

The Evaluation of Enhanced Academic Instruction in After-School Programs

Final Report

Executive Summary

The Evaluation of Enhanced Academic Instruction in After-School Programs

Final Report

Executive Summary

September 2009

Authors:

Alison Rebeck Black
Marie-Andrée Somers
Fred Doolittle
Rebecca Unterman
MDRC

Jean Baldwin Grossman
Public/Private Ventures

Project Officer:

Elizabeth Warner
Institute of Education Sciences

NCEE 2009-4078
U.S. Department of Education

U.S. Department of Education

Arne Duncan

Secretary

Institute of Education Sciences

John Q. Easton

Director

National Center for Education Evaluation and Regional Assistance

John Q. Easton

Acting Commissioner

September 2009

This report was prepared for the National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, under contract no. ED-01-CO-0060/0004 with MDRC.

This report is in the public domain. Authorization to reproduce it in whole or in part is granted. While permission to reprint this publication is not necessary, the citation should read: Black, A. R., Somers, M.-A., Doolittle, F., Unterman, R., and Grossman, J. B. (2009). *The Evaluation of Enhanced Academic Instruction in After-School Programs: Final Report* (NCEE 2009-4077). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions or recommendations or views with regard to actions policymakers or practitioners should take in light of the findings in the report.

To order copies of this report,

- Write to ED Pubs, Education Publications Center, U.S. Department of Education, P.O. Box 1398, Jessup, MD 20794-1398.
- Call in your request toll free to 1-877-4ED-Pubs. If 877 service is not yet available in your area, call 800-872-5327 (800-USA-LEARN). Those who use a telecommunications device for the deaf (TDD) or a teletypewriter (TTY) should call 800-437-0833.
- Fax your request to 301-470-1244 or order online at www.edpubs.org.

This report is also available on the IES website at <http://ncee.ed.gov>.

Alternate Formats

Upon request, this report is available in alternate formats, such as Braille, large print, audiotape, or computer diskette. For more information, call the Alternate Format Center at 202-205-8113.

Acknowledgments

This study represents a collaborative effort among the authors and the staff from the participating school districts and schools; the curriculum developers, Harcourt School Publishers and Success for All; our colleagues at MDRC and Public/Private Ventures (P/PV); and Institute of Education Sciences (IES) staff. The study has benefited especially from the time, energy, and commitment put forth by staff in the participating school districts and community-based organizations to implement the two enhanced after-school programs used in the Evaluation of Enhanced Academic Instruction in After-School Programs Study, to allow access to after-school classrooms, and to respond to requests for data.

Susan Bloom and staff at Bloom Associates and on-site district coordinators provided invaluable support to the school districts and community-based organizations in their efforts to implement the enhanced after-school programs and meet the demands of the evaluation.

The study's technical working group provided valuable insights on the evaluation design, data analysis, and early versions of the report. We thank Megan Beckett, Thomas Dee, Carolyn Denton, Larry Hedges, Nancy Jordan, and Rebecca Maynard for their expertise and guidance.

The listed authors of this report represent only a small part of the team involved in this project. Linda Kuhn and the staff at Survey Research Management and local data collection coordinators managed and conducted the baseline and follow-up testing and survey data collection effort, as well as classroom observations and interviews with school staff working in the regular after-school program. And Laurie Kotloff at P/PV processed and managed the interview data.

At MDRC, Alixandra Barasch coordinated production of this report and worked tirelessly to prepare the final versions of tables, figures, and other supporting documents. Rebecca Kleinman and Adam Wodka assisted with data collection and provided programming and analysis support. Gordon Berlin, James Kemple, Cynthia Miller, Corinne Herlihy, and John Hutchins provided substantive expertise through their thoughtful comments on, and reviews of, this report. And Mario Flecha, Julia Gomez, Diane Singer, Setha Sean, and Genevieve Williams assisted with report production and fact-checking.

Finally, the authors would like to thank Robert Weber and John Hutchins for their thoughtful editing of the report and Stephanie Cowell and Inna Kruglaya for preparing the final text for publication.

The Author

Disclosure of Potential Conflicts of Interest¹

The research team for this evaluation consists of a prime contractor, MDRC, Inc., of New York City, NY, and three subcontractors, Public/Private Ventures (P/PV) of Philadelphia, PA, Survey Research Management (SRM) Corporation of Boulder, CO, and Bloom Associates, Inc. of New York. None of these organizations or their key staff has financial interests that could be affected by findings from the evaluation of the two enhanced after-school interventions considered in this report. No one on the Expert Advisory Panel, convened by the research team to provide advice and guidance, has financial interests that could be affected by findings from the evaluation.

¹Contractors carrying out research and evaluation projects for IES frequently need to obtain expert advice and technical assistance from individuals and entities whose other professional work may not be entirely independent of or separable from the particular tasks they are carrying out for the IES contractor. Contractors endeavor not to put such individuals or entities in positions in which they could bias the analysis and reporting of results, and their potential conflicts of interest are disclosed.

Executive Summary

The primary purpose of this study is to determine whether providing structured academic instruction in reading or math to students in grades two to five during their after-school hours — instead of the less formal academic supports offered in regular after-school programs — improves their academic performance in the subject. This is the second and final report from the Evaluation of Enhanced Academic Instruction in After-School Programs — a two-year demonstration and random assignment evaluation of structured approaches to teaching math and reading in after-school settings. The study is being conducted by MDRC in collaboration with Public/Private Ventures and Survey Research Management.

The study was commissioned by the National Center for Education Evaluation and Regional Assistance at the U.S. Department of Education's Institute of Education Sciences (IES), in response to growing interest in using out-of-school hours as an opportunity to help prepare students academically (Bodilly and Beckett, 2005; Ferrandino, 2007; Miller, 2003). The federal government has been making an investment toward this goal through its 21st Century Community Learning Centers (21st CCLC) funding.¹ A distinguishing feature of after-school programs supported by 21st CCLC funds has been the inclusion of an academic component. Yet, findings from the National Evaluation of the 21st CCLC program indicate that, on average, the 21st CCLC program grants had limited effects on students' academic achievement (Dynarski and others, 2003; Dynarski and others, 2004; James-Burdumy et al., 2005). One possible explanation for this finding is that academic programming in after-school centers is typically not sufficiently intensive, usually consisting primarily of sessions in which students received limited additional academic assistance (such as reading/math tutoring or assistance with homework). In response, IES decided to fund the development, implementation, and evaluation of instructional resources for core academic subjects that could be used in after-school programs.

As part of this study, enhanced after-school programs providing instruction in either reading or math were implemented in after-school centers during two school years. In the first year of the demonstration (2005-2006), the enhanced programs were implemented in 50 after-school centers — with 25 after-school centers offering the enhanced math program and 25 centers offering the enhanced reading program. The study was then extended to include a second year of operations (2006-2007). This report focuses on the 27 after-school centers that

¹The 21st CCLC program is a state-administered discretionary grant program in which states hold a competition to fund academically focused after-school programs. Under the No Child Left Behind Act of 2001, the program funds a broad array of before- and after-school activities (for example, remedial education, academic enrichment, tutoring, recreation, and drug and violence prevention), particularly focusing on services to students who attend low-performing schools, to help meet state and local student academic achievement standards in core academic subjects (U.S. Department of Education, 2007).

agreed to participate in the study for both years — 15 of which implemented an enhanced after-school math program, and 12 of which offered the enhanced after-school reading program.²

The purpose of this report is to address questions that are relevant to both years of implementation, such as whether one-year impacts are different in the second year of program operations and whether students benefit from being *offered* two years of enhanced after-school academic instruction. Therefore, this report presents findings from the 27 centers that have data to address all these study questions.

Key Findings

Enhanced Math Program

- **One year of enhanced instruction produces positive and statistically significant impacts on student achievement.** The impacts in the 15 centers on SAT 10 total math scores are 3.5 scaled score points in the first year (which is statistically significant) and 3.4 scaled score points in the second year of operations (which is not statistically significant). However, the difference in impacts between implementation years is not statistically significant. The impact of 3.5 scaled score points represents approximately one month's worth of extra math learning.
- **Two years of the enhanced program produces no additional achievement benefit beyond the one-year impact.** Several different analyses support this conclusion. An experimental analysis using the two-year sample finds that the estimated impact of offering students the opportunity to enroll in the enhanced program for two consecutive years (2.0 scaled score points, p-value = 0.52) and the estimated impact on these students of their first year of enrollment in the enhanced program (5.2 scaled score points, p-value = 0.07) are not statistically significantly different (p-value = 0.28). A nonexperimental analysis finds that this remains the case after adjustments are made for students in the enhanced program group who did not attend the enhanced program at all in the second year.
- **There was program fidelity across both years of implementation.** Certified teachers were hired, trained, and provided paid preparation time as in-

²Findings from all 50 centers are summarized in Appendix A of this report and are presented in the first-year report (Black et al., 2008). The 27 continuing centers are not statistically representative of all 50 centers so the findings from the 27 sites should not be generalized to all 50 centers.

tended; class sizes were approximately 9 students per instructor (intended ratio was 10 students per instructor); and reports from teachers and district coordinators (i.e., locally based technical assistance staff) indicated that teachers were able to cover the expected material in a class session.

- **Students in the enhanced program received math instruction that was more structured and intensive than regular after-school program students.** Students in the enhanced program group were offered formal instruction in math for three hours per week, and students in the regular program received a mix of homework help and other services not focused on math — although 17 percent of regular program group students in the first year, and 27 percent in the second, received some form of math instruction. Overall, during their first year of participation, enhanced program students received between 42 and 48 more hours of after-school math instruction than did students in the regular after-school program, which converts to a 26 to 30 percent increase in formal instruction in math over the course of the school year.
- **No clear lessons emerge for program improvement or targeting the program in particular types of schools.** Analysis exploring the associations between center-level impacts and the characteristics of schools in which centers operated and the implementation of the program produced no strong associations with clear programmatic implications.

Enhanced Reading Program

- **The enhanced program has no impact on total reading test scores after one year of participation.** This is true in both implementation years in these 12 centers.
- **Two years of participation produces significantly fewer gains in reading achievement for students in the enhanced program group.** Experimental analysis finds that offering students two years of the enhanced reading program has a negative and statistically significant impact on their total reading scores. Nonexperimental analysis suggests that this remains the case even after statistical adjustments are made for students in the enhanced program group who did not actually attend the enhanced program in the second year.
- **Though the reading program was staffed and supported as planned, implementation issues — especially related to the pacing of lessons — occurred in both years.** As with math, certified teachers were hired, trained, and provided paid preparation time as intended, and class sizes were approx-

imately 9 students per instructor (intended ratio was 10 students per instructor). However, lesson pacing was a problem in the first year and continued to be in the second year in at least four of the districts.³

- **Students in the enhanced program received reading instruction that was more structured and intensive than regular after-school program students.** Students in the enhanced program group were offered formal reading instruction for three hours per week, and most students in the regular program received a mix of homework help and other services not focused on reading — although 17 percent of regular program group students in the first year, and 12 percent in the second, received some form of reading instruction. Overall, during their first year of participation, enhanced program students received between 54 and 56 more hours of after-school instruction in reading than did students in the regular after-school program, which converts to 22 to 23 percent more formal instruction in reading over the course of the school year.
- **No systematic relationship exists between center-level impacts and program implementation or the local school context.**

Research Questions

The overarching purpose of this evaluation is to determine whether providing students with enhanced after-school academic instruction improves their math or reading achievement above and beyond what they would have achieved had they remained in a regular after-school program. In particular, the study examines whether making the enhanced program available to students for one year improves student achievement, and whether that impact differs when the program is in its second year of operation and, thus, more mature, compared to the first implementation year. Therefore, the following impact questions are examined in this report:

- **What is the impact on student achievement of offering students the opportunity to participate in the enhanced after-school program for one school year?**
- **Is this impact different in the second year of implementation than in the first year?**

³In the second year, district staff who helped in implementing the model were asked if pacing continued to be a problem for staff. Of all 10 district staff interviewed, four said it was a problem, four said it was not, and two did not answer the question, so it's not clear whether it was or was not a problem in those last two districts.

The study can also examine whether making the enhanced program available to students for two school years — thereby potentially lengthening students’ average level of exposure to the program — improves student achievement. Hence, the following question is also addressed in this report:

- **What is the impact of offering students the opportunity to participate in the enhanced after-school programs for two consecutive years?**

To help interpret and understand the magnitude of the impact findings, the study also examines how well the academic services received by the enhanced after-school program group were implemented, whether the implementation differed across implementation years, and whether there is a measurable difference between the services received by students assigned to the enhanced program and the services received by students assigned to the regular after-school program.

The report also examines two questions that cannot be answered based on the experimental design of the study. First, in order to provide information about the treatment for those who actually received it in both years (rather than the effect of offering two years of programming, which includes students who did not actually participate both years), this report examines the relationship between achievement and program participation for those students who participated in both years of the enhanced after-school services. Second, because the enhanced program was offered in a variety of settings, this report also examines the association between impacts on achievement and the variation in the local school context, as well as variation in program implementation. These nonexperimental findings can then be used to help interpret the generalizability of the overall experimental findings, as well as generate possible avenues for program improvement.

Study Design

After-School Centers in the Study

At the start of the study, after-school centers were chosen based on their expressed interest and their ability to implement the program and research design. Assignment of centers to either the reading or the math enhanced program was based on a combination of local preferences, including knowledge of their student needs, sufficient contrast between current academic offerings in the subject area and the enhanced program, and their ability to meet the study sample needs. The 27 after-school centers that voluntarily agreed to participate in the study for a second year are located in 11 sites within 10 states and include schools and community-based organizations in a variety of municipalities (rural, urban, and suburban) across the country. They provided the same type of enhanced after-school program (math or reading) as they had provided in the first year of the study.

Student Sample and Random Assignment

The research design uses a lottery-like process (random assignment) to offer students one of two alternative types of academic support during a 45-minute block of time: the enhanced after-school academic services being tested in this project or the regular after-school services offered in their center. Regular after-school services consisted most commonly of help with homework — although, across both years of implementation, 22 percent of regular program staff in math centers reported providing some form of academic instruction in math and 14 percent of regular program staff in reading sites reported providing some form of academic instruction in reading.

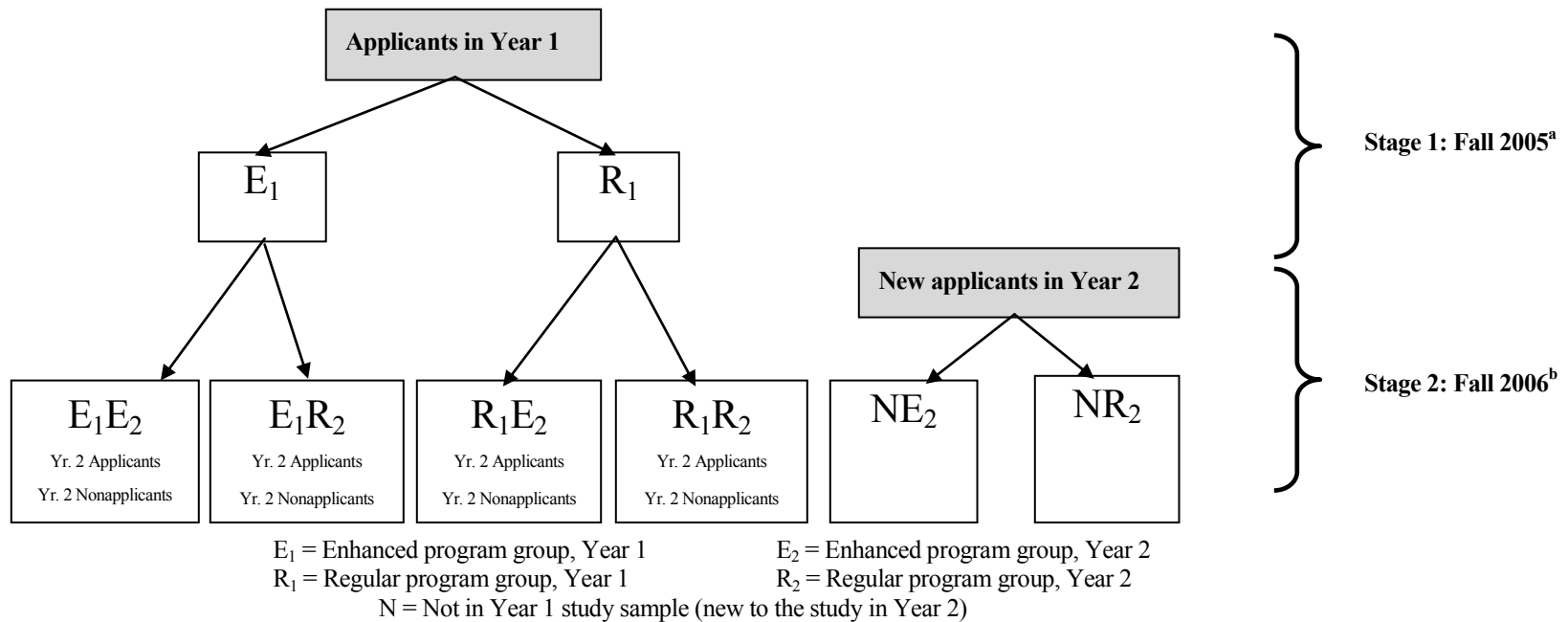
The target population for the study is students in second through fifth grades who are behind grade level in reading or math but not by more than two years. The study sample was recruited from students enrolled in after-school programs and identified by local staff as in need of supplemental academic support to meet local academic standards. Those whose parents then consented to be part of the study and applied for their children to participate in the enhanced program were included in the study sample. Given that instruction in these programs is provided in a small-group format and is not specifically developed to address special needs, students with severe learning disabilities and behavior problems or who could not receive instruction in English were excluded from the sample.

This study is based on a two-stage random assignment design of students, in which students were randomly assigned by grade within each after-school center on two separate occasions — once at the beginning of the first year of the study (first stage in fall 2005, see Stage 1 of Figure ES.1) and then again at the beginning of the second study year (second stage in fall 2006, see Stage 2 of Figure ES.1). (For more details on this two-stage random assignment design, see Box ES.1.) As a result, the sample includes: students who applied to the first year of the study (as described above) and were randomly assigned to either the enhanced program group (E_1) or the regular program group (R_1) and are referred to throughout this report as Cohort 1; students who were not offered the enhanced program in the first year and were applicants in the second year who were either offered the enhanced program (R_1E_2 and NE_2 applicants) or the regular program (R_1R_2 and NR_2 applicants) and are referred to throughout this report as Cohort 2; and students who, through the two-stage random assignment design, were randomly assigned to the enhanced program in both implementation years (E_1E_2 group in Figure ES.1) or assigned to the regular program in both years (R_1R_2 group) and are referred to as the two-year sample. Cohort 1 and Cohort 2 student samples are used to estimate the one-year intent-to-treat impact of the program in the first and second implementation years, respectively. The two-year sample is used to estimate the intent-to-treat impact of offering students the enhanced program for two consecutive years.

The Evaluation of Academic Instruction in After-School Programs

Figure ES.1

The Two-Stage Random Assignment Process



NOTES:

^aIn Stage 1 of random assignment, all identified low-performing students who applied to the study were randomly assigned, stratified by grade within each after-school center, to either the enhanced after-school program or the regular after-school program.

^bStage 2 of random assignment consisted of two groups, applicants and nonapplicants. Applicants in the second year consisted of newly identified low-performing student applicants in Year 2 and students from Year 1 who applied to the second year of the study. Both of these groups of second year student applicants were randomly assigned, stratified by grade and their first year treatment status (whether they were part of the enhanced or regular after-school program group, or not part of the study in its first year) within each after-school center, to either the enhanced after-school program or the regular after-school program. Nonapplicants are those students from Year 1 who had participated in the first year of the study, but did not apply to the second year of the study. They too were randomly assigned (separately from applicants) by grade and their first year treatment status within each after-school center.

Box ES.1

Two-Stage Random Assignment Design

The study is based on a two-stage random assignment design. At the beginning of the first study year (1st stage in fall 2005, see Stage 1 of Figure ES.1), identified low-performing students who applied to the study were randomly assigned by grade within each after-school center to either the enhanced program group (E_1) or the regular program group (R_1), and are referred to as Cohort 1.

At the end of the first study year, IES decided to extend the study for a second study year to assess both: (1) the one-year impact of the enhanced program and whether that impact changes over time once the site and staff have experience with the program (i.e., a comparison of the one-year impact of the program between the first and second study year), and (2) the impact of extended exposure to the enhanced program (i.e., an estimate of the two-year cumulative effect of being offered the enhanced program both years compared to being offered the regular program both years).

In order to address both these goals for the second study year, a second round of random assignment was conducted consisting of two groups of students, applicants and nonapplicants (2nd stage in fall 2006, see Stage 2 of Figure ES.1). The application process in the second year of the study was conducted the same as in the first year of the study. Applicants in the second year consisted of newly identified low-performing students who were *new applicants in year 2* and students from Cohort 1 who voluntarily applied to the second year of the study. Both of these groups of student applicants in Year 2 were randomly assigned by grade within each after-school center to either the enhanced program group or the regular program group; applicants from Cohort 1 were also randomly assigned by their first year treatment status (whether they were part of the enhanced or regular after-school program group). Randomly assigning for a second time students who participated in the first year, rather than allowing them to maintain their initial randomly assigned grouping, ensured that those who were offered the enhanced program the first year did not receive special treatment once the study was extended.

Nonapplicants are the remaining Cohort 1 students who had participated in the first year of the study, but did not apply to the second year of the study. They too were randomly assigned (separately from applicants) by grade and their first year treatment status within each after-school center. Randomly assigning both the applicants and nonapplicants from Cohort 1 maintains an intent-to-treat sample of Cohort 1 students who are cumulatively offered two years of the program or never offered the program. (Note, fifth-graders from Cohort 1 were excluded from the second stage of the random assignment in fall 2006 because, as sixth-graders, they were no longer eligible for the program and thus did not reapply.)

Impact findings are based on data collected from students, regular-school-day teachers, and school records. The Stanford Achievement Test, Tenth Edition (SAT 10), abbreviated battery for math or reading (depending on the intervention implemented), was administered to students at the beginning and end of the school year to measure the gains in achievement. For second- and third-grade students in the reading sample (and all students in the second year), the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) was also administered to measure fluency.

When estimating the impact of one year of exposure to the enhanced instruction separately for each implementation year, the study is equipped to detect an impact of 0.10 standard deviation in math and 0.11 standard deviation in reading in the first year of implementation, and an impact of 0.15 standard deviation in math and 0.14 standard deviation in reading in the second year of implementation.⁴ The study is also equipped to detect the impact of offering students two consecutive years of the program that is as small as a 0.21 standard deviation for the math program and 0.23 standard deviation for the reading program.

The following two sections present findings for the enhanced math and reading programs, respectively, based on the 27 after-school centers that participated in both years of the study.

Overview of the Interventions

The two interventions being tested in this evaluation involve providing 45 minutes of formal academic instruction during after-school programs to students who need help meeting local academic standards. The model includes the use of research-based instructional material and teaching methods that were especially designed to work in a voluntary after-school setting. Two curriculum developers — Harcourt School Publishers and Success for All — were selected through a competitive process to adapt their school-day materials to develop a math model and a reading model, respectively. The developers were asked to create material that is engaging for students, challenging and tied to academic standards, appropriate for students from diverse economic and social backgrounds, and relatively easy for teachers to use with a small amount of preparation time.

⁴The number of students in the sample is a crucial factor that determines the degree to which the impacts on student achievement and other outcomes can be estimated with enough precision to reject with confidence the hypothesis that the program had no effect. In general, larger sample sizes provide more precise impact estimates. A common way to represent statistical precision is through the “minimum detectable effect size” (MDES). Formally, the MDES is the smallest true program impact (scaled as an effect size) that can be detected with a reasonable degree of power (80 percent) for a given level of statistical significance (5 percent).

- **Harcourt School Publishers** adapted and expanded its existing school-day materials to develop Harcourt Mathletics, in which students’ progress through material at their own rate, with pretests at the beginning of each topic to guide lesson planning, “skill packs” for each topic to provide instruction on the skill in small groups and opportunities for individual practice in its application, and posttests to assess mastery or the need for supplemental instruction. The model also includes games to build math fluency; hands-on activities; projects; and computer activities for guided instruction, practice, or enrichment.
- **Success for All Foundation (SFA)** adapted its existing school-day reading programs to create Adventure Island, a structured reading model with daily lessons that involve switching quickly from one teacher-led activity to the next. It includes the key components of effective reading instruction identified by the National Reading Panel and builds cooperative learning into its daily classroom routines, which also include reading a variety of selected books and frequent assessments built into lessons to monitor progress.

Sites hired certified teachers and operated the enhanced programs with the intended small groups of students, approximately 10 students per instructor. The implementation was supported by the following strategies related to staffing, training and technical assistance, and attendance that were managed and supported by Bloom Associates, Inc.:

- Instructors received upfront training, multiple on-site technical assistance visits, continued support by locally based staff, and daily paid preparation time.
- Efforts were made to support student attendance through close monitoring of attendance, follow-up with parents and students when absences occurred to encourage attendance and address issues preventing it, and incentives to encourage and reward good attendance.

Findings for the Math Program

As mentioned earlier, the math findings presented in this report pertain to the 15 centers that participated in two years of program operations and data collection.

Implementation of the Enhanced After-School Math Program

Overall, the enhanced math program was largely implemented as intended in both years of program operations. Each center was expected to hire certified teachers and to operate with 10 students per instructor. In the first year, for example, 98 percent of instructors were certified

teachers, and the programs operated with the intended small groups of students — on average, in the first year, eight students attended per instructor. The goal was to offer the program for approximately 180 minutes per week, and average offerings were 189 minutes in the first year (a statistically significantly greater amount than intended, $p\text{-value} = 0.00$) and 171 minutes in the second (which does not statistically differ from the amount intended, $p\text{-value} = 0.45$). Instructors were trained by Harcourt staff at the beginning of the year and were provided ongoing assistance.⁵ They also received paid preparation time.

Impacts from Offering One Year of the Enhanced Math Program

The impact of enrollment in one year of the enhanced math program on student outcomes is estimated by comparing the outcomes of students who were randomly assigned to enroll in the enhanced after-school math program for one school year with the outcomes of students who were randomly assigned to remain in the regular after-school program during that same school year.⁶ This is estimated separately for each implementation year (Cohorts 1 and 2).

On average, students in the enhanced program group in Cohort 1 received 48 more hours of academic instruction in math during the school year than students in the regular program group. This difference — which is statistically significant ($p\text{-value} = 0.00$) — represents an estimated 30 percent increase in total math instruction over and above what is received by these students during the regular school day. In Cohort 2, enhanced program students received 42 more hours — also a statistically significantly greater amount of time ($p\text{-value} = 0.00$) than received by those in the regular program group, and an estimated 26 percent increase in total math instruction. However, the added hours of math instruction was statistically smaller in the second year of implementation (42 hours) than in the first year of implementation (48 hours) ($p\text{-value} = 0.00$).

One year of enrollment in the enhanced after-school program had a positive and statistically significant impact on students' math achievement in Cohort 1 (3.5 scaled score points or 0.09 standard deviation) as measured by SAT 10 total math scores. This statistically significant impact represents a 10 percent improvement over what students in the enhanced group would have achieved had they not had access to the enhanced program, or about one month's extra

⁵Enhanced math program staff received two full days of upfront training on how to use the math materials, including feedback from the developers in practice sessions using the materials. Ongoing support given to the enhanced program staff consisted of multiple on-site technical assistance visits (in the first year by Harcourt and Bloom Associates and in the second year by Bloom Associates) and continued support by locally based staff.

⁶Referring back to Figure ES.1, the analysis compared E_1 versus R_1 in the Cohort 1 sample and, in the Cohort 2 sample, R_1E_2 versus R_1R_2 (applicants who had not received the program in the first year) and N_1E_2 versus N_1R_2 (new students in the second year). An overall F-test indicates there is no systematic difference in the baseline characteristics of students in the enhanced and regular program groups in either of the cohort-specific samples.

learning over the course of a nine-month school year. The estimated impact of the enhanced math program on SAT 10 total math scores is not statistically significant for students in the second year of implementation (p-value = 0.07). However, the difference in impacts between implementation years (Cohort 1 and Cohort 2 samples) is not statistically significant. Thus, it cannot be concluded that the enhanced after-school math program was more effective in one implementation year than the other.

One year of enrollment in the enhanced math program also had a positive and statistically significant impact on students' performance on locally administered standardized math tests for Cohort 2 (0.18 standard deviation, p-value = 0.01), and the difference in one-year impacts across cohorts is not statistically significant (p-value = 0.16), so it cannot be concluded that the impact of the enhanced program on locally administered tests differed from one implementation year to the other. However, one year of enrollment did not produce impacts on regular-school-day teacher reports of academic behaviors (homework completion, attentiveness in class, and disruptiveness in class).

Impacts from Offering Two Years of the Enhanced Math Program

The impact of offering students the opportunity to participate in the enhanced program for two consecutive years is estimated using the two-year sample by comparing the outcomes of students who were randomly assigned to either the enhanced after-school program or the regular after-school program for two consecutive school years.⁷ However, as mentioned above, to maintain the experimental design, all Cohort 1 students were randomly assigned — both those Cohort 1 students who reapplied in the second year (applicants) and those Cohort 1 students who did not (nonapplicants). Thus, 42 percent of students in the math sample who were offered two years of the enhanced program did not reapply for, and did not receive, the second year of the program services. Hence, the impact findings presented in this section are of a two-year *offer* of services (an intent-to-treat analysis), rather than the impact of *receipt* of two years of the enhanced program — a nonexperimental analysis that is discussed later in this summary.

The estimated impact of offering students the opportunity to participate in the enhanced after-school program for two consecutive years is not statistically significant (2.0 scaled score points on the SAT 10 total score, p-value = 0.52). To place these results into context, the impact of these students' first year in the enhanced program was also estimated and compared to their cumulative two-year impact. Their first-year impact is not statistically significant (5.2 scaled score points, p-value = 0.07). And the estimated impact of assigning students to two years of enhanced services is not statistically different from the impact on these students of their first

⁷An overall F-test indicates there is no systematic difference in the baseline characteristics of students in the enhanced and regular program groups in the two-year sample.

year of access to the program (p -value = 0.28). Hence, for this sample, there is no evidence that offering the enhanced math instruction a second year provides an added benefit.

Figure ES.2 places these impact estimates in the context of the actual and expected two-year achievement growth of students in the enhanced program group. It shows the two-year growth for students in the enhanced program and what their expected growth would have been had they been assigned to the regular program. It also shows the test score growth for a nationally representative sample of students. The test scores of students in the enhanced program group grew 66.3 points over the two years (44.5 points in the first and 21.8 points in the second). Test scores of students in the regular program group grew by 64.3 points (39.4 points in the first year and 24.9 points in the second). These growth rates for the two program groups produce the estimated (not statistically significant) impacts mentioned above, a five-point difference in test scores for this sample after one year and a two-point difference after two years.

Because not all students in the enhanced program group actually received a second year of enhanced services, a nonexperimental analysis was conducted to examine whether longer exposure to the enhanced program is associated with improved math achievement. This analysis is based on instrumental variables estimation, which makes it possible to statistically adjust for the 42 percent of students in the enhanced program group who never attended the enhanced program in the second year. These findings do not establish causal inferences and thus should be viewed as hypothesis-generating. However, such an analysis may help with interpreting the two-year impacts and provide useful information to program developers.

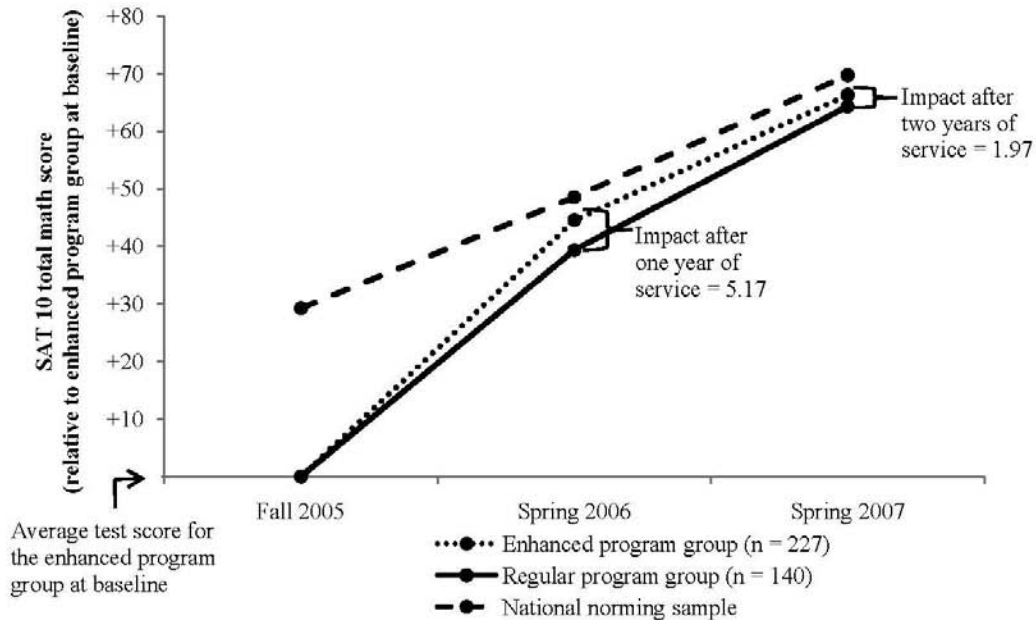
The findings from this nonexperimental analysis suggest that there is no additional benefit to a second year of enhanced services, even after adjustments are made for students who did not attend a second year. The nonexperimental estimate of receiving two years of enhanced after-school services (3.7 scaled score points for SAT 10 total math scores, p -value = 0.36) does not statistically differ from the 5.2 scaled score points estimated impact of one year of enhanced services (p -value = 0.40). Thus, across both the experimental and nonexperimental analyses, there is no evidence that a second year of the enhanced program — whether offered or received — improves math achievement, over and above the gains produced by the first year of enrollment.

Because the effectiveness of enhanced after-school instruction may be related to factors associated with program implementation or what the students experience during the regular school day, the study also examined whether characteristics of schools and program implementation are correlated with center-level impacts. The analysis is based on center-level impacts in both years of the study (i.e., 30 center-level impacts) and examines whether the impact of one year of enhanced services on SAT 10 total math scores in each after-school center is associated

The Evaluation of Academic Instruction in After-School Programs

Figure ES.2

SAT 10 Total Math Scores from Baseline to Follow-Up and the Associated Impact of the Enhanced Math Program After One Year and Two Years of Service



SOURCES: MDRC calculations are from baseline and follow-up results on the Stanford Achievement Test Series, 10th ed. (SAT 10) abbreviated battery. National norming sample calculations are from the SAT 10 (2002 norming sample): Stanford Achievement Test Series: Tenth Edition: Technical Data Report (Harcourt Assessment, 2004, pp. 312-338).

NOTES: The growth line for the enhanced program group is based on the observed mean baseline and follow-up test scores of students assigned to the enhanced after-school program for two consecutive years (baseline is Fall 2005; follow-ups are Spring 2006 and Spring 2007). The growth line for the regular program group represents the test scores that students in the enhanced program group would have obtained had they not been assigned to the enhanced program (calculated as the mean test score for the enhanced program group minus the estimated impact at a given time point). The growth line for the national norming sample is based on the average SAT 10 total math scores for a nationally representative sample of students with the same grade composition in each period as the two-year sample. Specifically, at each point in time (the fall baseline, the first spring, and the second spring), the SAT 10 national norm scores for second-, third-, and fourth-graders are averaged weighting each grade average score according to their proportion in the two-year study sample at baseline. This creates an expected two-year improvement of nationally representative students at the same grade levels as this study's sample. The baseline for the national norming sample is set relative to the average baseline score of the enhanced program group.

Estimated impacts on follow-up results are regression-adjusted using ordinary least squares, controlling for indicators of random assignment, baseline math total scaled score, race/ethnicity, gender, free-lunch status, age, overage for grade, single-adult household, and mother's education.

Among those who did not reapply to the study in the second year, nonresponse weights are used to account for those students for whom follow-up data was not collected. Statistical significance is indicated by (*) when the p-value is less than or equal to 5 percent.

with (1) the characteristics of the school that housed the after-school center and (2) the characteristics of a center's implementation of the enhanced program.

Though center-level program impacts on total math scores are correlated jointly with the overall set of school context and implementation measures included in the analysis, as well as with some individual measures, no clear lessons emerge for program operations. Program impacts were larger in after-school centers that offered the enhanced program for a greater number of days during the school year, suggesting a positive association between impacts and program dosage. However, this finding is inconsistent with the nonexperimental estimates of two versus one year of enhanced program participation. Program impacts were also larger in centers where one or more teachers left the enhanced program during the school year and in schools that made their Adequate Yearly Progress goals. With the available information, it is not possible to explain the reason for these relationships.

Findings for the Reading Program

Again, the reading findings presented in this report pertain to the 12 centers that participated in two years of program operations and data collection.

Implementation of the Enhanced After-School Reading Program

The enhanced reading program was staffed as intended and offered the intended amount of instruction in both years of program operations. Each center was expected to hire certified teachers and to operate with 10 students per instructor. In the second year, for example, all instructors were certified teachers, and the programs operated with the intended small groups of students — on average, in the second year, nine students attended per instructor. The goal was to offer the program for approximately 180 minutes per week, and average offerings were 177 minutes in the first year and 175 minutes in the second. Instructors were trained by SFA staff at the beginning of the year and were provided ongoing assistance.⁸ They also received paid preparation time.

However, in both years of the study, instructors found it challenging to maintain the intended pace of instruction. In the first year of the study, 79 percent of instructors reported that it was consistently or sometimes difficult to include all aspects of the reading program and maintain the intended pace of the daily lesson plan. In the second year of the study, half of the responding district coordinators reported that pacing continued to be a problem.

⁸Ongoing support given to the enhanced program staff consisted of multiple on-site technical assistance visits (by SFA and Bloom Associates) and continued support by locally based staff.

Classroom observations conducted by district coordinators were used to assess the fidelity with which instructors implemented the enhanced reading program. In the classes with students at the first- and second-grade reading levels (in Adventure Island, students are grouped by their initial reading level, not by grade), average fidelity scores did not statistically differ across the first and second years of implementation;⁹ in the classes with students reading above the second-grade level, average scores were lower in the second year, by a statistically significant amount (p-value = 0.00).¹⁰ It was also found that, in any given year, implementation of the program lacked consistency, as indicated by variation in the number of program components implemented by teachers.¹¹ In particular, in the second implementation year, returning teachers in both the lower and upper levels of the program had statistically significantly higher implementation fidelity scores than teachers who were new to the program (p-value = 0.00).

Impacts from Offering One Year of the Enhanced Reading Program

This analysis focuses on the impact of one year of enrollment in the enhanced reading program on student outcomes.¹² The difference between the background characteristics of students in the enhanced and regular program groups, both in Cohort 1 and Cohort 2, was greater than what would be predicted by chance, especially as related to baseline reading achievement test scores and household composition.¹³ Measures of student characteristics (including students' baseline test score) were included in the impact model to control for observed differences between the two program groups at baseline. Sensitivity analyses were conducted to gauge whether these covariates adequately control for baseline differences between students in the two program groups. These tests confirm that controlling for students' baseline characteristics — and particularly their pretest scores — produces internally valid estimates of the impact of the enhanced program.

⁹In both years, the average fidelity score was 5.1 out of a total possible score of six components.

¹⁰The average fidelity score was 4.2 out of a total possible score of five components in the first year; in the second year, it was 3.9.

¹¹For example, in the first implementation year, 9 percent of lower-level Adventure Island classes included between three and four of the six measured components; 68 percent included between four and five and 23 percent included between five and six.

¹²As was the case for math, this question is answered by comparing the outcomes of students who were randomly assigned to enroll in the enhanced after-school reading program for one school year and the outcomes of students who were randomly assigned to remain in the regular after-school program during that same school year. Referring back to Figure ES.1, the analysis compared E_1 versus R_1 in the first year sample, R_1E_2 versus R_1R_2 (returning students who had not received the program in the first year) and N_1E_2 versus N_1R_2 (new students) in the second year.

¹³Students in the enhanced group had statistically significantly lower baseline test scores and were more likely to come from a single-adult household.

On average, students in the enhanced program reading group in Cohort 1 received 54 more hours of academic instruction in reading during the school year than students in the regular program group. This difference — which is statistically significant (p-value = 0.00) — represents an estimated 22 percent increase in total reading instruction over and above what is received by these students during the regular school day. In Cohort 2, enhanced program students received 56 more hours — also a statistically significantly greater amount of time (p-value = 0.00) than received by those in the regular program group, and an estimated 23 percent increase in total reading instruction. And the net difference in added hours of instructional reading between implementation years is not statistically significant (p-value = 0.63).

One year of enrollment in the enhanced after-school reading program did not have a statistically significant impact on students' reading achievement (as measured by SAT 10 total reading scores), whether in the first or second year of implementation. It also did not have a significant impact on students' performance on locally administered standardized reading tests, nor did it produce impacts on the DIBELS measures of fluency or on regular-school-day teacher reports of academic behaviors (homework completion, attentiveness in class, and disruptiveness in class).

Impacts from Offering Two Years of the Enhanced Reading Program

The impact of offering students the opportunity to participate in the enhanced reading program for two consecutive years is estimated using the two-year sample in the same way as for the math sample, by comparing the outcomes of students who were randomly assigned to either the enhanced after-school program or the regular after-school program for two consecutive school years.¹⁴ The difference between the background characteristics of students in the enhanced and regular program groups in the two-year sample was greater than what would be predicted by chance, especially related to baseline reading achievement test scores and household composition.¹⁵ Measures of student characteristics (including students' baseline test scores) were included in the impact model to control for observed differences between the two program

¹⁴Referring back to Figure ES.1, this analysis involves comparing students in E_1E_2 versus R_1R_2 . As noted in the discussion of the math findings, the two-year sample includes “nonapplicants” from the first-year study sample who did not reapply to second year of the study. These nonapplicants — who constitute 43 percent of students in the enhanced program group for this analysis — did not actually receive a second year of enhanced after-school services as intended. Hence, the impact findings presented in this section are of a two-year *offer* of services (an intent-to-treat analysis), rather than the impact of two years of *receiving* the enhanced program, which is a nonexperimental analysis discussed later in this summary.

¹⁵Students in the enhanced program group have lower baseline test scores on average and are more likely to come from a single-adult household.

groups at baseline. Sensitivity analyses were conducted to gauge whether these covariates adequately control for baseline differences between students in the two program groups. These tests confirm that controlling for students' baseline characteristics — and particularly their pre-test score — produces internally valid estimates of the impact of the enhanced program.

The estimated impact of offering students the opportunity to enroll in the enhanced after-school program for two consecutive years is negative and statistically significant (-5.6 scaled score points on SAT 10 total reading scores; p -value = 0.04). To place these results into context, the estimated impact on these students of their first year of program enrollment (-3.6 points) was not statistically significant. And the estimated impact of assigning students to two years of enhanced services does not statistically differ from the impact on these students of their first year of access to the program (p -value = 0.46). Hence, while it can be said that being assigned to two years of enhanced services produces significantly fewer gains on test scores, it cannot be concluded that assigning students to enroll in the enhanced program for two years has a different impact on their reading achievement than assigning them to enroll in one year of the enhanced program.

Figure ES.3 places these impact estimates in the context of the actual and expected two-year achievement growth of students in the enhanced program group. It shows the two-year growth for students in the enhanced program and what their expected growth would have been had they been assigned to the regular program. It also shows the test score growth for a nationally representative sample of students. The test scores of students in the enhanced program group grew 25.1 points in the first year and 17.7 points in the second, for a total of 42.8 points. However, the test scores of students in the regular program group also grew, by 28.7 points in the first year and 19.7 points in the second, for a total of 48.4 points. The difference in growth rates between the two program groups produces the two-year impact estimate mentioned above, a -5.6-point difference after two years (in favor of the regular program group).

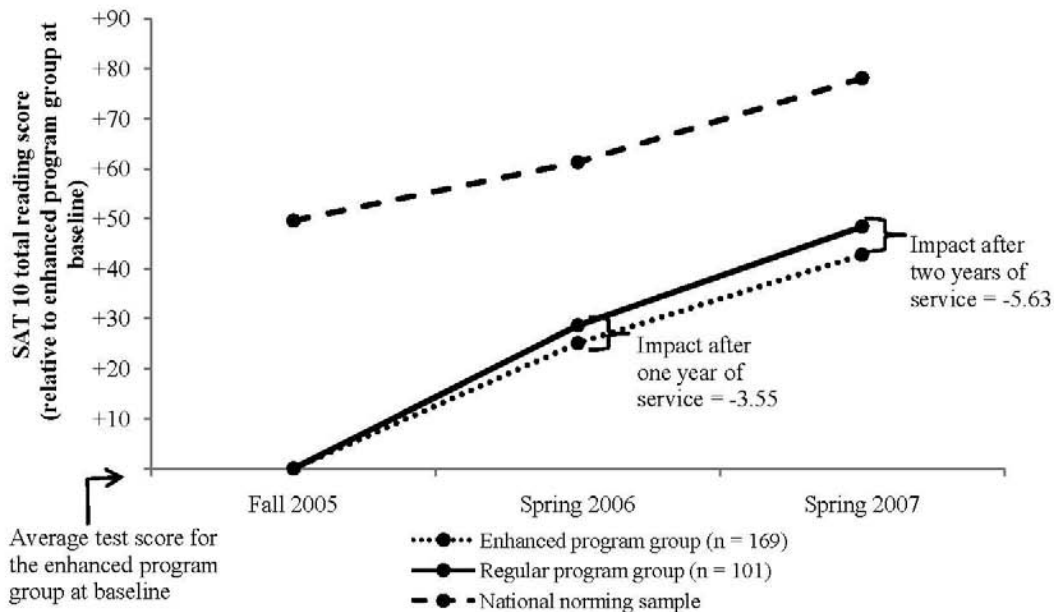
As in the math analysis, the association between receiving two years of enhanced services and reading achievement was estimated using nonexperimental methods, by statistically adjusting for the 43 percent of students in the enhanced program group who did not attend the program in the second year.¹⁶ Consistent with the experimental estimate for the impact of offering students two years of enhanced services, the association between receiving enhanced academic services for two consecutive years and SAT 10 total reading scores is negative and statistically significant (-7.5 scaled score points, p -value = 0.04). These findings suggest that two years of enhanced after-school services — whether offered or received — produces significantly fewer gains on reading achievement than two years in the regular program group.

¹⁶The association between receiving two years of enhanced services and reading achievement is estimated using instrumental variables estimation.

The Evaluation of Academic Instruction in After-School Programs

Figure ES.3

SAT 10 Total Reading Scores from Baseline to Follow-Up and the Associated Impact of the Enhanced Reading Program After One Year and Two Years of Service



SOURCES: MDRC calculations are from baseline and follow-up results on the Stanford Achievement Test Series, 10th ed. (SAT 10) abbreviated battery. National norming sample calculations are from the SAT 10 (2002 norming sample): Stanford Achievement Test Series: Tenth Edition: Technical Data Report (Harcourt Assessment, 2004, pp. 312-338).

NOTES: The growth line for the enhanced program group is based on the observed mean baseline and follow-up test scores of students assigned to the enhanced after-school program for two consecutive years (baseline is Fall 2005; follow-ups are Spring 2006 and Spring 2007). The growth line for the regular program group represents the test scores that students in the enhanced program group would have obtained had they not been assigned to the enhanced program (calculated as the mean test score for the enhanced program group minus the estimated impact at a given time point). The growth line for the national norming sample is based on the average SAT 10 total reading scores for a nationally representative sample of students with the same grade composition in each period as the two-year sample. Specifically, at each point in time (the fall baseline, the first spring, and the second spring), the SAT 10 national norm scores for second-, third-, and fourth-graders are averaged weighting each grade average score according to their proportion in the two-year study sample at baseline. This creates an expected two-year improvement of nationally representative students at the same grade levels as this study's sample. The baseline for the national norming sample is set relative to the average baseline score of the enhanced program group.

Estimated impacts on follow-up results are regression-adjusted using ordinary least squares, controlling for indicators of random assignment, baseline reading total scaled score, race/ethnicity, gender, free-lunch status, age, overage for grade, single-adult household, and mother's education.

Among those who did not reapply to the study in the second year, nonresponse weights are used to account for those students for whom follow-up data was not collected. Statistical significance is indicated by (*) when the p-value is less than or equal to 5 percent.

The analysis also explored whether the one-year impact estimates for each of the 12 centers are correlated with factors related to program implementation or what the students experience during the regular school day. The analysis is based on center-level impacts in both years of the study (i.e., 24 center-level impacts) and examines whether the impact of one year of enhanced services on SAT 10 total reading scores in each after-school center is associated with (1) the characteristics of the school that housed the after-school center and (2) the characteristics of a center's implementation of the enhanced program. Program impacts on total reading scores are not systematically correlated jointly with either the set of school context and implementation characteristics or with any of those characteristics individually. Thus, the measured local characteristics do not highlight any lessons for settings in which the program will be more effective than average.

Conclusion

This project found that it is possible to implement structured instruction in math and reading for second- through fifth-graders in an after-school setting. The provision of four days of training, ongoing on-site technical assistance, and local program coordinators supported implementation. In both years, math instructors reported few problems implementing Mathletics; teachers implementing the Adventure Island reading program found it challenging to maintain the intended pace of instruction in both years of the study.

It also proved possible to recruit certified teachers who will commit to participate for the full school year. Despite staff turnover across the two years of service offerings, there was growing experience in implementing the programs in the centers. Students also could be recruited each year and retained within each year in the program. The enhanced programs included a combination of extra monitoring of attendance and incentives and encouragement to attend, and students attended the enhanced program as much or more than regular after-school activities, despite initial concerns the program would not be appealing to students or their parents. However, as with most after-school programs (Dynarski and others, 2003; Dynarski and others, 2004), there was substantial dropoff in enrollment across school years (i.e., 42 and 43 percent of students who participated in the enhanced math and reading programs, respectively, in the first year did not attend the enhanced program for a second year).

The enhanced program produced a 26 to 30 percent increase in hours of academic instruction for math and 22 to 23 percent increase for reading, over the school year. For math, this produced one-month worth of extra learning, as measured by math standardized tests. Further, for math, the findings suggest that the benefits of the after-school academic instruction are captured in students' first year of participation, as a second year of math instruction did not

produce any additional benefits for students. However, for reading there were no positive effects on achievement after one year of the program, and findings after two years indicated the enhanced reading program led to slower progress in reading than did the regular after-school programming. In conclusion, these findings are consistent with a growing body of research that finds some evidence of improving achievement through after-school activities (Vandell, Reisner, and Pierce, 2007; Zief, Lauver, and Maynard, 2006).

References

- Black, A. R., Doolittle, F., Zhu, P., Unterman, R., and Grossman, J. B. (2008). *The Evaluation of Enhanced Academic Instruction in After-School Programs: Findings After the First Year of Implementation* (NCEE 2008-4021). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Bodilly, S., and Beckett, M. K. (2005). *Making Out-of-School Time Matter: Evidence for an Action Agenda*. Santa Monica, CA: RAND Corporation.
- Dynarski, M., James-Burdumy, S., Moore, M., Rosenberg, L., Deke, J., and Mansfield, W. (2004). *When Schools Stay Open Late: The National Evaluation of the 21st Century Community Learning Centers Program: New Findings*. U.S. Department of Education, National Center for Education Evaluation and Regional Assistance. Washington, DC: U.S. Government Printing Office.
- Dynarski, M., Moore, M., Mullens, J., Gleason, P., James-Burdumy, S., Rosenberg, L., Pistorino, C., Silva, T., Deke, J., Mansfield, W., Heaviside, S., and Levy, D. (2003). *When Schools Stay Open Late: The National Evaluation of the 21st Century Community Learning Centers Program, First-Year Findings*. Report submitted to the U.S. Department of Education. Princeton, NJ: Mathematica Policy Research, Inc.
- Ferrandino, Vincent L. (2007). *A New Day for Learning*. Denver, CO: Time, Learning, and After-school Task Force.
- Harcourt Assessment. *Stanford Achievement Test Series, Tenth Edition – Technical Data Report*. (2004). San Antonio, TX: Harcourt Assessment, Inc.
- James-Burdumy, S., Dynarski, M., Moore, M., Deke, J., and Mansfield, W. (2005). *When Schools Stay Open Late: The National Evaluation of the 21st Century Community Learning Centers Program: Final Report*. U.S. Department of Education, National Center for Education Evaluation and Regional Assistance. Washington, DC: U.S. Government Printing Office.
- Miller, B. M. (2003). *Critical Hours: Afterschool Programs and Educational Success*. Brookline, MA: Miller Midzik Research Associates.
- U.S. Department of Education, 21st Century Community Learning Centers. (2007). *Program Description*. Retrieved September 7, 2007, from <http://www.ed.gov/programs/21stcclc/index.html>.
- Vandell, D. L., Reisner, E. R., and Pierce, K. M. (2007). *Outcomes linked to high-quality after-school programs: Longitudinal findings from the study of promising after-school programs*. California After-School Network. Web site: <http://www.afterschoolnetwork.org/node/759>.
- Zief, S.G, Lauver, S., Maynard, R.A. (2006). *Impacts of After-School Programs on Student Outcomes: A Systematic Review for the Campbell Collaboration*. Campbell Collaboration Policy Review (Oslo, Norway). Web site: http://www.campbellcollaboration.org/review_list/categoryPrinterPage.shtml