Huang, Denise; Leon, Seth; La Torre Matrundola, Deborah

Exploring the relationships between LA's BEST program attendance and cognitive gains of LA's BEST students

Journal for educational research online 6 (2014) 3, S. 34-53

Journal for Educational Research Online
Journal für Bildungsforschung Online
Volume 6 (2014), No. 3, 34–53

Denise Huang, Seth Leon & Deborah La Torre Matrundola

Exploring the relationships between LA's BEST program attendance and cognitive gains of LA's BEST students

Abstract

The purpose of this study was to examine the impact of participation in the Los Angeles Better Educated Students for Tomorrow (LA's BEST) after-school program on positive achievement outcomes in math and English-language arts. A quasi-experimental design was utilized, and hierarchical linear modeling (HLM) was employed to examine the relations between intensity of program participation and achievement outcomes across four years of data using two cohorts of students. Results revealed that regular attendance (over 100 days per year) in the LA's BEST after-school program led to higher achievement in California Standards Test (CST) math performance, but not in CST English-language arts performance. Therefore, LA's BEST can improve their program outcomes in math by setting program structures, activities, and policies to encourage all students to attend regularly.

Keywords

After-school programs; Attendance dosage; Academic outcomes; Cognitive gains; Quasi-experimental

Untersuchung der Beziehung zwischen der Teilnahme am LA's BEST-Programm und dem kognitiven Gewinn von LA's BEST-Schülern

Zusammenfassung

Die Absicht dieser Studie war es, das Ausmaß zu erforschen, in dem die Teilnahme an dem Nachmittagsförderprogramm „Los Angeles Better Educated

Dr. Denise Huang (corresponding author) · Seth Leon, B.S. · Deborah La Torre Matrundola, M.A., National Center for Research on Evaluation, Standards, and Student Testing (CRESST); University of California, Los Angeles, 300 Charles E. Young Drive North, 91604, Los Angeles, United States
e-mail: dhuang@cse.ucla.edu
leon@cse.ucla.edu
latorre@cse.ucla.edu

Students for Tomorrow" (LA's BEST) zu Leistungssteigerungen in den Fächern Mathematik und Englisch geführt hat. Dabei wurden ein quasi-experimentelles Studiendesign und Hierarchical Linear Modeling (HLM) genutzt, um den Zusammenhang zwischen der Intensität der Teilnahme am Programm und der Leistungen der Schüler, über einen Zeitraum von vier Jahren und unter Einbezug von zwei Schülerkohorten, zu untersuchen. Die Ergebnisse zeigten, dass eine regelmäßige Teilnahme am LA's BEST-Programm (über 100 Tage pro Jahr) beim California Standards Test zu besseren Ergebnissen in Mathematik, nicht aber in Englisch führte. Daher kann LA's BEST die Ergebnisse seines Programms im Fach Mathematik dadurch verbessern, dass es durch die Struktur des Programms, angebotene Aktivitäten und Leitsätze alle SchülerInnen dazu anregt, regelmäßig teilzunehmen.

Schlagworte
Nachmittagsförderprogramm; Teilnahmehäufigkeit; Akademische Resultate; Kognitive Gewinne; Quasi-experimentell

1. Introduction

The number of after-school programs in the United States has grown tremendously in the past few decades. Between 1994 and 2000, the number of public schools that offered after-school programs doubled (De Kanter, 2001). By 2009 the After-school Alliance estimated that about 8.4 million school age children (or 15 %) in the United States were participating in after-school programs. On average, students in these programs participate three days per week for an average of eight hours, and 41 % of the 8.4 million children in these programs qualify for free or reduced lunch-an indicator of low family income (After-school Alliance, 2009). As indicated, after-school programs are serving a high need population, and the need keeps on expanding. Furthermore, 73 % of children who participate regularly in the 21st Century Community Learning Center (21st CCLC) programms receive free or reduced lunch (After-school Alliance, 2014).

The 21st Century Community Learning Centers (21st CCLC) initiative is the only federal funding source dedicated exclusively to after-school programs. Its appropriation has increased from 40 million in 1998 to over a billion per year since 2010 (After-school Alliance, 2014). The purpose of the 21st CCLC initiative is to provide support of the creation of community learning centers that provide academic enrichment opportunities during non-school hours for students, particularly those who attend high-poverty and low-performing schools. The program intends to help students meet state and local standards in core academic subjects, such as reading and math. In addition, learning centers provide non-academic enrichment activities to complement the school curriculum as well as literacy and oth-

er educational services to the families of participating children (U.S. Department of Education, 2014).

2. Purpose of the study

When reviewing research on the academic benefit of after-school participation, results are mixed. Bergin, Hudson, Chryst, and Resetar (1992) found positive associations between after-school participation and higher achievement scores. Their study followed a group of kindergartners who attended an after-school program and compared them to a control group. Initially, the standardized test scores of both groups were below national average. However, by the spring of first grade, the treatment group was outperforming the control group and was performing above national norms. After-school participation is also associated with higher classroom grades, higher math and reading scores, increased day school attendance, lower dropout rates, higher homework completion rates, and higher graduation rates (Goerge, Cusick, Wasserman, & Gladden, 2007; Grossman et al., 2002; Riggs, 2006; Riggs & Greenberg, 2004; Vandell, Reisner, & Pierce, 2007).

Other studies have reported mixed, insignificant or even negative outcomes regarding issues such as academic performance, school retention, feelings of safety, and behavior (Cooper, Charlton, Valentine, & Muhlenbruck, 2000; Dynarski et al., 2004; James, 1997; Vanderhaar & Muñoz, 2006). In particular, two large-scale national evaluations of 21st CCLC after-school programs have generated controversies. These evaluations of elementary school (James-Burdumy, Dynarski, Moore, Deke, & Mansfield, 2005) and middle school students (Dynarski et al., 2004) did not find any significant gains on achievement test scores. These evaluation findings concerning the 21st CCLC led some to suggest drastic reductions in levels of federal support for after-school programs (Mahoney & Zigler, 2006). Furthermore, others have suggested that this was the result of methodological shortcomings in the studies (Kane, 2004; Mahoney & Zigler, 2006).

After-school studies face a number of common challenges: there are wide variations of goals across programs; it is difficult to obtain valid control groups as well as clean records of data; there is an inherent potential of selection bias in the after-school population; and there are high transience rates for participating students. In particular, after-school studies have been challenged by the fact that programs have failed to measure and evaluate the dosage of participation that students receive (Lauer et al., 2003). For any intervention project, it is necessary for subjects to receive adequate treatment in order to demonstrate effects. Furthermore, it was not until recently that after-school studies have begun to examine the importance of dosage. This is due, in part, to the fact that the growth of after-school programs has been so rapid that efficient data management and recordkeeping, such as student attendance, have only recently become productive, making the examination of dosage difficult. Thus, the purpose of this study is to fill this research gap by com-

paring students with different dosage levels and to examine the students' achievement trends over a period of four years.

The research questions for this study are as follows:

- Does participation in the LA's BEST after-school program lead to positive achievement outcomes in math and English-language arts?
- Do the achievement outcomes of LA's BEST students' vary as a function of their different intensity levels of LA's BEST participation?

3. The importance of dosage

Dosage (defined as intensity of participation) is a critical factor to examine when assessing the effect of an intervention. In medical research, dosage refers to the quantity of a drug or other agent administered for therapeutic purposes. Factors taken into account in these studies include the strength, duration, and/or application of the dose of treatment (e.g., all at once or gradually over time) necessary to impact an organism biologically. The dosage, concentration, and division over time may all be critical considerations in the administering of drugs (Pinkel, 1958). Within the medical field, recognizing the importance of dosage has made a major contribution in the ability to determine variability in the effectiveness of treatment (Modi & Keay, 1983). Since afterschool programs are also considered as interventions – for example to reduce delinquency or improve academic performance – dosage ought to be considered as a significant factor influencing outcomes.

Despite this, attendance is generally examined as the days, weeks, or hours that students spend in activities (Fiester & Policy Studies Associates, 2004). When measuring dosage, attendance can be further separated intensity (frequency of attendance during one program year), duration (years of attendance), and total exposure (frequency of attendance over multiple years). The current study focuses on the effects of these indicators.

In general, studies that examine dosage have found a positive relationship between intensity of participation and positive student outcomes (Frankel & Daley, 2007; Lauer et al., 2003; McComb & Scott-Little, 2003). For instance, Frankel and Daley (2007) found that higher after-school participation is associated with higher academic achievement, while Goldschmidt, Huang, and Chinen (2007) found that medium (10–14 days per month) and high (15 or more days per month) participation in an after-school program is associated with lower juvenile crime rate.

Furthermore, multiple studies have found a relationship between intensity of after-school participation and day school attendance (Frankel & Daley, 2007; Huang, Gribbons, Kim, Lee, & Baker, 2000; Munoz, 2002; Welsh, Russell, Williams, Reisner, & White, 2002). More specifically, in 2007, Frankel and Daley (2007) released a report that found an association between high dosage of after-school participation and higher math assessment scores, English-language arts assessment scores, and day school attendance. They created four attendance level categories:

(a) 1–20 days, (b) 21–50 days, (c) 51–100 days, and (d) more than 100 days per year. The study found that, in order to benefit academically, the elementary school students needed to attend the after-school program for at least 100 days per year, and middle school students needed to attend at least 50 days annually. Similar findings have also been reported by Munoz (2002) and Jenner and Jenner (2007).

Intensity of after-school participation can also predict social outcomes. Goldschmidt and colleagues (2007) examined the long-term effectiveness of participation in after-school programs in lowering juvenile crime rates. They found that students who consistently participated in the LA's BEST after-school programs demonstrated a substantive significant reduction in juvenile crime as compared to students with inconsistent attendance and no attendance.

Finally, in reviewing research on participation and outcomes in after-school programs, McComb and Scott-Little (2003) concluded that students who participate in after-school programs more frequently and for longer periods benefit the most. They suggest that after-school programs should be an integral part of schools' academic and developmental programs. They stated that in all cases where data was examined using *intensity* level (or dosage), results favored students who had participated at higher levels.

4. Reducing selection bias

Another frequent critique of after-school studies is selection bias (Fashola, 2002; Hollister, 2003; Little & Harris, 2003; Scott-Little, Hamann, & Jurs, 2002). Studies commonly compare students who participated in the program to those who did not, with the implication being that any differences are caused by the program. Self-selection bias occurs when certain unaccounted for characteristics, such as whether family environment impacts student motivation and/or participation, have the potential to influence apparent effects on academic outcomes. In other words, since after-school program participation is voluntary, students (or their parents) self-select themselves into participation and non-participation groups. Thus, in comparing participating students to non-participating students in the same school, there are inherent biases that researchers need to balance or control statistically in order for the findings to be valid.

For example, when researchers are studying low performing schools with an intervention program, using "average students" across the country as the comparison group may be overly unfavorable to the program since students from better off schools tend to score higher academically. However, using non-participants of the program as the comparison may be overly favorable to the program because the participants may be more motivated to do better academically since they elected to participate in the program.

In 2004, Dimsdale and Kutner (2004) stated that the *gold standard* in after-school research is for a study to meet all the standards of scientifically based

research as called for in the No Child Left Behind Act (NCLB; 2001). This includes the use of experimental designs that involve randomization and control groups. Randomization is believed to be the single most reliable indicator that the study's finding can be interpreted without the fear of selection bias (Slavin, 2002). However, as mentioned above, due to the social context of after-school programs, reaching the gold standard of research is often difficult. In reality, it is often challenging, and potentially unethical, for most after-school programs to randomize their participants unless the programs are oversubscribed (so that the programs will not intentionally refuse to service the population in need just so that students can serve as controls). The students who are refused enrollment may end up being unsupervised and without the homework help that they desperately need. Consequently, many after-school studies lack either a true experimental control or a valid comparison group. Thus, most studies in this field are quasi-experimental, with researchers using a comparison group and making use of statistical controls. In these quasi-experimental studies, one needs to be cautious when inferring causality. Because of this, it is important for researchers to demonstrate that the comparison group shares most of the same characteristics as the participant or intervention group other than participation in the program. A high degree of similarity between the two student groups at the beginning of the program is critical to the study design.

In summary, dosage and self-selection bias are two major criticisms on after-school studies. The measures necessary to determine the extent of students' participation in after-school programs are often not collected (Fiester & Policy Studies Associate, 2004). Most studies just compare the outcomes of after-school participants to non-participants and do not consider the attributions of participation levels to the outcomes. Bodilly and Beckett (2005) explained that one reason might be the overshadowing of selection bias on participation levels. Students who attended after-school programs more often or for longer differed from both non-participants and those participating at lower levels in observable and non-observable ways. This study intends to help address these two research gaps by taking selection bias and dosage into consideration within the study design.

The study examines the LA's BEST after-school program to investigate the impact of differential intensity of participation on achievement and uses a comparison group of students who had some exposure to the program at baseline to reduce self-selection bias. Sample size, adjustments for pre-test differences, duration of program participation, and the use of standardize scores as the unbiased outcome measure were all carefully considered. First, a brief description of the LA's BEST program is provided.

5. The LA's BEST program

LA's BEST was first implemented in the fall of 1988. The program operates under the auspices of the Mayor of Los Angeles, the Superintendent of the Los Angeles Unified School District (LAUSD), a board of directors, and an advisory board consisting of leaders from business, labor, government, education, and the community.

LA's BEST seeks to provide a safe haven for at-risk students in low-income neighborhoods where gang violence, drugs, and other types of anti-social behaviors are common. The program is housed at selected LAUSD elementary schools and is designed for students in kindergarten through fifth or sixth grade, depending upon the school. The LA's BEST sites are chosen based on certain criteria, such as low academic performance and their location in low-income, high-crime neighborhoods.

LA's BEST is a free program open to all students in the selected sites on a first-come, first-served basis. Students who sign up for the program are expected to attend five days a week in order to reap the full benefits of the program. In an effort to emphasize the development of the whole child (Hodgkinson, 2006; Schaps, 2006) daily program offerings at these sites include academic assistance, enrichment, and physical activities.

Currently, LA's BEST serves a student population of approximately 28,000 with about 60 % Hispanic and about 10 % African American elementary students. English learners comprise at least half of the student population at most sites. Of this population, the majority's primary language is Spanish, while the other percentage of the English Learner population is composed of those whose first language is of Asian/Pacific origin. The overall demographic characteristics of the LA's BEST participants are very similar to the demographics of the 21st CCLC program participants.

6. Study design and methods

This study employed a quasi-experimental design that consisted of a longitudinal sample of both academic and LA's BEST program attendance. The original sample was comprised of two cohorts of second grade students who attended schools hosting an LA's BEST site (2002–03 and 2003–04). The students in each cohort were then separated into four attendance categories based on their dosage (intensity of attendance) in the program. Finally, hierarchical linear modeling (HLM) was applied to examine academic outcomes.

This methodology enabled the study to utilize the longitudinal nature of the data and follow students' academic development over time. By utilizing this structure, the study was able to move beyond traditional pre/post analysis, which is limited by data requirements and explanatory possibilities (Rogosa, Brandt, & Zimowski, 1982; Raudenbush & Byrk, 2002). This study not only accounted for

student level variation, but also for variation across schools that may have influenced student achievement outcomes.

6.1 Defining the study sample

The basis for this study sample is the LAUSD student database that the research team has collected and stored since the 1992–93 school year. The first step in constructing a study sample is to generate a sampling frame. This task was accomplished by going back through the historical records and tracking four-years of background and California Standards Test (CST) achievement data for the students in the two cohorts. Students who were in second grade during the 2002–03 and 2003–04 school years who attended the same school through their projected fifth grade year (2005–06 and 2006–07) were included in these cohorts. It should be noted that the second cohort was larger than the first because the after-school program increased from 85 to 133 sites during the 2003–04 school year.

Examination of attendance patterns indicates that students participate in after-school programs with varying regularity. Furthermore, students who voluntarily enroll in after-school programs and those who do not may be very different in some of their background characteristics that are not directly observable. For example, families or students may elect not to participate for a variety of reasons such as a lack of interest in the program's activities, participation in other after-school activities, or the availability of an adult at home to supervise and provide help with homework. These variations make interpretations for study findings somewhat complex. Therefore, the study was designed to take into account whether students changed their enrollment in the after-school program between baseline and follow up when defining the student groups.

Within each cohort, the students who participated in the after-school program over 20 days at baseline but did not participate during the follow-up period were classified as Group 1. The students who had no participation in the after-school program at baseline, but participated 1–20 days on average during the follow-up period were classified as Group 2, and those who participated 21–99 days during follow-up were classified as Group 3. Regular participation (Group 4) referred to those students who averaged 100 days or more per year during the follow-up period. Using the pooled math sample as an example, Figure 1 illustrates the manner in which students were included in the four groups.

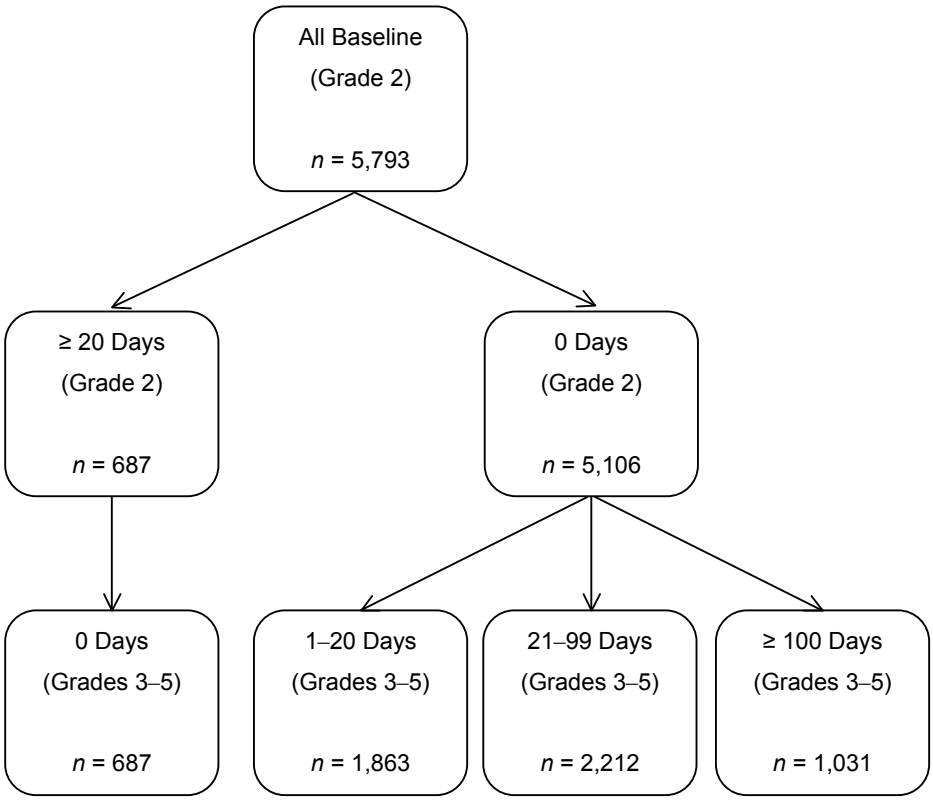

Figure 1: Flow chart showing the math sample for the two cohorts

6.2 Student population characteristics

The two cohorts included in this study had similar population characteristics. Students in the 2002–03 cohort were predominantly Latino (90.5 %), half were female (50.0 %), and about one-third had parents who either graduated from high school or attended some level of college. Furthermore, over three-quarters of the students in this cohort were classified as Limited English Proficient (LEP) (79.6 %). Participants in the 2003–04 cohort were also primarily Latino (88.7 %), slightly more than half were female (51.3 %), and about one-third (32.7 %) had parents who either graduated from high school or attended some level of college. Furthermore, about two-thirds (66.4 %) were classified as LEP.

6.3 Measures

6.3.1 Student demographic variables

We obtained student demographic data from district administrative files, including gender (female), parent education (high school graduate or more, less than high school graduate), race/ethnicity (African-American/Black, Hispanic/Latino), and regular school attendance.

6.3.2 English language learner status

Information on whether each student was an English language learner was also included in the district administrative file. In California, all children whose parents report that their home language is not English are required to take the California English Language Development Test (CELDT), which measures skills in listening, speaking, reading, and writing English (California Department of Education, 1999). Students who score in the lowest three levels (beginning, early intermediate, or intermediate) are then classified as English learners (ELs). Similarly, students who score in the top two levels (early advanced or advanced) during their initial testing and have teacher and parent recommendations, are classified as Initially Fluent English Proficient (IFEP).

6.3.3 Standardized achievement scores

The CSTs are administered each spring to students in second through eleventh grade as part of the Standardized Testing and Reporting Program. The CSTs are criterion-referenced tests, which were specifically developed to assess students' performance on California's academic content standards. The English-language arts tests have between 65 and 75 multiple-choice questions depending upon the grade level being assessed. At the elementary school level, the fourth grade test also includes an 8-point writing section. The mathematics test also includes 65 standards-based questions. Individual student test scores and school averages were also obtained each year from the district administrative files.

6.4 Data analysis

To examine the effects of the after-school program on achievement, the study employed an HLM design that not only accounted for student-level variation, but also the variation across schools that may influence student achievement outcomes. The following details the two HLM approaches used.

6.4.1 Residual gain analysis

For this study a residual gain approach was used to address the first research question concerning whether attendance in the after-school program led to positive achievement outcomes in math and English-language arts. Residual gain approaches are often described as a value added analysis because the performance at the end of the follow-up period is compared to the best prediction for each student based on their baseline characteristics. Since each student's after-school group membership is not considered in the HLM model that produces the best prediction, any difference between actual performance and the best prediction that varies substantially from zero for any given group would represent the value added by the group membership.

There were two necessary steps in the residual gain approach. In Step 1, the achievement outcome was predicted using covariates for CST achievement at baseline, regular school attendance, gender, parent education level, language proficiency status, and race/ethnicity. Note that in Step 1, the different levels of intensity of participation were not included in the model. This was done so that the prediction of achievement at the follow-up year could be obtained without this knowledge. In Step 2, the difference between actual achievement at the end of the follow-up period and the model-based predictions (obtained at Step 1 without the knowledge of intensity of attendance) were compared between each group descriptively.

6.4.2 Mixed model hierarchical linear modeling

In order to explore the second research question concerning intensity of after-school participation, a mixed model HLM approach was employed. Since this approach was applied to the pooled data from across the two cohorts, a factor was used to represent the four groups. As with the residual gain analysis, achievement outcomes at the end of the follow-up period were predicted using covariates for CST achievement at baseline, regular school attendance, gender, parent education level, language proficiency status, and race/ethnicity. Since the data was pooled, an additional covariate for cluster (2003–04) was included. Estimated means for each group were computed controlling for the baseline variables and 95 % confidence intervals for the differences between the comparisons were displayed.

7. Results

To provide more clarity to our analyses, the modeling results will be presented by content area: math and English-language arts. The synthesis of the results will be presented in the *Discussion* and *Conclusion* sections.

7.1 Math results

For math achievement, residual gain analyses were conducted for each cohort and then mixed modeling was employed to the pooled data (across the two cohorts) using a factor to represent the four groups.

7.1.1 Residual gain analyses

Table 1 displays the difference between predicted and actual math achievement outcomes for the 2002–03 cohort. Predicted outcomes were obtained from the Step 1 model that accounted for baseline achievement and background characteristics, but ignored group membership (intensity of participation). Means of each student's actual minus predicted scores shows that students in Group 4 scored an average of almost seven scale points higher than the model predicted. In contrast, students in Groups 1 and 2 on average scored about two scale points lower than the model predicted. As can be seen in Table 2, residual gains for the 2003–04 cohort follow a similar pattern. While students in Group 1 had an average score of almost eight scale points higher than predicted, students in Groups 1 and 2 scored about three scale points lower than the model predicted. In both cohorts, students in Group 3 had similar actual and predicted scale scores.

Table 1: Residual gains by group, math achievement of the 2002–03 cohort

		Actual scale score		Predicted scale score		Residual gains (actual-predicted)	
Group	*n*	Group means	*SD*	Group means	*SD*	Group means	*SD*
Group 1	257	329.08	76.62	331.25	54.27	-2.18	58.31
Group 2	809	323.60	75.94	326.08	48.95	-2.49	55.81
Group 3	883	327.78	74.00	327.79	49.25	-0.01	55.07
Group 4	383	345.72	84.36	338.98	55.52	6.74	58.90

Table 2: Residual gains by group, math achievement of the 2003–04 cohort

		Actual scale score		Predicted scale score		Residual gains (actual-predicted)	
Group	*n*	Group means	*SD*	Group means	*SD*	Group means	*SD*
Group 1	430	335.94	76.81	338.96	50.04	-3.02	55.21
Group 2	1,054	332.71	74.67	336.26	50.09	-3.55	53.85
Group 3	1,329	337.58	74.35	337.60	50.62	-0.01	53.86
Group 4	648	345.57	76.03	337.76	48.23	7.81	56.22

7.1.2 Analyses of pooled data

To test the significance of the differences found in the residual gain analyses, a mixed model was performed controlling for baseline variables with the LA's BEST grouping variable included in the model. Estimated means were produced for each group, and differences between the regular participation group (Group 4) and the three other groups were tested for significance. In Table 3, descriptive statistics for math achievement are presented at the baseline and at the follow-up time points for the pooled data. The mixed model results in Table 4 reveal that students in Group 4 (students with regular participation) performed significantly better than did students in the other three groups ($p < .05$). A Bonferroni correction was applied to the significance values shown in Table 1 to account for multiple comparisons.

This suggests that students who participated in LA's BEST regularly (Group 4) during the follow-up period performed better than expected when compared to students who did not participate in the program during the follow-up year (Group 1) or who participated in the program with less regularity (Groups 2 and 3). The estimated mean at the end of the follow-up period was about 12 scale points higher for the regular participants than for the group with no LA's BEST participation and the group with low average participation in the follow-up period. It should be noted, however, that the effect size of these differences was small when compared to the standard deviations of the actual group means at follow-up (see Table 3).

Table 3: Baseline and follow-up CST scale score means and standard deviations, math achievement of the pooled cohorts

		Baseline		Follow-up	
Group	*n*	Mean	*SD*	Mean	*SD*
Group 1	687	337.93	67.64	333.37	76.76
Group 2	1,863	333.46	66.39	328.75	75.34
Group 3	2,212	336.38	67.97	333.67	74.35
Group 4	1,031	339.34	68.95	345.62	79.18

Table 4: Multiple comparisons estimated mean differences (to Group 4), math achievement of the pooled cohorts

	Mean estimated scale score difference	SE	Significance value	95 % confidence interval	
				Lower bound	Upper bound
Group 4 - Group 1 (342.58 - 330.74)	-11.84	2.88	.000	-18.743	-4.930
Group 4 - Group 2 (342.58 - 330.81)	-11.77	2.29	.000	-17.261	-6.278
Group 4 - Group 3 (342.58 - 330.74)	-8.23	2.20	.001	-13.492	-2.972

7.2 English-language arts results

As with math achievement, residual gain analyses were conducted for each cohort and then mixed modeling was employed against the pooled data using a factor to represent the four groups.

7.2.1 Residual gain analyses

Tables 5 and 6 present the differences between predicted and actual English-language arts outcomes for both the 2002–03 and 2003–04 cohorts. The means of each student's actual minus predicted scores show that average differences for students in each group were within approximately one scale score point. This suggests that the prediction model produced fairly accurate predictions of actual scores for each group and that participation intensity did not have an effect on the predicted score.

Table 5: Residual gains by group, English-language arts achievement of the 2002–03 cohort

		Actual scale score		Predicted scale score		Residual gains (actual-predicted)	
Group	*n*	Group means	*SD*	Group means	*SD*	Group means	*SD*
Group 1	256	317.53	44.68	316.71	33.38	0.82	31.54
Group 2	808	315.90	44.91	316.43	31.26	-0.53	31.97
Group 3	883	318.75	43.28	319.01	30.55	-0.26	28.86
Group 4	383	327.77	46.47	326.59	33.52	1.18	29.68

Table 6: Residual gains by group, English-language arts achievement of the 2003–04 cohort

		Actual scale score		Predicted scale score		Residual gains (actual-predicted)	
Group	*n*	Group means	*SD*	Group means	*SD*	Group means	*SD*
Group 1	423	321.78	46.07	321.16	31.23	0.62	31.22
Group 2	1,044	318.11	42.70	319.19	39.57	-1.08	30.32
Group 3	1,327	322.56	43.21	322.40	30.98	0.16	30.09
Group 4	648	326.08	44.17	325.07	30.82	1.01	31.11

7.2.2 Analyses of pooled data

As with the math achievement data, findings from the residual gain analyses for English-language arts were confirmed using a mixed model that controlled for baseline variables with the LA's BEST grouping variable included in the model. Estimated means were produced for each group, and differences between the regular participation group (Group 4) and the three other groups were tested for significance. In Table 7, descriptive statistics for English-language arts achievement are presented at the baseline and at the follow-up time points for the pooled data. Results in Table 8 indicate that there is no significant difference among the groups ($p > .05$), confirming that program participation and intensity of participation did not influence English-language arts performance.

Table 7: Baseline and follow-up CST scale score means and standard deviations, English-language arts achievement of the pooled cohorts

		Baseline		Follow-up	
Group	*n*	Mean	*SD*	Mean	*SD*
Group 1	679	308.85	48.90	320.18	45.57
Group 2	1,852	307.51	46.33	317.15	43.68
Group 3	2,210	311.36	47.36	321.03	43.27
Group 4	1,031	317.36	48.10	326.71	45.03

Table 8: Multiple comparisons estimated mean differences (to Group 4), English-language arts achievement of the pooled cohorts

				95 % confidence interval	
	Mean estimated scale score difference	SE	Significance value	Lower bound	Upper bound
Group 4 - Group 1 (321.76 - 320.68)	-1.08	1.59	1.000	-4.884	2.721
Group 4 - Group 2 (321.76 - 319.29)	-2.47	1.26	.149	-5.484	0.544
Group 4 - Group3 (321.76 - 320.37)	-1.39	1.20	.743	-4.277	1.492

8. Discussion

This study set out to reduce a research gap by using rigorous methodology to study the effects of dosage (intensity of after-school attendance) on students' academic outcomes. It was hypothesized that the regular participation group (of 100 days or more per year) would benefit more from the program than the low participation or the no participation groups. The current findings confirm this hypothesis.

Results of the analysis show that regular participation in the LA's BEST program is significantly associated with small but positive gains in CST math achievement when compared to students with low or no participation in the program during the follow-up period. This finding is supported by current literature concerning that participation in after-school programs, particularly those like LA's BEST that offer both enriching youth development activities and a strong academic component, can lead to small gains in academic outcomes. Also in support of the current literature, students who had higher intensity of participation benefitted more than students who participated in lower intensity. These findings support our theory that dosage is an important indicator of program effectiveness and should be considered in future studies on after-school program effects.

Given the importance of "dosage" in the examination of program participation outcomes, it is apparent that simply examining the differences between participants and non-participants overlooks many of the important aspects of program participation. Thus, it is important for researchers to understand the subtle variations in levels of participation and to help programs build data-driven arguments of program effectiveness. Furthermore, understanding the attendance patterns of students can also provide insights to program leaders for program quality improvement. Thus, after-school programs should collect meaningful attendance data and use the data to construct a system of accountability and program improvement. Fiester (2004) asserts, "the 'right' methods for collecting, organizing, and analyzing data depends on how program leaders expect to use the data – what questions they need to answer, and for whom" (p. i.). Program participation intensity, duration, and breadth are all important indicators to be considered. Though beyond the scope and not examined in this study, as a reminder, when discussing the relations between participation in after-school programs and outcomes, the issues of student participants' engagement and involvement should also be addressed in studies for program quality.

As an intention to reduce selection bias, the statistical tests were performed on pooled data across two cohorts of students. By doing this, the study has demonstrated that the participating and non-participating students were initially equivalent in regards to academic performance, socio-economic status, and other measures. More specifically, the differences in achievement for the comparison groups as measured by residual gains are consistent across the two separate cohorts of students that the study followed over a period of four years. It is important to note that these results were obtained after carefully accounting for existing differences in students' background characteristics so that the most plausible explanation of these statistical differences is in the regularity of LA's BEST participation. However, it should be noted that the data used for the study did not include family background characteristics such as foster child status, single parent households, student motivation, and student attribution characteristics.

In terms of students' English-language arts achievement, the results reveal that students with regular participation (100 days or more per year) during the follow-up period did not have significantly higher CST English-language arts residu-

al gains when compared to students who did not participate in the program or who had low (1–20 days) or moderate (21–99 days) intensities of participation during the follow-up period. This finding is also consistent across both cohorts, suggesting that additional programmatic focus in this content area will be needed in order to generate significant reading gains for the participants. At the same time, it is also important to keep in mind that two thirds or more of LA's BEST students are English learners and speak a second language. Additional factors such as language spoken at home, opportunities to read and communicate in English, student motivation/engagement, parental/peer support, community environment, and so forth may all influence students' language development and should be examined further in future studies.

McComb and Scott-Little (2003) suggest that after-school programs should be an integral part of a school's academic and developmental programs. There are some practice and policy implications that can be derived from this study as well. Federal and state policymakers can greatly enhance the operation and effectiveness of after-school programs by directing funding and providing regulations that encourage continuous participation of the students. State governments can also take the lead in coordinating resources by blending multiple public and private sources to ensure that families and students who need the programs have continued access to them. To ensure after-school programs meet local needs so that students will participate regularly, federal and state policymakers should provide opportunities, such as local conferences and workshops, for schools, local businesses, and families to plan, build, and improve after-school programs in collaboration. Emphasis should also be placed on bringing in a wide range of local resources, perspectives, and talents to enhance after-school opportunities and provide a broader range of human and social resources to participating families, thus enhancing family and student motivation, engagement, and retention.

9. Conclusion

In conclusion, this study provides evidence that LA's BEST is a program that shows promise in improving students' CST math performance. However, as with any intervention project, students need to participate regularly in order to reap the program benefits. This study suggests that 100 or more days of annual participation are necessary. Implications from this study also highlight that simple indicators of program participation are inadequate to capture program effects fully. For a program to have impact on student achievement, the students need to receive sufficient exposure. Participation level would be a better indicator of program effects until the field can find methodologies that control the unobservable self-selection biases that are inherent for both participants and non-participants.

References

After-school Alliance. (2008). *21st Century Community Learning Centers providing after-school supports to communities nationwide*. Washington, DC: Author.

After-school Alliance. (2009). *After-school programs making a difference in America's community by improving academic achievement, keep kids safe and helping working families*. Washington, DC: Author.

After-school Alliance. (2014). *21st Century Community Learning Centers; providing after-school and summer learning supports to community nationwide*. Washington, DC: Author.

Bergin, D. A., Hudson, L. M., Chryst, C. F., & Resetar, M. (1992). An after-school intervention program for educationally disadvantaged young children. *The Urban Review, 24*(3), 203–217.

Bodilly, S. J., & Beckett, M. K. (2005). *Making out-of-school time matter: Evidence for an action agenda* (report No. MG-242-WF). Santa Monica, CA: RAND.

Cooper, H., Charlton, K., Valentine, J. C., & Muhlenbruck, L. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development, 65*(1), 1–118.

De Kanter, A. (2001). After-school programs for adolescents. *NASSP Bulletin, 85*(626), 12–21.

Dimsdale, T., & Kutner, M. (2004). *Becoming an educated consumer of research: Look at the basics of research methodologies and design*. Washington, DC: The American Institutes for Research.

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

Ed. Gov., U.S. Department of Education. (2002). *No Child Left Behind Act of 2001, Pub. L. No. 107–110, 115 Stat. 1425*. Retrieved from http://www.ed.gov/legislation/ESEA02/

Fashola, O. (2002). *Building effective after-school programs*. Thousand Oaks, CA: Corwin Press, Inc.

Fiester, L., & Policy Studies Associates. (2004). *After-school Counts! A guide to issues and strategies for monitoring attendance in after-school and other youth programs*. New York, NY: The After School Project, Robert Wood Johnson Foundation.

Frankel, S., & Daley, G. (2007). *An evaluation of after school programs provided by Beyond the Bell's partner agencies*. Los Angeles, CA: Beyond the Bell Branch, LAUSD.

Goerge, R., Cusick, G. R., Wasserman, M., & Gladden, R. M. (2007). *After-school programs and academic impact: A study of Chicago's After School Matters*. Chicago, IL: Chapin Hall Center for Children.

Goldschmidt, P., Huang, D., & Chinen, M. (2007). *The long-term effects of after-school programming on educational adjustment and juvenile crime: A study of the LA's BEST after-school program*. Los Angeles, CA: UCLA Center for the Study of Evaluation.

Grossman, J. B., Price, M. L., Fellerath, V., Jucovy, L. Z., Kotloff, L. J., Raley, R., & Walker, K. E. (2002). *Multiple choices after-school: Findings from the extended-services schools initiative*. Philadelphia, PA: Public Private Ventures.

Hodgkinson, H. (2006). *The whole child in a fractured world*. Alexandria, VA: Commission on the Whole Child, convened by the Association for Supervision and Curriculum Development.

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

Huang, D., Gribbons, B., Kim, K. S., Lee, C., & Baker, E. L. (2000). *A decade of results: The impact of the LA's BEST after school enrichment program on subsequent student achievement and performance*. Los Angeles, CA: UCLA Center for the Study of Evaluation.

James, D. W. (1997). *Some things do make a difference for youth: A compendium of evaluations of youth programs and practices* (ERIC Document Reproduction Service No. ED 409 462). Washington, DC: American Youth Policy Forum.

James-Burdumy, S., Dynarski, M., Moore, M., Deke, J., & Mansfield, W. (2005). *When schools stay open late: The national evaluation of the 21st Century Community Learning Centers program*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Jenner, E., & Jenner, L. W. (2007). Results from a first-year evaluation of academic impacts of an after-school program for at-risk students. *Journal of Education for Students Placed At Risk, 12*(2), 213–237.

Kane, T. J. (2004). *The impact of after-school programs: Interpreting the results of four recent evaluations*. New York, NY: William T. Grant Foundation: Publications and reports.

Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. A., Snow, D., & Martin-Glenn, M. (2003). *The effectiveness of out-of-school-time strategies in assisting low-achieving students in reading and mathematics*. Aurora, CO: Mid-continent Research for Education and Learning.

Little, P. M. D., & Harris, E. (2003). A review of out-of-school time program quasi-experimental and experimental evaluation results. *Out-of-School Time Evaluation Snapshots, 1*(1).

Mahoney, J. P., & Zigler, E. F. (2006). Translating science to policy under No Child Left Behind Act of 2001: Lessons from the national evaluation of the 21st Century Community Learning Centers. *Journal of Applied Developmental Psychology, 27*(4), 282–294.

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

Modi, N., & Keay, A. J. (1983). Phototherapy for neonatal hyperbilirubinaemia: the importance of dose. *Disease in Childhood, 58*, 406–409.

Munoz, M. A. (2002). *Outcome-based community–schools partnerships: The impact of the after-school programs on non-academic and academic indicators* (ERIC Document Reproduction Service No. ED 468 973). Louisville, KY: Jefferson County Public Schools.

Pinkel, D. (1958). The use of body surface area as a criterion of drug dosage in cancer chemotherapy. *Cancer Research, 18*, 853–856.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Riggs, N. R. (2006). Attendance in after-school program and the social development of migrant Latino children. *Journal of Community Psychology, 34*(1), 75–87.

Riggs, N. R., & Greenberg, M. T. (2004). Moderators in the academic development of migrant Latino children attending after-school programs. *Journal of Applied Developmental Psychology, 25*(3), 349–367.

Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 92*(3), 726–748.

Schaps, E (2006). *Educating the Whole Child*. Alexandria, VA: Commission on the Whole Child, convened by the Association for Supervision and Curriculum Development.

Scott-Little, C., Hamann, M. S., & Jurs, S. G. (2002). Evaluations of after-school programs: A meta-evaluation of methodologies and narrative summary findings. *American Journal of Evaluation, 23*(4), 387–419.

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher, 31*, 15–21.

U.S. Department of Education. (2014). *Programs: 21st Century Community Learning Centers*. Retrieved from: http://www2.ed.gov/programs/21stcclc/index.html

Vandell, D., Reisner, E., & Pierce, K. (2007). *Outcomes linked to high-quality afterschool programs: Longitudinal findings from the study of promising practices*. Flint, MI: Charles Stewart Mott Foundation.

Vanderhaar, J., & Muñoz, M. A. (2006). *Educating at-risk African American males: Formative and summative evaluation of the Street Academy Program* (ERIC Document Reproduction Service No. ED 495 958).

Welsh, M. E., Russell, C. A., Williams, I., Reisner, E. R., & White, R. N. (2002). *Promoting learning and school attendance through after-school programs: Student-level changes in educational performance across TASC's first three years*. Washington, DC: Policy Studies Associates, Inc.