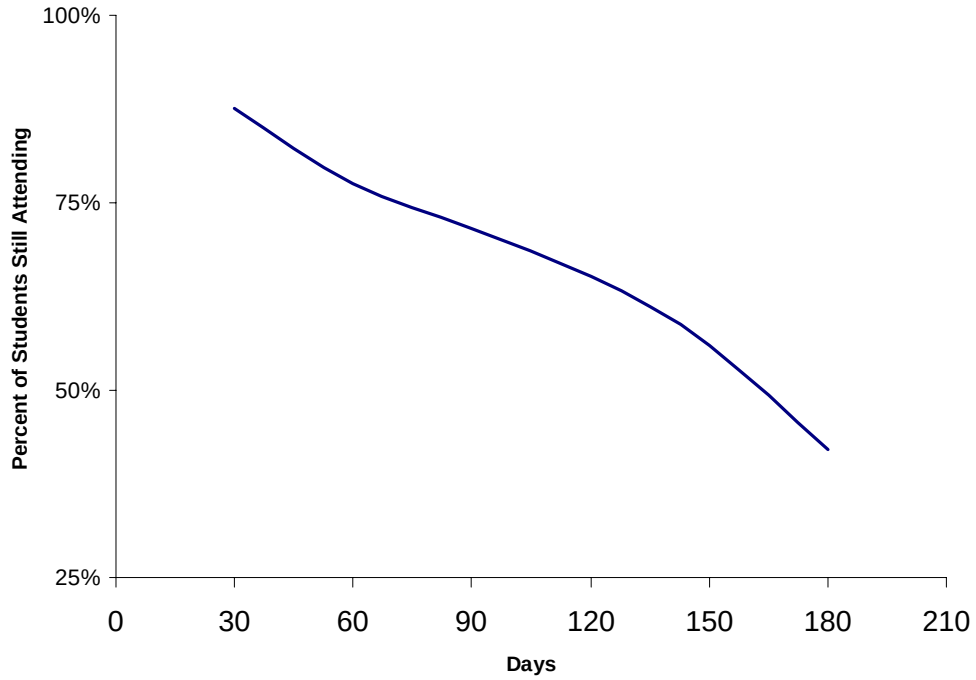
average attendance during the school year is evident, as are the sharp declines around the Thanksgiving and Christmas holidays.[17]

Figure II.2 shows the distribution of days between the first and last calendar day that a middle school student in the sample attended a center. The figure indicates that 60 days after students had started attending centers, 25 percent of them had stopped attending (they had no day of attendance in the center's records more than 60 days after their first day of attendance), and 50 percent of students had stopped attending after 160 days. The figure suggests that a group of students may have been trying out centers and stopped attending after a month or two. Students who continued to attend beyond the tryout period were likely to attend for a longer time, and 25

[17]The figure is based on the number of weeks starting from the week when a center opens, so any particular calendar date can fall in a different week, depending on when centers opened. For this reason, holidays that fall on a specific date can be in different weeks on the figure. The Christmas holiday may fall in the 15th week for a center that opens early in a year and in the 10th week for a center that opens later in a year.

Figure II.2

Distribution of Length of Time Attending Centers
(Middle School Centers, 2000-2001 School Year)



Source: Center attendance records.

percent of middle school students were still attending centers more than 180 days after the first day they attended.

The steady erosion of attendance would have contributed to the need to recruit students throughout the school year to fill slots that became vacant when students stopped attending. Center staff recognized the recruitment and retention issue, with almost 60 percent of program directors and 64 percent of coordinators of middle school centers rating "recruiting students" as a challenge. Staff members told site visitors that they believed the program was less appealing than other things students could do after school, such as sports, other organized activities, or activities at home. However, center policies also made it easy to attend for a limited number of days. Site visitors noted that most centers (82 percent) allowed students to enroll at any time, and many centers allowed students to participate on a drop-in basis, choosing each day whether

to participate. In addition, as noted earlier, many centers scheduled activities in cycles. This may have led to students attending only during the cycles when the activities they wanted to participate in were offered.

Student perceptions about the centers also suggest reasons why recruiting would have been challenging. A survey of students who did not participate in centers at six middle school grantees indicated that, while many nonparticipants thought centers were a fun place to go (86 percent) and a good place to get homework done (87 percent), nearly half (46 percent) considered centers to be "mostly a place kids go when their parents are at work," and 27 percent considered them to be "just for kids who need help in school" (Table II.7). Rather than go to centers, 64 percent of nonparticipants said that they wanted to "hang out" after school, 42 percent said that centers did not have activities they were interested in, and 39 percent said they went to other organized activities after school. Others cited responsibilities, such as doing chores around the house (50 percent) or caring for siblings (28 percent), that may have made participation difficult regardless of how centers were perceived or what they offered.

Twenty percent of nonparticipants also said it was too hard to get a ride home after the program.[18] Nonparticipants said they would be more likely to go to centers if they could choose what to do there (81 percent); if more of their friends attended (78 percent); if centers were less like school (68 percent); and, paradoxically, if they could get their homework done there (67 percent).

[18]Center staff members and teachers also told site visitors that transportation could be a major obstacle to attendance. Some centers did not provide students with transportation home. For some that did, students had to wait too long for buses to leave, bus rides were too long, students got home too late, or buses dropped students too far from home.

Table II.7

Nonparticipant Views about 21st-Century Centers
(Middle School Centers in Six Sites)

What They Think about the 21st-Century Center	
It's good for getting your homework done	86.7
It's a fun place for anyone to go	86.4
It's mostly a place where kids go when their parents are at work	46.8
It's just for kids who need help with school	25.3
It's a punishment for kids who misbehave	13.5
Reasons for Not Participating in Centers	
Want to just "hang out" after school	64.4
Have to do chores around the house	49.8
Program doesn't have activities I want to participate in	42.2
Go to other organized activities after school	38.7
Think the program is for other kinds of kids	22.2
Have to take care of younger brothers or sisters	28.1
Too hard to get a ride home from the program	20.2
Didn't think I would be safe in the program	10.4
Would go to the After-School Program if…	
They could choose what they did there	81.2
More of their friends went there	78.4
It were less like school	67.7
They could get their homework done there	67.4
It were easier to get a ride home	46.4
The teachers there paid more attention to them	40.6
More of their regular teachers were there	41.3

SOURCE: Nonparticipant Survey, Students not attending at all during the year, sample size is 427. Some items had smaller sample sizes due to nonresponse.

C. District and Local School Staff Shared Administrative Responsibilities for the Centers and Used Teachers to Staff Centers

Under the terms of their grants, school district officials had responsibility for developing after-school programs, hiring staff, and overseeing center activities at selected schools within the district. They could use different approaches to carrying out this responsibility. For example, they could administer centers as a part of the host school (which most chose to do), or they could subcontract the center's administration and operation to local organizations (which they did at only 15 percent of schools in the sample). Centers could rely either on teachers as staff (which

was the most frequent approach) or on community members and staff from local organizations (used by a smaller fraction).

Overall, administrators had successfully set up the administrative functions necessary to operate centers. Role delineation was clear, and school centers had the autonomy to customize their offerings. Furthermore, they were able to recruit the staff necessary to lead activities, drawing largely from the ranks of teachers at the host school. Hiring staff, however, consumed considerable administrative attention throughout the school year, possibly limiting time available for other functions such as developing collaborations or planning for sustainability. The preferences for hiring teachers also made investments in professional development appear less necessary, as will be discussed in section C.4 below.

1. Roles and Responsibilities of District and Center Staff

Within the district office, project directors usually administered the centers with little intervention from senior district administrators. In addition, center coordinators (one level below project directors) often were responsible for planning and managing their own centers. This devolution of authority and responsibility stemmed from the view that host school staff members could best understand staff capabilities and student needs and interests. Moreover, many project directors oversaw more than one center and had responsibilities beyond the 21st-Century grant, so decentralization was necessary. Center coordinators were not completely free agents, however. On key decisions, they often consulted with their project director and sought input or approval from school principals or assistant principals.

Project directors usually tried to involve representatives of key organizations or constituencies, including school staff, district staff, community agencies and businesses, and parents when establishing after-school centers. Some project directors used advisory boards as part of these outreach efforts. The responsibilities of advisory boards varied widely. Some

boards met formally to make key decisions, such as how much grant money would be allocated to each center and what activities would be offered. Some centers had boards that played an advisory role but had no decision-making responsibilities. Finally, some centers had no advisory boards at all, either because project directors had been unable to use a convened board effectively or because they relied on informal networks for ideas and feedback.

District and school staff members who were part of the 21st-Century program typically had diverse responsibilities. Project directors often had supervisory and administrative roles—overseeing center operations, hiring center coordinators, dealing with budget issues, and sometimes serving as the key liaison between centers and collaborating partners (Table II.8). Coordinators usually handled the day-to-day details of running the centers, such as recruiting students, setting activity schedules, recruiting and assigning staff members, monitoring attendance, and, sometimes, leading activities.

Table II.8

Project Director Roles in Middle School Centers

	Percentage Reporting		
	Major Role	Minor Role	No Role
Having Final Decision-Making Power	77	23	0
Planning the Program	69	25	6
Hiring Center Coordinators	69	19	13
Hiring Center Staff	47	38	16
Meeting with Center Staff on a Regular Basis	55	29	16
Supervising Center Staff	41	31	28
Preparing Grant Application	41	16	44

SOURCE: Project Director Survey. Sample size is 31.

To gain a sense of the priority given to the development of programs, we were particularly interested in how coordinators—those closest to the day-to-day operations of centers—spent their time. Daily operations occupy much of coordinators' available time. Coordinators devoted over half of their time performing administrative tasks and directly dealing with students (Table II.9). Planning activities ranked third, consuming 22 percent of coordinators' time.

Other staff members typically had roles more narrowly focused than those of coordinators, spending 81 percent of their time working directly with students. Survey responses (not shown in tables) revealed that 59 percent of other staff members reported playing a role in homework help or tutoring. More than a third supervised recreational activities or games, and a similar proportion played a role in math or science instruction.

Table II.9
Staff Time Use in Middle School Centers

	Average Percentage of Time	
	Coordinators	Other Staff
Administrative Tasks	31.0	4.0
Working Directly with Students	29.0	81.0
Planning Activities	22.0	8.0
Consulting with School-Day Teachers	10.0	3.0
Interacting with Parents	8.0	3.0
Other	1.0	2.0

SOURCE: Staff Survey. Sample size for coordinators is 60 and for other staff is 518. Percentages may not add to 100 due to rounding.

2. Qualifications and Work Schedules of Center Staff

About two-thirds of coordinators and of other staff had experience as a classroom teacher, and 34 percent of coordinators and 60 percent of other staff currently worked as a teacher during the school day (Table II.10). Some staff members who did not work at the host school or another

Table II.10

Center Staff Hours, Pay, and Roles

	Coordinators	Other Staff
Average Days a Week Worked at the Center	4.4 days	3.0 days
Average Hours a Day Worked at Center	5.1 hours	2.8 hours
Average Hourly Pay	$16.50/hour	$15.80/hour
Employed by (Percent)		
21st-Century Program	77.0	82
Some other organization	23.0	18
Involvement in Student Activities or Instruction (Percent)		
Lead teacher, tutor, coach, of student activity	52.0	77.0
Assist in student activities	21.0	18.0
Not directly involved in student activities	26.0	5.0
Currently Have Another Job (Percent)		
Yes	66.0	84.0
No	34.0	16.0
Work as Teacher During Regular School Day (Percent)		
Yes	34.0	60.0
No	66.0	40.0

SOURCE: Staff Survey. The sample size is 61 for coordinators and 524 for other staff. Some items had smaller sample sizes as a result of nonresponse. Percentages may not sum to 100 due to rounding.

school during the day worked for another organization, such as a collaborating partner. This group of staff was a minority overall, although they were more of a presence in particular centers. Across all centers, 23 percent of coordinators and 18 percent of other staff reported being employed by an external organization.

Many 21st-Century administrators favored hiring teachers for reasons that they viewed as more important than pedagogical skills. Whereas 45 percent of coordinators rated teaching experience as "very important," 95 percent rated rapport with students as very important, and 91 percent rated experience working with children as very important. Site visit interviews also revealed that administrators wanted staff members to have personal assets, such as a positive attitude, an outgoing personality, and strong interpersonal skills.

3. Hours and Pay

Coordinators worked at centers most days of the week, averaging about five hours a day (Table II.10). Other staff averaged nearly three hours each day, three days a week. Most middle school center staff worked at another job in addition to working at the center. Among center coordinators, 66 percent had another job, often as teachers but also as assistant principals, guidance counselors, or teacher's aides.

Middle school coordinators paid by the hour reported an average hourly wage of $16.50. Other staff members, nearly all of whom were paid by the hour, reported an average hourly wage of $15.80. These pay rates exceed other reported levels of compensation for child care providers and after-school programs. For the Making the Most of Out-of-School Time (MOST) after-school initiative, pay for coordinators ranged from $8 to $15 an hour, and other staff received between $5 and $9 an hour (Halpern et al. 2001). If the focus is on teachers' customary pay, however, these hourly rates are lower than compensation levels for teaching in the regular school day. We estimate that regular teachers received an average of $25 an hour for working a 10-month schedule in the 2001-2002 school year, considerably more than the $15.80 earned, on average, by staff in the programs.[19] Because of provisions in union contracts in some districts, not all teachers were reimbursed at this lower level. We also do not have data to make comparisons with what schools paid teachers for other work outside the regular school day.

4. Professional Training

A common view among project directors and coordinators was that heavy doses of training were unnecessary because staff with teaching backgrounds were sufficiently trained for performing their after-school roles, and training would only increase the demands on teachers'

[19]Based on 1998-1999 data presented in the *Digest of Education Statistics*, 2001, adjusted for inflation.

after-school time commitments. During the 2000-2001 school year, for example, about three-quarters of center coordinators reported receiving training but only a quarter of all other staff reported receiving training. When they received training, coordinators reported an average of 23 hours of training and other staff reported an average of 19 hours of training.

Two kinds of training were common: (1) orientation and (2) skill training. Orientation gave staff members information they needed to know to work in the centers. Often, centers had a general hour- or daylong orientation meeting before the start of the school year (or before the start of each program session during the year) to discuss general issues related to the program, such as its objectives, paperwork requirements, policies and procedures, and plans for the upcoming year or session. Skill training taught staff how to perform tasks critical to their roles. Examples included discipline and classroom management techniques, remedial reading and math instruction, and first aid and CPR techniques. Some center coordinators and project directors also received training addressing broader issues, such as sustaining programs, managing volunteers, and providing high-quality services. These opportunities were often affiliated with state, regional, or national conferences.

D. Programs Established Modest Links to the Regular School Day and Weak Links to Community Organizations

Most centers in the evaluation operated in supportive atmospheres within their host schools, although the programs functioned in tandem with the schools and not as integrated components. Outside organizations from the community, while serving at the behest of the center and school staff, were a major new presence in many of the host schools through their involvement with the after-school program. However, the roles these organizations played did not appear likely to expand substantially in most centers, because both parties had limited interest and inclination in bringing this about.

1. School Support for, and Links with, Centers

Most principals were supportive of the centers. Principals of host schools visited with center coordinators frequently and interacted with center staff in a variety of roles. About half the principals played a major role in planning the program and getting it started, and more than three-fourths were advisers to it (Table II.11). Despite the breadth of interactions on center issues, however, most principals (60 percent) reported that they spent only an hour a week or less on these issues. Principals also believed having teachers working in the centers had led to positive outcomes for centers and for teachers. For example, 76 percent of principals who had a teacher working in a center reported that this strengthens "to a great extent" the alignment of curriculum and instruction between the school day and the center, and 65 percent of principals indicated that it was "not at all" true that teachers working in centers caused students and teachers to spend too much time together. Principals were more moderate in their views of how centers had improved student outcomes. Just over a third reported that homework completion had increased to a great extent, a quarter felt that school attendance had improved to a great extent, and fewer than a fifth felt that classroom behavior had improved to a great extent.

Not surprisingly, centers received their greatest support and had the most communication with teachers from host schools who worked at the centers. These teachers believed that their teaching had been improved by working in centers.[20] Nearly all agreed or strongly agreed that their relationships with some students had improved because they got to work with those students after school, and about three-quarters agreed or strongly agreed that they had had the chance to try new activities and teaching strategies. Teachers who had never worked for the

[20]The statement is based on teachers that worked in centers and did not have a student in the sample. To reduce burden, we excluded from the staff survey any teachers that had students in the sample, because these teachers were being asked to complete a different questionnaire as part of the evaluation.

Table II.11

Middle School Principals' Involvement In Centers

	Percentage
Principal Had a Major Role in:	
Planning the center	48.0
Hiring center staff	36.0
Administering the center	26.0
Meeting with center staff on a regular basis	26.0
Preparing the grant proposal	23.0
Supervising center staff	21.0
Principal Has Position of	
Advisor to staff running the center	75.0
Member of center management and planning committee	53.0
Advisor to outside organization running the center	19.0
Other	10.0
Center coordinator	5.0
No role	3.0
Time Principal Spends on Tasks Related to the Center	
Less than one hour a month	12.0
A few hours a month	28.0
About one hour a week	20.0
A few hours a week	27.0
More than a few hours a week	13.0
How Often Principal Visits the Center	
Never	0.0
Less than once a semester	0.0
One to two times a semester	7.0
About once a month	12.0
A few times a month	15.0
At least once a week	67.0

SOURCE: Principal Survey. Sample size ranges from 57 to 60.

program were generally supportive, at least to the extent of referring students (Table II.12). However, this group of teachers communicated much less with center staff—only about half reported discussions on topics related to students in the after-school program. Consistent with this finding, interviews during site visits indicated that host school teachers not working for centers often knew little about centers and were not involved with them.

Table II.12

Host School Teacher Interactions with 21st-Century Middle School Centers

	All Teachers	Teachers Currently Working for the Program	Teachers That Had Worked for the Program	Teachers That Had Not Worked for the Program
Percentage of Teachers Who Referred Students to Program	93	94	99	91
Percentage of Teachers Reporting They Occasionally or Frequently Communicated with Center Staff to:				
Exchange information about students' assignments	65	77	76	54
Discuss students' academic needs or progress	71	81	78	63
Discuss students' behavior	62	82	66	51
Discuss or identify learning issues exhibited by students	60	77	66	50
Discuss any other items related to coordinating in- and after-school learning	64	84	73	50

SOURCE: Teacher Survey. Sample size is 327 for all teachers, 79 for teachers currently working for the program, 79 for teachers that had worked for the program, and 169 for teachers that had not worked for the program.

Centers linked with their schools by way of offering homework sessions and hiring teachers as staff. Staff and principals perceived that teachers' familiarity with the skills of particular students and with the homework assigned to students were an effective way of tying the after-school program to the regular school's program. More extensive links were evident in some centers where concepts or skills from state achievement tests and state assessment learning standards determined the content of academic assistance sessions. Despite the solid representation of school staff, however, recreation and enrichment activities seldom linked with regular school day instruction. In fact, program designers often wanted to provide a clear distinction with school-day activities to pique students' interest and encourage enrollment.

2. Collaborations with Community Agencies and Organizations

A stated objective in the legislation creating the 21st-Century program was to foster closer ties between schools and communities through collaborative partnerships with a range of local organizations. The evaluation explored the nature and extent of collaborative efforts between centers and local organizations and the tensions and barriers that arose.

Collaborations in the sense of programs and community organizations working together were common, according to information that programs provided on their annual performance reports (Table II.13). Among the most frequent collaborators were community-based organizations, county or municipal agencies, Boys and Girls Clubs, the YMCA, colleges and universities, and local businesses. Libraries and museums, health institutions, and faith-based organizations were noticeably less common. Center coordinators reported collaborations with an average of four organizations.

Information from annual performance reports asks programs to indicate only that they were working with various organizations but did not ask about the nature of the collaboration. A closer examination by site visitors found that the dominant type of collaboration was for local organizations to provide services for hire to the 21st-Century program. For example, a center might pay a local martial arts studio, a dance studio, or a theater company to provide weekly after-school classes, but these outside instructors and their organizations would have limited additional involvement with the program. Collaborations in the sense of shared governance and integrated operations were rare.

The tensions and challenges of collaborating were evident in some programs, where several school staff members considered the staff of other organizations as inexperienced in school settings and expressed concern about their lack of reliability. Instances when outside staff failed to show up when expected to lead activities buttressed this concern. Against this backdrop of

Table II.13

Organizations Working with 21st-Century Centers

	Percentage of Grantees Reporting Participation
Community-Based Organizations	89
County or Municipal Agencies (Such as Police, Parks and Recreation, Health and Social Services)	89
National Organizations (Such as Boys and Girls Clubs, YMCA/YWCA, Big Brothers/Big Sisters)	80
Colleges or Universities	74
Businesses	61
Libraries or Museums	40
Hospitals/Clinics/Health Providers	40
Faith-Based Organizations	29

SOURCE: Tabulations from Annual Performance Reports. Sample size is 225.

perceived disadvantages, and given the ample 21st-Century grant, some centers concluded that they could easily function with little to no involvement of outside organizations.

Some grantees collaborated effectively with community organizations, and their experiences suggest practices that may facilitate stronger collaborations. The grantees provided sufficient staff time to develop and maintain collaborations, and they often had full-time directors or center coordinators who could spend more time working with organizations. They involved organizations at the grant-writing stage, spelling out roles, responsibilities, and budgets there. These successful grantees also fostered open and frequent communications between centers and organization staff through regular planning and coordination meetings, as well as informal communications in the hallways.

E. Programs Depended on Federal Grants and Had Not Prepared for Sustainability

The resources that grantees and centers used to support program operations provide insights into efforts that will be needed to sustain centers after the federal grants expire. The relatively short period entailed in the three-year federal grants puts immediate pressure on districts to begin considering how they will support centers after the grant ends. Information from annual performance reports and from site visits provides a gauge of resources that would need to be replaced and the prospects and challenges districts face in sustaining their after-school programs.[21]

1. Grantee Expenditures and Funding Sources

Performance report data indicate that grantees had budgeted to spend about $196,000 per center, with about $135,000 per center (about 69 percent) from their 21st-Century grants. With centers enrolling about 200 students a year (see Table I.1), planned expenditures amounted to almost $1,000 per enrolled student.[22] Public school expenditures per enrolled student averaged about $6,100 in 1999, so centers were supplementing regular school expenditures by roughly 16 percent.[23]

Grantees also tapped a range of other sources for funds, including grants from other federal, state, or private sources, Title I funds, the USDA National School Lunch Program, private foundations, and cash gifts from private citizens or philanthropies. A handful of grantees in the

[21]More information on sustainability will be available next year, when second-cohort grantees, for which the grant period ended in December 2001, and third-cohort grantees, for which the grant period ends in summer 2002, face the challenge of continuing their centers.

[22]Planned expenditure per program *slot* is higher because students do not attend the full year. For example, a center may have 100 slots that are filled by 200 students over the course of a year (this would be true, for example, if the average participant stayed in the program for half a year). In this case, planned expenditure per slot would be twice as much as planned expenditure per enrolled student.

[23]According to the 1998 *Condition of Education* public-school expenditure per student was $5,734 in the 1996-1997 school year. Inflating this amount by the increase in the Consumer Price Index from 1996 to 1999 (the index value was 156.9 in 1996 and 166.6 in 1999) yields the figure in the text.

national evaluation also charged fees for some or all students to help cover the cost of particular activities in which students participated. However, fees were seldom a major revenue source. Many grantees also received in-kind donations that helped stretch their financial resources. Common in-kind donations included schools providing free access to and use of facilities, equipment, and supplies; community members and teachers volunteering to work in the centers; outside organizations covering the cost of their staff working for centers; businesses, clubs, and other organizations donating equipment, materials, and supplies; and public agencies and private organizations granting reduced-price access to local facilities or events.

2. Approaches to Sustainability

With grantees relying heavily on 21st-Century grant funds, a key issue for their future is their ability to find funding from other sources after the grant ends. Without comparable resources, from either direct finding or in-kind support, service reductions are inevitable.

Project directors and district staff often expressed their desire to keep centers open after the 21st-Century grants end. However, site visitors observed few concrete actions leading toward sustainability. At the time of the visits, about one-third of grantees had made no plans and taken no actions to sustain their programs; half had developed some plans but had not yet taken any action. In surveys of center coordinators and host school principals, only 10 percent of principals and 12 percent of coordinators reported that funding sources had been identified or secured. And, 34 percent of project directors and 40 percent of principals expected that lack of funds would prevent them from having a similar after-school program after the 21st-Century grant ended.

First-cohort grantees that had received the earliest grant awards were no farther along in their planning than grantees with later awards. The fall 2001 status of the first-cohort grantees, whose grants ended in the previous spring, illustrates the varying success that these grantees

enjoyed. Of the nine first-cohort grantees in the national evaluation, one stopped providing services entirely, one had secured a new 21st-Century grant, four were using carryover funds (funds from the grant that had not been expended by the end of the grant period), and three had scaled back their after-school programs significantly.

Grantees were considering a wide variety of potential funding sources. Project directors and other officials could offer few specifics but mentioned some possible sources.[24]

- ***Federal Funds or Grants.*** Several grantees indicated that they were hoping to win another 21st-Century grant (perhaps being overoptimistic; few first-cohort grantees had been successful in winning an additional grant).
- ***State Funds or Grants.*** About one-third of grantees also mentioned the possibility of securing state funding, such as average daily attendance funds and Temporary Assistance for Needy Families funds.
- ***District Funds.*** About one-third of the grantees cited the school district as a potential funding source but noted that their districts were hard-pressed to provide basic services and would not have much money to spare.
- ***Community Funds.*** About one-third of middle school grantees and most of the elementary grantees were looking to the community for future support. They anticipated that community-based organizations, public agencies, and businesses would help sustain after-school programs. The vendor model of relationships commonly in place with many of these external organizations is likely to limit such prospects, however.
- ***Fees.*** A very few grantees indicated they would consider fees—either charging them for the first time or raising current fees—to help pay for services, although some project directors were skeptical that families could afford higher fees.

Three factors appear to have impeded sustainability efforts. First, the responsibility for securing resources to sustain programs typically fell to project directors, but more than half of the directors had responsibilities beyond overseeing the 21st-Century grant and had limited time available for sustainability planning. Second, grantees in poor or remote areas described having limited local resources, such as corporations and foundations, to draw on for future funding.

[24]These data were gathered before the reauthorization of the 21st-Century program, which has charged states with administration of the competitive grants.

Third, the 21st-Century grant's lack of a matching requirement for grantees may have contributed to limited sustainability planning. Many grantees reported that they received matching funds, often in-kind, but the 21st-Century grant did not require matching as a condition. Without a matching requirement, grantees had a full three years of funding at the outset of their grants—enough to offer a full program to a significant number of students; they did not have to find other funding sources during that time, as would be the case for grants that required a larger match in later years.

The newly authorized 21st-Century program that transfers administration to the states addresses sustainability in two ways. Grants can be given for up to five years, and states can set matching requirements for the grants. Whether relaxing these aspects of grants makes sustainability efforts more successful nonetheless depends on other sources of funds being available at the state and local levels. More states and local school districts have begun to provide funding for after-school programs, but even these have been created as time-limited grants rather than as ongoing funding that would be better suited for sustaining programs (Langford 2001).

The lack of a clear path to sustainability has implications for program implementation. Centers (and other after-school programs) in principle may be concerned about developing attributes of their programs such as staff skills, interesting and engaging activities for participants, and connections with local community organizations and agencies, but a lack of assured funding would be a strong disincentive for investing in these activities. If districts planned for centers to be operating for longer periods and if funding were more assured,

investments in staff and curriculum could have greater returns and would be more appealing to districts wanting to raise the quality of center offerings.[25]

We will obtain more evidence about how grantees approach sustainability during the second year of the evaluation, which will see the end of the 21st-Century grants for most grantees in the evaluation.[26] If second- and third-cohort grantees are similar to first-cohort grantees, however, few will have secured funding to sustain their centers.

F. Designers of More Academic After-School Programs in Middle Schools Will Have to Resolve Challenging Issues

The requirement to provide academic services and support from within a school distinguishes 21st-Century centers from after-school programs in general. Site visitors noted that providing academic services increased support for centers from districts and host schools interested in helping students raise grades and test scores. In addition, as noted earlier, parents believed that attending centers would help students do better in school.

Providing academic services also created issues, however. Programs wanted to attract students by being seen as more than just an extension of school, and providing academic services was counter to that desire. Furthermore, the low levels of attendance limited the kinds of academic services centers could offer. With participants attending an average of one or two days a week and sometimes not attending for weeks, teachers in centers would have had just as much difficulty following a curriculum scope and sequence to improve academic outcomes as they would have had if students attended regular school only one or two days a week. Activities that started and ended on the same day, and that reinforced or expanded on what was being taught

[25]Under the new legislation, states will be able to use a portion of their allotted funds to provide technical assistance and training and may undertake some of the investments that individual grantees did not.

[26]Some grantees in the national evaluation may have unexpended funds they can carry over, but the carryover period typically lasts less than a year. Some grantees also may have received additional 21st-Century grants, but these funds usually would apply to centers not studied as part of the national evaluation.

during the regular school day, were probably best suited for centers, although their limited nature would likely not contribute as much to academic improvement.

The nature of middle schools themselves also affected the academic activities middle school centers could offer. With subjects divided into departments and taught at different levels, and with students possibly having many teachers during a school day, linking academic activities in centers with course subject matter would be challenging even if students attended centers frequently.

Attracting teachers to provide or oversee academic activities also created challenges for programs. (Although programs were not required to use certified teachers for academic activities, nearly all of them did.) Teachers are experienced in curriculum and instruction and familiar with the demands of maintaining classroom order, and simply having them as center staff created links to the regular school program. However, center directors and coordinators noted that many teachers had little desire to teach after school when they had already taught a regular school day. Centers accommodated teachers' limited enthusiasm by being flexible in the hours and days that they asked teachers to commit to working in the centers. Although teachers made up the majority of center staff they were also likely to be working the fewest days and hours in the centers (47 percent of all paid staff members, but 56 percent of teachers, worked one or two days a week).

In the end, providing homework sessions was a middle ground that fit within the constraints. Nearly all students said they did homework after school anyway, so homework sessions provided a service that fit their needs. Sessions were not compromised by infrequent attendance, because regular classroom teachers were determining the content of the sessions by the homework they assigned. Even an inexperienced teacher could oversee homework sessions and help students who had questions or needed assistance with their homework.

The broad appeal of homework sessions is counterbalanced by their effectiveness, however. Of all the activities that centers offered, site visitors typically observed that homework sessions engaged the students least, and the teachers who oversaw the sessions did little teaching. Furthermore, improvement in academic outcomes is limited by the extent to which homework would improve academic outcomes in general. Whether homework has these effects is an area where research has been far from unanimous (Cooper and Valentine 2001). The evidence is clearer that homework improves academic achievement for high school students but less clear for middle and elementary school students. Getting more benefit from homework sessions is a challenge that after-school staff will need to address.

III. Impacts Of Middle School Centers

Estimating the impacts of centers was a major goal of the national evaluation. This chapter presents estimates of impacts of middle school centers on a range of student outcomes. These outcomes include activities after school, grades, test scores, behavior in and out of school, and perceptions of safety after school. We also present estimates for the impacts of different types of centers and of centers for different types of students.

Results from the analysis suggest that middle school centers were associated with small increases in school attendance, classroom effort, and math grades, as well as with greater parental involvement in school-related events. On the other hand, centers had no effects on student classroom performance, student disciplinary problems, grades in English, science, and history and students' social development. These were all about the same for center participants and comparison students. Participants were also somewhat more likely to engage in negative behavior outside the classroom and to have been victimized in some way.

This pattern of mixed results has been observed in other studies of after-school programs for middle school students. For example, an evaluation of the Big Brothers and Big Sisters Program found that participants earned slightly higher grades and had slightly better attendance than control group students but that there were no differences between the groups on homework completion (Tierney et al. 1995). An evaluation of the LA's BEST after-school program (Brooks et al. 1995) found that the program improved grades but that its effects on test scores depended on the method used to estimate effects (Brooks et al. 1995). An evaluation of a 21st-Century program in San Francisco found no effects on test scores or grades, although the program did affect some types of students (Trousdale 2000).

A. Characteristics of Participant and Comparison Students Were Similar

As noted in Chapter I, to estimate impacts of middle school centers, a random sample of grantees serving middle school students was selected, and for each grantee, comparison students were identified using propensity score matching (Rosenbaum and Rubin 1983). Propensity score matching selects comparison group students from a pool of potential students based on how closely their characteristics resemble those of participants. Potential comparison group members are matched to participants with similar propensity scores.

The score matching procedure ensures that the overall group of comparison students resembles participants on most observed characteristics, but some differences can arise for two reasons. First, the matching process is inexact and can yield participant and comparison groups that differ on some characteristics. Second, parents had to give their consent to be in the evaluation after matches were identified, and whether or not parents gave their consent could be related to the students' characteristics. We used regression modeling techniques to adjust for differences between participants and comparison students on characteristics that could be observed. Appendix B provides more details about the matching process and the regression techniques used to estimate impacts.

The success of the matching method is suggested by the small differences of the participants and comparison groups on a variety of characteristics (Table III.1). For example, 47.1 percent of the participant group and 46.5 percent of the comparison group were male, 12.3 percent of the participant group and 11.9 percent of the comparison group were Hispanic, and 52.0 percent of the participant group and 58.0 percent of the comparison group lived in a two-parent household.

On the other hand, because information from parent questionnaires was available only after matching was completed, the design could not use information about parents to match students.

Table III.I

Characteristics of Center Participants and Comparison Group Students:
Middle School Centers

Characteristic	Percentage of Program Participants	Percentage of Comparison Group Members	p-Value[a]
Gender			
Male	47.1	46.5	0.70
Female	52.9	53.5	0.70
Race/Ethnicity			
White (non-Hispanic)	38.2	40.6	0.15
Black (non-Hispanic)	27.8	24.8	0.05*
Hispanic	12.3	11.9	0.73
Other	15.5	15.8	0.79
Mixed race	6.3	6.8	0.53
Grade Level			
6	20.5	21.3	0.63
7	37.6	38.2	0.62
8	34.3	34.4	0.82
Other or ungraded	7.7	6.0	0.03**
Primary Language at Home Is Not English	17.7	18.8	0.38
Household Structure			
Two parents	52.0	58.0	0.44
Single mother	28.5	24.3	0.09*
Single father	2.9	2.7	0.87
Other	16.6	15.0	0.16
Parental Education			
Mothers have at least a high school degree/GED	79.9	78.5	0.30
Mothers have a four-year college degree	10.8	14.4	0.00***
Fathers have at least a high school degree/GED	76.9	76.7	0.90
Fathers have a four-year college degree	11.7	16.0	0.00***
Employment Status			
Both parents with a full-time job	26.6	31.0	0.00***
Mothers with a full-time job	59.7	60.5	0.57
Mothers with a part-time job	14.3	15.8	0.21
Fathers with a full-time job	81.1	84.6	0.01***
Fathers with a part-time job	4.5	3.6	0.53
Household Income			
Less than $10,999	16.1	15.0	0.35
$11,000-$24,999	27.5	24.8	0.07*
$25,000-$39,999	24.7	21.0	0.01***
$40,000-$59,999	17.5	18.4	0.49
More than $60,000	14.1	20.7	0.00***
Receipt of Government Assistance			
Food stamps	18.3	17.3	0.44
TANF	7.4	6.8	0.51

Table III.1 (*continued*)

Characteristic	Percentage of Program Participants	Percentage of Comparison Group Members	p-Value[a]
Student-Reported Baseline Grades			
Mostly A's	30.3	34.1	0.02**
Mostly B's	35.7	36.4	0.31
Mostly C's	23.3	21.2	0.03**
Mostly D's or below	8.8	7.6	0.34
Not graded	1.8	0.7	0.00***
Sample Size[b]	**1,752**	**2,437**	

SOURCE: Student survey, parent survey, school records.

[a]The p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and comparison group members equals zero can be rejected. If the p-value is less than .01, the difference is significant at the 1 percent level; if the p-value is less than .05, the difference is significant at the 5 percent level, and so on.

[b]Sample sizes for some characteristics differ because of missing values.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Data gathered from parents after matching found that a larger proportion of participants lived in households with low annual incomes, and participants were more likely to be from a single-parent household and to have a mother or father with less than a four-year college degree.

The differences between the participant and comparison groups underscore the importance of adjusting for student characteristics in measuring impacts. The regression models (discussed in Appendix B) used to estimate impacts adjust for observed participant-comparison differences in characteristics such as those shown in Table III.1.[27]

B. How Did Middle School Centers Affect Students?

The conceptual framework presented in Chapter I links participation in centers with a range of outcomes. For example, participation might immediately influence students' after-school activities, location, and supervision. These factors could in turn influence other outcomes, such as school attendance, participation in class, and completion of homework. They could also extend to long-term outcomes such as course grades and test scores. Similarly, participation could influence student behavior and personal development through the activities and services centers provide.[28]

[27]We tested the extent to which regression models were able to reduce baseline differences by estimating models in which the baseline outcome was regressed on student characteristics. Results from these models indicate that the baseline differences were substantially reduced by the adjustment. In practice, the regression adjustment approach is stronger than what is found in the test because the baseline value of the outcome variable is included in the model.

[28]Because impacts can represent different units and the absolute magnitude of the impacts is not always informative, the text sometimes refers to an impact's "effect size," which is the impact expressed as a percentage of the outcome's standard deviation. Effect sizes of 10 to 20 percent are common in program evaluation and effect sizes of 30 percent or more are considered large and are relatively uncommon.

1. Centers Increased Adult Supervision but Did Not Affect Self-Care

Centers decreased how often parents or older siblings cared for students and increased how often adults who were not their parents (for example, program staff members) cared for these students. Table III.2 shows that a parent cared for 53.2 percent of participants at least three days during a typical week, compared with 59.2 percent of comparison students.[29] It also shows that an older sibling cared for 4.6 percent of participants, compared with 7.2 percent of comparison group students. On the other hand, 20.2 percent of participants and 11.7 percent of comparison group students were being cared for by an adult who was not their parent. However, centers did not affect whether students were in self-care (defined as students being by themselves, with friends, or with younger siblings after school, and not being cared for by an adult). About 17 percent of both participants and comparison group students were in self-care at the time of the follow-up survey. The net effect was to increase the proportion of students being cared for by an adult (either a parent or a non-parent adult) by reducing the proportion being cared for by older siblings.[30]

Not surprisingly, centers also increased the time students spent at school during the after-school hours. Participants were more likely to remain at school after the regular school day ended, with 30 percent spending three or more days in a typical week at school, compared with 18 percent of comparison students (Table III.2). During the average week, participants spent twice as many days at school as comparison group members (1.2 versus 0.6 days, on average, an effect size of 10 percent). Participants also were less likely than comparison students to be in the

[29]The results reported in Table III.2 are based on student reports of their after-school location and care arrangements. We also examined impacts based on parent reports of location and care arrangements. The patterns of impacts were similar, though parents were more likely to report their child was cared for by a parent.

[30]Defining self-care in other ways yielded similar results. There was no significant difference between treatment and comparison groups in whether students were home alone for three or more days a week, or for one or more days a week.

Table III.2

Center Impacts on Location, Supervision, and Student Activities After School:
Middle School Centers

Outcome	Student-Reported Supervision and Location			
	Center Participants	Comparison Group	Estimated Impact of Participation	p-Value[a]
Percentage of Students in the Following Types of Supervision at Least Three Days after School in a Typical Week:				
Self-care	17.5	17.3	0.1	.92
Parent care	53.2	59.2	–6.0***	.00
Nonparent adult care	20.2	11.7	8.5***	.00
Sibling care	4.6	7.2	–2.6***	.00
Mixed care (not in any one category for at least three days)	4.6	4.6	0.0	.94
Percentage of Students in the Following Locations after School at Least Three Days in a Typical Week:				
Own home	67.5	78.4	–10.9***	.00
Someone else's home	8.9	13.3	–4.4***	.00
School or other place for activities	30.0	18.0	12.0***	.00
Somewhere to "hang out"	9.0	10.3	–1.3	.24
Mixed location (not in one location for at least three days)	8.7	7.2	1.4	.16
Mean Number of Days Stayed after School for Activities in Typical Week	1.2	0.6	0.6***	.00
Mean Number of Days Students Participate in Activity after School:				
Homework	3.2	3.1	0.1	.12
Tutoring	0.6	0.3	0.3***	.00
Non-homework reading, writing, or science activities	1.4	1.2	0.2**	.01
School activities (band, drama, etc.)	1.0	0.6	0.4***	.00
Lessons (music, art, dance, etc.)	0.6	0.5	0.1	.19
Organized sports	1.4	1.2	0.2***	.00
Clubs (Boy and Girl Scouts, Boys and Girls Club, etc.)	0.4	0.2	0.1***	.00
Activities at church, temple, mosque	0.5	0.5	0.0	.83
Watched TV or videos	3.5	3.5	0.0	.91
Surfed the Internet or did other things on a computer	2.0	1.8	0.2***	.01
"Hung out" with friends	2.7	2.4	0.3***	.00
Volunteered or did community service	0.4	0.3	0.1	.11
Worked at a job	0.5	0.4	0.0	.51
Did chores around the house	3.0	3.1	-0.1	.17
Took care of a brother or sister	1.6	1.6	0.0	.89
Sample Size[b]	**1,750**	**2,437**		

SOURCE: Student survey.

NOTE: The percentages and mean values of outcomes for participants and comparison group members have been regression-adjusted for baseline differences between the groups. The control variables in the regression included 26 different student and household characteristics, such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated impacts shown in the table do not always equal the difference between center participants and the comparison group.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

[a]The p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and control group members equals zero can be rejected. If the p-value is less than .01, the difference is significant at the 1 percent level. If the p-value is less than .05, the difference is significant at the 5 percent level, and so on.

[b]Sample sizes differ for some outcomes due to missing values.

homes of others after school. However, centers did not affect the amount of time students spent "hanging out" away from school or home—participants were as likely as comparison students to report going somewhere after school to "hang out."[31]

2. Centers Affected After-School Activities

Center participants were more likely to report doing academic activities after school. Students reported on whether they had engaged in up to 15 different activities after school during a recent typical week, including doing homework, being tutored, watching television, and doing chores around the house. Participants were significantly more likely than comparison students to have engaged in seven of these activities (Table III.2), and, in particular, participants were more likely to say that they received tutoring after school (with the average participant being tutored twice as often as the average comparison student during a typical week, an effect size of almost 30 percent). Participants also were more likely to say that they participated in some reading, writing, or science activity not related to homework.[32]

Centers also affected nonacademic activities. During a typical week, participants spent an average of 1.0 days participating in school activities such as band or drama, compared with 0.6 days for comparison students, an effect size of 27 percent (Table III.2). Participants also were more likely to engage in organized sports or activities at clubs, possibly because some after-school programs linked with clubs to provide activities during program hours. In addition, while the two groups spent the same amount of time watching television after school, participants spent

[31]The term "hanging out" is used in the questionnaires completed by students and they were free to interpret the expression in their own way.

[32]As with location and care, we also examined student and parent reports about after-school activities. There were some differences between students and parents in reported activities after school, but student reports are presented here based on the assumption that student reports of their after school activities are more accurate.

more time doing things on a computer, such as surfing the Internet (2.0 versus 1.8 days, effect size is 10 percent).[33]

The location and care information in Table III.2 provides some information about how students spend after-school time that may have a bearing on impacts. Interestingly, most students in the comparison group reported that they were at home for at least part of the after-school time, and many also reported that they were with a parent. Additional analyses not reported here found that the majority (53 percent) of comparison-group students reported being at home *and* being with a parent after school. Centers and center staff can have impacts by being a possibly more academically oriented form of care, but may not be a better form of care in other dimensions.

Also, differences in location and care arrangements (shown in Table III.2) were related to differences in participation in after-school activities, but the largest differences were for activities that are more likely to take place in school (for example, band and drama, and sports). Being cared for by other adults and being at school were associated with higher levels of these kinds of after-school activities for all students. However, activities that could occur either at school or elsewhere, such as music lessons, often were similar for students in various locations or types of care. For example, 24 percent of comparison students in the care of other adults reported participating in lessons, compared with 23 percent of comparison students being cared for by a parent and 22 percent of comparison students in self-care. And participation in homework generally was high for all care and location categories.[34] The data are not detailed

[33]Participants possibly considered time at the center as time spent "hanging out with friends." In this case, even if they reported that one of their activities was "hanging out with friends," they may have reported their location as being at school.

[34]Across location and care categories, the lowest rate of doing homework (77 percent) was for students who said they went somewhere to "hang out" after school. The highest rate of doing homework (88 percent) was for

enough to describe after-school activities of students who attend other types of formal after-school programs such as church-based programs or Boys and Girls Clubs.

3. Centers Improved Some Academic Outcomes

a. Centers Improved Attendance and Classroom Effort but Not Homework Completion

Centers were associated with slightly higher school attendance. School records indicated that center participants were absent an average of 9.0 days during the 2000-2001 school year and comparison students an average of 10.1, an effect size of 11 percent (Table III.3). Similarly, participants were less likely to be late to school (5.0 days late, versus 6.2 days late, effect size 11 percent).[35]

Centers did not increase homework completion, with 83.4 percent of both groups reporting that they "often" or "always" did the homework their teachers assigned (Table III.3). (Teachers reported that about half of participants and comparison students completed their homework.) Participants, however, were more likely to complete assignments to the teacher's satisfaction (58.0 versus 53.3 percent, effect size 9 percent).[36] Another indication of greater classroom effort among participants is that a composite measure of student effort as reported by teachers was

(continued)

students in parent care. Students in parent care were more likely to do homework (88 percent) than students in self-care (81 percent), but the difference is not as large as might be expected.

[35]We also examined impacts on attendance and tardiness based on teachers reports. Attendance impacts were consistent with impacts based on records data. Tardiness impacts based on teacher reports were small and statistically insignificant.

[36]The difference between teacher evaluations of whether students completed homework and whether they completed assignments to teachers' satisfaction may arise because the more general term "assignments" includes work done in class as well as work done at home.

Table III.3

Impacts on Homework Completion and Level of Effort and Behavior in the Classroom: Middle School Centers

Outcome	Center Participants	Comparison Group	Estimated Impact of Participation	p-Value[a]
Percentage of Students Who Report That They "Often" or "Always" Do the Homework Teachers Assign	83.4	83.4	0.1	.97
Percentage of Students Whose Teachers Report That They "Often" Complete Their Homework	50.3	49.6	0.7	.69
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That They Complete Assignments to the Teacher's Satisfaction	58.0	53.3	4.8***	.01
Percentage of Students Whose Teachers Report That They "Usually Try Hard" in English Class	51.3	48.4	2.9	.12
Teacher-Reported Level of Effort Composite[b] (Mean)	3.6	3.5	0.1***	.00
Percentage of Students Whose Parents "Agree" or "Strongly Agree" That Their Child Works Hard at School	75.1	75.3	-0.2	.91
Student-Based Discipline Problem Composite[c] (Mean)	1.4	1.4	0.0	.91
Teacher-Based Discipline Problem Composite[d] (Mean)	1.4	1.4	0.0	.26
Mean Number of Days Student Was:				
Absent from class	9.0	10.1	-1.1***	.00
Late for class	5.0	6.2	-1.1***	.00
Sample Size[e]	**1,752**	**2,437**		

SOURCE: Student survey; teacher survey.

NOTE: The percentages and mean values of outcomes for program participants and comparison group members have been regression-adjusted for baseline differences between the groups. The regression's control variables included 26 different student and household characteristics, such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated impacts shown in the table do not always equal the difference between center participants and the comparison group.

[a]The p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and control group members equals zero can be rejected. If the p-value is less than .01, the difference is significant at the 1 percent level. If the p-value is less than .05, the difference is significant at the 5 percent level, and so on.

[b]The level of effort composite is based on five items reported by teachers: whether the student (1) usually tries hard, (2) often performs at or above his or her ability level, (3) is attentive in class, (4) participates in class, and (5) volunteers in class. The composite is equal to the mean of the five variables. Values on these items range from 1 to 5; a value of 1 on the composite indicates a low level of effort, and a value of 5 indicates a high level of effort.

[c]The student-based discipline problem composite is based on four items: the extent to which students report (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having their parents called to school about a problem they are having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems.

[d]The teacher-based discipline problem composite is based on four items: the extent to which the teacher reports that the student is (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having their parents called to school about a problem they are having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems.

[e]Sample sizes may differ for some outcomes due to nonresponse.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

higher for participants, although this evidence is tempered by two factors.[37] First, as with the impact on attendance, the impact on classroom effort was small. For example, one of the items contributing to the effort composite is whether the student "usually tries hard" in English class. Teachers were more likely to report that participants usually tried hard, but the difference is only 3 percentage points (51.3 percent among participants, versus 48.4 percent among comparison students) and the effect size is only 7 percent. Second, according to their parents, participants were no more likely to "work hard in school": about three-fourths of the parents of both groups reported that their child works hard.

b. Centers Did Not Affect Classroom Behavior

According to students and their teachers, participants were equally likely to skip school, be sent to the office, get a detention for misbehaving, or have their parents called because of a behavior problem. Table III.3 shows that participants and comparison group students had the same value of a composite measure of disciplinary problems.

c. Centers Increased Grades in Math but Not in Other Subjects

Centers improved grades in math but not in other subjects (Table III.4). Participants and comparison students had about the same English, science, and social studies or history grades (in each case, the average grade was about 81). In math, however, participants had a marginally higher grade (80.3 points, compared to 79.5 points). This effect was statistically significant at the 10 percent level, but the effect size is only 6 percent.[38]

[37]The five items on the effort composite are whether the student (1) usually tries hard, (2) often performs at or above ability level, (3) is attentive in class, (4) participates in class, and (5) volunteers in class.

[38]Impacts on test scores also were estimated but not reported here, as the magnitude and direction of the estimated impacts was sensitive to the method used to impute missing baseline test scores. Only a fourth of the sample had a follow-up score available and only a tenth of the sample had both a baseline and a followup score available.

Despite reporting slightly higher levels of classroom effort among participants, teachers were no more likely to report that program participants performed better than comparison students academically. Teachers reported that about one-third of each group achieved at an "above-average" or "very high" level (Table III.4), and 52 percent of teachers reported that the two groups "get good grades on tests."

4. Centers Did Not Improve Behavioral and Youth Development Outcomes

a. Centers Did Not Improve Social and Personal Development

In general, centers did not improve developmental outcomes measured by the evaluation (Table III.5). For example, participants and comparison students had about the same values of a social engagement composite variable based on a set of variables that reflect how the students get along with others and how easily they can make and keep friends (however, the high value of the variable, 3.5 on a scale of 4, may have made impacts difficult to generate). In addition, no effect was found on a peer interaction and empathy composite, which reflects the extent to which students work well with others, have empathy for others, and believe the best about others. Centers had a small positive effect on educational expectations, with 83 percent of participants versus 80 percent of comparison students reporting that they expect to graduate from college, an effect size of 9 percent.[39]

b. Centers Increased Parent Involvement

Parents of participants were more likely to report that they regularly participated in school events and school-related activities. For example, 27 percent of the parents of participants

[39]We also estimated insignificant impacts on student confidence in their reading ability, whether students had a positive attitude toward learning according to their teacher, whether students go along well with others, students' ability to plan and solve problems, and the extent to which students helped their parents.

Table III.4

Impacts on Teacher-Reported Achievement and Grades:
Middle School Centers

Outcome	Center Participants	Comparison Group Members	Estimated Impact of Participation	p-Value[a]
Percentage of Students Whose Teachers Report That They Achieve at an "Above-Average" or "Very High" Level	33.6	32.9	0.7	.67
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That They Get Good Grades on Tests	52.2	52.2	0.0	.99
Teacher-Reported Achievement Composite (Mean)[b]	3.4	3.4	0.0	.58
Mean Grade				
Math	80.3	79.5	0.7*	.06
English	80.9	80.9	0.1	.87
Science	81.3	81.1	0.1	.81
Social studies/history	81.0	80.5	0.4	.33
Sample Size	**1,752**	**2,437**		

SOURCE: Teacher survey; school records.

NOTE: The percentages and mean values of outcomes for program participants and comparison group members have been regression-adjusted for baseline differences between the groups. The control variables in the regression included 26 different student and household characteristics, such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores, attendance, disciplinary problems, and self-reported grades.

[a]The p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and control group members equals zero can be rejected. If the p-value is less than .01, the difference is significant at the 1 percent level. If the p-value is less than .05, the difference is significant at the 5 percent level, and so on.

[b]The teacher-reported achievement composite is based on teacher responses to five questions: (1) At what level is this student performing in reading? (2) Does this student get good grades on tests? (3) Does this student complete assignments to my satisfaction? (4) Does this student have good communication skills? (5) Is this student a proficient reader? Values on these items range from 1 to 5; a value of 1 on the composite indicates low achievement, and a value of 5 indicates high achievement.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Table III.5

Impacts on Social Engagement, Educational Expectations, and Parental Involvement: Middle School Centers

Outcome	Center Participants	Comparison Group Members	Estimated Impact of Participation	p-Value[a]
Student-Reported Social Engagement Composite[b] (Mean)	3.5	3.5	0.0	.43
Peer Interaction/Empathy Composite[c] (Mean)	3.0	3.1	0.0	.12
Percentage of Students Who Rate Themselves as "Good" or "Excellent" at Working Out Conflicts with Others	60.7	65.0	–4.2**	.01
Percentage of Students Who Rate Themselves as "Good" or "Excellent" on Using a Computer to Look Up Information	36.8	34.4	2.3	.18
Percentage of Students Who Think They Will:				
Graduate from college	82.9	79.7	3.2**	.01
Graduate from high school but not college	15.3	18.0	–2.7**	.03
Attend high school but not graduate	1.9	2.3	–0.5	.34
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year				
Attended an open house at the school	27.4	19.1	8.4***	.00
Attended parent-teacher organization meetings	33.8	27.6	6.1***	.00
Attended an after-school event	47.0	40.2	6.8***	.00
Volunteered to help out at school	17.8	14.5	3.3**	.02
Sample Size[d]	**1,752**	**2,437**		

SOURCE: Student survey; teacher survey; parent survey.

NOTE: The percentages and mean values of outcomes for program participants and comparison group members have been regression-adjusted for baseline differences between the groups. The control variables in the regression included 26 different student and household characteristics, such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated impacts shown in the table do not always equal the difference between center participants and the comparison group.

[a]The p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and control group members equals zero can be rejected. If the p-value is less than .01, the difference is significant at the 1 percent level. If the p-value is less than .05, the difference is significant at the 5 percent level, and so on.

[b]The social engagement composite is based on five items: the extent to which students report that they (1) have friends to "hang out with," 2) are never lonely, (3) get along with others their age, (4) find it easy to make new friends, and (5) never feel left out of things. The composite is equal to the mean of the five variables. Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level of social engagement, and a value of 4 indicates a high level of engagement.

[c]The peer interaction/empathy composite is based on three items: students' rating of their ability to (1) work with others in a team or group, (2) feel bad for other people who are having difficulties, and (3) believe the best about other people. Values on these items range from 1 to 4; a value of 1 on the composite indicates poor peer interactions, while a value of 4 indicates excellent peer interactions.

[d]Sample sizes may differ for some outcomes due to nonresponse.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

attended an open house at the school three or more times during the year, compared with 19 percent of comparison parents, an effect size of 21 percent (Table III.5). Parents of participants were more likely to attend parent-teacher organization meetings (34 versus 28 percent, effect size 14 percent), to attend after-school events (47 versus 40 percent, effect size 14 percent), and to volunteer to help out at school (18 versus 15 percent, effect size 8 percent).

c. Centers Did Not Affect Feelings of Safety and Increased Some Negative Behaviors

Centers did not affect feelings of safety after school. More than 60 percent of students reported that they felt "very safe" after school, and only 3 percent reported that they feel "not at all safe" (Table III.6). The rest of the students reported feeling "somewhat safe." However, center participants were somewhat more likely to report having had their property (such as clothing or books) damaged on purpose, with 17 percent of participants and 14 percent off comparison students reporting having been victimized in this way, an effect size of 7 percent (Table III.6).

Students were asked about different types of negative behavior, including the extent to which they break things on purpose, punch or hit others, argue with or lie to their parents, give a teacher a "hard time," steal from a store, sell illegal drugs, or get arrested by police.[40] Participants were more likely than comparison students to report selling illegal drugs (Table III.6), with 3 percent of participants and 2 percent of comparison students reporting that they did this "some" or "a lot" (an effect size of 6 percent). Participants also had a higher value of the negative behavior composite variable, which reflects the frequency with which students reported engaging in these behaviors.

[40]Giving teachers a "hard time" was the expression used in the questionnaire completed by students and they were free to interpret the expression in their own way.

Table III.6

Impact on Student Safety, Negative Behavior, and Victimization: Middle School Centers

Outcome	Center Participants	Comparison Group	Estimated Impact of Participation	p-Value[a]
Percentage of Students Who Report Feeling the Following Levels of Safety After School up Until 6:00 P.M.:				
Very Safe	60.5	62.1	–1.6	.35
Somewhat safe	36.2	34.7	1.6	.37
Not at all safe	3.2	3.2	0.1	.92
Percentage of Students Who Report That They Do the Following "Some" or "A Lot":				
Break something on purpose	8.8	7.8	1.0	.31
Punch or hit someone	20.9	18.9	2.1	.15
Steal from a store	4.8	3.7	1.2	.11
Sell illegal drugs	3.3	1.8	1.5***	.01
Get arrested or detained by police	3.8	3.3	0.5	.48
Negative Behavior Composite[b] (Mean)	1.53	1.49	0.04***	.01
Percentage of Students Who Report the Following Happened to Them "Some" or "A Lot":				
Been offered, sold, or given an illegal drug	15.1	13.2	1.9	.13
Been "picked on" after school	32.4	30.7	1.7	.31
Been threatened or hurt with a weapon	7.7	6.3	1.4	.15
Been threatened by a gang or gang member	8.2	8.1	0.1	.88
Had property damaged on purpose	16.9	14.1	2.8**	.03
Percentage of Students Who Report That They Do the Following "Some" or "A Lot":				
Smoke cigarettes	2.6	2.3	0.3	.90
Have at least one alcoholic drink	6.6	6.7	–0.1	.90
Smoke marijuana	3.7	2.7	1.0*	.10
Student-Reported Tobacco, Alcohol, and Drug Use Composite[c] (Mean)	1.1	1.1	0.0	.68
Sample Size[d]	**1,752**	**2,437**		

SOURCE: Student survey; parent survey.

NOTE: The percentages and mean values of outcomes for program participants and comparison group members have been regression-adjusted for baseline differences between the groups. The control variables in the regression included 26 different student and household characteristics, such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated impacts shown in the table do not always equal the difference between center participants and the comparison group.

[a]The p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and control group members equals zero can be rejected. If the p-value is less than .01, the difference is significant at the 1 percent level. If the p-value is less than .05, the difference is significant at the 5 percent level, and so on.

[b]The negative behavior composite is based on student responses to eight questions regarding how frequently they do the following: (1) break something on purpose, (2) punch or hit someone, (3) argue with their parents, (4) lie to their parents, (5) steal from a store, (6) give a teacher a hard time, (7) sell illegal drugs, and (8) get arrested or detained by police. Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level of negative behavior, while a value of 4 indicates a high level of negative behavior.

[c]The tobacco, alcohol, and drug use composite is based on seven items: the extent to which students (1) smoke cigarettes, (2) use smokeless tobacco, (3) have at least one drink of alcohol, (4) have five or more drinks of alcohol in a row, (5) smoke marijuana, (6) use inhalants, and (7) use any other illegal drug. Values on these items range from 1 to 4; a value of 1 on the composite indicates no substance abuse, while a value of 4 indicates frequent substance abuse.

[d]Sample sizes may differ for some outcomes due to nonresponse.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

C. Center Participation Affected Some Students More than Others

Center participation may affect some students more than others. Effects were estimated and compared for six subgroups defined by (1) grade level, (2) race or ethnicity, (3) gender, (4) high versus low baseline grades, (5) high versus low baseline disciplinary problems, and (6) single-parent versus two-parent households.[41]

Findings from the subgroup analysis indicate that participation improved a range of student outcomes for black students and for students who had had fewer disciplinary problems in the prior year. Among black students, for example, centers increased effort in the classroom, reduced lateness for school, and increased math grades (Table III.7A).[42] Impacts on math grades and being on time to class were evident for Hispanic students. None of these impacts were evident for white students. For students with fewer behavioral problems (in the baseline year), centers increased effort in the classroom and math and social studies grades (Table III.7B). None of these impacts were evident among students that had had more disciplinary problems. Participation also increased the extent to which female students were victimized either by being "picked on" after school or by having their property damaged (Tables III.8A and III.8B).[43] Among males, participation did not significantly affect either of these outcomes. And, although centers increased parental involvement in school-related activities for nearly all groups of students, increases in involvement for parents in two-parent families were larger than for parents in single-parent families. For example, participation led to a 14 percentage point increase in parents from two- parent households attending open houses but only a 6 percentage point increase for single parents.

[41]In addition, we estimated subgroup impacts based on the mother's education, whether a student's teacher worked in an after-school program, and whether a student had attended an after-school program in the previous year. There were almost no statistically significant impacts for these subgroups.

[42]As with score results for the full sample, subgroup score impacts were estimated but found to be sensitive to alternative estimation methods and not reported.

[43]Being "picked on" was the expression used in the questionnaire completed by students and they were free to interpret the expression in their own way.

Table III.7A

Impacts on Homework Completion, Level of Effort, and Classroom Behavior, by Subgroup: Middle School Centers

Outcome	Estimated Impact						
	Grade Level		Race/Ethnicity			Gender	
	5-6	7-8	White	Black	Hispanic	Male	Female
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That:							
Student completes assignments to my satisfaction	2.5	5.3**	4.6*	7.3**	2.5	3.9	5.0**
Student comes prepared and ready to learn	1.6	5.2**	3.5	6.2*	5.4	5.6**	2.9
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That:							
The student is attentive in class	0.9	3.8*	**1.3**	**11.0*** **	**1.6**	4.2	1.8
The student participates in class	4.8	4.5**	2.4	7.4**	5.0	4.5*	4.4*
Teacher-Reported Level of Effort Composite (Mean)	0.0	0.1***	**0.0**	**0.2*** **	**0.1**	0.2***	0.1
Teacher-Reported Disciplinary Problems Composite (Mean)	**0.1***	**–0.1***	0.0	–0.1	–0.0	–0.0	0.0
Number of Days of School Student Was:							
Absent from class	**–2.0*** **	**–0.8** **	–0.7	–1.6**	–1.3**	–0.9**	–1.3***
Late for class	–0.7	–1.3***	**0.1**	**–1.8*** **	**–1.6** **	**–1.8*** **	**–0.6**
Percentage of Students Whose Teachers Report That They Achieve at an "Above-Average" or "Very High" Level	–3.4	1.7	–2.4	2.1	2.6	**4.4***	**–3.1**
Mean Grades							
Math	1.2*	1.6	**–0.3**	**1.7** **	**1.5***	0.6	0.9*
English	0.7	0.3	–0.2	0.8	1.0	0.0	0.4
Science	1.0	0.0	0.2	0.4	0.4	–0.5	0.8
Social studies	**2.2*** **	**0.1**	–0.3	0.7	1.0	0.2	0.6
Sample Size							
Student-reported outcomes	1,138	2,893	1,391	940	1,091	1,868	2,156
Parent-reported outcomes	987	2,588	1,304	819	937	1,657	1,916
Teacher-reported outcomes	939	2,238	1,152	721	884	1,494	1,679
School records outcomes (attendance)	1,095	2,818	1,345	939	1,067	1,834	2,078
School records outcomes (grades)	889	2,685	1,214	890	935	1,702	1,871

SOURCE: Parent survey; student survey; teacher survey; school records.

NOTE: Subgroup impacts in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other subgroup(s) at the .10 level or higher.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Table III.7B

Impacts on Homework Completion, Level of Effort, and Classroom Behavior, by Subgroup: Middle School Centers

Outcome	Estimated Impact					
	Baseline Grades		Baseline Disciplinary Problems Composite		Number of Parents in Household	
	Low	High	Low	High	Two	One
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That:						
Student completes assignments to my satisfaction	6.1**	4.5**	**7.1*** **	**–0.2**	6.1**	5.1**
Student comes prepared and ready to learn	4.1	4.6**	**7.6*** **	**–2.7**	3.8	4.0*
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That:						
The student is attentive in class	6.0*	2.3	4.9**	–0.2	4.5	3.3
The student participates in class	8.3**	3.5*	5.9***	1.4	3.9	6.9***
Teacher-Reported Level of Effort Composite (Mean)	0.2***	0.1*	0.1***	0.0	0.1*	0.1**
Teacher-Reported Disciplinary Problems Composite (Mean)	0.0	0.0	0.0	0.0	0.0	0.0
Mean Number of Days Student Was:						
Absent from class	–1.1*	–1.1***	–0.9***	–1.3**	–0.9*	–0.4
Late for class	–1.6**	–0.9	–0.9***	–1.5**	–0.3	–1.1
Percentage of Students Whose Teachers Report That They Achieve at an "Above-Average" or "Very High" Level	**6.1*** **	**–1.4**	0.7	0.8	0.9	0.37
Mean Grades						
Math	0.1	1.0**	**1.4*** **	**–0.3**	1.0	0.4
English	0.5	0.3	0.5	–0.4	**1.6** **	**–0.3**
Science	**–1.5***	**1.0**	**1.0**	**–1.1**	1.1	–0.5
Social studies	0.4	0.6	0.9*	–0.4	1.0	–0.1
Sample Size						
Student-reported outcomes	1,219	2,729	2,613	1,307	1,248	2,045
Parent-reported outcomes	1,066	2,427	2,343	1,137	1,470	2,297
Teacher-reported outcomes	943	2,154	2,090	1,010	1,044	1,664
School records outcomes (attendance)	1,184	2,641	2,508	1,301	1,266	2,040
School records outcomes (grades)	1,110	2,382	2,276	1,205	1,162	1,858

SOURCE: Parent survey; student followup survey; teacher survey; school records.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .10 level or higher.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Table III.8A

Impacts on Other Student and Parent Outcomes, by Subgroup:
Middle School Centers

Outcome	Estimated Impact						
	Grade Level		Race/Ethnicity			Gender	
	5-6	7-8	White	Black	Hispanic	Male	Female
Student-Reported Delinquent Behavior Composite (Mean)	0.04	0.04**	0.06***	0.07**	0.00	0.05**	0.03*
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:							
Attended an open house at school	7.9***	8.8***	8.2***	11.1***	7.7**	8.6***	8.6***
Attended a parent-teacher organization meeting	4.6	6.5***	4.6	3.4	9.1**	6.8***	5.5**
Attended an after-school event	8.0**	6.8***	5.9*	7.3*	6.7**	7.8***	6.4**
Volunteered to help out at school	–0.7	4.4***	3.2	2.0	3.0	2.1	4.3**
Percentage of Students Who Report the Following Happened to Them "Some" or "A Lot"							
Been offered, sold, or given an illegal drug	0.9	1.9	**1.4**	**5.8****	**–0.7**	1.4	1.9
Been "picked on" after school	–0.2	2.5	3.3	5.1	0.1	**–1.2**	**4.7****
Been threatened or hurt with a weapon	1.0	1.3	1.6	0.6	–0.2	1.1	1.4
Been threatened by a gang member	–2.3	0.7	1.9	–1.0	0.2	0.7	–0.6
Had property damaged on purpose	0.4	3.1**	3.8*	3.9	0.5	**–0.2**	**4.9*****
Student-Reported Tobacco, Alcohol, Drug Use Composite (Mean)	–0.01	0.01	**0.02**	**0.01**	**0.0**	0.00	0.01
Sample Size							
Student-reported outcomes	1,138	2,893	1,391	940	1,091	1,868	2,156
Parent-reported outcomes	987	2,588	1,304	819	937	1,657	1,916

SOURCE: Parent survey; student followup survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .10 level or higher.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Table III.8B

Impacts on Other Student and Parent Outcomes, by Subgroup:
Middle School Centers

Outcome	Estimated Impact					
	Baseline Grades		Baseline Disciplinary Problems Composite		Number of Parents in Household	
	Low	High	Low	High	Two	One
Student-Reported Delinquent Behavior Composite (Mean)	0.04	0.04**	0.03**	0.06*	0.05*	0.03
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:						
Attended an open house at school	8.9***	8.5***	8.9***	8.5***	**13.5*** **	**5.5*** **
Attended a parent-teacher organization meeting	7.6**	5.6***	**7.4*** **	**0.8**	7.5***	5.0**
Attended an after-school event	9.6***	5.1**	5.6**	9.2***	10.2***	4.5**
Volunteered to help out at school	6.7***	1.9	3.2*	3.4	1.7	4.4**
Percentage of Students Who Report the Following Happened to Them "Some" or "A Lot":						
Been offered, sold, or given an illegal drug	1.9	1.0	**–0.6**	**5.6****	1.2	0.7
Been "picked on" after school	1.7	1.7	1.7	0.8	4.0	–0.4
Been threatened or hurt with a weapon	**–3.1***	**3.2*** **	0.8	0.6	1.1	0.9
Been threatened by a gang member	–1.1	–0.1	–1.4	0.8	–1.7	–0.4
Had property damaged on purpose	1.6	2.6*	1.6	2.4	2.0	0.7
Student-Reported Tobacco, Alcohol, Drug Use Composite (Mean)	**–0.02**	**0.01**	0.00	0.01	0.00	0.00
Sample Size						
Student-reported outcomes	1,219	2,729	2,613	1,307	1,248	2,045
Parent-reported outcomes	1,066	2,427	2,343	1,137	1,470	2,297

SOURCE: Parent survey; student followup survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .10 level or higher.

*Significantly different from zero at the .10 level, two–tailed test.
**Significantly different from zero at the .05 level, two–tailed test.
***Significantly different from zero at the .01 level, two–tailed test.

D. Programs with More Academic Focus Did Not Have Larger Effects

We estimated impacts separately for grantees based on a variety of characteristics, including proxies for their academic focus, the extent to which project directors felt that getting support from school day teachers was a challenge, the extent to which project directors felt retaining staff was a challenge, the extent to which they felt that getting support from the community was a challenge, whether grantees were from the first, second, or third cohort, and whether comparison group students were drawn from the school where the centers operated or from other schools. Few patterns emerged from the analysis.

In particular, we estimated impacts separately by grantee to assess whether more academic programs had larger effects. A center's academic focus was assessed in two ways. First, program directors reported whether academic enrichment was a major objective of the program. Second, centers were considered to have an academic focus according to the extent to which they offered math classes to students.[44] These are rough proxies for academic focus, but more detailed data about program academic activities are not available. No patterns emerged between the two measures of academic emphasis and impacts (Table III.9). Only two impacts were statistically significant at the 10 percent level, and the differences indicate that English and science grades increased more when grantees did not have academic improvement as a major objective. However, that finding is not supported when grantees' academic focus is assessed by whether they provide math classes. The small number of significant differences and the

[44]Whether centers offered other types of academic classes was highly correlated with math offerings—centers that frequently offered math classes also tended to frequently offer other academic classes.

Table III.9

Impacts by Grantee Characteristics:
Middle School Grantees

Outcome	Grantee Characteristic			
	Grantee's Assessment of Whether Academic Enrichment Is Major Objective		Whether Math Classes Were Offered Frequently	
	Major Objective	Not Major Objective	Frequent Math Classes	Infrequent Math Classes
Teacher-Reported Level of Effort Composite (Mean)	0.1	0.1	0.1	0.1
Mean Class Grade				
Math	–0.2	0.9	0.9	–0.3
English	–0.4	1.1**	0.4	–0.1
Science	–0.5	1.1*	0.8	–0.5
Social studies	–0.0	1.2	1.1	–0.1
Percentage of Students Whose Parents Attended a Parent-Teacher Organization Meeting at Least Three Times Last Year	5.0	6.0	9.0	3.0
Student-Reported Negative Behavior Composite (Mean)	0.02	0.05	0.01	0.05
Number of Grantees[a]	**19**	**13**	**13**	**19**

SOURCE: Parent survey; student followup survey; teacher survey; school records; site visitor assessments.

[a]Not all outcomes were available for all grantees, so sample sizes vary.

* The difference is statistically significant at the .10 level, two-tailed test.
** The difference is statistically significant at the .05 level, two-tailed test.

inconsistency of the differences suggests that any relationship between the academic proxies and impacts is weak or nonexistent.

E. Attendance Was Not Related to Most Outcomes

Many students attended centers sporadically but some attended more regularly. If impacts of centers are related to attendance, it may be possible to use attendance data to estimate a relationship between impacts and attendance. This relationship would provide insights into the kinds of students most likely to benefit from center participation and the potential gains from efforts to improve attendance. We looked at the relationship in two ways, first at the level of the program and then at the level of the individual student.

At the program level, we divided grantees into three attendance categories to assess whether higher-attendance programs showed larger effects. The average participant attended fewer than 20 days during the school year for grantees that were considered "low-attendance," 20 to 40 days for medium-attendance grantees, and more than 40 days for high-attendance grantees. No relationship is evident between average attendance and impacts (Table III.10). The only impact that attendance appeared to influence was parental involvement. High-attendance grantees had a larger impact (14.6 percentage points) than low- and medium-attendance grantees (0 to 4 points) on attendance at parent-teacher organization meetings.

Measurement issues are more complex at the individual student level. The relationship between impacts and student attendance (the "dosage" effect) is difficult to measure correctly because students and parents can choose their dosage and the factors affecting their choice of dosage (such as whether a student likes to be at school) may also affect outcomes. An ideal scheme for measuring the "dosage" effect would be to assign participants randomly to various dosages (for example, short, medium, and long), which would ensure that the groups experiencing the different dosages were statistically equivalent at the outset. However, for

Table III.10

Impacts by Average Attendance:
Grantee-Level School

Outcome	Estimated Impact: Average Grantee Attendance: Low (Less than 20 days)	Medium (20 to 40 days)	High (More than 40 days)
Teacher–Reported Level of Effort Composite (Mean)	0.13	0.06	0.18
Mean Class Grade			
Math	–0.1	0.2	0.1
English	1.5	0.3	–0.9
Science	0.1	0.0	–0.2
Social Studies	0.8	0.3	0.1
Percentage of Students Whose Parents Attended a Parent-Teacher Organization Meeting At Least Three Times Last Year	4.3**	–0.4**	14.6**
Student-Reported Delinquent Behavior Composite (Mean)	0.01	0.05	0.03
Number of Grantees	**12**	**11**	**9**

SOURCE: Parent survey, student followup survey, teacher survey, school records, program attendance records.

**The differences are jointly statistically significant at the .05 level, two-tailed test.

practical reasons it is not possible to assign participants to center dosages. Instead, participants (implicitly) choose their own dosage, and those choosing large dosages could differ systematically from those choosing small ones. Statistical methods can adjust for observed characteristics that differ between the groups, but unobserved characteristics (such as motivation) also may affect outcomes. Outcome differences between low- and high-dose participants therefore do not have a strict causal interpretation as the effect of the dosage difference.[45]

Comparisons of the characteristics of participants who attended centers frequently (in the top third of days attended) and participants who attended them infrequently (in the bottom third of days attended) confirm that the groups differed (Table III.11).[46] Participants who attended frequently were more often black (37 percent, compared to 20 percent) and living in single-parent households (33 percent, compared to 27 percent). They also had lower average household incomes and higher rates of public assistance receipt. However, mothers of frequent participants were less likely to have dropped out of high school.

Under the assumption that students who are more motivated attend centers more often, outcome differences can provide some information about dosage effects. For example, under the assumption that those who frequently attend are more motivated, negligible outcome differences suggest that centers are having no dosage effect on outcomes (or affecting them negatively), and negative outcome differences suggest that dosage may be reducing outcomes. Positive outcome

[45]Instrumental variables techniques can be used to adjust for unobserved differences if a reasonable instrumental variable or set of such variables can be identified. In this context, an instrument is a variable that would be correlated with attendance but not with the outcomes. Several variables were tried as instruments but failed statistical tests as a result of their correlation with outcomes.

[46]To ensure that only unobserved characteristics affect outcome differences between frequent and infrequent attenders, the comparisons reported here have all been adjusted for observed characteristics using regression models. The adjustment variables are the same as those in the impact regression models used earlier in the chapter.

Table III.11

Baseline Differences between Frequent and Infrequent Participants:
Middle School Centers

Baseline Characteristic	Infrequent Participants (Bottom Third)	Frequent Participants (Top Third)	Difference[a]	p-Value[b]
Percent of Hispanic Students	30.8	24.0	6.7**	0.01
Percent of Black Students	20.0	37.3	–17.3***	0.00
Percent of Students Whose Parents Received Public Assistance	30.1	36.0	–5.9**	0.02
Percent Whose Mother Dropped Out of High School	23.2	16.3	6.9***	0.00
Household Income (in Thousands of Dollars)	35.8	33.5	2.3*	0.07
Percent of Students in a Single-Parent Household	27.1	33.4	–6.3**	0.01
Number of Absences Last Year	8.4	7.1	1.3***	0.00
Sample Size	**548**	**602**		

SOURCE: Student baseline survey; parent survey; school records; program attendance records.

[a]Due to rounding, estimated impacts shown in the table do not always equal the difference between center infrequent participants and the frequent participants.

[b]The p-value is the smallest level of significance at which the null hypothesis that the impact equals zero can be rejected. If the p-value is less than .01, an impact is significant at the 1 percent level; if the p-value is less than .05, the impact is significant at the 5 percent level, and so on.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

differences may arise because of the assumed motivation difference and are not interpreted here as evidence that more frequent attendance would improve the outcome.

Turning to the outcome differences, frequent participants spend more time at school after the regular school day ends (Table III.12).[47] They also spend less time with parents and more time in the care of other adults. There are some indications that frequent participants behave better in school than less frequent participants. Frequent participants were less likely to say that they give their teacher a "hard time" "some" or "a lot" of the time (Table III.12). Teachers also report that frequent participants were less likely to be disruptive, and frequent participants were absent 1.5 fewer days than infrequent participants. However, academic achievement of frequent participants is about the same as for less frequent participants. Grades were not statistically different, and teachers reported no differences in overall academic achievement. Frequent participants appear to be somewhat less able to interact socially—they were more likely to feel lonely, more likely to feel "picked on," and less likely to find it easy to make new friends (Table III.12). Furthermore, teachers were less likely to say that frequent participants were good at getting along with others.

The picture that emerges from this analysis suggests that frequent participants are more likely to be from disadvantaged households and to want to improve in school, as their better behavior in school and more frequent attendance indicate. However, the analysis does not suggest that higher levels of center attendance lead to improved outcomes.

[47]Appendix B provides details about the calculations of outcome differences for frequent and infrequent participants.

Table III.12

Outcome Differences by Attendance:
Middle School Centers

Outcome	Median Participation (44 Days)	Frequent Participation (104 Days)	p-Value [a]
Percentage of Students in the Following Locations after School at Least Three Days in a Typical Week:			
Own home	71.7	53.2	0.00***
Someone else's home	9.7	4.3	0.00***
School or other place for activities	27.3	48.2	0.00***
Percentage of Students in Following Types of Supervision at Least Three Days After School in a Typical Week:			
Self-care	18.3	13.2	0.01***
Parent care	55.5	43.5	0.00***
Nonparent adult care	17.2	36.3	0.00***
Percentage of Students Who Report That They Give a Teacher a "Hard Time" "Some" or "A Lot"	18.9	15.1	0.04**
Percentage of Students Whose Teachers Report That They Are "Often" Disruptive	8.4	5.3	0.05**
Teacher-Reported Level of Effort Composite	3.6	3.6	0.51
Teacher-Reported Academic Achievement Composite	3.4	3.3	0.36
Delinquent Behavior Composite	1.5	1.5	0.40
Tobacco, Alcohol, Drug Use Composite	1.1	1.1	0.64
Percentage of Students Who Report That They Do the Following "Some" or "A Lot":			
Sell illegal drugs	3.0	1.9	0.16
Get arrested or detained by police	3.6	1.4	0.02**
Percentage of Students Who Report the Following Happened to Them "Some" or "A Lot":			
Been offered, sold, or given an illegal drug	15.4	13.4	0.27
Been picked on after school	31.5	37.3	0.02**
Mean Number of Days Student Was:			
Absent from class	9.6	8.1	0.00***
Late for class	5.4	5.3	0.91
Mean Grade in English	80.7	79.5	0.05**
Percentage of Students Who Report the Following Are "Somewhat True" or "Very True":			
They are lonely	12.2	15.5	0.09*
They find it easy to make new friends	89.2	86.0	0.07*
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That They Get Along Well with Others	78.6	74.3	0.08*

SOURCE: Student survey; parent survey; school records; program attendance records.

NOTE: The percentages and mean values of outcomes have been regression-adjusted to account for baseline differences between the groups. The control variables in the regression included 26 different student and household characteristics such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores, attendance, disciplinary problems, and self-reported grades. Estimated impacts shown in the table do not always equal the difference between center participants and the comparison group due to rounding.

Table III.12 (*Continued*)

[a]The p-value is for the significance of the coefficient on attendance in the regression model. The p-value is the smallest level of significance at which the null hypothesis that the impact equals zero can be rejected. If the p-value is less than .01, an impact is significant at the 1 percent level; if the p-value is less than .05, the impact is significant at the 5 percent level, and so on.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

IV. Implementation and Impacts of Elementary School Centers

To study the implementation of elementary school centers, the evaluation used the same procedures and instruments it used to study middle school centers. It also analyzed a range of outcomes for elementary school centers, including activities after school, grades, test scores, behavior, and feelings of safety after school. Unlike the approach for middle school centers, the evaluation used experimental designs to estimate whether elementary school centers affected outcomes. Students and their parents applied to participate in the centers in the fall of the 2000-2001 school year, and applicants were randomly assigned either to a treatment group that was offered the chance to enroll in a center or to a control group that was not. (Dynarski et al. [2001] describe the operational aspects of random assignment in more detail.) In the first year, the experimental design was implemented in 18 centers operated by seven grantees. As noted in Chapter I, another set of elementary school grantees was included in the national evaluation in its second year and a future analysis will examine findings from the full set of grantees.

The first-year findings indicate that centers had improved grades in social studies and had improved some aspects of involving parents with their child's schooling, including attendance at after-school events. However, treatment group students scored about the same as control students on a standardized reading test and, according to student and teacher reports, were as likely to complete their homework. Treatment students were as confident as control students in reading and in interpersonal skills, such as working with others in a team or as a group. Treatment students also felt as safe as control students during after-school hours and were as likely to engage in (and be disciplined for) bad behavior.

A. Implementation of Elementary School Centers

The seven grantees and 18 elementary school centers in the national evaluation formed a select group, because oversubscription allowed them to support an experimental evaluation design. Because high demand for the after-school program was a major factor in their selection, considerations about appealing to students and encouraging enrollment in the after-school programs were generally secondary for program designers. Poor academic performance was the primary consideration that shaped the content and design of these elementary programs, as nearly all the schools were performing below their states' average achievement level (Table IV.1). Program designers did include enrichment and recreation activities but mainly to keep students productive and energized.

Table IV.1

Objectives of 21st-Century Elementary School Centers in the National Evaluation

	Percentage		
	Major	Minor	None
Help Children to Improve Academic Performance	83	17	0
Help Children to Develop Socially	39	33	28
Provide a Safe Environment for Children After School	44	56	0
Provide Recreational Opportunities for Children	28	56	17
Provide Cultural Opportunities for Children	17	56	28
Help Parents and/or Other Adults with Literacy or Other Skills (Such as Parenting)	28	11	61

SOURCE: Site visitor assessments based on visits to 18 centers.

Most of the elementary centers followed a similar approach to structuring after-school activities, but there were notable exceptions. Some centers, for example, focused on building skills for state assessment tests. A handful emphasized skills for parents and other adults in the community. Four centers that were part of one district's grant built their programs on a model of serving adults in the school's community. Children attended the center only when they accompanied their parent or grandparent. By using computer-oriented instruction and requiring adult participants to volunteer in the schools for a specified period, these centers expected to improve student performance and improve the lives of students out of school.

The elementary centers tended to be open most of the after-school hours (about 15 each week). A typical after-school schedule in an elementary school center included time for a snack and homework, followed by one or two sessions of academic activities or enrichment and recreation. In general, students had 45 minutes to an hour to work on homework, an hour of another academic activity, and one to two hours for other activities. Several centers gave students little opportunity to select after-school activities. At four grantees that focused on academic support, centers required students to attend both the homework period and the cognitive activity. In another two grantees, center staff members allowed students more choice.

The elementary center schedules changed little during the school year. The schedule for academic activities changed hardly at all, and the schedule for nonacademic activities often followed cycles. For example, students would rotate through 10 activities during a 10-week period and begin the rotation again at the beginning of the next 10-week cycle.

At eight centers, the purpose of the academic program was to help students complete homework; at four, it was to help them master material taught during the day; and at six, it was to help them perform better on state or district tests. Some of the academic activities used small-

group tutoring, self-paced computer-aided instruction, and purchased curricula. Students often were grouped by grade level during academic activities.

Centers had different approaches to homework sessions. Some insisted that students work on homework (or other cognitive activities if they had no homework) during the session. Others did not emphasize completing homework (for example, they combined the homework session with snack time). Consistent with these differences, the quality of homework help varied in its orderliness and supportiveness. Site visitors witnessed several chaotic homework sessions in which few students focused on their work and few staff members were engaged with students. Although some other centers maintained better control, centers did not give students much help with their homework. Homework help was particularly limited if the after-school staff members were not regular day teachers and there were no procedures for logging students' assignments or tracking their progress on them.

Additional academic content provided by centers focused on reading and writing. In some cases, the content was linked to areas that schools had identified from achievement or assessment test scores as needing improvement. Five centers used purchased curricular materials to help students with their reading and writing. Another center used hands-on activities designed to complement instruction during the regular school day.

Students could work on computers, either as an elective or as a component the center required. Four centers at one grantee required second- through fifth-graders to use a computer-based instructional program in reading and spelling. Centers at another grantee provided at least 30 minutes a week of computer instruction. In other centers, students could choose to spend time in the computer lab as part of their nonacademic activity time.

Centers at three grantees had strong links with the regular school program. The links arose from shared curricula or lessons that focused on the same standardized tests, and natural links

created when centers hired staff members from the regular school to run or supervise activities. A common element among grantees that did not have strong links was that few teachers from the regular school worked in the after-school centers.

All grantees placed some emphasis on activities beyond academic ones. Center staff viewed providing recreation activities as a major objective at 5 of the 18 elementary centers, and promoting social opportunities was a major objective for 7. The nonacademic activities were often used as incentives or rewards for the time students spent in the academic activities. While some nonacademic activities had cognitive aspects, most were recreational and designed to be fun (see box).

Illustrative Recreation and Enrichment Activities

Art
Dance
Drama
Free gym time
Karate
Leadership

Other aspects of the implementation of these selected elementary school centers were analogous to those of middle school centers. School districts administered elementary school centers, and center staff had characteristics similar to those of middle school centers. In general, they enjoyed support from their host schools but were weakly linked with community organizations and had not made much progress in planning for sustainability.

B. Student Participation Was Moderate

Analysis of data from records of elementary school centers shows that attendance was moderate during the 2000-2001 school year (Table IV.2). Students who were randomly assigned to the program and attended at least one day attended an average of 58 days during the year. While about one-quarter attended the program for fewer than 25 days, nearly 40 percent did so for more than 75 days.

Table IV.2

21st-Century Elementary School Center Attendance
(2000-2001 School Year)

Average Days Attended in School Year[a]	58.3
Number of Days Attended (Percent of Participants)	
1 to 25 Days	27.0
26 to 50 Days	19.8
51 to 75 Days	14.6
76 to 100 Days	21.7
101 to 125 Days	14.8
More than 125 Days	2.1
Attendance Rate[b] (Percent of Participants)	
10 Percent or less	23.8
11 to 25 Percent	15.7
26 to 50 Percent	22.3
51 to 70 Percent	22.3
71 to 85 Percent	14.9
86 to 100 Percent	1.0

SOURCE: Center Attendance Records. Sample size is 395 participants.

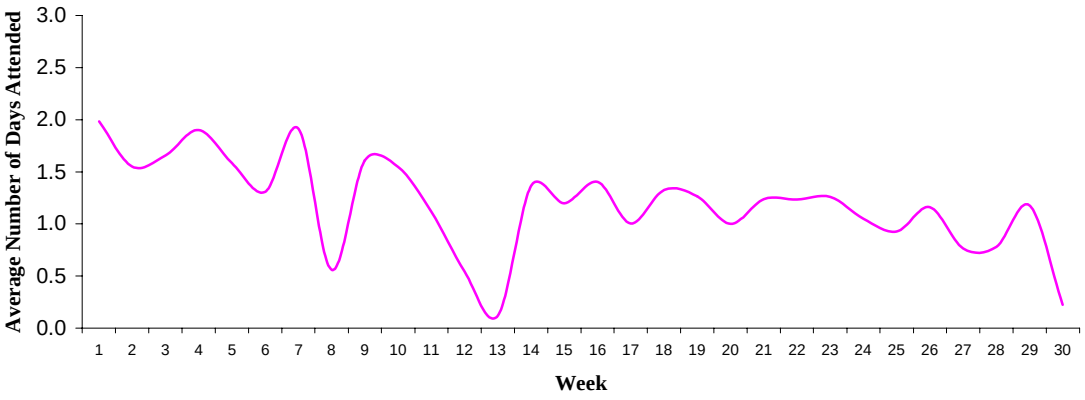
[a]Average number of days is calculated for program participants who attended the program at least one day after being randomly assigned to the program. Participation data from one grantee could not be used.

[b]The attendance rate is the number of days participants attended as a proportion of the number of days centers were open, which is obtained from grantee annual performance reports.

The pattern of attendance during the school year shows a slow but steady decline in attendance, with sharper (and temporary) declines around major holidays. Figure IV.1 plots average days attended a week for the school year for elementary school students. The downward trend of average attendance during the school year is evident, as is the sharp decline around the Thanksgiving and Christmas holidays. Most of the decline is due to students who no longer attend. Figure IV.2 plots the decline in attendance throughout the school year. A sizable group of elementary school students attended centers for long periods, although they may not have attended often within the time period. Figure IV.2 shows the distribution of calendar days between the first day and last day an elementary school participant attended a center. The relatively straight aspect of the line indicates that students were not more likely to stop attending

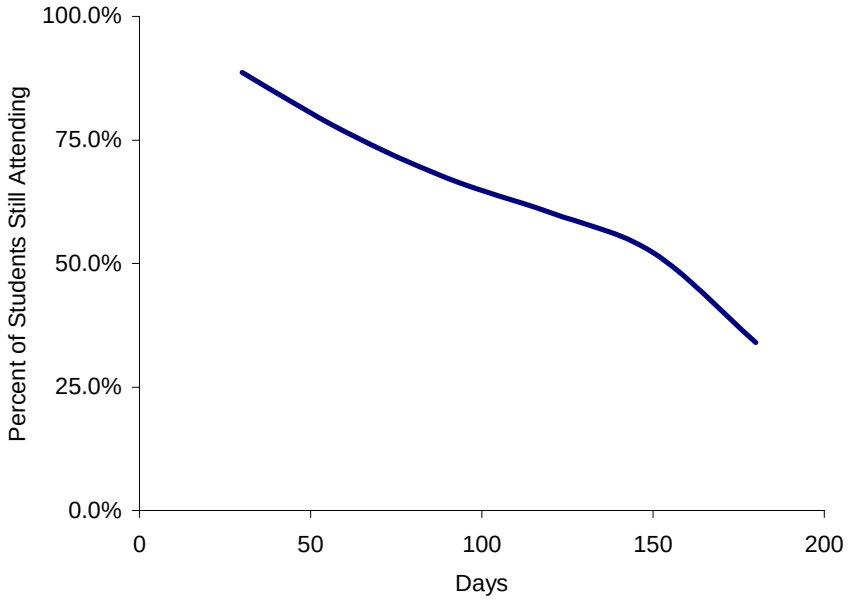
Figure IV.1

Average Days Attended Per Week
(Elementary School Centers, 2000-2001 School Year)



Source: Center attendance records.

Figure IV.2
Distribution of Length of Time Attending Centers
(Elementary School Centers, 2000-2001 School Year)



Source: Center attendance records.

during any particular times of the year. About 25 percent of participants had stopped attending after about two calendar months, which suggests that a group of students may be trying out centers before they stop attending. The median time period of attendance was about 160 days for participants, which suggests that students who continue to attend beyond the first few months were likely to attend for a much longer time. Twenty-five percent of elementary school students were still attending centers more than 190 days after the first day they attended.

C. Centers Did Not Affect Most Outcomes

Centers could affect a wide range of outcomes, as noted in the conceptual framework in Chapter I. Immediate effects might be observed in student activities, location, and supervision after school. Effects on some measures of academic performance also might be observed, especially those that are not cumulative by nature, such as completing homework and trying hard in class. Further, effects on personal development may be observed because of the nature of many of the enrichment activities the centers provided.

Because the groups were created through random assignment, the baseline characteristics of students in the groups should in principle be similar, and statistical tests indicate this generally to be the case. No statistically significant differences were found for most characteristics. When they were found, analyses revealed no pattern indicating that one group was more or less disadvantaged than the other.

Table IV.3 shows baseline characteristics of the groups. Gender, race and ethnicity, grade level (a proxy for student age), and parent age distributions are similar, with treatment group students being somewhat less likely to be white and more likely to be black. School absences, tardy arrivals, and incidences of suspension are similar, as are standardized test scores. Treatment students are somewhat less likely than control students to feel safe walking in their neighborhoods, but the views of their parents do not differ much in this area. However, although

Table IV.3

Characteristics of Treatment and Control Group Students at Baseline:
Elementary School Centers

Characteristic	Treatment Group	Control Group	p-Value[a]
Gender			
Male	46.4	50.6	0.24
Female	53.6	49.4	0.24
Race/Ethnicity			
White (non-Hispanic)	9.7	8.3	0.55
Black (non-Hispanic)	66.7	72.5	0.13
Hispanic	18.9	14.1	0.13
Other	1.6	2.3	0.57
Grade Level (percentages)			
Kindergarten	10.5	12.9	0.25
1st grade	18.1	17.1	0.68
2nd grade	17.0	18.0	0.67
3rd grade	13.4	10.5	0.17
4th grade	17.4	18.5	0.67
5th grade	19.5	16.5	0.24
6th grade	4.0	6.5	0.08*
Mother's Average Age (Years)	36.4	35.5	0.21
Father's Average Age (Years)	37.5	39.0	0.09*
Number of Tardy Arrivals During 1999-2000 School Year	3.4	3.6	0.78
Number of Absences During 1999-2000 School Year	7.2	7.5	0.62
Parent Feels It is Safe for Child to Walk in Neighborhood	59.8	54.9	0.16
Student Feels It is Safe to Walk in Neighborhood	72.3	78.5	0.14
Mean Reading Test Score (Percentile)	36.6	36.3	0.94
Sample Size[b]	**587**	**381**	

SOURCE: Student Survey, Parent Survey, School Records.

NOTE: The percentages and mean values of outcomes for treatment and control students have been regression-adjusted for baseline differences between the groups. The control variables in the regression included 14 different student and household characteristics such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores and attendance.

[a]The p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and control group members equals zero can be rejected. If the p-value is less than .01, the difference is significant at the 1 percent level. If the p-value is less than .05, the difference is significant at the 5 percent level, and so on.

[b]Sample sizes differ depending on the data source.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

random assignment appears to have done what it should in creating similar groups, we estimated impacts using regression adjustment techniques to remove remaining differences and to improve the precision of the impact estimates.

1. Centers Affected Student Location and Supervision after School

Centers increased time spent at school or in another place outside the home (Table IV.4).[48] For example, 48 percent of treatment students and 37 percent of control students spent at least three days at school or in a similar place after school (an effect size of 22 percent). Conversely, centers reduced time spent at home after school, with 65 percent of treatment students and 77 percent of control students being cared for in their own homes after school at least three days a week (an effect size of 27 percent). However, centers did not change the percentage of students in other locations, such as someone else's home or somewhere to "hang out."

Consistent with the changes in student locations, centers reduced time that students were in parent or sibling care and increased time spent in the care of non-parent adults by an offsetting amount (Table IV.4). For example, parents cared for 63 percent of treatment students and 70 percent of control students at least three days a week after school (effect size 16 percent). Centers reduced sibling care by more than half, with 2 percent of treatment and 5 percent of control students being cared for by siblings at least three days a week after school (effect size 12

[48]Two impacts are estimated for elementary school students. The first and most formal is the impact for the full treatment and control groups, known as the "intent to treat" estimate. The second, shown in the following tables in the column titled "impacts on participants," is the impact adjusting for the proportion of treatment group members that did not participate in the program ("no-shows") after being randomly assigned to the treatment group. Across grantees, the proportion of no-shows ranged from 5 percent to 30 percent of the treatment group. To estimate impacts adjusted for no-shows, the estimated impact for the full group was divided by the proportion of treatment students that participated. This adjustment assumes that no-show students receive no benefits from the program, which seems reasonable in this context. However, the text focuses on the intent-to-treat estimator, which relies on fewer assumptions.

Table IV.4

Impacts on Location, Supervision, and Activities after School, Elementary School Centers

Outcome	Treatment Group	Control Group	Estimated Impact	p-Value [a]	Estimated Impact on Participants
Percentage of Students in the Following Locations after School at Least Three Days in a Typical Week according to Parent Reports:					
Self-care	1.6	2.2	–0.6	0.62	–0.9
Parent care	62.8	70.2	–7.4*	0.09	–6.2
Non-parent adult care	31.2	20.7	10.5***	0.01	11.0*
Sibling care	2.0	4.6	–2.7*	0.09	–4.0
Mixed care (not in any one category for at least three days)	2.5	2.3	0.2	0.90	0.1
Percentage of Students in the Following Types of Supervision at Least Three Days after School in a Typical Week, according to Parent Reports:					
Own home	65.4	76.7	–11.2***	0.01	–13.4**
Someone else's home	14.9	16.2	–1.3	0.70	–0.5
School or other place for activities	47.7	37.3	10.4**	0.02	12.9***
Somewhere to "hang out"	4.2	4.2	0.1	0.96	0.7
Mixed location (not in one location for at least three days)	1.6	2.2	–0.6	0.62	–0.2
Mean Number of Days Stayed after School for Activities in Typical Week, according to Parent Reports	1.8	0.9	0.9***	0.00	1.1***
Percentage of Students Who Reported Participating in the Following Activities Yesterday after School:					
Homework	80.6	76.0	4.7	0.36	7.6
Tutoring	8.2	10.9	–2.7	0.42	–3.0
Non-homework reading, writing, or science activities	34.3	30.5	3.8	0.50	5.6
School activities (band, drama, etc.)	28.9	22.7	6.3	0.24	7.9
Lessons (music, art, dance, etc.)	30.2	21.9	8.4	0.11	11.3
Organized sports	28.7	27.3	1.4	0.80	0.9
Clubs (Boy and Girl Scouts, Boys and Girls Club, etc.)	18.2	27.8	–9.6**	0.05	–13.0*
Activities at church, temple, or mosque	21.8	22.1	–0.2	0.97	0.8
Watched TV or videos	76.5	68.5	7.9	0.13	10.7
Surfed the Internet or did other things on the computer	37.6	35.2	2.4	0.69	0.0
"Hung out" with friends	55.3	46.2	9.1	0.14	12.9
Did chores around the house	56.3	51.5	4.8	0.43	6.8
Took care of a brother or sister	44.8	28.1	16.7***	0.01	25.6***
Played	8.5	8.9	–0.5	0.90	–1.4
Mean Time Students Reported Watching Television in the Past Day	2.0	1.9	0.1	0.63	0.2
Mean Time Students Reported Reading for Fun in the Past Day	0.4	0.4	0.0	0.93	–0.1
Sample Size[b]	**403**	**226**			

SOURCE: Parent Survey, Student Survey.

NOTE: The percentages and mean values of outcomes for treatment and control students have been regression-adjusted for baseline differences between the groups. The control variables in the regression included 14 different student and household characteristics such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores and attendance.

[a]The p-value is the smallest level of significance at which the null hypothesis that the impact equals zero can be rejected. If the p-value is less than .01, an impact is significant at the 1 percent level; if the p-value is less than .05, the impact is significant at the 5 percent level, and so on.

[b]Sample sizes differ for some outcomes due to nonresponse. Sample sizes for student-reported outcomes are 285 for the treatment group and 156 for the control group. Only students in third grade and above completed a student survey.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

percent). Centers increased non-parent adult care, with 31 percent of treatment students and 21 percent of control students being cared for by non-parent adults at least three days a week after school (effect size 26 percent). Centers did not affect self-care, with 2 percent of both treatment and control students in self-care at least three days a week after school. Centers also had mostly insignificant effects on student activities after school.[49]

2. Centers Did Not Improve Most Grades and Did Not Affect Behavior in School

Centers increased grades in social studies significantly (the effect size is 30 percent), but while grades in other subjects generally appeared higher for treatment students, the differences were not significant.[50] Centers did not have significant effects on reading test scores (Table IV.5). For example, treatment students had a percentile average reading score of 34.3, and control students had an average score of 34.2.

According to teacher and student reports, treatment students were as likely as control students to complete their homework (Table IV.5). Eighty-five percent of the treatment group and 87 percent of the control group reported doing the homework their teachers assigned. Teachers reported that treatment students were as likely as control students to complete their homework often and to be prepared and ready to learn in class. According to teachers, treatment students were more likely than control students to try hard in reading. However, according to

[49]We used parent-reported data on location and care arrangements, as students in grades K-2 did not complete questionnaires. We used student-reported data on after-school activities, for students in grades 3 to 5.

[50]We estimated impacts on reading test scores using two samples: (1) the set of students who had a follow-up test score (regardless of whether they had a baseline score), and (2) the set of students who had both follow-up and baseline test scores. The results were not affected by the method used. We also collected baseline and follow-up math test scores for small number of students when the scores were available from school records but samples were too small (only 170 students) to support reliable impact estimates.

Table IV.5

Impacts on Academic and Other In-School Outcomes, Elementary School Centers

Outcome	Treatment Group	Control Group	Estimated Impact	p-Value [a]	Estimated Impact on Participants
Mean Grade:					
Math	81.0	79.6	1.3	0.23	1.7
English/language arts	82.6	81.7	0.9	0.29	1.3
Science	84.0	83.5	0.5	0.61	0.5
Social studies/history	83.0	80.3	2.7**	0.01	3.3**
Mean Reading Test Score	34.3	34.2	0.1	0.96	0.4
Percentage of Students Who Report That They "Often" or "Always" Do the Homework Teachers Assign	85.0	86.6	–1.6	0.72	–2.5
Percentage of Students Whose Teachers Report That They "Often" Complete Their Homework	55.3	60.9	–5.5	0.20	–8.1
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That They Complete Assignments to the Teacher's Satisfaction	56.1	59.3	–3.2	0.44	–5.9
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That the Student Comes to School Prepared and Ready to Learn	57.8	63.3	–5.5	0.18	–8.8
Percentage of Students Whose Teachers Report That They "Usually Try Hard" in Reading or English	57.0	49.5	7.5*	0.08	9.7
Percentage of Students Whose Parents "Agree" or "Strongly Agree" That Their Child Works Hard at School	80.7	87.4	–6.7*	0.06	–11.0**
Level of Effort Composite[b] (Mean)	3.6	3.6	0.0	0.72	0.0
Percentage of Students Whose Teachers Report Doing the Following "Two or More Times":					
Disciplining the child for misbehaving	48.0	42.7	5.3	0.23	8.2
Sending child to the office for misbehaving	12.5	10.7	1.7	0.55	2.4
Giving child detention	18.9	14.3	4.6	0.16	7.2
Calling parents about child's behavior	29.8	23.6	6.2	0.12	8.2
Percentage of Students Who Were Suspended During 2000-2001 School Year	7.1	5.2	1.9	0.41	2.3
Mean Number of Days					
Absent from class	8.2	8.2	0.1	0.91	0.0
Late for class	5.0	4.4	0.5	0.42	0.6
Reading Confidence Composite[c] (Mean)	3.1	3.0	0.1	0.24	0.1
Sample Size[d]	**403**	**226**			

SOURCE: Student Survey, Parent Survey, School Records, Teacher Survey.

Table IV.5 (C*ontinued*)

NOTE: The percentages and mean values of outcomes for treatment and control students have been regression-adjusted for baseline differences between the groups. The control variables in the regression included 14 different student and household characteristics such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores and attendance.

[a]The p-value is the smallest level of significance at which the null hypothesis that the impact equals zero can be rejected. If the p-value is less than .01, an impact is significant at the 1 percent level; if the p-value is less than .05, the impact is significant at the 5 percent level, and so on.

[b]The level of effort composite is based on five teacher-reported items regarding student (1) effort, (2) performance at ability level, (3) attentiveness, (4) participation, and (5) volunteering. Values on these items range from 1 to 5; a value of 1 on the composite indicates a low level, and a value of 5 indicates a high level.

[c]The reading confidence composite is based on students' reports on three items: (1) reading is hard to learn, (2) they are a good reader, and (3) they would read better if they had more help. Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level, and a value of 4 indicates a high level.

[d]Sample sizes differ for some outcomes. For teacher-reported outcomes, the sample sizes are 409 treatment group members and 253 control group members; for student-reported outcomes, the sample sizes are 285 treatment group members and 156 control group members; for records outcomes, the sample sizes are 537 treatment group members and 317 control group members.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

parents, treatment students were less likely than control students to work hard in school (Table IV.5). Treatment students also had the same value as control students on a composite variable combining various measures of effort such as performing at the student's ability level, attentiveness, participation, and volunteering in class (Table IV.5).

Suspensions, absences, tardy arrivals, and teacher reports of discipline problems showed no significant differences for treatment and control students (Table IV.5). For example, treatment and control students were both absent 8.2 days a year on average, and treatment and control students were late for class 5.0 and 4.4 times a year, respectively. In addition, 7 percent of treatment and 5 percent of control students were suspended during the school year, an insignificant difference.

3. Center Had Few Effects on Other Outcomes

Centers had no effects on student interpersonal skills (Table IV.6). Treatment students were as likely as control students to report getting along with others their age, as likely to report feeling left out of things, and as likely to rate themselves highly on working with others on a team, using a computer to look up information, or setting a goal and working to achieve it (Table IV.6). For example, 76 percent of both treatment and control students rated themselves as "good" or "excellent" on working with others on a team. Treatment students also were as likely as control students to rate themselves highly on sticking to what they believe in even if their friends do not agree.

Centers did not affect how safe or unsafe students felt after school. For example, 74 percent of treatment group students and 76 percent of control students reported feeling "very safe" after school (Table IV.6). Two percent of treatment students and 5 percent of control students did not feel safe at all after school, but the difference between the groups was not significant (Table IV.6).

Table IV.6

Impacts on Other Outcomes, Elementary School Centers

Outcome	Treatment Group	Control Group	Estimated Impact	p-Value [a]	Estimated Impact on Participants
Percentage of Students Who Report the Following Are "Somewhat True" or "Very True":					
They get along with others their age	81.2	84.2	–3.1	0.54	–6.4
They feel left out of things	32.4	29.2	3.2	0.58	2.8
Percentage of Students Who Do the Following "Some" or "A Lot":					
Help another student in school	78.0	79.2	–1.1	0.84	–2.8
Help another student after school	60.8	49.5	11.3*	0.07	15.0
Percentage of Students Who Rate Themselves as "Good" or "Excellent" on the Following:					
Working with others on a team or group	76.0	76.2	–0.3	0.96	–0.8
Feeling bad for other people who are having difficulties	71.4	75.4	–4.0	0.47	–7.7
Believing the best about other people	80.4	77.6	2.7	0.61	4.0
Percentage of Students Who Rate Themselves as "Excellent" on the Following:					
Using a computer to look up information	51.9	42.3	9.6	0.12	13.9
Setting a goal and working to achieve it	56.6	61.2	–4.6	0.46	–7.5
Percentage of Students Who Rate Themselves as "Excellent" on Sticking to What They Believe in, Even if Their Friends Don't Agree	60.0	52.4	7.7	0.21	10.2
Percentage of Students Who Report Feeling the Following Levels of Safety after School up until 6 p.m.:					
Very safe	74.3	75.5	–1.1	0.83	–2.0
Somewhat safe	24.0	19.3	4.6	0.37	6.9
Not at all safe	1.7	5.2	–3.5	0.14	–4.9
Negative Behavior Composite[b]	1.6	1.6	0.0	0.77	0.0
Percentage of Students Whose Parents Report Doing the Following:					
Helped their child with homework at least three times last week	68.4	58.3	10.1**	0.02	12.9**
Checked on their child's homework completion at least three times last week	91.7	91.9	–0.2	0.93	–0.4
Asked their child about things they were doing in class at least seven times last month	73.1	65.4	7.7*	0.07	10.1*
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:					
Attended an open house at the school	37.2	39.1	–1.9	0.68	–3.6
Attended parent-teacher organization meetings	53.3	50.5	2.7	0.55	3.8
Attended an after-school event	51.5	42.2	9.2**	0.04	11.5*
Volunteered to help out at school	31.5	37.2	–5.7	0.19	–8.4
Sample Size	**403**	**226**			

Table IV.6 (*Continued*)

SOURCE: Student Survey, Parent Survey.

NOTE: The percentages and mean values of outcomes for treatment and control students have been regression-adjusted for baseline differences between the groups. The control variables in the regression included 14 different student and household characteristics such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores and attendance.

[a]The p-value is the smallest level of significance at which the null hypothesis that the impact equals zero can be rejected. If the p-value is less than .01, an impact is significant at the 1 percent level, if the p-value is less than .05, the impact is significant at the 5 percent level, and so on.

[b]The negative behavior composite is based on student responses to five questions regarding how often they do the following: (1) break something on purpose, (2) punch or hit someone, (3) argue with their parents, (4) lie to their parents, and (5) give a teacher a "hard time." Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level while a value of 4 indicates a high level.

[c]Sample sizes differ for outcomes depending on the source. For teacher-reported outcomes, the sample sizes are 409 treatment group members and 253 control group members; for student-reported outcomes, the sample sizes are 285 treatment group members and 156 control group members; for records outcomes, the sample sizes are 537 treatment group members and 317 control group members.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Centers did not affect student behavior outside school. A composite variable measuring delinquent behavior—and the items upon which the composite are based—did not differ significantly (Table IV.6).

Centers increased the percentage of parents helping their child with homework at least three times in the last week, with 68 percent of parents of treatment students and 58 percent of parents of control students doing so, an effect size of 21 percent (Table IV.6). Centers also increased the percentage of parents asking their child about classwork: 73 percent of parents of treatment students and 65 percent of parents of control students asked about classwork at least seven times in the past month (effect size 16 percent). Centers also increased parent attendance at after-school events: 52 percent of parents of treatment students and 42 percent of parents of control students attended at least three after-school events in the past year (effect size 19 percent). Centers did not affect several other indicators of parent involvement, such as whether parents checked that their child had completed homework, attended school events such as open houses and parent-teacher organization meetings, or volunteered in the school.

D. No Effects Were Evident for Student Subgroups

We estimated effects for four subgroups based on characteristics that programs might use to target services: (1) gender, (2) grade level, (3) high versus low reading test scores, and (4) a high versus a low number of discipline problems. For example, centers could focus on serving particular grade levels, students with low test scores, students with high discipline problems, or a combination of these characteristics. The findings indicate few subgroup impacts (Table IV.7 and IV.8). Some differences appear large but statistical tests of the differences are not significant. Small sample sizes may be a factor in the lack of significance.

Table IV.7

Impacts on Academic Outcomes by Subgroup,
Elementary School Centers

	Estimated Impact								
	Gender		Grade Level			Baseline Reading Test Scores [a]		Baseline Disciplinary Problems Composite [b]	
Outcome	Male	Female	K-2	3-4	5-6	Low	High	Low	High
Mean Class Grade									
Math	0.3	0.7	–0.5	3.1	0.2	1.4	–0.4	0.1	2.5
English	0.5	–0.4	1.0	1.3	0.3	2.5	–0.1	–1.3	1.0
Science	0.5	–0.4	0.6	0.1	–0.2	0.9	–0.3	–0.1	–0.3
Social studies	2.4	2.0	2.7**	9.6***	0.7	3.0	0.3	5.0*	2.0
Reading Test Score	0.2	0.2	–1.7	–0.3	2.3	–1.3	–0.4	–1.9	–2.5
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That:									
Student completes assignments to my satisfaction	–4.2	–0.3	1.1	–2.7	–18.5	14.6	–10.1	–21.9***	15.6
Student comes prepared and ready to learn	0.8	–8.4	–3.4	–14.9*	–3.9	9.0	–12.8*	–7.9	0.8
Percentage of Students Whose Teachers Report That They "Usually Try Hard" in Reading or English	13.1**	–0.9	10.1	–0.5	–2.4	24.3**	–1.1	7.0	–0.4
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That:									
The student is attentive in class	0.6	–1.1	2.2	–7.1	–7.8	19.5**	–13.8**	–11.8	4.0
The student participates in class	4.1	–3.0	2.7	–5.8	4.7	5.7	1.9	–5.4	15.3
Percentage of Students Whose Parents "Agree" or "Strongly Agree" That Child Works Hard at School	–8.0	–2.0	–0.8	–12.5*	–7.7	4.7	–8.7	–11.0	–31.6**
Percentage of Students Whose Teachers Report Disciplining for Misbehaving "Two or More Times"	0.7	6.2	1.8	5.3	13.1	3.5	6.4	13.2	2.3
Teacher-Reported Disciplinary Problems Composite (Mean)	.02	.07	–.01	.15	.28	–.01	.13	.17	.16
Percentage of Students Whose Teachers Report That They Are "Often" Disruptive	0.0	0.2	1.2	1.6	4.9	–4.7	–1.3	8.2	–9.8
Percentage of Students Who Were Suspended During the 2000-2001 School Year (School Records)	3.7	3.3	–0.9	5.8	1.5	5.5	1.7	–0.2	8.3

Table IV.7 (*Continued*)

	Estimated Impact								
	Gender		Grade Level			Baseline Reading Test Scores [a]		Baseline Disciplinary Problems Composite [b]	
Outcome	Male	Female	K-2	3-4	5-6	Low	High	Low	High
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That Student Is a Proficient Reader	–6.5	–1.9	0.1	–13.5*	8.3	–1.2	–4.1	–1.7	–1.1
Sample Size									
Parent-reported outcomes	191	228	290	187	150	219	194	203	98
Teacher-reported outcomes	304	337	311	201	139	211	194	208	104
School records outcomes (suspensions)	384	397	330	229	155	266	227	223	135
School records outcomes (grades)	234	261	203	154	132	221	204	195	96
School records outcomes (reading scores)	296	302	300	156	136	236	220	178	82

SOURCE: Parent Survey, Teacher Survey, School Records.

NOTE: The percentages and mean values of outcomes for treatment and control students have been regression-adjusted for baseline differences between the groups. The control variables in the regression included 14 different student and household characteristics such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores and attendance.

[a]Students were classified into the "low baseline test scores" subgroup if their baseline test scores were less than the median baseline test score among all sample members with valid test scores. Those whose baseline test scores were above the median were classified into the "high baseline test scores" subgroup. Students with missing baseline test scores were not included in either of these subgroups.

[b]Students were classified into the "low number of disciplinary problems" subgroup if their baseline disciplinary problems composite score was less than the median composite score among all sample members. Those whose baseline disciplinary problems composite scores was above the median were classified into the "high number of disciplinary problems" subgroup. Students with missing baseline disciplinary problems composite scores (including all kindergarten through second grade students) were not included in either of these subgroups.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Table IV.8

Impacts on Other Outcomes by Subgroup,
Elementary School Centers

	Estimated Impact								
	Gender		Grade Level			Baseline Test Scores [a]		Baseline Disciplinary Problems Composite [b]	
Outcome	Male	Female	K-2	3-4	5-6	Low	High	Low	High
Student-Reported Social Engagement Composite (Mean)	.05	.07	n.a.	-.04	-.00	.17	-.23	-.01	.04
Student-Reported Negative Behavior Composite (Mean)	.13	–.14	n.a.	.01	–.01	.05	.08	–.04	.03
Percentage of Students Whose Parents Report That They Break Something on Purpose "Some" or "A Lot"	–3.9	–0.2	–5.1*	–0.6	5.7*	–1.2	2.3	2.7	2.9
Percentage of Students Whose Parents Report Helping Them with Homework at Least Three Times Last Week	11.2*	12.0**	12.1**	6.8	4.3	–9.3	1.4	6.2	3.0
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:									
Attended an open house at school	10.9*	–7.7	–4.3	–9.5	20.4*	12.4	13.4	2.1	9.8
Attended a PTO meeting	1.2	6.2	–0.5	3.0	10.2	21.6	1.9	–3.7	12.3*
Attended an after–school event	15.7**	0.5	6.1	–0.6	10.0	13.6	19.1**	4.9	16.5
Volunteered to help out at school	0.2	–5.5	–5.3	–8.3	–3.4	–0.0	4.0	–10.4	12.7
Sample Size									
Student–reported outcomes	195	229	0	241	199	189	135	264	146
Parent–reported outcomes	285	322	290	187	150	216	194	203	98

SOURCE: Parent Survey, Student Followup Survey.

NOTE: The percentages and mean values of outcomes for treatment and control students have been regression-adjusted for baseline differences between the groups. The control variables in the regression included 14 different student and household characteristics such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores and attendance.

[a]Students were classified into the "low number of baseline test scores" subgroup if their baseline test scores were less than the median baseline test score among all sample members with valid test scores. Those whose baseline test scores were above the median were classified into the "high number of baseline test scores" subgroup. Students with missing baseline test scores were not included in either of these subgroups.

[b]Students were classified into the "low number of disciplinary problems" subgroup if their baseline disciplinary problems composite score was less than the median composite score among all sample members. Those whose baseline disciplinary problems composite scores was above the median were classified into the "high number of disciplinary problems" subgroup. Students with missing baseline disciplinary problems composite scores (including all kindergarten through second grade students) were not included in either of these subgroups.

*Significantly different from zero at the .10 level, two–tailed test.
**Significantly different from zero at the .05 level, two–tailed test.
***Significantly different from zero at the .01 level, two–tailed test.

E. Greater Attendance Was Not Related to Higher Outcomes

Treatment students may have experienced larger impacts if they attended centers more frequently. As noted in Chapter III, frequent attenders may differ from infrequent attenders in ways that affect outcomes but are not controlled for by the experimental design. For example, if more-motivated students attend centers more often and also have higher outcomes because of their motivation, comparing their outcomes with those of less-motivated students may create an incorrect perception of the impacts of greater attendance. Under the assumption that more-motivated students attend centers more often, positive outcome differences may arise because of the assumed motivation difference and are not interpreted here as causal evidence that frequent attendance improves outcomes.[51]

Frequent participants in after-school programs spent more time in a variety of activities at school (Table IV.9). Parents of frequent participants reported that their children spent less time with them and more time in the care of other adults: 66 percent of children who attended moderately were in parent care for at least three days a week, versus 57 percent of frequent participants, and non-parent care was higher by about 12 percentage points for frequent participants. Parents of frequent attenders also reported that their children spent more time in organized activities such as tutoring, band, music lessons, and clubs. The additional time spent in these activities came at the expense of activities that participants could have done while at home, such as household chores, volunteer work, and taking care of siblings.

[51]To ensure that only unobserved characteristics affect outcome differences between frequent and infrequent attenders, we used regression models to adjust for observed characteristics. The adjustment variables are the same as those in the regression models used elsewhere in the chapter.

Table IV.9

Relationship between Attendance and Outcomes
Elementary School Centers

Outcome	Moderate Participation (49 Days)	Frequent Participation (105 Days)	p-Value [a]
Mean Grade:			
Math	81.4	80.8	0.50
English	82.3	82.0	0.65
Science	84.3	83.8	0.50
Social studies/history	83.6	82.8	0.31
Reading Test Score (Percentile)	31.4	30.9	0.80
Percentage of Students in the Following Locations after School at Least Three Days in a Typical Week:			
Own home	66.7	52.2	0.00***
Someone else's home	14.0	4.0	0.00***
School or other place for activities	44.5	59.5	0.00***
Somewhere to "hang out"	3.0	0.0	0.01***
Mixed location (not in one location for at least three days)	1.5	0.9	0.54
Percentage of Students in The Following Types of Supervision at Least Three Days after School in a Typical Week (According to Parents):			
Self-care	1.8	1.4	0.76
Parent care	66.1	57.4	0.01**
Non-parent adult care	26.8	38.6	0.00***
Mean Number of Days Stayed after School for Activities in Typical Week	1.7	2.8	0.00***
Percentage of Students Participating in Each Activity for at Least One Day After School (According to Parents):			
Homework	87.6	86.8	0.75
Tutoring	24.1	31.8	0.02**
School activities (band, drama, etc.)	26.9	34.1	0.03**
Lessons (music, art, dance, etc.)	25.8	35.2	0.00***
Volunteered or did community service	6.4	3.3	0.11
Did chores around the house	74.2	68.6	0.10*
Took care of a brother or sister	19.3	11.6	0.01**
Percentage of Students Who Report That They "Often" or "Always":			
Do the homework teachers assign	85.8	86.1	0.91
Do homework in same place each day	54.7	54.3	0.92
Do homework at same time each day	38.3	34.5	0.41
Write down homework assignments	65.0	67.1	0.64
Percentage of Students Who Receive Homework Help from:			
An adult besides their mother or father	31.6	39.6	0.05*
Their mother or father	61.2	58.2	0.50
Another child	15.2	9.0	0.08*
Percentage of Students Who Report the Following Are "Somewhat True" or "Very True":			
They have friends to play with	91.2	94.6	0.18
They are lonely	21.6	14.7	0.07*
They get along with others their age	82.9	87.9	0.16
Percentage of Students Who Do the Following "Some" or "A Lot":			
Help another student in school	77.2	74.3	0.46
Help another student after school	55.5	65.3	0.03**

Table IV.9 (C*ontinued*)

Outcome	Moderate Participation (49 Days)	Frequent Participation (105 Days)	p-Value [a]
Percentage of Students Who Rate Themselves as Good or Excellent at Working with Others on a Team or in a Group	73.7	84.4	0.01***
Percentage of Students Who Rate Themselves as "Excellent" on the Following:			
Using a computer to look up information	48.7	52.2	0.44
Setting a goal and working to achieve it	50.4	55.5	0.26
Planning for things in the future	51.3	55.8	0.33
Planning/Problem Solving Composite (Mean)	3.2	3.3	0.02**
Percentage of Students Whose Parents Report They Do the following "Some" or "A Lot:"			
Break something on purpose	3.5	0.8	0.08*
Punch or hit someone	14.2	11.1	0.23
Percentage of Students Whose Parents Report Doing the Following:			
Helped their child with homework at least three times last week	70.2	72.8	0.48
Checked on their child's homework completion at least three times last week	92.9	96.3	0.10*
Asked their child about things they were doing in class at least seven times last month	70.4	72.6	0.54
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:			
Attended an open house at the school	37.3	37.9	0.87
Attended parent-teacher organization meetings	52.8	57.8	0.19
Attended an after-school event	46.5	53.2	0.08*
Volunteered to help out at school	28.4	29.7	0.73
Percentage of Students Whose Parents "Agree" or "Strongly Agree" with the Following:			
My child's school is academically challenging	73.7	72.3	0.68
I am satisfied with class sizes	77.4	74.9	0.41
I am satisfied with teacher quality	75.9	74.1	0.57
I feel welcome at my child's school	87.6	89.2	0.54

NOTE: The percentages and mean values of outcomes for treatment and control students have been regression-adjusted for baseline differences between the groups. The control variables in the regression included 14 different student and household characteristics such as indicators of students' demographic characteristics, household socioeconomic status, and students' baseline test scores and attendance.

[a]The p-value is for the significance of the attendance coefficient in the regression model. The p-value is the smallest level of significance at which the null hypothesis that the impact equals zero can be rejected. If the p-value is less than .01, an impact is significant at the 1 percent level; if the p-value is less than .05, the impact is significant at the 5 percent level, and so on.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Frequent participants were not significantly different from infrequent participants on most academic outcomes. Homework habits, grades, test scores, and regular school attendance were similar for frequent and less frequent attenders. The one difference between frequent and infrequent attenders was that frequent attenders received more help on their homework from a non-parent adult and less help from another child, which possibly is related to homework assistance the centers provided.

Frequent attenders reported feeling better about themselves and interacting better with others (Table IV.9). They were less likely than infrequent participants to report feeling lonely and more likely to believe they got along well with others their age. Frequent participants also were more likely to feel that they worked well in a group, more likely to help other students after school, and more likely to say they had good problem-solving skills.

Appendix A

Response Rates and Data Quality

The results presented in this report were based on two waves of data collected from a variety of respondents and from school and program records. In the fall of 1999-2000, we conducted baseline surveys with middle school students, elementary school students, and elementary school parents, and administered standardized reading tests to elementary school students. In the spring of 1999-2000, we surveyed students, parents, teachers, school principals, and center directors, coordinators, and staff. We also collected students' school records and program attendance, and we again administered standardized reading tests to elementary school students.

We collected data on schools and centers (from principals, after-school program directors, center coordinators, and staff) in 41 sites (34 middle and 7 elementary school sites). At two middle school sites, the baseline administration of student surveys was delayed because of the time needed for obtaining parental consent. When reporting data collected on individual students (from students, parents, teachers, and school records), we excluded those two sites. As part of the enhanced study supported by the grant from the C. S. Mott Foundation, another group of students not participating in the 21st-Century program at six middle school sites completed a questionnaire on their after-school activities.

A. Data Collection Procedures for Middle School Sites

1. Baseline

We surveyed 21,156 students in 32 middle school sites at the baseline. Questionnaires were generally self-administered during the school day. The questionnaire covered family background, after-school activities, school experiences, in-school and out-of-school behavior, and experiences in and knowledge of after-school programs. In most schools, students completed questionnaires two to six weeks after the 21st-Century program began operating for the year. A few weeks before administering the questionnaires at a site, we either mailed (or

gave to students to take home, depending on the preference of the school) letters explaining the study and requesting parental consent. Parents who declined to participate returned the consent form in the postage-paid, addressed envelope provided. Three sites required parents to return the forms to approve their child's participation. We also asked middle school students to assent to participate in the study. The questionnaire cover served as a detachable assent form and explained the study components, its voluntary nature, and the confidentiality of data. Only students who signed the cover completed the questionnaire.

Using after-school program attendance records collected for a four-week period at the start of each program, we classified surveyed students as participants (those who attended the 21st-Century center three or more times), under-attenders (students that had attended one to two times), or potential comparisons (all students at comparison schools, or students at the host schools that had not attended centers). We excluded under-attenders from the rest of the study. We then used propensity score methods to match participants with students in the potential comparison group (see Appendix B, Technical Methods, for a description of the matching process). Table A-1 shows the number and types of students that were surveyed at baseline.

Table A.1

Classification of Students Completing the Baseline Survey:
Middle School Sites

Student Status	Number	Percentage
Participant Group	2,472	11.7
Potential Comparison Group	17,596	83.1
Matched Comparisons	3,921	18.5
Nonmatched Comparisons	13,675	64.6
Under-attenders	1,088	5.1
Total Surveyed	**21,156**	**100.0**

NOTE: Because of rounding, percentages do not sum to 100.0.

After matching was completed, we requested parental consent for participants and matched comparison students to participate in the study. Two-thirds of the 6,393 parents who were asked to participate (67 percent) gave their consent (Table A-2), with the proportion ranging by site from 45 to 100 percent (Table A-3).

Table A.2

Percentage of Parents Consenting to Participate in the Study: Middle School Sites

	Sample Size					
	Total		Treatment		Comparison	
Parental Consent	N	%	N	%	N	%
Asked to Consent	6,393		2,472		3,921	
Consented	4,264	67	1,782	72	2,482	63

Table A.3

Consent Rates by Site: Middle School Sites

Percentage of Parents Consenting	Number of Sites
90 to 100	2
80 to 89	2
70 to 79	11
60 to 69	11
50 to 59	5
40 to 49	1
Total	**32**

2. Follow-Up

Approximately six weeks before the end of each school's 2000-2001 academic year, field staff returned to middle school sites to administer the follow-up questionnaire. These were nearly identical to the baseline questionnaires except that items on demographics and after-

school program participation in the previous (1999-2000) school year were dropped. Ninety-five percent of the 4,264 students in the study completed the questionnaire (Table A-4), and response rates were more than 90 percent at all but two sites (Table A-5). Nearly all students completed the survey in school (84 percent). The others (16 percent), who were primarily transfer students, completed the questionnaire with computer-assisted telephone interviewers.

B. Data Collection Procedures for Elementary School Sites

1. Baseline

We surveyed 90 percent of the 522 third- to sixth-grade elementary school students at baseline (Table A.4). Response rates ranged from 84 to 96 percent across the six sites (all students at the seventh elementary school site were in kindergarten through second grade and were not surveyed) (Table A.5). Questionnaires were generally self-administered during the school day (in a few instances teachers read the questions aloud to their class).

Like middle school students, elementary students were asked to assent to participate in the study by signing the cover of the questionnaire, and only students who gave their assent completed the questionnaire. MPR interviewers conducted telephone questionnaires with a small number of students who were not surveyed at school (primarily transfer students).

MPR field staff also administered the reading component of the Stanford Achievement Test 9 (SAT-9) in school to 70 percent of students in kindergarten through sixth grade who had not completed a district-administered version of the SAT-9 that fall or the previous spring (Table A.4). Response rates across sites ranged from 44 to 93 percent (Table A.5), excluding one site that provided SAT-9 test scores for students in grades 2 to 5 but did not allow kindergarteners and first-graders to be tested. MPR field staff administered tests in student homes to a small number of students who were not tested in school.

Table A.4

Sample Sizes and Response Rates for Student Data

	Sample Size					Response Rate					
	Total	Treatment		Comparison		Total		Treatment		Comparison	
Instrument	N	N	%	N	%	N	%	N	%	N	%
Middle School Follow-Up											
Student Survey	4,264	1,782	42	2,482	58	4,059	95	1,700	95	2,359	95
Elementary School Baseline											
Student Survey[a]	522	333	64	189	36	467	90	304	91	163	86
Student Test[b]	798	497	62	301	38	561	70	358	72	203	67
Parent Survey	973	589	61	384	39	861	88	528	90	333	87
Elementary School Follow-Up											
Student Survey[a]	522	333	64	189	36	441	85	285	86	156	83
Student Test[b]	621	394	63	227	37	522	85	342	87	180	79
Combined Elementary and Middle School Follow-Up											
Parent Survey	5,237	2,371	45	2,866	55	4,224	81	1,898	80	2,326	81
Middle school	4,264	1,782	42	2,482	58	3,595	84	1,495	84	2,100	85
Elementary school	973	589	61	384	39	629	65	403	68	226	59
Teacher Survey[c]	5,237	2,371	45	2,866	55	3,969	76	1,834	77	2,135	74
Middle school	4,264	1,782	42	2,482	58	3,307	78	1,425	80	1,882	76
Elementary school	973	589	61	384	39	662	68	409	69	253	66
School Records	5,237	2,371	45	2,866	55	4,923	94	2,253	95	2,670	93
Middle school	4,264	1,782	42	2,482	58	4,069	95	1,716	96	2,353	95
Elementary school	973	589	61	384	39	854	88	537	91	317	83

[a]Sample includes only grades 3 to 6.

[b]SAT-9 tests were administered only to students for whom districts did not have test scores.

[c]Sample size and response rates are based on number of students, not teachers; 82.5 percent of the 939 teachers in the sample completed surveys.

Table A.5

Response Rates by Site for Student Data

Instrument	Number of Sites: Total	Percentage: 90 to 100	80 to 89	70 to 79	60 to 69	50 to 59	Less than 50
Middle School Follow-Up							
Student Survey	32	30	2	0	0	0	0
Elementary School Baseline							
Student Survey[a]	6	2	4	0	0	0	0
Student Test[b]	7	1	1	2	0	1	2
Parent Survey	7	3	4	0	0	0	0
Elementary School Follow-Up							
Student Survey[a]	6	1	2	3	0	0	0
Student Test[b]	7	2	3	1	0	1	0
Combined Elementary and Middle School Follow-Up							
Parent Survey	39	8	18	6	6	1	0
Middle school	32	8	18	5	1	0	0
Elementary school	7	0	0	1	5	1	0
Teacher Survey[c]	39	11	9	8	4	2	5
Middle school	32	11	9	5	2	1	4
Elementary school	7	0	0	3	2	1	1
School Records	39	31	5	1	1	1	0
Middle school	32	27	4	0	0	1	0
Elementary school	7	4	1	1	1	0	0

[a]Surveys were administered only to third- to sixth-grade students; one elementary school site had no sample in those grades.

[b]SAT-9 tests were administered only to students for whom districts did not have test scores.

[c]Sample size and response rates are based on number of students, not teachers; 82.5 percent of the 939 teachers in the sample completed surveys.

We also asked elementary school parents to complete a baseline questionnaire about their academic expectations for their child, safety concerns, interactions with their child, and their child's social and behavioral outcomes, school experiences, and after-school activities the previous spring. Almost 9 of 10 parents (88 percent) completed a questionnaire (Table A.4).

Response rates across sites ranged from 82 to 95 percent (Table A.5). About one-fourth (26 percent) returned questionnaires by mail, and three-fourths (74 percent) completed them by telephone.

2. Follow-Up

Approximately six weeks before the end of each school's 2000-2001 academic year, field staff returned to elementary school sites to administer follow-up questionnaire, which were identical to the baseline questionnaires except for one item on language that was not asked again. Eighty-five percent of students in grades 3 to 6 completed the follow-up questionnaire (Table A.4), and response rates across sites ranged from 72 percent to 92 percent (Table A.5). Nearly all the students who completed the questionnaire did so in school. MPR interviewers administered telephone questionnaires to the rest (primarily transfer students).

We again administered the reading component of the SAT-9 in school to students in kindergarten through sixth grade who would not be given a district-administered version of the SAT-9 that spring. Eighty-five percent of students completed the test (Table A.4), with response rates ranging by site from 57 to 95 percent (Table A.5). MPR field staff administered make-up tests at students' homes to a small number of children who were not tested in school.

C. Elementary and Middle School Sites Combined: Follow-Up

Beginning in the late spring of 2000-2001, we collected data on individual students from parents, teachers, and school records. We also collected data on schools and centers from principals and program staff.

1. Data Collected on Individual Students from Parents, Teachers and Records

The parent follow-up questionnaire included many items from the baseline questionnaire administered to elementary school parents, as well as items on family and child characteristics, academic expectations for their child, safety concerns, interactions with their child, and their child's social and behavioral outcomes, after-school activities, school experiences, and after-school program experiences. Eighty-one percent of parents completed the follow-up questionnaire—84 percent of middle school and 65 percent of elementary school parents (Table A.4). Slightly more than half (54 percent) responded to a mail survey; we interviewed the rest by telephone (46 percent). Response rates ranged by site from 51 to 96 percent (Table A.5).

We asked the English teacher of students at middle school schools and the homeroom teacher of students at elementary schools to complete a questionnaire on the student's classroom behavior and academic performance, teacher views of the after-school program and the school environment, and teacher demographics. About 83 percent of the teachers completed questionnaires, which provided data on 76 percent of the students—78 percent of middle school students and 71 percent of elementary school students (Table A.4). Most teachers responded by mail (70 percent) or telephone (28 percent), though a few completed the survey via the Web (2 percent). Response rates across sites ranged from 0 to 100 percent (Table A.5).

At the end of the 2000-2001 school year, we collected student records, which contained information on the students' demographics, attendance, suspensions, retention, academic services received, disabilities, standardized test scores, and grades. We obtained school records for 94 percent of students—95 percent of middle school students and 88 percent of elementary school students (Table A.4). We collected more than 80 percent of records at all but three sites, with response rates ranging from 52 percent to 100 percent (Table A.5). Generally, students for whom we were unable to collect school records had transferred to other schools.

Table A.6 summarizes the data collected on individual students in the follow-up, showing the percentage of students for whom data were obtained from one instrument (school records), two instruments (school records and student survey), three instruments (school records, student survey, and parent survey), and four instruments (school records, student survey, parent survey, and teacher survey). At the middle school sites, for example, we collected data for 95 percent of students for one instrument, 91 percent for two instruments, 78 percent for three, and 62 percent for four. Response rates for students at elementary school sites are divided between those to whom student surveys were and were not administered (grades 3 to 6 and kindergarten to grade 2, respectively).

Table A.6

Follow-Up Response Rates for Individual Student Data

Students	Total	School Records		School Records and Student Survey		School Records and Student and Parent Surveys		School Records and Student, Parent, and Teacher Surveys	
		N	%	N	%	N	%	N	%
Middle School	4,264	4,069	95	3,895	91	3,339	78	2,653	62
Elementary School Grades 3-6	522	459	88	403	77	279	53	205	39
Elementary School Grades K-2	451	395	88	(a)	(a)	258	57	203	45

[a]A survey was not administered to students in kindergarten through second grade.

2. Data Collected from Center and School Staff

Principals completed questionnaires on the relationship between the school and the 21st-Century program and their views of the program's objectives, facilities and resources, sustainability, benefits, and challenges. Ninety-five percent of principals completed a questionnaire (Table A.7)—79 percent by telephone and 21 percent by mail.

Table A.7

Sample Sizes and Response Rates: Data Collected from School and Center Staff

Instrument	Sample Size	Response Rate	
		N	%
Principal Survey[a]	80	76	95
Project Director Survey[a]	41	39	95
Center Coordinator Survey[b]	89	77	87
Staff Survey[a]	894	609	68
Program Attendance Records	75	69	92

[a]Includes 41 sites

[b]Nine after-school programs had two center coordinators; both coordinators returned surveys at five after-school programs.

We asked all 21st-Century program staff to complete a questionnaire that included items on staff roles and responsibilities, program objectives, experiences, interactions with school-day teachers and administrators, interactions with parents, professional development, professional background, and demographics. Center coordinators responded to those questions, as well as to another module that asked about interactions with parents, size of program, staff recruitment and retention, program challenges, facilities and resources, sustainability, and additional items on their role and responsibilities in the program. Like the principal questionnaire, the project director questionnaire covered program objectives, sustainability, benefits, and challenges. Project directors also answered questions on their role and responsibilities in the program and on their experience.

Questionnaires were mailed to project directors to distribute to center coordinators and to (paid) staff that were age 19 and older. We conducted follow-up telephone interviews with nonrespondents. Ninety-five percent of project directors, 87 percent of center coordinators, and

68 percent of staff completed a questionnaire (Table A.7). Most responded by mail (70 percent of center coordinators and 65 percent of staff).

We collected program attendance records from 92 percent of 21st-Century program centers (Table A.7). The centers provided copies of their records in whatever form they typically maintained attendance, such as by day or by activities offered each day. In a few cases, centers provided the total number of days students attended, rather than the daily attendance records.

Although the elementary school study design precluded attendance by students in the control group, records showed that 8 percent attended the 21st-Century program at least once. There were a variety of reasons for controls being able to attend the program. For example, because of changes in program staff, some staff were not aware of the students who should have been excluded from the program. Of those controls that attended the program, about three-fourths (76 percent) attended from 1 to 25 days, and the average attendance was 17 days.

The middle school study design did not bar any students from attending the 21st-Century program. About 14 percent of students in the comparison group attended the program at least once. Most (89 percent) attended from 1 to 25 days, and the average attendance was 10 days.

3. Data Collected from Nonparticipants

As part of the enhanced study, we surveyed students not participating in the 21st-Century program in six sites. In these sites, we drew a random sample of nonparticipating students in the schools that had centers. Comparison students participating in the larger study had lower probabilities of selection than did other nonparticipating students. The nonparticipant questionnaire asked about students' after-school activities, self-concepts, homework, and demographics. We included a module of questions on awareness of and familiarity with the after-school program, reasons for not attending, and ways nonparticipants would be encouraged

to attend. Eighty-two percent of sampled students completed the survey (868 of 1,062 students). We surveyed most students by telephone, and obtained parental permission before beginning the interview. A small number of students completed the survey by mail.

D. Tests for Response Bias

Not all consenting middle school students completed the follow-up questionnaire, which introduces the possibility of response bias. Table A.8 shows means for a range of characteristics for the sample of students that consented to be in the study and for the sample of students that completed a follow-up questionnaire.

Comparing characteristics that differ significantly for the comparison group and for participants indicates that participants generally were at higher risk of academic difficulty. For example, participants had lower average grades and test scores, more disciplinary incidents, less parental education, and less parental income. Parental characteristics and test scores were not part of the matching process, and the differences evident in the table indicate that matching did not yield groups that were equivalent on these characteristics. However, essentially the same differences are evident in the sample for which follow-up questionnaires were obtained, which indicates that the process of responding to the questionnaire did not introduce further differences. This is an expected result considering the high follow-up response rate of 95 percent (see Table A.4).

In addition, not all consenting elementary school students completed the follow-up questionnaire. Table A.9 shows baseline characteristics of the treatment and control group students both at baseline and at follow-up. The few differences in baseline characteristics compared to the middle school sample is attributable to the random assignment design. As in middle school sites, the follow-up process evidently did not introduce bias, as four characteristics

Table A.8

Participant and Comparison Group Characteristics,
Middle School Centers

Variable	Baseline Characterics of Students Consenting to Be in the Study			Baseline Characteristics of Students Responding to Follow-Up Questionnaire		
	Participants	Comparison Group	p-value[a]	Participants	Comparison Group	p-value [a]
Race						
Black	27.6	24.0	0.01***	27.1	24.4	0.05*
White	32.7	35.9	0.03**	32.9	36.1	0.03**
Hispanic	27.9	26.2	0.22	28.1	25.9	0.12
Other	7.4	8.3	0.29	7.6	8.2	0.51
Grade						
6	17.6	17.8	0.85	17.7	18.2	0.69
7	41.5	42.0	0.74	41.8	42.0	0.94
8	34.9	35.6	0.64	34.6	35.3	0.65
Other	6.0	4.6	0.04**	5.9	4.6	0.07*
Average Grades	83.0	84.1	0***	83.19	84.14	0.01***
Homework						
The student does the homework teachers assign	3.48	3.54	0.07*	3.48	3.53	0.15
Mother or father helps student with homework	63.5	63.2	0.86	63.3	63	0.86
Mean of homework habits index (Low=Does Not Do Homework)	2.83	2.88	0.03**	2.83	2.88	0.02**
Number of Hours Read for Fun Yesterday	0.30	0.32	0.05*	0.30	0.32	0.06*
Number of Hours Watched TV Yesterday	2.14	2.01	0.01***	2.16	1.99	0***
Mean of Index of Confidence in Reading Skills (Low=Little Confidence)	3.12	3.21	0***	3.13	3.22	0***
Student Expects to Drop out of High School	2.5	1.9	0.14	2.5	1.6	0.04**
At Least One Parent Is a College Graduate	29.0	32.5	0.02**	29.5	33.0	0.02**
At Least One Parent Is a High School Dropout	17.3	15.5	0.13	17.0	15.6	0.25

Table A.8 (c*ontinued*)

Variable	Baseline Characterics of Students Consenting to Be in the Study			Baseline Characteristics of Students Responding to Follow-Up Questionnaire		
	Participants	Comparison Group	p-value[a]	Participants	Comparison Group	p-value [a]
Mean of Index of Discipline Problems (Low=Few Problems)	1.40	1.33	0***	1.39	1.32	0***
Mean of Index of Bad Behavior (Low=Never)	1.54	1.51	0.04**	1.54	1.51	0.03**
Mean of Index of Good Behavior (Low=Never)	3.03	3.02	0.61	3.03	3.03	0.9
Mean of Index of Using Drugs/Alcohol	1.12	1.1	0.06*	1.11	1.09	0.01**
Mean of Index of Empathy (Low=Poor)	3.10	3.09	0.92	3.10	3.10	0.81
Mean of Index of Controlling Destiny (Low=Poor)	3.02	3.02	0.96	3.02	3.03	0.71
Mean of Parental Discipline Index (Low=Least Strict)	2.94	2.95	0.52	2.94	2.95	0.5
Mean of Social Position Index (Low=Least Engaged/High Isolation)	3.43	3.46	0.04**	3.43	3.46	0.03**
Mean of Safety Index (Low = Not Safe)	3.33	3.36	0.05**	3.33	3.36	0.06*
Mean of Index of Been Harmed or Threatened (Low=Little Harm)	1.52	1.49	0.11	1.51	1.48	0.06*
Mother's Education						
Eighth grade or less	8.3	7.5	0.41	8.5	7.6	0.3
Some high school (did not graduate)	11.3	14.0	0.02**	11.3	14.0	0.02**
High school equivalence (GED)	5.8	5.3	0.47	5.9	5.2	0.38
High school graduate	25.3	22.2	0.03**	25.0	22.1	0.05*
Vocational, trade, or business school after leaving high school	8.8	7.7	0.21	8.8	7.7	0.25
Some college	18.3	18.6	0.82	18.2	18.5	0.79
Graduated from a two-year college	10.5	9.5	0.36	10.4	9.5	0.35
Four-year college degree or other advanced degree	10.5	14.4	0***	10.6	14.7	0***

Table A.8 (c*ontinued*)

Variable	Baseline Characterics of Students Consenting to Be in the Study			Baseline Characteristics of Students Responding to Follow-Up Questionnaire		
	Participants	Comparison Group	p-value[a]	Participants	Comparison Group	p-value [a]
Household Income						
Less than $10,999	14.2	13.9	0.79	13.5	13.9	0.72
$11,000 to $24,999	24.8	22.5	0.11	24.9	22.3	0.07*
$25,000 to $39,999	22.4	19.5	0.03**	22.6	19.6	0.04**
$40,000 to $59,999	16.7	18.3	0.22	16.9	18.3	0.27
More than $60,000	14.0	19.6	0***	14.2	19.7	0***
Sample Size	**1,483**	**2,090**		**1,431**	**2,024**	

SOURCE: Student Survey and Parent Survey, School Records.

[a]The p-value is the smallest level of significance at which the null hypothesis that the impact equals zero can be rejected. If the p-value is less than .01, an impact is significant at the 1 percent level; if the p-value is less than .05, the impact is significant at the 5 percent level, and so on.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Table A.9

Treatment Group and Control Group Characteristics,
Elementary Schools

Variable	Baseline Characteristics			Baseline Characteristics of Students Responding to Follow-Up		
	Treatment Group	Control Group	p-value	Treatment Group	Control Group	p-value [a]
Gender						
Male	46.4	50.6	0.24	43.2	55.3	0.02**
Female	53.6	49.4	0.24	56.9	44.7	0.02**
Race						
Black	66.7	72.5	0.13	70.3	74.9	0.40
White	9.7	8.3	0.55	9.6	6.1	0.30
Hispanic	18.9	14.1	0.13	17.5	14.3	0.49
Other	1.6	2.3	0.57	1.6	1.7	0.97
Grade						
Kindergarten	10.5	12.9	0.25	0	0	
1	18.1	17.1	0.68	0	0	
2	17.0	18.0	0.67	0	0	
3	13.4	10.5	0.17	24.2	17.7	0.11
4	17.4	18.5	0.67	32.2	34.1	0.68
5	19.5	16.5	0.24	35.5	34.3	0.80
6	4.0	6.5	0.08*	8.1	14.0	0.05**
Mother's Average Age (Years)	36.4	35.5	0.21	38.9	38.4	0.66
Father's Average Age (Years)	37.5	39.0	0.09*	39.2	41.4	0.07*
Number of Tardy Arrivals During 1999-2000 School Year	3.4	3.6	0.78	4.0	3.0	0.30
Number of Absences During 1999-2000 School Year	7.2	7.5	0.62	6.6	6.5	0.83
Parent Feels It Is Safe for Child to Walk in Neighborhood	59.8	54.9	0.16	64.8	62.0	0.57
Parent Feels It Is Safe to Walk in Neighborhood	72.3	78.5	0.14	71.4	77.7	0.16

Table A.9 (c*ontinued*)

Variable	Baseline Characteristics			Baseline Characteristics of Students Responding to Follow-Up		
	Treatment Group	Control Group	p-value	Treatment Group	Control Group	p-value [a]
Sample Size for Items Above	**587**	**381**		**285**	**153**	
Baseline Reading Test Score (Percentile)	36.3	36.6	0.94	31.1	27.0	0.24
Sample Size for Reading Test Score	**378**	**206**		**227**	**102**	
Sample Size[b]	**278**	**148**		**158**	**71**	

SOURCE: Student Survey and Parent Survey, School Records.

[a]The p-value is the smallest level of significance at which the null hypothesis that the impact equals zero can be rejected. If the p-value is less than .01, an impact is significant at the 1 percent level; if the p-value is less than .05, the impact is significant at the 5 percent level, and so on.

[b]Sample sizes for the control group range from 148 to 381, sample sizes for the consenting treatment group range from 278 to 587. Sample sizes for control group members and treatment group members who completed followup surveys (third through fifth graders) range from 71 to 153 and from 158 to 285, respectively.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

are significantly different at baseline and the same four and one other were significantly different at follow-up. This also is an expected result, considering the high follow-up response rate of 85 percent for the student questionnaire (see Table A.4).

Appendix B

Technical Methods

This appendix describes the technical approach for estimating impacts of middle and elementary school centers. Section A provides details for the methods used to estimate impacts of middle school centers, presented in Chapter IV. Section B provides details for the methods used to estimate the impacts of elementary school centers, presented in Chapter V.

A. Methods for Estimating Impacts: Middle School Centers

As described in Chapter III, we estimated impacts for middle school centers using a comparison group design. We used propensity score matching techniques to select a comparison group in the 34 middle school grantees, of which 32 provided follow-up data used in the analysis. We used regression models to estimate impacts, and then weighted the site impacts to represent all first- through third-cohort grantees that served middle school students. We conducted a separate analysis to explore the relationship between attendance and outcomes.

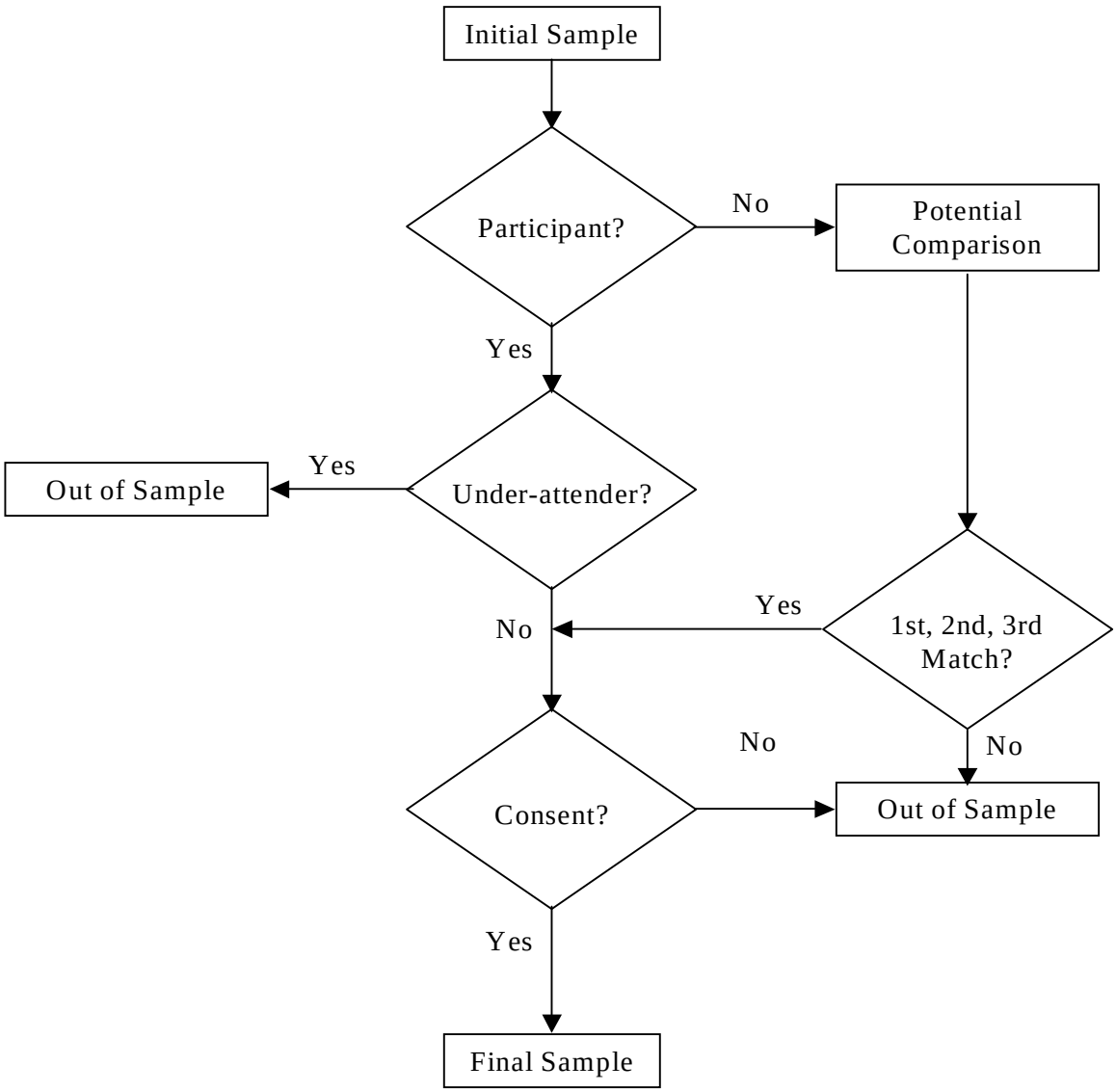
1. Identifying a Comparison Group Using Propensity Score Matching

To implement the comparison group design, we had a large group of students from schools operating centers and from similar schools complete questionnaires at the beginning of the 2000-2001 school year. Appendix A describes this part of the data collection in more detail and how the potential comparison group was identified from within the sample of students completing questionnaires. Ultimately, the potential comparison group was about seven times as large as the participant group. Having such a large potential comparison group provided a basis for the matching method to identify good matches. Figure B.1 depicts the various steps in the process of creating the comparison groups.

Using propensity score matching (PSM), we selected from the potential comparison group the students whose characteristics most closely resembled those of participants. PSM techniques were most prominently developed by Rosenbaum and Rubin (1983; 1985), who showed that the

Figure B.1

Overview of Process for Creating Middle School
Participant and Comparison Groups



technique yields a comparison group that is equivalent to a control group created using random assignment under the assumption that unobserved student characteristics are not correlated with outcomes. We applied the PSM method separately for each of the 34 middle school grantees, following five steps:

1. ***Logistic Regression Model Estimated.*** We estimated a logistic regression model in which the dependent variable was participation status and the independent variables were student demographic characteristics, indicators of student social development, measures of academic performance, and measures of student behavior (see Table B.1 for a listing of matching variables). Data on student characteristics used as a basis for matching were drawn from the baseline student questionnaire, the only data available at the time matching was conducted.[52] In most sites, 38 student characteristics were used in the matching.

2. ***Propensity Scores Calculated.*** For participants and potential comparison group students, we calculated propensity scores using data on each student's characteristics and parameter estimates from the logistic model estimated in the first stage. Presumably, students who actually were participants would have on average higher propensity scores than potential comparison group students; however, not every participant had a higher propensity score than every potential comparison group student.

3. ***Matching to Identify the "Closest" Comparison Students.*** For each participant, we selected, as that participant's "first-best" match, the potential comparison group student whose propensity score was numerically closest to the participant's score. To allow for possible attrition if parental consent was not received for the first-best match, we also identified potential comparison group students whose propensity scores would rank them as the second- and third-best match. We repeated the process for each participant, allowing individual students in the potential comparison group to be selected as matches for more than one participant.[53]

[52]Because school records data and parent questionnaire data were not available at the time matching was done, we could not include family income, parents' education and employment status, and students' baseline scores on standardized tests as matching variables.

[53]The number of participants to which a single potential comparison group student could be matched as a first-best match was limited to 10. The restriction was needed after it became evident that in some sites, a large proportion of participants were being matched to a single potential comparison group student, which resulted in a small comparison group in that site. Allowing such a comparison group to be selected would have dramatically reduced the statistical power of our analysis.

Table B.1

Percentage of Grantees for Which Participants and First- and Third-Best Matches Had Statistically Equivalent Mean Characteristics

	Original Match	
Variable	First Best	Third Best
Race		
Black	93.9	90.9
White	100.0	100.0
Hispanic	97.0	93.9
Other	97.0	93.9
Student Is a Female	100.0	87.9
Grade		
6	90.9	100.0
7	93.9	90.9
8	97.0	93.9
Other	97.0	93.9
Average Grades	100.0	97.0
Homework		
The student does the homework teachers assign	100.0	97.0
Mother or father helps student with homework	97.0	93.9
Mean of homework habits index (Low=Does Not Do Homework)	100.0	90.9
Number of Hours Read for Fun Yesterday	97.0	90.9
Number of Hours Watched TV Yesterday	97.0	93.9
Mean of Index of Confidence in Reading Skills (Low=Little Confidence)	93.9	90.9
Mean of Index of Helping Students Learn (Low=Little Help)	100.0	93.9
Overall Grade Student Gives School, 5=A through 1=F	97.0	100.0
Student Expects to Drop out of High School	97.0	93.9
Student Expects to Graduate from High School	100.0	93.9
At Least One Parent Is a College Graduate	100.0	93.9
At Least One Parent Is a High School Dropout	90.9	97.0
Child Doesn't Know Parents' Education Level	97.0	93.9
Mean of Index of Friends Encouraging Bad Behavior (Low=Never)	97.0	97.0
Mean of Index of Discipline Problems (Low=Few Problems)	100.0	97.0
Mean of Index of Bad Behavior (Low=Never)	97.0	93.9
Mean of Index of Good Behavior (Low=Never)	100.0	93.9

Table B.1 *(continued)*

Variable	Original Match	
	First Best	Third Best
Mean of Index of Using Drugs/Alcohol (Low=Does Not Use Drugs/Alcohol)	97.0	100.0
Mean of Index of Empathy (Low=Poor)	100.0	97.0
Mean of Index of Controlling Destiny (Low=Poor)	100.0	100.0
Mean of Parental Discipline Index (Low=Least Strict)	97.0	93.9
Mean of Social Position Index (Low=Least Engaged/High Isolation)	100.0	93.9
Overall Safety After School (1=Not Safe)	100.0	97.0
Mean of Safety Index (Low=Not Safe)	100.0	97.0
Mean of Index of School Climate (Low=Low Engagement)	97.0	97.0
Mean of Index of Serious School Problems (Low=Few Problems)	100.0	90.9
Mean of Index of Altering Behavior Because of Fear (Low=Never Alter Behavior)	97.0	97.0
Mean of Index of Been Harmed or Threatened (Low=Little Harm)	100.0	97.0
Average Percentage of Sites Matching Across All Characteristics	**97.8**	**95.0**

SOURCE: Student Survey.

NOTE: For each of 32 sites, t-tests are performed for the difference in means between participants and nonparticipants. This table shows the percentage of sites for which there is no significant difference between participants and nonparticipants. For example, in 100 percent of sites there was no difference in the percent of participants who were female and the percent of nonparticipants who were female for all samples except for the "third-best match" sample (in which there were no significant differences in 87.9 percent of sites).

4. ***Quality of Match Tested.*** Once we identified matching students for each participant, we could test the equivalence of the overall participant and matched comparison groups. We conducted an F-test of the equality of the set of characteristics for participants and their first-best matches a and used the *p*-value from the F-test as the indicator of match quality.[54] Higher *p*-values indicated that the groups were more similar.

5. ***Alternative Matching Specifications Assessed.*** To assess whether other comparison groups could be more equivalent, we used an algorithm to generate 2,000 logistic regression specifications by drawing randomly from the much larger set of all combinations of characteristics and second-order interactions of characteristics (squared terms and interacted terms). For each specification, we carried out steps 1 through 4 and used the comparison groups from the models with the highest, second-highest, and third-highest *p*-values from the F-test.

The matching process resulted in 3,921 comparison group students matched to 2,472 participants. Statistical tests verified the similarity of the matching students. The *p*-values for the joint test of equality of the 38 matching characteristics averaged 0.96, whereas the *p*-value of the F-test that these characteristics were the same for the participant group and the entire potential comparison group was 0.08. In other words, whereas participants and the potential comparison group had significantly different characteristics, participants and the matched comparison group were statistically indistinguishable at reasonable levels of significance.

We also examined the mean values of individual characteristics for participants and matched comparison students. Table B.1 consists of 38 rows corresponding to the 38 matching characteristics, showing the percentage of the 32 grantees for which there were significant differences between the mean value among participants and mean value among the first- (or third-) best matches. For each characteristic, the mean value in nearly all sites among participants and among matched comparison students was not significantly different at the 10

[54]The *p*-value indicates the probability that differences in the values of the participant and comparison group means that were observed could have resulted by chance under the null hypothesis that the means were jointly equal. Higher *p*-values imply that observed differences were more likely to have resulted from chance.

percent level. For example, the average characteristic had a significantly different mean value among participants and first-best matches in only 1 of 32 sites.

2. Impact Estimation

The basic approach for estimating the impact of middle school centers consisted of comparing the follow-up outcomes of participants and matched comparison group members, using regression models to adjust for baseline characteristics that may have influenced the outcomes. Outcomes of interest were regressed on an indicator of whether sample members were in the participant or comparison group, as well as a set of other explanatory variables. The basic regression model was:

$$(1) \quad Y = \alpha + P\beta + X\delta + \varepsilon$$

where $\alpha, \beta, \& \delta$ are coefficients that were estimated; P is the indicator of whether a student was in the participant group, and X represents the set of explanatory variables assumed to affect the outcome Y. (See Table B.2 for a list of the explanatory variables included in the model.) The estimated value of the coefficient β is an estimate of impact (that is, the difference in means between the participant group and the comparison group after adjusting for other characteristics). Because we used a complex sample design in selecting the sites and weights in the analysis, we used the SUDAAN® statistical package to estimate the standard errors of the coefficients of the model.

After estimating the regression models, we estimated "regression-adjusted" mean values of outcomes to facilitate interpreting the estimated impacts. Conceptually, the regression-adjusted mean value of an outcome for participants is the value of the outcome that the estimated model predicts for the average characteristics of the full sample. Similarly, the regression-adjusted

Table B.2

Explanatory Variables Included in the Basic Regression Model

Variable
Race/Ethnicity (White Excluded)
Black
Hispanic
Other race
Mixed race
English is not student's native language
Student is a female
Grade (grade 5 excluded)
Student is overage for grade level
Average grades
Average grades squared
Student-reported confidence in reading skills composite variable
Student-reported peer interaction/empathy composite variable
Overall grade student gives school, 5=A through 1=F
Student expects to drop out of high school or graduate from high school but not attend college
Student-reported discipline problems composite variable
Student-reported index of controlling destiny
Student-reported parental discipline composite variable
Student-reported social position composite variable
Student-reported safety index
Parent-reported variables
Family receives Food Stamps/Temporary Aid to Needy Families/Medicaid/housing assistance
Household income
Whether student's mother has a two- or four-year college degree
Whether student moved during previous year
Household structure (two-parent households excluded)
Student lives with single parent and no other adults
Student lives in other household arrangement
Student suspended during 1999-2000 school year
Number of times suspended during 1999-2000 school year
Number of absences during 1999-2000 school year
Number of times late to class during 1999-2000 school year
Student retained in grade prior to current year

SOURCE: Baseline Student Survey, Followup Parent Survey, Baseline School Records.

NOTE: All student-reported variables were drawn from the baseline student survey. Parent-reported variables were drawn from the follow-up parent survey but were limited to those variables that were unlikely to have been influenced by program participation. Variables based on school records data were limited to those that measured baseline outcomes. In addition to the variables listed, the explanatory variables also included missing value flags—binary indicators of observations in which a particular characteristic was missing and its value was imputed.

mean for comparison group students is the value of the outcome that the estimated model predicts for a comparison student who had the average characteristics of the full sample. The difference between the regression-adjusted mean outcome of participants and comparison group students is the estimated impact of participating in centers. The regression-adjusted mean values are calculated as follows:

1. ***Regression Model Estimated.*** Using data from the full sample, the regression model is estimated and coefficient estimates generated.
2. ***Predicted Outcome Values Calculated for Each Student.*** For every student in the sample, the coefficient estimates and the student's actual characteristics are used to calculate a predicted value of the outcome for that student under two different scenarios. A predicted value is calculated using all the student's characteristics except participation status, which is set to one under the assumption that the student is a participant. Another predicted value is calculated under the assumption that the student is a comparison group member (the participation status variable is set to zero).
3. ***Mean Values of the Two Predicted Values Calculated.*** Among all students in the sample, the mean values of the two predicted values are calculated using sample weights to ensure that the resulting mean value is representative of the population. The mean of the predicted values calculated under the assumption that each student was a participant is the regression-adjusted mean among participants. The mean of the predicted values calculated under the assumption that each student was a comparison group member is the regression-adjusted mean among comparison group students. The difference between these two regression-adjusted mean values is the estimated impact of participation and should be equal to the estimated coefficient β.

We used a variant of the basic regression model to estimate impacts for subgroups of students, supplementing the model with an interaction term between the treatment indicator and an indicator of whether sample members were in the subgroups being considered. An example of a subgroup model is:

(2) $Y = \alpha + P * S1\beta_1 + P * S2\beta_2 + X\delta + \varepsilon$

where the terms are defined as in equation (1) except that S1 is a binary variable denoting membership in a particular subgroup and S2 is a binary variable denoting membership in its

complementary subgroup. For example, to estimate the differential impact of program participation on outcomes for males and females, S1 might equal 1 for males and S2 would equal 1 for females. In this model, the estimated coefficient β_1 would be the model's estimate of the effect of center participation for males and β_2 would be the model's estimate of the effect of participation for females.

3. Analysis of the Relationship between Center Attendance and Outcomes

The approach to estimate the relationship between center attendance and outcomes was to adjust for observable differences between frequent and infrequent participants using regression models. The regression model that was estimated was:

$$Y_i = X_i b + a_1 P_i + a_2 D_i + u_i$$

where Y_i is the outcome for student i, X_i is a set of student characteristics, P_i is an indicator variable for whether a student is a center participant, and D_i is a variable indicating the number of days the student attended the center during the year. The observable characteristics (X_i) were the same as those used to estimate impacts in the basic model. The estimate of the coefficient a_2 represents the "effect" of attending the program for additional days.

To calculate regression-adjusted mean outcomes, we estimated the coefficient estimates from the above model to calculate predicted outcome values for an assumed level of attendance. We then averaged the predicted values to generate the predicted value of the outcome, which was termed the "moderate participation" value. We used similar techniques to estimate mean regression-adjusted outcomes for infrequent and frequent participants, which are presented in the text.

Because students could differ in unobserved characteristics that were not accounted for in the regression model, the estimated differences in outcomes do not represent the causal effect of the difference in attendance. An alternative approach to estimate causal effects would be to identify characteristics that are related to attendance but unrelated to outcomes, termed "instrumental variables," and apply well-known methods to estimate the effects of attendance. We considered several potential instrumental variables from the student and parent surveys but ultimately rejected them. One was mother's employment status, because mothers who work may have a greater need to place their child in an after-school program. However, mother's employment status proved to be nearly uncorrelated with attendance. We rejected other potential instruments for similar reasons.

4. Sample Weights

Because grantees included in this evaluation were sampled from among all grantees in cohorts one, two, and three that served middle school students, weights needed to be applied so that impact estimates could be applied to the full population of middle-school students served by centers. The construction of sample weights had two parts. First, a basic weight was constructed as the inverse of a grantee's probability of selection in its stratum. Second, the basic weight was modified so that the number of students in the sampled grantees represented the number of students in the stratum from which the grantees were sampled. The formula for the sample weight of a student was:

$$(3)\quad W_{ij} = \frac{1}{p_{j1} * p_{j2}}$$

where

$$p_{j1} = \frac{\#\, grantees \quad sampled \quad from \quad stratum}{total\,\# \quad grantees \quad in \quad stratum}$$

$$p_{j2} = \frac{n_p + n_c}{N_p}$$

where n_p and n_c are the number of treatment and comparison students for each grantee.

Table B.3 shows the 16 strata from which grantees were selected, along with the associated selection probabilities of grantees sampled from those strata. The effect of the first part of the sample weight is to make each student's data representative of students in grantees in the sample stratum that were and were not selected. The second part of the weight ($1/p_{j2}$) ensures that the weight given to all sampled students in a particular grantee depended on the number of eligible students served by the grantee (N_p,) rather than the number of treatment and comparison students included in the sample ($n_p + n_c$). When the weights from sample members within a site are summed, the site's cumulative weight is:

$$\sum_j W_{ij} = \sum_j \frac{1}{p_{j1} * p_{j2}} = \frac{n_p + n_c}{p_{j1} * p_{j2}} = \frac{n_p + n_c}{p_{j1}} * \frac{N_p}{n_p + n_c} = \frac{N_p}{p_{j1}}$$

B. Methods for Estimating Impacts: Elementary School Centers

We used an experimental design to estimate the impact of centers that served elementary school students. Students who were eligible to participate in elementary school centers were randomly assigned into a treatment group that was allowed to participate in the centers or a control group that was not allowed to participate. The experimental design ensured that the treatment group and control group were statistically similar in their baseline characteristics (both observed and unobserved). Thus, any outcome differences between the two groups at follow-up could be attributed to participation in the center.

The elementary school design was distinct from the middle school design in another respect. Whereas middle school grantees were selected for the evaluation at random from among all

Table B.3

Middle School Grantee Selection Probabilities

Stratum Number	Stratum Name	Number of Grantees in Population	Number of Grantees in Sample	Selection Probability
1	Northeast, Rural	16	2	0.13
2	Northeast, Urban	38	2	0.05
3	East, Rural	27	2	0.07
4	East, Urban	22	2	0.09
5	Southeast, Rural	32	2	0.06
6	Southeast, Urban	25	2	0.08
7	North Central, Rural	28	2	0.07
8	North Central, Urban	35	3	0.09
9	Midwest, Rural	22	2	0.09
10	Midwest, Urban	11	2	0.18
11	Mid-South, Rural	34	2	0.06
12	Mid-South, Urban	23	2	0.09
13	Northwest, Rural	37	3	0.08
14	Northwest, Urban	7	2	0.29
15	Southwest, Rural	27	2	0.07
16	Southwest, Urban	33	2	0.06

grantees serving middle school students, elementary school grantees were selected for the evaluation purposefully in order to ensure that they would be able to successfully implement random assignment. The impact findings for the elementary school centers in the evaluation have high internal validity but do not generalize to all grantees serving elementary school students.

If random assignment is correctly implemented, a comparison of average outcomes for the treatment and control groups is an estimate of the impact of participation in elementary school centers. However, the variance of the estimates can be reduced by estimating impacts using regression models to adjust for chance differences in baseline characteristics. The model used to estimate elementary school impacts was slightly different from that used to estimate middle school impacts. The model is

$$Y = \alpha + X\delta + \beta_1 P*G1 + \beta_2 P*G2 + ... + \beta_7 P*G7 + \varepsilon$$

In this model, we estimated separate impacts for each elementary school grantee (G1 through G7). To generate the estimate of the overall impact of elementary school centers, we calculated a simple mean of the seven site-specific impacts. We calculated the standard error of the overall impact estimated using the information from the variance-covariance matrix of the estimates of the seven coefficients representing the site-specific impacts. The regression model included whether students were overage for grade, race, parental education, parental income, household structure, whether the family received public assistance, the number of times the family moved in the past year and, when possible, the baseline value of the outcome variable.

For elementary schools, we used the same procedures to estimate regression-adjusted treatment and control means, subgroup impacts, and the attendance-outcome relationship described for middle schools. We also weighted students to offset differential probabilities of

selection that arose because centers had different numbers of applicants for slots and therefore required different ratios of treatment to control group assignments. In this case the weight simply was the inverse of the selection probability, suitably normalized to sum to the number of students that were randomly assigned.

References

Bissell, J., and J. Malloy. "Evaluation of California's After-School Learning and Safe Neighborhoods Partnerships Program: 1999-2001." Irvine, Calif., February 2002.

Brooks, Pauline, *et al*., "Longitudinal Study of LA's BEST After-School Education and Enrichment Program, 1992-1994." Los Angeles, Calif.: UCLA Center for the Study of Evaluation, 1995.

Cooper, Harris and Jeffrey Valentine. "Using Research to Answer Practical Questions about Homework." *Educational Psychologist,* 36(3), 2001, pp. 143-153.

Dynarski, Mark, Susanne James-Burdumy, Daniel Mayer, Mary Moore, John Mullens, Tim Silva, Carol Pistorino, and Douglas Hermond. *A Broader View: The National Evaluation of the 21st-Century Community Learning Centers Program.* Design Report submitted to the U.S. Department of Education. Princeton, N.J.: Mathematica Policy Research, March 2001.

Grossman, J., M. Price, V. Fellerath, L. Jucovy, L. Kotloff, R. Raley, and K. Walker. "Multiple Choices After School: Findings from the Extended-Service School Initiative." Philadelphia, Pa.: Public/Private Ventures, June 2002.

Halpern, Robert, Julie Spielberger, and Sylvan Robb, "Evaluation of the MOST (Making the Most of Out-of-School Time) Initiative: Final Report. Summary of Findings." Discussion Paper PS-32. Chicago: Chapin Hall Center for Children, 2001.

Langford, Barbara Hanson. "State Legislative Investments in School-Age Children and Youth." Washington, D.C.: The Finance Project, June 2001.

Moore, Mary, Mark Dynarski, John Mullens, Susanne James-Burdumy, and Linda Rosenberg. "Enhancing the 21st-Century Community Learning Centers Program Evaluation: A Concept Paper." Submitted to the Charles Stuart Mott Foundation. Princeton, N.J.: Mathematica Policy Research, November 2000.

Reisner, E., R. White, J. Birmingham, and M. Welsh. "Building Quality and Supporting Expansion of After-School Projects: Evaluation Results from the TASC After-School Program's Second Year." Washington, D.C.: Policy Studies Associates, February 2001.

Rosenbaum, Paul and Donald Rubin. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, vol. 70, 1983, pp. 41-55.

Tierney, Joseph, Jean Baldwin Grossman, and Nancy Resch. "Making a Difference: An Impact Study of Big Brothers and Big Sisters." Philadelphia, Pa.: Public/Private Ventures, November 1995.

Trousdale, Donna. "First Year Evaluation of the After-School Program for Middle School Youth." Educational Research Service, Arlington, Va., Summer 2000. [http://www.ers.org/spectrum/ sum00a.htm].

U.S. Department of Education. "Safe and Smart: Making the After-School Hours Work for Kids." [http://www.ed.gov/pubs/SafeandSmart/index.html]. Washington, D.C., June 1998.

U.S. Department of Education. "Working for Children and Families: Safe and Smart After-School Programs." [http://www.ed.gov/pubs/parents/SafeandSmart/index.html]. Washington, D.C., May 2000.

Warren, Constancia, P. Brown, and N. Freudenberg. "Evaluation of the New York City Beacons." Phase 1 Findings. New York: Academy for Educational Development, 1999.