JOURNAL OF EDUCATION FOR STUDENTS PLACED AT RISK, *12*(2), 213–237


# Results from a First-Year Evaluation of Academic Impacts of an After-School Program for At-Risk Students

Eric Jenner and Lynne W. Jenner
*The Policy & Research Group,*
*New Orleans, LA*

This article presents the research findings of an evaluation of the academic impacts of 21st Century Learning Centers (CCLC) in Louisiana. Using quasi-experimental design, the article operationalizes academic achievement as core and subject test performance on nationally standardized pre- and posttests (Iowa Test of Basic Skills; ITBS). Based on previous research and evaluation requirements, the article (a) employs outcomes of interest to policymakers (standardized test scores); (b) uses program attendance as a key independent variable; (c) uses efficient methods to control for extraneous impact on the dependent variable; and (d) focuses the evaluation on a specific group of student—at-risk children in Louisiana. Findings indicate that the 21st CCLC program is having a positive academic impact on participants who attend the program for 30 days or more. Further, impacts are shared across specific grantee programs, specific subjects, and subgroupings of students. Finally, the study finds that intensity of attendance is positively related to academic impact.

More than ever before, after-school programs are being viewed by parents and educators as vital to the health, growth, and education of children and adolescents, especially those at risk. Public policy research has tended to add support to these convictions. Shifts in economic dynamics and changes in the social landscape have weakened community networks and the fabric of social capital, thereby increasing the risk that children and adolescents will encounter academic and social problems (Putnam, 1995). After-school time is prime time for children to become engaged in and exposed to risky behavior. Juvenile crime, and especially juvenile violent crime, peaks in the hours immediately after school (Fight Crime: Invest in Kids,

Correspondence should be addressed to Eric Jenner, Ph.D., The Policy & Research Group, 8131 Oak St., Suite 300, New Orleans, LA 70118. E-mail: ejenner@policyandresearch.com

1997; U.S. Department of Justice, 1999). Teens that are left unsupervised after school are far more likely to use drugs, alcohol, and tobacco and to engage in sexual activity and other risky behavior (U.S Department of Education & U.S. Department of Justice, 1998).

With the increased popularity of after-school programs, policymakers and funding agencies have directed more money toward programming. In turn, they have also increased their expectations of the impacts of after-school programs. In particular, federal funding for some after-school programs is now directly contingent upon academic growth.

Whereas the preponderance of research on after-school programs finds evidence of emotional, social, psychological, and safety benefits (Barker, 1998; Gaynor & Horrowitz, 1998; Marshall et al., 1997; Posner & Vandell, 1994), research on the academic outcomes of after-school programming has been more equivocal (Huang, Gribbons, Kim, Lee, & Baker, 2000; Riggs & Greenberg, 2004; U.S. Department of Education, 2003). The inconsistent findings are less a flaw of the individual evaluations than the result of the "very rudimentary state" of outcome research as a whole (Scott-Little, Hamann, & Jurs, 2002, p. 411). Put another way, research has yet to specify who, among the heterogeneity of American children, may benefit from after-school programming (Riggs & Greenberg, 2004).

In all, a limited base of research and shifting demands on after-school programming have resulted in erratic policy. One well-known instance of this inconsistency is exhibited by the federal government's attention toward the 21st Century Community Learning Centers (CLCC) program. Originally initiated under the Elementary and Secondary and Education Act of 1998 and reauthorized under the No Child Left Behind (NCLB) Act of 2001, funding for the program increased from $40 million to $1 billion in 2002. Just 1 year later, and coincident with first-year findings of a Department of Education-sponsored evaluation (U.S. Department of Education, 2003) that concluded that 21st CLCCs were having no academic impact on after-school participants, the President proposed cutting the program's funding by 40%, citing in part "disappointing findings from a rigorous evaluation" (Harvard Family Research Project, 2003b, p. 1).

Since the Department of Education first released their initial report in 2003, scholars have been strident in their criticism of its legitimacy. Taken together, critics (e.g., Bissell et al., 2003; Kane, 2004; Riggs & Greenberg, 2004 ) have contended that the study's limitations did not warrant the broad inferences drawn and the policy decisions made in its wake. The weaknesses of this one study point to a more general problem: Where there is a dearth of high-quality research on educational interventions, policymakers struggle to make appropriate, well-informed funding decisions.

As policymakers increasingly focus on the academic outcomes of after-school programming, standardized test scores have become the chosen measure of success (Huang et al., 2000). Many scholars argue that the benefits of after-school pro-

grams transcend the narrow measure of standardized tests. Still, for good or ill, policy decisions are made on the basis of a very limited number of evaluations that employ standardized tests as a dependent variable. It strikes us that this is a void that needs filling. In conjunction with literature that documents and explains other benefits, after-school program outcome research needs more published findings that directly address the current demands of policymakers.

The purpose of this article is to present the research findings of a statewide evaluation of 21st CCLCs in Louisiana. The article's research design and findings, which were conducted for the Louisiana Department of Education, contribute to nascent outcome research literature on after-school programs and will hopefully assist decision-makers in the policy arena.

The study uses methods and presents findings in a way that speaks to the weaknesses in the literature outlined in sections to follow. Outcomes are measured as the relative change in standardized test scores over a single year. The evaluation focuses on a specific population—at-risk children in Louisiana—and includes further specifications to allow for grantee-level analysis. Key baseline measures are identified and used to control for selection bias. Thanks to a rigorous attendance measure, required by the Louisiana State Department of Education, exposure may be measured. The study's design addresses the demands of policymakers and contributes to evaluation research in a way that can be easily replicated by others.

## LITERATURE REVIEW

In recent years, educators, community leaders, and parents have begun to think of after-school programs as a means to address a wide array of challenges that confront young people and at-risk children in particular (Harvard Family Research Project, 2003a). Research, however, has yet to catch up with their enthusiasm. Empirical evaluation literature so far offers little more than equivocal results to guide funding organizations or policy decisions (Harvard Family Research Project, 2003a; Scott-Little et al., 2002). Two recent surveys of the literature concluded that research is at a rudimentary stage and further specification of research design is necessary before findings will provide clear documentation of program effects (Fashola, 1998; Scott-Little et al., 2002).

Research that evaluates the effectiveness of after-school programs generally falls into one of two categories. Process and implementation evaluations (Acosta-Tello, 1998; Beck, 1999; Fashola, 1999; King, Lipsey, Shayne, & Hoskins, 1998; Lamare, 1997; O'Donnell & Michalak, 1997; Reisner, White, Birmingham, & Welsh, 2001; Scales, Morris, & George, 1998; Torre, 1997; Walker, Grossman, Raley, Fellerath, & Holton, 2000; Warren, Brown, & Freudenberg, 1999) analyze the dynamics of programs in an attempt to provide insight into their operation. Outcome analyses (Hamilton & Klein, 1998; Huang et al., 2000;

Lamare, 1997) examine the impacts of after-school programs to gauge program efficacy.

Both types of analyses are critical in the development of effective after-school programs. Practitioners are attentive to process studies because they provide guidance on the particulars of program implementation and operation. A recent survey of the literature (Scott-Little et al., 2002) contended that the emergent findings from the collection of process evaluations suggests that the relationship between the after-school and regular school programs (Fashola, 1999; Lamare, 1997; Reisner et al., 2001), staff issues and training (O'Donnell & Michalak, 1997; Scales et al., 1998), and transportation (Walker et al., 2000) were the key challenges facing program success.

For policymakers and funding agencies, outcome evaluations are of particular interest. In general, evaluations tend to find positive effects associated with after-school attendance, including the reduction of criminal offenses (Baker, 1998), improving race relations and reducing substance abuse (LoSciuto & Hilbert, 1999), developmental and behavioral issues (Hudley, 1999; Phillips, 1999), and generally positive social and emotional outcomes (Grossman, Price, Fellerath, Jucovy, & Kotloff, 2002; Marshall et al., 1997; Pettit, Laird, Bates, & Dodge, 1997; Posner & Vandell, 1994; see also Carlini-Cotrim & Aparecida de Carvalho, 1993).

Notwithstanding these positive outcomes, policymakers remain narrowly focused on test scores as the singular measure of success. As Huang et al. (2000) stated, "Policymakers reject any evaluation efforts that cannot be directly translated into gains on standardized test scores" (p. 3). Although this zeal for standardized test scores has, indeed, produced research efforts in the area—academic achievement is the most commonly evaluated program outcome (Scott-Little et al., 2002)—the research as a collection is still perceived of as suffering from a pervasive weakness of design (Fashola, 1998; Riggs & Greenberg, 2004; Scott-Little et al., 2002).

By far, the most influential evaluation of academic outcomes has been the Mathematica Policy Research study, in which the researchers found no empirical evidence of positive academic outcomes (U.S. Department of Education, 2003). There has, however, been considerable criticism of the limitations of the evaluation and the danger in generalizing and inferring firm conclusions from it (Bissell et al, 2003; Kane, 2004; Riggs & Greenberg, 2004).

Bissell et al. (2003) contended that the study suffered from problems such as nonequivalency of treatment and control groups, sample size limitations, extremely low levels of dosage, and crossover effects resulting from the diffusion of treatment effect within the control groups. In addition, they also pointed out that the results of the preliminary study were premature because the outcomes studied did not reflect the focus of the programs at the time.

Kane (2004) likewise outlined weaknesses in the evaluation design in his survey of four outcome evaluations. In general, he found that measures were biased in favor of finding no effects. The study defined a participant at extremely low dosage levels. Control for differences in baseline measures (and especially test scores) was poor. Finally, he contended that the expected impact scores are much too high for the type of intervention being evaluated. The common practice of researchers to identify an impact as one-tenth to three-tenths of a standard deviation is unrealistic and arbitrary—especially considering the relative magnitude of the intervention that after-school programming provides.

In addition to these concerns, Riggs and Greenberg (2004) argued that after-school research is equivocal because of wide variation in children and programs studied. In this way, the weakness of the Mathematica Policy Research study can be considered a distillation of the literature as a whole. It is unlikely that every after-school program will have a positive effect or that all different types of children across all levels of attendance will be positively impacted. In their words, "the heterogeneity of American children makes it very unlikely that all children need after-school programming or that there is but one brand of after-school programming suitable for all youth" (Riggs &Greenberg, 2004, p. 177).

Beyond specific criticisms leveled at the Mathematica Policy Research evaluation, there is an emerging consensus on the limitations of earlier research designs and the direction that future research should take. From a policymaking standpoint, research on after-school programming should (a) address the outcome measures with which policymakers are concerned, (b) include key independent variables such as attendance, and (c) establish or control for baseline parity (within the strictures of evaluative research). Finally, because after-school programming is not likely to impact all students, research should (d) examine specific groups of students, not heterogeneous populations.

Policymakers are more responsive to certain outcome measures than others. In the case of 21st CCLCs, policymakers are now obligated to gauge program success on academic outcomes, specifically standardized test score measures. Although test scores alone are a narrow definition of success (Huang et al., 2000), they are considered compelling evidence by policymakers and thus tend to be more influential.

Second, program attendance is a variable worthy of specific attention (Harvard Family Research Project, 2004). Findings indicate that there is at least some evidence that higher dosage of after-school programs predicts higher academic achievement (Huang et al., 2000; Lamare, 1997; Ross, Lewis, Smith, & Sterbin, 1996; see also Evans & Marken, 1984).

Third, establishing baseline parity between treatment and control groups is essential for valid inferences to be made on program effects. Because most research on after-school programs has been quasi-experimental, evaluators need to care-

fully control for any extraneous variation that might impact the dependent variable. Research has not been satisfactory in this regard. Comparisons are made without controlling for baseline measures at all (University of Cincinnati Evaluation Services Center, 1999) or by using mean test scores (Klein and Bolus, 2002), instead of using individual-level data.

Finally, as Riggs and Greenberg (2004) pointed out, the sheer heterogeneity of school-age children and youth makes universal positive academic impact of after-school programs unlikely. The more heterogeneous the population, in other words, the more likely it will result in a Type II error. The disparate findings in the literature as a whole and the null findings in the Mathematica Policy Research study in particular might support this argument.

## PURPOSE OF RESEARCH

In the Spring of 2003, The Policy & Research Group was contracted by the Louisiana Department of Education to evaluate the academic impact of the first cohort of 21st CCLCs awarded by the state. In addition to some process and outcome measures not examined here, the evaluation sought to answer whether the after-school program was having a positive academic impact on participants. The following four research questions form the empirical core of the evaluation:

1. Do participants of 21st CCLC after-school programs demonstrate improved Spring core test scores (and in specific subject areas) when compared to nonparticipants?
2. Do different 21st CCLC grantees demonstrate differential evidence of academic impact on participants?
3. Do particular groupings of participants (i.e., gender, ethnicity, baseline achievement) show differential academic impact from after-school programming?
4. Does intensity of attendance (the number of days a youth attends the program over the school year) impact academic growth?

A confluence of state requirements and research design allowed the research team to construct an evaluation with an eye to the design weaknesses previously discussed. The particular aims of this evaluation project were specifically designed to meet the demands of the Louisiana Department of Education. As such, the outcome measures were of interest to policymakers (test scores) and the target population was necessarily homogenous (at-risk children in Louisiana). Reporting requirements at the state's Department of Education allowed the authors to further

construct a valid and reliable measure of attendance intensity. Finally, a stringent, yet efficient method to control for baseline parity was developed.

## METHODS

Because random assignment of participants was impracticable, this evaluation was designed as a nonequivalent control group design. We employed standardized pretest and posttest measures of treatment and control groups. Standardized tests were given during the Fall and Spring in 1 school year.[1] Ordinary Least Squares (OLS) regression was used for hypothesis testing and to statistically control for select extraneous and intervening variation that lay outside of experimental control.

The key measure of interest in this study is academic impact. Given policymakers' focus on standardized tests (Huang et al., 2000), we chose to operationalize academic impact as the relative performance of each student in an end of year standardized test.

### Dependent Variable

The principal dependent variable in the study, therefore, is the spring standardized test score of each student.[2] To ensure that it was the after-school program (treatment) that created the effect (higher or lower test score), the authors controlled for background variation, including demographic, social, and economic measures that were expected to influence the dependent variable and the fall test score (baseline) of each student.

### Independent Variable of Interest

The key independent variable was the experimental treatment itself—whether or not the student participated in the 21st CCLC after-school program. In accordance with federal guidelines, participants were defined as students who took part in the after-school program for 30 days or more.[3]

---

[1]The Fall test is administered only to program participants and control students. The Spring test is part of a statewide accountability program and, as such, is taken by all students in Grades 3, 5, 7, and 9. The use of the same academic year Fall-to-Spring impact measures is preferable to Spring-to-Fall academic impact measures (Kane, 2004).

[2]As measured by the Iowa Test of Basic Skills (ITBS). The ITBS forms A and B measure skills and achievements in reading, language arts, math, social studies, science, and information sources. The ITBS is used by a number of school districts nationwide to measure academic success. The Core score is the average of a student's performance in reading, language, and math.

[3]The U.S. Department of Education defines a participant as any child who has attended the program for 30 days or more (21st CCLCs).

Unlike earlier research studies that suffered from unreliable attendance data, this study benefited from an accurate daily measure of program attendance. Although various measures of attendance are included to investigate how intensity of attendance impacts academic achievement, the initial categorical division of 30-day attendees should, in contrast with some earlier research, ensure that the after-school treatment is meaningful.

For all grantee programs, nonparticipants were defined as students who attended the same school as participants but did not enroll in or attend the after-school program. In Grant, Baton Rouge, and Bienville parishes, additional students from nonparticipating schools were also included as part of the control group. The schools were selected on the basis of being economically and demographically similar to the experimental schools.

## Quasi-Experimental Design

Although there can be no doubt that true randomized experiments are a superior design, real-world constraints often make true experimentation impractical. Random assignment into treatment and control groups requires oversubscription to services—so that putatively similar subjects chosen at random can provide an experimental counterfactual to the treatment group. It also requires that the researchers have the authority to direct the programmatic decisions of the after-school service providers. In the limited scope of statewide or regional evaluations, especially in regions where the need for safe after-school environments is acute, this is not always possible or ethical.

Such was the case here. The challenge, therefore, was to control for all extraneous variation that might be expected to influence the dependent variable. Although some research has tried to construct a nonparticipant group that approximates the treatment group as closely as possible, this study relied much more heavily on statistical techniques to control for variation. Although this isn't always the best way to equate groups, precedent in after-school research shows that it is extremely difficult to preselect on the wide variety of observable and unobservable variation that will impact academic performance. A comparison group was selected that was expected to be very similar in the likely variables that would affect academic outcomes (classmates). It was reasoned that other variation is best controlled for through baseline measures.

It seems evident, in fact, that there is inherent difference in those who select into after-school programs and those who do not (Kane, 2004). Valid inference does not require that the two groups are identical, only that one controls for variation that may have an effect on the dependent variable. Kane (2004) stated that using statistical controls for the differences of participants and nonparticipants biases the study against finding an impact on after school care. Pretests, however, offer an ef-

fective way to essentially equalize the two groups at baseline for this very narrow definition of academic achievement.

## Key Control Variables

As previously described, this study measures academic impact as the student's Spring Iowa Test of Basic Skills (ITBS) test score, controlling for the Fall ITBS score. By using the pretest scores as a predictor of posttest scores, the study essentially captures (and controls for) the effect of past performance (and indirectly controls for all the variation that might impact academic performance) on the Spring test score. By statistically isolating this effect, a specific and more accurate measure of the academic impact of the 21st CCLC participation may be captured.

This is a pragmatic and efficient approach to controlling for extraneous variation. This technique has been used previously in other education evaluation studies (Mathematica Policy Research, 2004). Surprisingly, it appears to be novel in the evaluation of academic impacts of after-school programs. Previous research has collected and used pretest scores as part of the analysis (Huang et al., 2000; University of Cincinnati, 1999; U.S. Department of Education, 2003), but evaluations have yet to use pretest scores as regressors for the outcome variable. Using lagged values of the dependent variable avoids difficult theoretical and methodological questions of identification of which background characteristics might influence academic achievement, hence is of little interest to some academic researchers. The technique, however, is well suited for outcome evaluations because it arguably allows for the most accurate statistical control for deciphering program impact.

In addition to pretest scores, the study also controlled for other variables that theoretically could influence academic performance, including gender, race/ethnicity, and eligibility for free or reduced-price lunch, all of which were reported on the Spring answer sheets. The impact of 21st CCLC participation on Spring test scores may be measured after controlling for the impact of pretest scores, gender, race, and poverty on Spring test scores.[4]

## Analysis Design

The academic impact of the 21st CCLC after-school program was estimated using OLS regression. Several specifications were used to determine overall program impact and the impact of each of the individual grantee programs. For the impact of all four after-school programs, the key independent variable was a dummy variable of program attendance (0 = *nonparticipant*; 1 = *participant*). Specifically,

[4]In addition to the independent variables of interest discussed later in this article, the following control variables were used: student gender (0 = *boy*, 1 = *girl*); minority status (0 = *White*, 1 = *other*); recipient of free or reduced-priced lunch (0 = *no*, 1 = *yes*).

Spring ITBS test scores were regressed on Fall ITBS test scores, a collection of control variables, and a dummy participation variable. The basic regression equation was as follows:

$$Y_i^{spring} = a + \beta_1 Z_i^{fall} + \beta_2 X_i^{part} + \beta_N X_i^N + \varepsilon_i$$

where:

$Y$ = Spring ITBS test score;

$Z$ = Fall ITBS test score;

$X^{part}$ = dummy variable indicating whether a student attended for 30 days;

$X^N$ = control variables assumed to effect the outcome of $Y$;

$\beta$ = coefficients to be estimated; and

$\varepsilon$ = random error term.

The estimated value of the coefficient $\beta_2$ is an estimate of the program's impact (i.e., the difference in means between the participant group and the nonparticipant group after adjusting for other characteristics).[5]

## Sampling Procedure

The first cohort of state-administered 21st CCLC grants was awarded in the Summer/Fall of 2003. Sixteen grants were made. For this study, 4 grantees were selected to participate: 2 in rural areas (Bienville Parish School System and Grant Parish School System) and 2 in urban areas (Big Buddy in Baton Rouge and University of New Orleans in New Orleans).[6]

Through collaboration with each parish's department of school accountability, 1192 students in Grades 3 and 5 took an ITBS pretest in the Fall of 2003. The Spring ITBS test was taken by all third- and fifth-grade students in Louisiana in March 2004. The ITBS test includes sections on reading, language, math, science, social studies, and sources of information. The Fall 2003 test was used as the pretest and the Spring 2004 as the posttest in this study.

---

[5]To measure the impact of the regional programs, we constructed a model with dichotomous regional variables and interacted them with the participation variable. The coefficient of the interaction term can be interpreted as the academic impact of participants (vs. nonparticipants) in that particular region.

[6]See the Appendix for a qualitative overview of grantee programs.

## FINDINGS

### Sample Population

The students in this study were drawn from four areas of Louisiana: Baton Rouge, New Orleans, Grant Parish, and Bienville Parish. The specific grantees were selected because together (a) they offered a rough cross-section of the children in the state who could possibly benefit from the 21st CCLC program and (b) they represented the cultural and geographic variation of the region (north and south; rural and urban).

Table 1 illustrates some of the background characteristics and composition of the treatment and control groups. As expected, baseline measures for the treatment (participant) and control (nonparticipant) groups are nonequivalent. In line with previous literature (Scott-Little et al., 2002), minority students and recipients of free or reduced-price lunch tend to participate in after-school programs at a higher rate.

The specific findings presented in the sections that follow cannot be inferred to apply generally to all children in the United States or even across Louisiana. The specific purpose of this study was to determine what academic effects the program was having on the federally specified beneficiaries—low-income and at-risk children—in the state of Louisiana. Demographic indicators in Table 1 (80% free or reduced-price lunch recipients and 80% minority) suggest that the study's sample approximates this population quite well. Although it was not the authors' intention at the time, this design offers a version of the specification that Riggs and Greenberg (2004) proposed.

Baseline test scores (Fall ITBS) as reported in Table 2 corroborate the demographic evidence that the study's sample is reflective of the at-risk child population. The mean score for a participant was around 10 points below the national av-

TABLE 1
Background Characteristics of Students in Study

| | *Participants* | | *Non-participants* | | *Total* | |
|---|---|---|---|---|---|---|
| *Characteristic* | *N* | *%* | *N* | *%* | *N* | *%* |
| Female | 146 | 57 | 432 | 46 | 578 | 49 |
| Male | 112 | 43 | 501 | 54 | 613 | 51 |
| Grade 3 | 105 | 40 | 412 | 44 | 517 | 43 |
| Grade 5 | 154 | 60 | 521 | 56 | 675 | 57 |
| Minority* | 205 | 79 | 506 | 54 | 711 | 60 |
| Non-minority | 54 | 21 | 427 | 46 | 481 | 40 |
| Free/Reduced | 207 | 80 | 604 | 65 | 811 | 68 |
| Not F/R | 51 | 20 | 323 | 35 | 374 | 32 |
| Total | 259 | | 933 | | 1,192 | |

*Note*. * Minority = 703 African American, 6 Asian, 1 American Indian, 1 Hispanic.

TABLE 2
Baseline Average Subject Test Scores for Treatment and Control Groups

| *Subject* | *Participant* | *Non-participant* |
|---|---|---|
| Math | 41.58 | 45.69 |
| Language | 42.21 | 47.52 |
| Reading | 38.29 | 46.06 |
| Science | 39.80 | 45.98 |
| Social Studies | 39.16 | 46.31 |

erage across all subject areas. Participants, moreover, had lower academic achievement than nonparticipants across the board. In each of the subject areas, program participants had a lower average score than nonparticipants. Still, nonparticipants as a group were similarly below the national average, which suggests that the control group is similarly at-risk academically, if not entirely equivalent to the participant group. Among the test subjects, reading stands out as the subject with the greatest achievement gap. Nearly eight points separate the average scores of the two groups.

*Research Question 1: Do participants of 21st CCLC after-school program demonstrate improved Spring core test scores (and in specific subject areas) when compared to nonparticipants?* A simple comparison of mean core (reading, math, and language) scores on the pre- and posttests indicates that, in an observable and real way, participants have closed the achievement gap with nonparticipants. Whatever advantage nonparticipants had prior to the after-school program, that advantage had been narrowed by after-school activities. Figure 1 illustrates the change of mean core test scores from Fall to Spring for attendees and nonattendees.

A comparison of mean scores naturally offers poor empirical control over extraneous variation. As is illustrated in the first two columns of Table 3, the results of the regression confirm the simple comparison of means. Results from this first model show that participation in the program has a strong ($b = 2.087$) and statistically significant impact on the Spring academic achievement of students, as measured by standardized test scores. Positive coefficients representative of over 2 normal curve equivalent (NCE) points confirm that program participants did better than nonparticipants, controlling for background variation.

A closer look at some of the inferential statistics corroborates the original argument that pretest scores are an efficient and effective control for extraneous variation. First, the hypothesis testing statistic shows that the pretest is a strongly significant predictor of the posttest score ($t = 55.28$). Second, the strong coefficient of determination indicates that our model accounts for nearly 77% of all variation in the dependent variable ($R^2 = .768$).

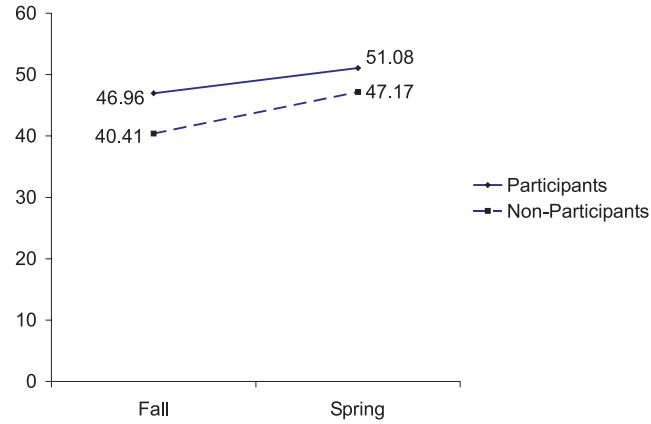


FIGURE 1 Pre and Post test scores for program participants and non participants.

Assuming that the study's controls are adequate and the model is plausible, the hypothesis testing statistic ($t = 3.59$) indicates that one may confidently reject the null hypothesis that program participation has no positive effect on academic achievement. But how much relative improvement is necessary before one may conclude that the impact of the after-school programs was meaningful? To answer this question, the results must be standardized and findings estimated here compared with research reported elsewhere. In short, academic improvement based on NCEs is not directly comparable with other research without standardization.

One common metric that allows for the comparison of results from studies employing different outcome measures is the effect size (ES). An impact score of 2.09 NCEs corresponds to an effect size of 0.13. The average participant, in other words, will perform 13% of a standard deviation better than an average nonparticipant. Although it initially appears modest, an ES of this magnitude is quite significant, especially when one considers that the time invested to achieve these results is limited. Kane (2004) argued that after-school programs, by their nature, should produce less substantial impacts than a change in the full-day classroom instruction. Even with that in mind, the results are comparable to other, more costly, time-consuming interventions, including the Teach for America impact on math scores ($ES = 0.15$) reported by Mathematica Policy Research (2004) and the impact of reducing a class size from 23 to 15 students as reported by Finn and Achilles (1999).

An effect of .13 is meaningful, moreover, because it represents a substantial academic gain for participating students. The typical student who takes part in the after-school program is expected to achieve a learning gain of one and one-third months over a counterpart who does not participate in the program. At the end of the school year, this student has a meaningful academic advantage over a nonparticipant. Furthermore, when examining the after-school program's effect within

TABLE 3
OLS Estimates of Impacts on Standardized Post Test Scores

| | *Model 1* | | *Model 2* | | *Model 3* | |
|---|---|---|---|---|---|---|
| *Independent Variable* | *b* | *t* | *b* | *t* | *b* | *t* |
| Pretest | **0.878** | **55.28** | **0.871** | **54.32** | **0.877** | **52.21** |
| Minority | 0.066 | 0.12 | –0.005 | –0.01 | 0.067 | 0.12 |
| Free/Reduced Lunch | **–1.951** | **–3.47** | **–1.815** | **–3.23** | **–1.969** | **–3.49** |
| Gender | 0.067 | 0.14 | 0.024 | 0.05 | 0.074 | 0.16 |
| Constant | 11.087 | 11.27 | 13.983 | 9.77 | 11.101 | 11.26 |
| Participant | **2.087** | **3.59** | — | — | — | — |
| Bienville | — | — | 21.56 | 1.13 | — | — |
| Baton Rouge | — | — | **1.750** | **2.04** | — | — |
| Grant | — | — | 1344 | 1.27 | — | — |
| Orleans | — | — | **4.708** | **3.28** | — | — |
| Baton Rouge** | — | — | **–2.818** | **–2.86** | — | — |
| Grant | — | — | **–2.607** | **–2.41** | — | — |
| Orleans | — | — | **–4.089** | **–3.55** | — | — |
| Attends 30–59 days | — | — | — | — | **1.781** | **1.98** |
| Attends 60–89 days | — | — | — | — | **2.009** | **2.14** |
| Attends 90 days and Up | — | — | — | — | **2.469** | **2.72** |
| *N* | 1185 | | 1185 | | 1185 | |
| $R^2$ | 0.768 | | 0.771 | | 0.768 | |

*Note***.** Boldface items indicate a level of statistical significance at or below .05 for the coefficient using a two-tailed test. * Standardized partial slope coefficients are included for comparative purposes only. The unstandardized coefficient (*b*) represents the OLS estimated impact of the independent variable. ** Bienville is captured by the intercept.

the context of the standardized test results themselves—keeping in mind the increased importance that these tests play as a metric of school performance—the impact of program participation has real import for principals, teachers, parents, and policymakers. In terms of ITBS scores, the observed effect of the program raises the regression-adjusted performance of the typical student from below the national mean (NCE = 48.9) to above the national mean (NCE = 50.9). In short, even modest program participation is enough to move students from below to above average on a well-respected national scale. The impact therefore, although apparently modest, has some very material and positive implications for the student, school, and school district.

To investigate the impact on academic achievement across the various subjects that make up the core test score, similar regressions were conducted, substituting the core pre- and posttest scores with subject-area scores. Findings are outlined in Table 4.

TABLE 4
OLS Estimates of Program Impact by Subject

| *Subject* | *b* | *t* | *N* | $R^2$ |
|---|---|---|---|---|
| Language | **1.619** | **2.22** | 1244 | 0.670 |
| Reading | **1.478** | **2.00** | 1251 | 0.626 |
| Math | 0.627 | 0.80 | 1233 | 0.547 |
| Social Studies* | **1.875** | **2.03** | 1132 | 0.389 |
| Science* | -0.226 | –0.25 | 1128 | 0.452 |

*Note*. * Not included in computation of core test NCE score. Boldface items indicate a level of statistical significance at or below .05 for the coefficient using a two-tailed test.

Results indicate that academic impact is positive in three of the five subjects tested in the ITBS battery. Participants experienced positive and statistically significant impacts in language, reading, and social studies tests. There was no such impact in either math or science.

*Research Question 2: Do different 21st CCLC grantees demonstrate differential evidence of academic impact on participants?* As outlined, a separate regression equation was constructed that included interaction effects to capture the impact of each program. Results are presented in the second two columns of Table 3.

The results are fairly straightforward; there is, indeed, a difference in academic effects across programs. The three coefficients at the bottom of Model 2 indicate the relative school district or parish-level effects. In this particular model, Bienville Parish is captured by the intercept. Negative coefficients for the remaining three parish/district variables indicate that students in each of these school districts are expected to achieve a lower core NCE score on the ITBS.

Coefficients presented just above the parish-level effects identify the predicted impact—as measured by core NCE scores on the Spring ITBS—of participating in each one of the programs. Participants in the New Orleans and Baton Rouge programs display positive and statistically significant effects. Participants are expected to do 4.7 and 1.75 points better than nonparticipants, respectively, controlling for parish effects, pretests, and demographics. Both Bienville and Grant parish show positive effects, but the impacts are not significantly different from zero—although there is reason to believe that the null result for Bienville is partly attributable to a small treatment group.[7]

[7]In the case of Bienville parish, the chances of finding significant effects appear to be moderated by the small number of participants.

TABLE 5
OLS Estimates of Program Impact by Participant Subgroups

| | *b* | *t* | *N* | *No.* |
|---|---|---|---|---|
| Full Ethnicity/race | 2.087 | **3.59** | 1185 | 259 |
| Minority | **2.488** | **3.58** | 711 | 205 |
| Non-Minority | 0.610 | 0.55 | 481 | 54 |
| Gender | | | | |
| Boy | **2.017** | **2.2** | 613 | 112 |
| Girl | **2.107** | **2.87** | 579 | 147 |
| Initial academic achievement* | | | | |
| Low | 1.218 | 0.91 | 269 | 83 |
| Middle-low | **3.049** | **2.66** | 284 | 75 |
| Middle-high | **3.504** | **3.35** | 257 | 51 |
| High | –0.448 | –0.32 | 288 | 34 |

*Note*. *Initial academic achievement groups have been categorized by quartiles.
Boldface items indicate a level of statistical significance at or below .05 for the coefficient using a two-tailed test.

*Research Question 3: Do particular groupings of participants (i.e., gender, ethnicity, baseline achievement) show differential academic impact from after-school programming?* Our third research question asks whether 21st CCLC programs have varied effects across different groups of students. There are several ways to estimate intragroup variation. In following the design employed by Mathematica Policy Research (2004), we decided to use the basic model and compare each of the different groups by running separate regressions for each subgroup (see also Mathematica, 2001). The method will be to use the participation variable as a measure of effects within the group of interest. In other words, test scores of boys who participated in the program were compared with those of boys who did not. As before, core scores are used as the basis of academic impact because these are the main areas of concern according the NCLB. The results are presented in Table 5.

Overall, looking at the numbers of attendees and impact scores, one sees that positive impacts are broadly shared across most of the children who participated in the program. The middle two rows of the table show that both girls and boys who participated in the after-school program evidenced strong, positive, and statistically significant growth (2.11 and 2.02, respectively). Although girls who attended a program for at least 30 days showed slightly better performance than boys, both did significantly better than nonattendees.

There are some notable exceptions, however. In both categories where there are small numbers of participants (nonminorities and high achievers), there is no statistically significant impact associated with program attendance. The second and third rows in Table 5 show that minority students are the principal beneficiaries of

program effects. Impact scores for minority students are strong ($b$ = 2.49) and strongly significant ($p < .01$). Nonminorities, however, display moderate impact scores (0.61) that are statistically indistinguishable from zero.

A similar disparity exists across various levels of prior academic achievement.[8] Coefficients indicate that significant effects of the 21st CCLC programs were concentrated among the two middle achievement levels (3.049 and 3.504, respectively). Both were positive and strongly significant. High and low achievers apparently did not enjoy a similar increase in performance over nonattendees. Although the estimated impact score is positive, it is much lower than the other two and remains statistically indistinguishable from zero. Although it lies beyond the scope of this study, the discrepancy of effects across initial achievement deserves further empirical analysis.

*Research Question 4: Does intensity of attendance (the number of days a youth attends the program over the school year) impact academic growth?* There are several methods by which the effects of dosage intensity in after-school programming across the four grantees can be modeled. Following previous research (Mathematica, 2001), we decided to compare academic impacts across categorical groupings of participants: those who participated between 30 and 59 days, those who participated between 60 and 89 days, and those who participated for more than 90 days.[9] To do this, the participant variable was divided into three discrete dummy groupings of participant dosage and entered together into the basic regression.[10] Although some information may be lost in the conversion from a continuous variable, this approach has the advantage of presenting results in an easily interpretable way that may also be tested for directionality across levels of attendance.[11] Results are presented in the final two columns of Table 3.

Findings are clear and unambiguous and offer empirical support for the hypothesis that intensity of attendance in after-school programs is associated with improved academic performance (Harvard Family Research Project, 2003a;

---

[8]To gauge program effects on students of varied achievement levels, we divided the entire student sample into equal quarters based on pretest scores. Although the grouping is somewhat arbitrary, it mirrors other education effects research (Decker et al., 2004).

[9]We also regressed attendance as a single continuous variable on the same regressors. Consistent with the results discussed here, the partial slope coefficient was positive and strongly significant (b = 0.029, t = 4.04). Results do appear to be linear. Functionally transforming the continuous attendance variable did not produce a better explanation of variance in the dependent variable.

[10]The method employed here is similar to the technique employed for checking for linearity (and testing for a directional hypothesis) within a single variable. We divided the attendance variable into four discrete groups (control, 30- to 59-day participants, 60- to 89-day participants, and participants who participated for more than 90 days). Each of these dummy variables was included in the basic regression equation. The control group, in this case, is captured by the intercept.

[11]For instance, attendance regressed as a continuous variable may produce a positive slope coefficient, but the effect may be disproportionately produced by medium or high attendees.

Huang et al., 2000; Kane, 2004). As in the first two models, the slope coefficients presented in the third model can be interpreted as the academic impact—in terms of NCE scores—of being a member of any particular group. The results show an increase of academic performance is associated with increased program attendance. Keep in mind, too, that previous performance on test achievement is still being controlled for, so the observed results maybe confidently assumed as effects of program participation.[12] The last three rows in Model 3 show that, on average, higher levels of attendance are associated with increasing academic achievement. The impacts increase as attendance increases and the results become more significant. Further, the positive impacts appear to be quite linear once a student has reached the threshold level of attendance. The results are presented visually in Figure 2.

The model estimates that the initial impact of being a 30- to 59-day participant produces a slightly larger impact than moving up to a 60-day participant. Initial participation, in other words, produces the biggest incremental impact, but increasing the intensity of attendance appears to continually improve academic performance in a way that is consistent with the linear model of attendance intensity (Harvard Family Research Project, 2004). The results become more certain, moreover, as is illustrated by the final column of Model 3.

## DISCUSSION

The data gathered for this evaluation unequivocally indicate that the 21st CCLCs examined within the state of Louisiana are having a positive academic impact on at-risk students who attend the program for at least 30 days. Although the core test score is a broad indicator of academic achievement, program-wide individual subject data indicate that program impacts are positive and significant in reading, language, and social studies. Participants also share academic benefits broadly, though minority students and moderate achievers do appear to exhibit a more robust growth. Finally, academic outcomes conform more closely to the linear model of attendance (Harvard Family Research Project, 2004).

From this straightforward summation of the original research questions, it is worthwhile to elaborate on a couple of points. First, as can be inferred from the aforementioned analysis, attendance should be considered a key methodological consid-

[12]It is possible, of course, that the effect observed here may be due to other influence, such as increased attendance in school. But as Huang et al. (2000) pointed out, attendance in school may be endogenously related to attendance in the program. Causality is complex in this regard and deserves further investigation. Because previous performance was controlled for, however, one may be relatively confident that the impact scores here are not due to a corresponding achievement effect that might accompany higher attendance.

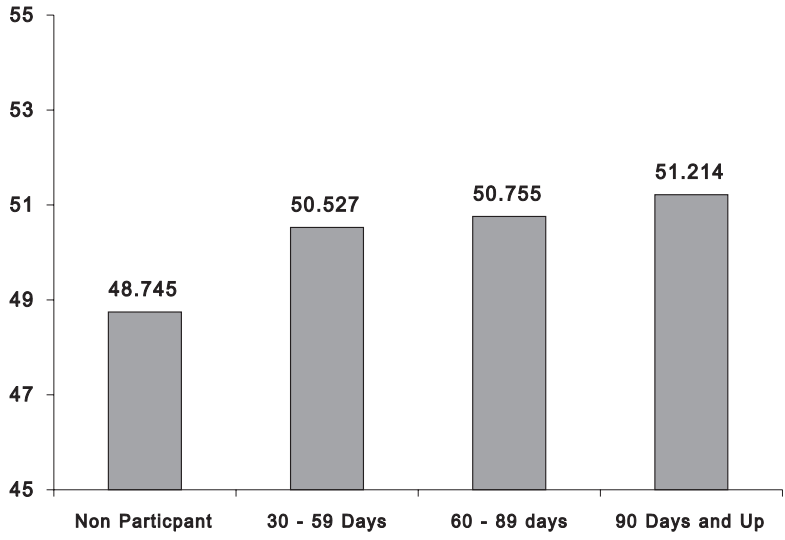


FIGURE 2 Regression-adjusted posttest scores by program attendance. Test scores are calculated for a male minority student who is eligible for free or reduced-price lunch and who achieved an average score of 45.093 on the pretest.

eration for academic outcome evaluation. Intensity of attendance has a real and observable influence on the academic achievement of at-risk children in Louisiana.

Quite simply, the data gathered indicate that the more a student attends 21st CCLC after-school activities, the more academic benefit he or she will reap. The idea that students benefit from increased attendance in after-school programming is not new (Harvard Family Research Project, 2004), but the results examined here offer strong empirical evidence that program attendance does positively impact the academic performance of at-risk children.

Positive and additive effects of a modest intervention such as this can, therefore, have substantial repercussions on students who have the opportunity to attend, especially if they can attend for an extended period of time. Consider, for instance, a typical hypothetical student in this study: a minority boy who is eligible for free or reduced-price lunch. If this student, Alex, attends after-school activities for 30 days or more, he could be expected to end the year with the equivalent of 1 month's worth of learning gain over and above what he would achieve without program participation. With 90 days of after-school programming, Alex would finish with the equivalent of 1-1/2 months' worth of gain. If one assumes that the effects are additive and would continue in a linear fashion similar to that observed after the initial program impact, one would expect that, after 3 years of moderate attendance

in the program (equal to 90 days per year), Alex would improve his academic capacity by half a grade equivalent.

If reading skills in particular are considered, the long-term effects of the after-school program are even more pronounced. If Alex attends the program for 30 days or more in an academic year, he will end that year with nearly 1 month's worth of added learning gain. If he attends for 90 days, his gain would be 2 months. Again, if the effects are linear over time, 3 years of moderate program attendance (90 days+) would mean that Alex would be reading at a level that is the equivalent of more than half of a grade better than if he did not attend the program.

At this point, multiyear effects are speculative. Future research should empirically examine this question. If, however, program effects are additive, as this study's findings suggest that they are, the cumulative academic impact of after-school programming for at-risk students appears to be considerable. At-risk students like Alex who attend after-school programs through elementary and middle school would have a substantial academic advantage over similarly at-risk children who do not.

The measures addressed in this evaluation report present an admittedly narrow gauge (Huang et al., 2000) of program outcomes. Standardized test scores are, however, of particular interest to policymakers, and in this case offer an unambiguous picture of the academic benefits of 21st CCLCs for a specific group of children.

Earlier research that found limited evidence of academic impact (U.S. Department of Education, 2003) has been questioned for defining participation at excessively low or sporadic levels of attendance (Kane, 2004). Empirical support for that hypothesis has been found.

From a programmatic standpoint, this underscores the importance of designing an approach that remains attractive to parents and students. The results of the student focus groups indicate that it is the enrichment and recreation activities that keep students coming. Instruction, tutoring, and other academic curricula will be influential only if the students continue to attend.

In particular, we have observed that programs could significantly boost academic outcomes and have a more broad impact on at-risk children by trying to improve the intensity of attendance of students who are attending the program for limited or sporadic periods of time. Data show that participants are more at risk academically than nonparticipants and would benefit from more exposure to the program. These students and their parents have already shown initial interest in the program. Improving their attendance should be a goal. Future evaluations could target low or sporadic attendees with instruments to determine what barriers are inhibiting improved attendance.

Finally, the imbalance of impacts across subjects may indicate that outcomes are the result of increased time spent in an enriched environment in addition to the academic curriculum. The feedback gathered during teacher and student focus groups would support this hypothesis. Although it is likely that the academic instruction helped the children to improve their test scores, it is also likely that (a) the

exposure to new ideas, (b) increased time engaged in reading or being read to, (c) enhanced vocabulary, and (d) interaction with adults contributed to the greater success of participants. Stronger gains seen in reading and language and lesser gains in math and science at least suggest this proposition.

This study has several limitations. We could not employ random assignment to treatment conditions; every effort was made, however, to statistically control for possible extraneous variation. Treatment and control groups were matched according to school-level, not individual-level, characteristics. Descriptive statistics previously discussed, however, indicate that the groups were similar in terms of social, economic, and educational status. Indeed, the treatment group compares unfavorably to the control group in terms of social and economic status and baseline educational achievement. Finally, the programs themselves were not randomly selected from all available state programs. As such, this study does not allow the authors to broadly generalize program effects within Louisiana, but rather serves as an indication of what impacts can result from participation in after-school programs.

Past research and national trends in reading and math test scores indicate that students should be more immediately receptive to math skills instruction than that of the language arts. Kane (2004) stated that:

> Evidence suggests that reading test scores may be slower to respond to instructional interventions than math scores. Reading is more likely to be driven by family background—the extent to which parents have read to their children outside of school—than math test scores, which depend more upon the specific topics the student has covered in school. (p. 20)

The strength of 21st CCLC programming, viewed in this light, may be that its influence transcends the structured and instructional interventions provided during the regular school day. After-school programs may provide academic benefits for students in much the same way that increased parental attention does.

Previous evaluations on broad and heterogeneous student populations have found little support for the hypothesis that students can benefit academically from after-school programming. This evaluation focused on a comparatively narrower group of students and found, to the contrary, strong evidence that after-school programs are having a positive, statistically significant impact on the academic performance of participating students. Controlling for a range of socioeconomic, demographic, and past academic performance variables, at-risk children from four rural and urban Louisiana parishes who participate in 21st CCLC programs are academically outperforming their peers who do not attend.

The principal policy objective of 21st CCLCs, as described in the NCLB Act, is to "provide opportunities for academic enrichment, including providing tutorial services to help students, particularly students who attend low-performing schools, to meet state and local student academic achievement standards in core academic subjects, such as reading and mathematics." This study finds ample evidence to suggest

that this objective is being achieved across the four programs examined here. This study cannot be generalized beyond the specific 21st CCLCs examined in this study, given its quasi-experimental design; however, we can say with confidence that, following the admittedly tough standards of academic outcomes, the results collected from four diverse programs in Louisiana show that students who participate in after-school programs are academically at-risk students according to descriptive statistics who show strong and statistically significant academic growth in core subject areas as a direct result of attending these programs.

## ACKNOWLEDGMENTS

The Louisiana Department of Education, Division of Student Standards and Assessments, provided the data for analysis. We gratefully acknowledge the assistance provided by Fen Chou at the Division of Standards and Assessments and Pamela Wall at the Department of Education's Division of School and Community Support. The Youth Services database developed by CitySpan Technologies was instrumental in our ability to use intensity of attendance as a key independent variable. Mark Min at CitySpan Technologies helped to provide attendance data. We also thank an anonymous reviewer for useful and constructive suggestions that helped to improve the manuscript.

## REFERENCES

Acosta-Tello, E. (1998). *A study of quality dimensions within selected school-age care programs in southern California*. Unpublished doctoral dissertation, University of California, Irvine and University of California, Los Angeles, CA.

Barker, N. C. (1998). Can specialized after-school programs impact delinquent behavior among African American youth? Child welfare and juvenile justice. In Liberton, C., Kutash, K., & Friedman, R. (Eds.), *Proceedings of the annual research conference of A System of Care for Children's Mental Health: Expanding the Research Base, Tampa, FL, 10* (pp. 345–350). Tampa, FL: The Research and Training Center for Children's Mental Health. Retrieved January 19, 2007, from http://www.fmhi.usf.edu/institute/pubs/pdf/cfs/rtc/10thproceedings/10thproctoc.html.

Beck, E. L. (1999). Prevention and intervention programming: Lessons learned from an after-school program. *The Urban Review, 31,* 107–124.

Bissell, J. S., Cross, C. T., Mapp, K., Reisner, E., Vandell, D. L., Warren, C., et al. (2003, May). Statement released by members of the scientific advisory board for the 21st Century Community Learning Center evaluation. Retrieved on December 25, 2006, from http://www.gse.harvard.edu/hfrp/content/projects/afterschool/resources/21stcclc_statement.doc

Carlini-Cotrim, B., & Aparecida de Carvalho, V. (1993). Extracurricular activities: Are they an effective strategy against drug consumption? *Journal of Drug Education, 23*, 310–314.

Evans, E., & Marken, D. (1984, April). *Longitudinal follow-up comparison of conventional and extended day public school kindergarten programs.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

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

Fashola, O. S. (1999). *The child first authority after-school program: A descriptive evaluation* (Report No. 38). Baltimore, MD: Center for Research on the Education of Students Placed at Risk, Johns Hopkins University.

Fight Crime: Invest in Kids. (1997). *After-school crime or after-school programs: Tuning in to the prime time for violent juvenile crime and implications for national policy.* Washington, DC: Author.

Finn, J. D., & Achilles, C. M. (1999). Tennessee's class size study: Findings, implications, misconceptions. *Educational Evaluation and Policy Analysis, 21*(2), 97–109.

Gaynor, F. K., & Horowitz, J. E. (1998). *Evaluation of the 4-H after-school activity program 1997-1998*. San Francisco: WestEd.

Grossman, J. B., Price, M. L., Fellerath, V., Jucovy, L. Z., & Kotloff, L. J. (2002). *Findings from the extended-services schools initiative*. Philadelphia: Public Private Ventures.

Hamilton, L. S., & Klein S. P. (1998). *Achievement test score gains among participants in the foundations school-age enrichment program* (RAND Project Memorandum, PM–858–EDU). Santa Monica, CA: RAND Corporation.

Harvard Family Research Project. (2003a). *A review of out-of-school time program quasi-experimental and experimental evaluation results* (Evaluation Snapshot No. 1). Cambridge, MA: Author.

Harvard Family Research Project. (2003b). *Why, when and how to use evaluation* (Brief No. 5). Cambridge, MA: Author.

Harvard Family Research Project. (2004). *Understanding and measuring attendance in out-of-school time programs* (Brief No. 7). Cambridge, MA: Author.

Huang, D., Gribbons, B., Kim, K. S., Lee, C., & Baker, E. L. (2000). *A decade of results: The impact of the LA's Best After School Enrichment Initiative on subsequent student achievement and performance.* Los Angeles, CA: University of California at Los Angeles, Graduate School of Education & Information Studies, Center for the Study of Evaluation.

Hudley, C. (1999, April). *Problem behaviors in middle childhood: Understanding risk status and protective factors*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec.

Kane, T. J. (2004). *The impact of after-school programs: Interpreting the results of four recent evaluations* (Working paper). New York, NY: William T. Grant Foundation. Retrieved January 19, 2007, from http://www.wtgrantfoundation.org/usr_doc/After-school_paper.pdf

King, R. D., Lipsey, M. W., Shayne, M. W., & Hoskins, A. (1998). *Final report on a formative evaluation of the first year of the Project for Neighborhood Aftercare (PNA): A school-based after-school program.* Nashville, TN: Vanderbilt Institute for Public Policy Studies.

Klein, S. P., & Bolus, R. (2002). *Improvements in math and reading scores of students who did and did not participate in the foundations after school enrichment program during the 2001–2002 school year.* Santa Monica, CA: Gansk & Associates.

Lamare, J. (1997). *Sacramento Start: An evaluation report.* Sacramento, CA: Sacramento California Neighborhoods Planning and Development Services Department.

LoSciuto, L., & Hilbert, S. (1999). A two-year evaluation of the Woodstock Youth Development Project. *Journal of Early Adolescence, 19*, 488–518.

Marshall, N. L., Garcia-Coll, C., Marx, F., McCartney, K., Keefe, N., & Ruh, J. (1997). After-school time and children's behavioral adjustment. *Merrill-Palmer Quarterly, 43*, 497–514.

Mathematica Policy Research. (2004). *The effects of Teach for America on students: Findings from a national evaluation*. Princeton, NJ: Author.

Mathematica Policy Research. (2001). *A broader view: The national evaluation of the 21st century community learning centers program* (Vol. I). Princeton, NJ: Author.

O'Donnell, J., & Michalak, E. A. (1997). Inner-city youths helping children: After-school programs to promote bonding and reduce risk. *Social Work in Education, 19*, 231–242.

Pettit, G. S., Laird, R. D., Bates, J. E., & Dodge, K. A. (1997). Patterns of after-school care in middle childhood: Risk factors and developmental outcomes. *Merrill-Palmer Quarterly, 43*, 515–538.

Phillips, R. S. (1999). Intervention with siblings of children with developmental disabilities from economically disadvantaged families. *Families in Society, 80*, 569–577.

Posner, J. K., & Vandell, D. L. (1994). Low-income children's after-school care: Are there beneficial effects of after-school programs? *Child Development, 65*, 440–456.

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *The Journal of Democracy, 6*(1), 65–78.

Reisner, E. R., White, R. N., Birmingham, J., & Welsh, M. (2001). *Building quality and supporting expansion of after school projects: Evaluation results from the TASC after-school program's second year.* Washington, DC: Policy Studies Associates.

Riggs, N. R., & Greenberg, M. T. (2004). After-school youth development programs: A developmental-ecological model of current research. *Clinical Child and Family Psychology Review, 7*(3), 177–190.

Ross, S. M., Lewis, T., Smith, L., & Sterbin, A. (1996). *Evaluation of the extended-day tutoring program in Memphis City Schools: Final report to CRESPAR*. Unpublished manuscript, University of Memphis, TN.

Scales, A. M., Morris, G. A., & George, A. W. (1998). A church operated after-school tutorial and enrichment program. *The Negro Educational Review, 49*, 153–164.

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

Torre, A. A. N. (1997). Influencing Latino education. *Education and Urban Society, 30*, 20–40.

U.S. Department of Education. (2003). *When schools stay open late: The national evaluation of the 21st-Century Community Learning Centers program*. Washington, DC: U.S. Government Printing Office.

U.S. Department of Education & U.S. Department of Justice. (1998). *Safe and smart: Making after-school hours work for kids*. Retrieved on January 19, 2007, from http://www.ncjrs.gov/pdffiles1/nij/grants/179991.pdf

U.S. Department of Justice. (1999). *Juvenile justice bulletin: Violence after school*. Retrieved January 19, 2007, from http://www.ncjrs.gov/html/ojjdp/9911_1/contents.html

University of Cincinnati Evaluation Services Center, College of Education. (1999). *1998–99 school-year program evaluation Urban School Initiative School Age Child Care Expansion*. Cincinnati, OH: Author.

Walker, K. E., Grossman, J. B., Raley, R., Fellerath, V., & Holton, G. I. (2000). *Extended service schools: Putting programming in place*. Philadelphia: Public/Private Ventures and Manpower Demonstration Research Corporation.

Warren, C., Brown, P., & Freudenberg, N. (1999). *Evaluation of the New York City beacons: Phase I findings*. New York: Academy for Educational Development.

APPENDIX
Overview of Programs

| | |
|---|---|
| *Hours of operation* | Days per week: Monday–Thursday or Monday–Friday<br>Hr per day: Programs start between 2:30 and 3:15 and end between 5:00 and 5:30 during the school year. |
| *Academic hours* | 3–4 hours per week |
| *Academic staff* | Primarily day teachers at the school. Some college students and volunteers. |
| *Overview of academic approach* | Varies by program—Homework assistance, group work, direct instruction, test preparation. Some programs use off-the-shelf curriculum and some use a customized curriculum based on state standards. |
| *Subject area focus* | Varies by program–Science, computers, language arts, math, social studies. |
| *Academic requirement* | Students are required to attend the academic component to participate in the program. |
| *Lesson plans* | Varies by program–Some require instructors to prepare lesson plans, others do not. |
| *Unique aspect of programs* | Big Buddy–A retired principal is used as an educational consultant for the program to ensure that lesson plans meet benchmarks and standards and to direct and monitor the quality of the after-school instruction.<br>Bienville–Remediation for standardized testing is done in conjunction with 21st CCLC so that each program can complement the other.<br>UNO/New Orleans–Extensive use of volunteers and other agencies even for academic component. For example, 2nd grade students who are below reading level are paired one-on-one with volunteers.<br>Grant Parish – *Highly skilled educators* outline what is being taught during regular school hours each week to inform after-school instruction. |
| *Nonacademic* | Varies by program–Includes sports, pep squad, tumbling, ball games, sewing, photography, cheerleading, Girl Scouts, stomp, choir, jump rope, dance, drama. |