Impacts of a Summer Learning Program:
A Random Assignment Study of Building Educated Leaders for Life (BELL)

Duncan Chaplin
Mathematica Policy Research/The Urban Institute

and

Jeffrey Capizzano
Teaching Strategies Inc.

August 2006

This report was prepared with support from the Smith Richardson Foundation and the William T Grant Foundation as part of an evaluation study of the BELL summer learning program. The study was conducted primarily by staff of the Urban Institute. We would like to gratefully acknowledge advice from Stephen Bell at Abt Associates; Rebecca Maynard at the University of Pennsylvania; Bob Granger and Ed Seidman at the Grant Foundation; Marv Eisen and Jane Hannaway at the Urban Institute; Charles Michalopoulos at MDRC; and Steven Glazerman, Mark Dynarski, Allen Schirm and Christina Tuttle at Mathematica Policy Research. Excellent help with data collection and site visits was provided by Laura Logerfo, Tim Dore, Maggie Stanislawski, Sheila Isanaka, David Chase, Megan Reamer, and Tiffany Miller at the Urban Institute. Linda Kuhn and Ashley Bronzan and their field staff at Survey Research Management (SRM) deserve praise for their success in obtaining high response rates during their data collection efforts in 2005. Tiffany Cooper and the rest of the staff of the BELL program provided invaluable help in all steps involved in designing the study, obtaining applicants for random assignment, and obtaining the baseline data. Any opinions, observations, findings, conclusions, and recommendations expressed in this report are our own and do not necessarily reflect the views of people who gave us advice, Mathematica Policy Research, the Urban Institute, the BELL Foundation, SRM, the Smith Richardson Foundation, or the Grant Foundation. Please direct any questions or correspondence to Duncan Chaplin at Dchaplin@mathematica-mpr.com.

Executive Summary

A growing body of evidence indicates that the test scores of low-income children drop significantly relative to their higher-income counterparts during the summer months. This study finds that a well-implemented summer learning program can improve reading skills and increase the extent to which parents encourage their children to read during the subsequent school year. These findings provide some support for investments in out-of-school time programming for low-income children during the summer, such as those currently coming from the 21st Century Community Learning Centers program and the Supplemental Services provisions of Title I of the *No Child Left Behind* Act.

This study used random assignment, the gold standard of evaluation methods, to evaluate the effectiveness of the Building Educated Leaders for Life (BELL) program—a summer program designed to improve academic skills, parental involvement, academic self-perceptions, and social behaviors among low-income children and families. Over 1,000 elementary school children who applied to BELL summer programs in New York and Boston in 2005 were randomly chosen to be in either a treatment group that was selected to participate in the BELL summer program, or a comparison group that was not. Independent researchers collected student reading tests (Gates-MacGinitie) and student and teacher surveys.

The study found that children in the BELL treatment group gained about a month's worth of reading skills more than their counterparts in the comparison group during the summer. This is a modest, yet notable increase in reading skills for a six-week program. The study also found evidence of positive impacts on the degree to which parents encouraged their children to read. No impacts were found on academic-self perceptions or social behaviors.

Overall, this study provides scientifically rigorous evidence regarding the ability of the BELL summer program to improve the reading skills of low-performing elementary school children. Few out-of-school time programs have produced evidence of effectiveness when evaluated in such a rigorous manner. The results are of particular importance given the long-standing public policy focus on raising achievement levels of low-income students.

Introduction

An emerging body of research has shown that low-income children may experience a decline in academic progress during the summer months, especially relative to higher-income children (Entwisle and Alexander 1992; H. Cooper et al. 1996). This learning loss may contribute to overall differences in educational achievement between low- and higher-income children. In addition, in recent years, a significant amount of public funding has been used to provide enriching summer programs for low-income children. This includes funds available because of provisions in the federal education legislation, the *No Child Left Behind Act of 2001*, which sets aside Title I money for supplemental educational services for low-income students. These funds can be used to fund summer programs provided by non-profit organizations, such as Building Educated Leaders for Life (BELL).

Given the evidence on academic losses during the summer, another body of research has attempted to test the effects of programs dedicated to mitigating this problem. H. Cooper et al. (2000) summarize much of this literature. They report on few random assignment evaluations in their summary of over 93 studies. In addition, the non-experimental studies they summarize are generally of low quality. For instance, many are based on a simple comparison of outcomes of participants before and after treatment, with no comparison group. Lacking a comparison group makes the interpretation of such findings difficult at best. In addition, those studies with comparison groups generally have very weak methods of ensuring that the two groups are comparable. For example, they generally lack information on outcomes before treatment; data that would help to rule out the possibility that differences observed after the program did not already exist before participation in the program.

More recent evidence is of much higher quality. For instance, Sunmonu et al. (2002) find that the summer academic program provided by the Montgomery County Public schools appears to have had substantial positive impacts on test scores for the students participating regularly. They compared changes in test scores during the summer for regular participants with the changes for non-participants, controlling for race, gender, and economic status of the children. While this evidence is a great improvement over that summarized by H. Cooper et al. (2000), it still has a significant weakness. In particular the results may be biased if regular participants are particularly motivated students compared to non-participants, even after controlling for past achievement and other observable characteristics. More generally the types of studies discussed above are all subject to the criticism that non-experimental results are often not replicated by careful experimental studies, which are considered to be far less likely to produce biased estimates (Michalopoulos 2005; Glazerman et al. 2003; Bloom et al. 2002, Agodini and Dynarski 2001; Wilde and Hollister 2002).

Another recent study of a summer school program in Chicago (Jacob and Lefgren 2002) addresses this issue of selection bias more rigorously by using the regression discontinuity method and finds impacts similar to around one month of regular school for 3rd graders, though no statistically significant impacts for 6th graders. While this is far better than the comparison of participants and non-participants done by Sunmonu et al., it is still not based on random assignment.[1]

1. While random assignment methods have many advantages, they have been criticized for a number of reasons—for example that they can be unethical or that they do not take into account the large number of factors that affect outcomes of interest (Murnane and Nelson 2005). These issues are discussed in Advisory Committee on Head Start Research and Evaluation (1999) and Michalopoulos (2005).

The most rigorous evidence in this area comes from Borman and Dowling (2006). They estimate the impacts of a summer learning program in Baltimore on student achievement using random assignment. They did not find a statistically significant impact of being assigned to the treatment group compared to the control group. However, there was considerable variability in student attendance across the three summers. Using quasi-experimental methods, they did find a statistically significant impact for those children who participated for at least two years. While this is encouraging it still leaves open the question of whether summer learning programs can improve outcomes for a larger subset of the target population and whether their impacts can be detected using more rigorous methods.

Roadmap

The next section of this paper presents a description of the BELL summer learning program. A conceptual model is used to explain how the program is expected to impact outcomes of interest. Details of our study design, the data we collected, the outcomes that are the focus of this study, and our analysis methods are explained to help set the stage for our results section. We finish with a discussion of our results and our conclusions. Appendices cover alternative methods we considered using, how we adjusted for the time our study group spent in school, robustness tests we conducted, detailed tables, and the parent survey instrument.

The BELL Summer Learning Program

Building Educated Leaders for Life (BELL) is a community-based organization that provides supplemental learning activities for low-income youth in Boston, New York, Washington, D.C. and Baltimore. It was founded in 1992 by a group of Black and Latino students at Harvard Law School. The overarching goals of the program are to "dramatically increase the academic achievements, self-esteem and life opportunities of elementary school

children living in underserved communities…" (BELL 2004). BELL's philosophy is to serve the whole child: it aims to not only increase academic success by improving basic math and literacy skills, but also works to assist in social and emotional development by exposing program participants to positive role models, and by building self esteem and encouraging parents to become more involved in their children's lives.

BELL operates both school year and summer programs. During the summer of 2005 they had programs at five locations (three in Boston, one in New York City and one in Washington, D.C.) These programs provided five to six weeks of classes for eight hours per day and five days per week for children entering grades 1–7.

For academic activities, students are clustered in groups of approximately 15 children with each cluster taught by 1 teacher (usually a regular teacher from the public school system) and one experienced teaching assistant (generally an Americorps volunteer or college student). Each week students receive approximately eight hours of literacy instruction (two hours per day, four days a week), four hours of math instruction (one hour a day, four days per week) and 6.5 hours of community time (0.5 hours, four days a week and a Friday 1/2 day field trip). In addition, every Friday students attend a speaker series where they hear from and ask questions of prominent citizens in their community.

What are BELL's academic components? BELL uses nationally recognized curricula in both math and reading, which were updated in 2005. In the previous year, the reading and writing portions of the program were drawn from a culturally sensitive curriculum developed by Voices for Love and Freedom (VLF), a non-profit educational organization affiliated with the New American Schools. The curriculum was designed to explore various themes, such as Democracy, through the use of multicultural literature. In addition, the reading content in the

curriculum was aligned with several national and state learning standards as well as some standardized assessments.

The math curriculum in 2004 was Math Steps, a research-based, sequential program produced by the Houghton-Mifflin publishing company, one of the major educational publishing companies in the United States. The curriculum was arranged so that teachers could follow a grade-specific, step-by-step plan that can be tailored to varying levels of achievement. Beyond the core lesson, Math Steps provided worksheets and materials for teachers to further assist students who are behind and also to provide additional, more challenging work for those who are advanced.

In 2005, BELL based its literacy intervention on a new curriculum—*Summer Success: Reading*—published by Houghton Mifflin. Additionally, to ensure that their program continued to provide a culturally relevant learning experience and assisted in developing social skills and values, they continued to use the multicultural literature they used with the VLF curriculum. BELL's 2005 literacy approach also included phonic instruction aligned to the recommendations of the National Reading Panel.

The changes to the literacy curriculum were driven largely by BELL's desire to better address teaching and learning needs in the classroom, including providing resources for English Language Learners, incorporating an explicit, and integrated phonics component. In addition, the *Summer Success: Reading* curriculum provided additional supports important to the BELL program including better alignment to national and state learning standards.

The program also made a small change to its math curriculum, moving from *Math Steps* to *Summer Success: Math.* The BELL staff explained that the two math curricula are very similar, but the latter curriculum is organized in a much more efficient way, laying out clear

daily and weekly plans for teachers to follow. In addition to this revised curricular approach, BELL's 2005 math intervention included math manipulatives and other activities to support mathematical development and to prepare children for state exams.

The BELL program also developed a guide for their teachers that provided guidance on a number of issues ranging from assessment and homework to how to fit a five-day-a-week curriculum into BELL's four-day instruction week.

Why evaluate BELL over other program models? There are a number of reasons why it is important to evaluate the BELL program model. First, it employs well-developed curricula in both reading and math and contains the features of positive developmental settings outlined by the National Research Council and Institute of Medicine in the report, *Community Programs to Promote Youth Development* (2002). The program has also demonstrated effectiveness in reducing summer learning loss among low-income children in ambitious internal evaluations.[2] For these reasons, the BELL program shows promise, making it a prime candidate for a rigorous evaluation.

Second, BELL is operating a program that has garnered strong multi-year support from many financial backers including New Profit, Inc.; The Charles Hayden Foundation; The Smith Family Foundation; Boston Red Sox; Fidelity; Bank of America; and others. Even more significantly, in April 1997, at the President's Summit for America's Future, BELL was awarded the President's Service Award. Thus, unlike many of the programs studied in the national evaluation of the 21st CCLC program (Dynarski et al. 2003), BELL is not only a "mature"

---

2. One internal evaluation reports that their summer scholars improve in math and reading by the equivalent of 4 months in reading and 5 months in math (T. Cooper 2002). In comparison, H. Cooper et al. (1996) report losses for low-income children over the summer. Additional internal evaluation reports

program, but has been judged to be a high quality, stable intervention that can likely be evaluated for a number of years.

Third, the BELL program has been approved by a number of states for funding for Title I supplemental services. As such, BELL represents the type of organization that state education agencies will be considering when deciding how to use their Title I funds.

Fourth, the BELL program appears to be well prepared to expand its program to additional sites as it provides substantial staff development, the importance of which has been noted by Patterson and Czajkowski (1979). Currently the national BELL program provides training to their teaching staff to ensure that they support their program objectives. Indeed, BELL has a Chief Program Officer at the national level hired explicitly to design recruitment, interviewing systems, and training frameworks. BELL also provides standardized training modules in the major program areas and staff handbooks. Every site utilizes the same curriculum and materials, including lesson plans. BELL also maintains a quality control system designed to ensure that all of their programs maintain high standards. This consists of regular site visits by the national staff (4 visits at each site per year), and a standardized site visit tool for local staff (weekly visits to each classroom).

Finally, the BELL staff has exhibited a strong interest in evaluation as shown by their own internal evaluations of both their summer and school year programs and the development of their Evaluation Advisory board.[3]

---

also produced promising evidence of academic improvements over the summer (T. Cooper 2003) and similar evidence for BELL's school-year program (T. Cooper 2004).

3. See T. Cooper (2002, 2003, and 2004).

The Conceptual Model

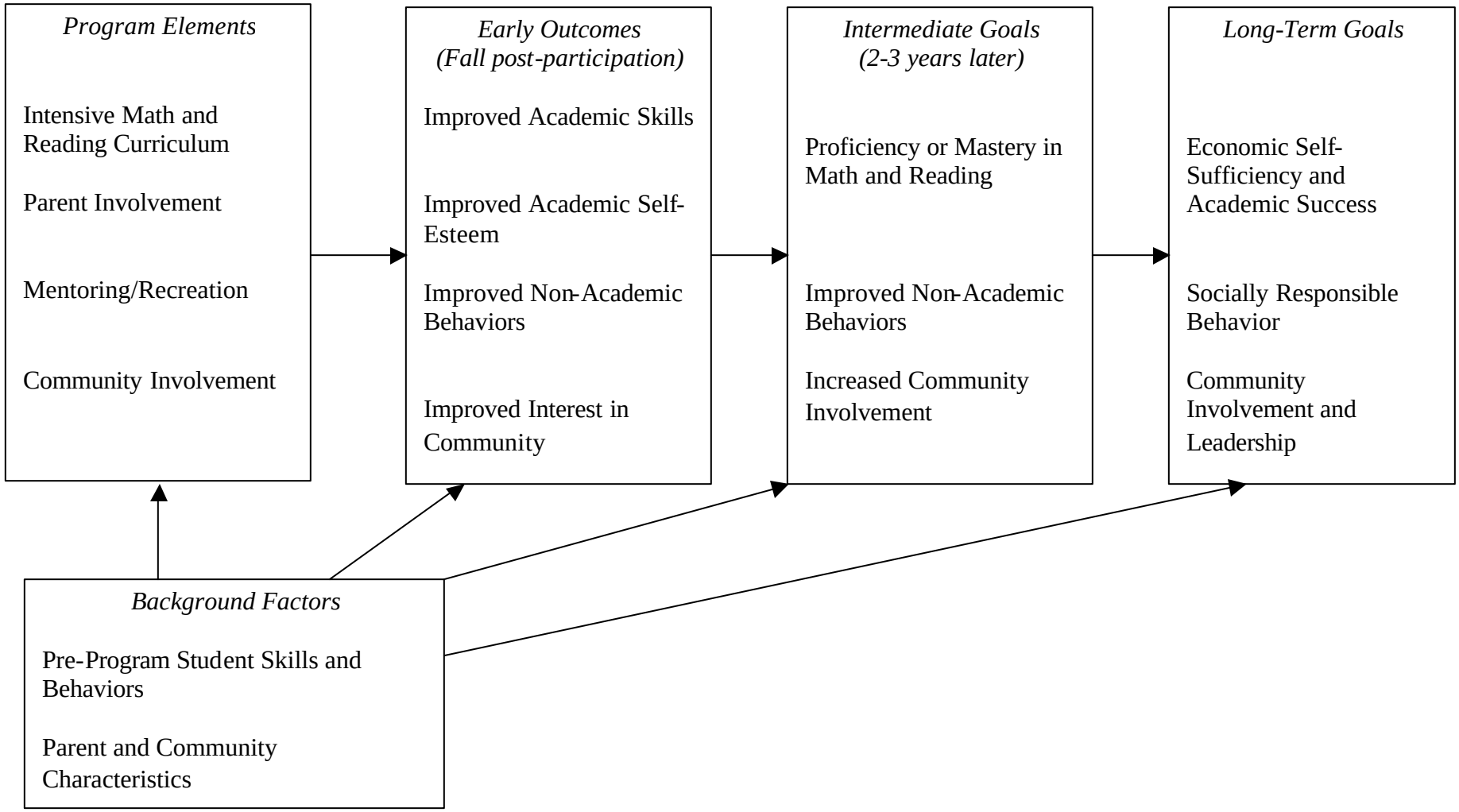
The BELL model can be thought of as producing positive academic and non-academic youth outcomes in stages, as shown in Exhibit 1. First, the treatment is theorized to have an impact on early outcomes, such as developing basic academic skills and positive academic self-concept, overall self-esteem, and improved relationships with adults and prosocial institutions, as well as awareness of the importance of community. This study focuses on early academic outcomes, as measured by test scores, and a set of youth development indicators that can be captured using student and parent surveys. In later years these early outcomes should translate into impacts on intermediate goals, such as the development of advanced reading and math proficiencies, improved behaviors while not under adult supervision (e.g. away from home or school), and meaningful involvement in the community. Finally, after many years, impacts on adult behaviors may become apparent. These would include economic self-sufficiency, academic success, socially responsible behaviors, community involvement and leadership.[4]

We also recognize that the students being assisted by the BELL program are facing serious challenges because of past experiences in their lives. Thus the background factors (also shown in the logic model) play a key role in affecting the outcomes analyzed in this report.

---

4. Estimating long-term program impacts will be far more difficult because of the influence of intervening experiences, including later participation in the BELL program by some members of the control group. However, it will be possible if we are able to increase our sample size sufficiently using additional funding in later years.

Exhibit 1: Logic Model for BELL Accelerated Summer Learning Program



Study Design

The BELL program agreed to allow the research staff to use random assignment to determine which applicants were accepted at three of their five sites—two in Boston and one in New York City. Two additional sites were excluded—one each in Boston and Washington, DC. Thus, the potential study participants are applicants for the BELL program at the three included sites. All students entering grades 1–7 in these cities are eligible to apply for the BELL summer program, but the program focuses its recruiting efforts on low-income minority students who are academically challenged but not in special education.

In total there were 1,917 applicants to the BELL 2005 summer program at the three included sites, more than double the 750 slots available (250 at each site). Of these applicants, 1,225 had parents who signed a form allowing their child to be in the study before being informed of whether or not they would be accepted into the program. No clear differences were found in the racial and gender distribution of those willing and not willing to participate in the study (Chaplin 2006).

In order to be fair, participation in the study was not a condition for participation in the program so even those who refused to participate in the study were included in the lottery to decide who would be accepted to the program. More details on participation rates in the study are given in Chaplin (2006).

All results presented here control for differences in the probability of getting into the program. This probability varied depending on the pool that an applicant was put in for randomization into the treatment or control group. These randomization pools varied by grade,

site, and when a family applied for the program. [5] The probability of getting into the program varied across randomization pools depending on the number of applicants and slots available. To control for these differences the treatment and control cases within each randomization pool are weighted up to the total number of applicants in the study in that pool. We also adjust for non-response and use three sets of weights depending on the outcome variable—one for student test scores, one for student academic self-perceptions from the student surveys and a third for the variables from the parent surveys. We also included indicator variables in our multivariate regressions to control for the randomization pool and thereby obtain more precise impact estimates. The different weights were used because the response rates varied noticeably by data source, especially between the data collected from parents and the data collected from students as those data were not always collected at the same time.

Within the group of 1,225 applicants whose parents agreed to let them be in the study, an additional 43 cases were left out because they were in a randomization pool where either no one was accepted or everyone was accepted. This happened when they applied in the last round for a grade-level and site where there were either no slots available or there were more slots than applicants. Another 95 cases with only one child per family and with a probability of being accepted close to 0 or 1 were excluded during the data collection process due to cost considerations. We did not exclude families with more than one child because we were collecting data in the homes and the cost of data collection per child was greatly reduced for families with more than one child. These 95 excluded cases were all in grades K–4 at the New York City site and in the last round of randomization. This left a sample of 1,087 applicants. Of

5. A random child was chosen in each family and randomization was done based on that child. Other children in the family were accepted or rejected based on that child's outcome.

these, we have complete test score data (including information on the day the test was taken) for 835 applicants for an overall weighted response rate of 78 percent for the control group and 78 percent for the treatment group on the student tests and surveys. Response rates on the parent survey were somewhat lower at 73 and 70 percent for the control and treatment groups respectively.

As is true for all random assignment evaluations, our estimates are most relevant to the types of individuals who agreed to be in the study, were included in the study, and complied with their random assignment status (Imbens and Angrist 2004).[6] The compliers were somewhat more likely to be black and male compared to other applicants (Chaplin 2006). The lack of statistically significant interactions of the estimated effect of the treatment with race, gender, site, or grade level (discussed below) suggests that similar results might hold for other children.

Data Collection Methods

The data used in this paper came from four sources—the BELL program, parent surveys, student surveys, and student tests.

BELL Program Data: The data obtained from the BELL program included the race, gender, and participation status of each applicant for the BELL program included in the study and was provided in the fall of 2005. These data were obtained from the application forms the parents filled out in the spring of 2005 and augmented with information from the staff on participants during the program.

Parent Surveys: The evaluation team administered the parent and student surveys and student tests. The parent surveys covered the types of learning activities the child engaged in

during the summer, the child's behaviors, parent involvement and other activities, and background information. We translated the parent survey to Spanish to assist those parents who were uncomfortable with English and hired Spanish-speaking interviewers to conduct phone interviews with non-English speaking parents.

Student Tests and Surveys: The student surveys focused on the children's perceptions of their academic skills. Separate surveys were used for children in grades 1 and 2 versus those in grades 3–6, as explained below. The student tests measure reading skills. We decided to focus on reading because there is evidence of a larger summer slide in reading than math for low-income students compared to higher income students (Entwisle and Alexander 1992; H. Cooper et al. 1996).

There was a small but important difference in how data were collected for the treatment and control groups. The parent and student data were collected over a 6 and a half-month period starting in mid-August of 2005 (at the end of the BELL program) and ending in January of 2006. Data collection began with testing on-site at the BELL program. Both the treatment and control groups were asked to participate in on-site testing but not surprisingly, the control group members were less likely to show up at the site. Consequently the control group was far more likely to be captured in the later data collection efforts that were done at the homes of the children. This resulted in a small, but important difference in the time that students were in school before being tested. This issue has important implications for our analyses of the test score data, as discussed below.

---

6. Thirty percent of the original sample did not agree to participate in the study and another twenty percent of those who remained did not provide data. About 34 percent of those in the treatment group did not participate in the BELL program and about 6 percent of the control group did.

Measures

The outcome data are designed to capture the major goals of the BELL program: to improve academic skills, student academic self-concept, student non-academic behaviors, and parent involvement. We also included a number of process outcomes related to activities that the child participated in during the summer.

Reading Tests: The evaluation team used the Gates-MacGinitie reading tests covering both vocabulary and comprehension. In grades 1 and 2 vocabulary refers to the "Word Decoding" sections of the Gates-MacGinitie tests. For grade 2 the Gates-MacGinitie reading test also has a "Word Knowledge" section that we did not use because of the extra testing time that would have been required. To estimate a total reading score in grade 2 we used the "Word Decoding" score to estimate a "Word Knowledge" score. Results for the sub-test scores (vocabulary and comprehension) are not affected by this adjustment.

We selected the Gates-MacGinitie Reading tests to measure student proficiency. We chose the Gates-MacGinitie tests over the Stanford Diagnostic tests and the SAT9 because BELL uses the Stanford diagnostic tests and Boston Public Schools used the SAT9 during the 2004-2005 school year. We were concerned that student performance might be affected if they had previously taken the test in question.

Students were given the test most appropriate for students entering the grade they were entering in the fall. We had tests for grades 1–6. Students repeating kindergarten were given the grade 1 test and students entering a grade above 6th grade were given the 6th grade test. The results suggest that these tests were at an appropriate level of difficulty for the vast majority of students tested. Only one student had a perfect score on the vocabulary test and none had a perfect score in comprehension. Only about 12 percent scored less than what they would get by

guessing on the questions they tried to answer in vocabulary and only 14 percent scored this low in comprehension. On average they answered around 40 percent of the questions correctly.

Measures of Student Academic Self-Concept: For participants entering grades 3–7, we used the Perception of Ability Scale Score (PASS), which is a self-evaluation instrument appropriate for students in grades three and higher (Hay et al. 1997). This survey took approximately 10 minutes to administer. BELL staff have administered the PASS survey to participants in earlier years and found little change over the summer. However, since they lacked a comparison group their evidence does not rule out the possibility that the BELL program had impacts on academic self-concept.

We divided the questions on the PASS survey into three types—those focused on English, Math, and General learning. The survey had a total of 70 yes/no questions, of which 37 focused on English, 12 on Math, and 21 on general learning skills.[7] Within each of these sets of questions about half were positive and half negative. We reverse coded the negative questions and then computed the mean value of the non-missing values within each set of questions so that higher responses represent more positive academic self-concept. The respondents to the PASS student surveys answered over 96 percent of the questions and about 5 percent were top-coded meaning that they answered all questions with a positive response.

For grades 1–2 we used the Academic Perceptions Inventory (API) in reading and arithmetic that has alpha reliability measures ranging from 0.59 to 0.91 for its K-3 instrument and takes about 10 minutes to complete.[8] This survey had 18 questions each of which was coded from –2 (for very sad) to +2 (for very happy). The total score is set to the sum of the scores on

7. The four questions on drawing were included in the math set.

each question so the scores range from –36 to +36. Missing answers are treated as 0s, which is half way between +1 (more happy than sad) and –1 (more sad than happy). As with the PASS survey, respondents to the API survey answered over 96 percent of the questions. None of the children answered all of the questions negatively (with a –2). About 22 percent were top-coded, meaning that they answered all questions positively (with a +2).

Non-Academic Behaviors and Parent Involvement: To measure non-academic child behaviors at home and parent involvement, we relied primarily on a subset of questions taken from the parent survey portion of the Social Skills Rating System (SSRS). SSRS measures positive social skills—cooperation, assertion, self-control and responsibility—as well as problem behaviors such as aggressive acts, poor temper control, and sadness and anxiety. The SSRS is currently being used as part of the Early Childhood Longitudinal Survey (ECLS) and was extensively researched and tested by researchers at the University of Michigan for the National Center for Education Statistics prior to its inclusion in the ECLS (U.S. Department of Education, 1996). Results using the SSRS have been published in number of refereed journal articles. We augmented the SSRS items with related measures from the Mathematica survey instrument used for the 21st Century Community Learning Centers study. The behavioral questions were all yes/no questions. To create summary measures of these behavior variables we computed the means of the positive and negative subsets of these questions. These are referred to as "Good Behaviors" and "Bad Behaviors" in our results section below.

---

8. This measure has received mixed reviews from Kessler (2000) but we were unable to identify an alternative.

These Gates-MacGinitie Reading tests, and the SSRS, API, and PASS surveys are proprietary so the question wording cannot be included in this paper. The rest of the Parent Survey is included in Appendix E.

Process Outcomes: We designed the parent survey to also elicit information describing the child, their activities during the summers of 2004 (briefly) and 2005 (in great detail), and parent activities. Our questions were based in part on existing surveys including the ones used by Mathematica to evaluate the 21st Century Community Learning Centers program and by the Urban Institute in their National Survey of American Families.

Analysis Methods

This analysis focuses on estimating the impacts of being in the treatment group rather than the control group using multivariate regressions where the key outcome is the post-program test score and the key independent variable is being randomly assigned to the treatment group. These are referred to as "Intent to Treat" impacts to distinguish them from impacts of participation in the program. The control variables consist of indicator variables for the randomization pools which vary by grade-level, city, and randomization round. The test scores are adjusted for the number of school days between when the program ended and when the student took the post-test (see Appendix B for details). Similar models are estimated for our other outcomes.

The basic model is described below.

$$Yij = aj + Xij'BX + eij \quad \text{(individual level)}$$

$$aj = a + Tj'BT + ej \quad \text{(family level)}$$

where Yij = the outcome (e.g. post-program test scores for individual i in family j),

Xij = control variables that vary by individual (randomization pool),

Tj= Treatment status (1 if in treatment group, 0 otherwise),

eij = unobserved factors at the individual level,

*e*j = unobserved factors at the family level and

BT and BX are parameters to be estimated.

This can be written in one equation as:

$$Yij = a + Tj'BT + Xij'BX + ej + eij$$

The primary research question is whether the BELL program improves reading test scores during the summer (i.e. whether BT>0). This hypothesis is tested using the full sample and standardized reading scores. Observations for which we are missing the reading test or treatment status are dropped. Similar methods are used to estimate impacts of the program on academic self-esteem from the student surveys and child behaviors as reported by parents. Exploratory analyses test for interactions with the background characteristics and for the robustness of the findings to variations in model specification.

One interaction of particular interest is for the treatment status with the grade level of the student. Recent evidence has come to our attention suggesting that impacts of learning interventions might be much larger in the early grades than in later grades. For example, Logerfo et al. (2006) find that test score growth is much smaller in later grades than it is in earlier grades, both in terms of scale scores and relative to the standard deviation in test scores. If these growth rates are taken as estimates of the impacts of regular schooling this suggests that school has much larger impacts on these types of test scores in earlier grades. The test norms for the Gates-MacGinitie tests suggest that similar patterns are found for this test. Finally, impacts of small class size also appear to be larger in the earlier grades than they are in the later grades based on

results from the Tennessee Star experiment (Krueger and Whitmore 2002). All of this suggests that the BELL program might also have larger impacts in earlier grades compared to later grades.

These data are clustered within families and the weights vary across observations. The 835 students with test score and date of test data come from 689 families. The weight for the student tests has a standard deviation of 0.72 when standardized to have a mean of 1. Hence Proc SurveyReg in SAS is used to control for clustering within families and for differential weighting across observations. Without these adjustments the standard error for the estimated impact of being in the treatment group on the total reading test score in table 5 decreases by around 11 percent.

In order to interpret the impact estimates it is also important to understand how the activities of the control and treatment groups differ. If, for example, few impacts of the BELL program were found, it could have been the case that the control group students participated in activities that were very similar to the activities promoted by the BELL program. Thus, finding out how the activities of the control and treatment group differ provides a context for understanding and interpreting the impact evaluation estimates. Methods similar to those described above are used to answer this set of questions by analyzing how the control and treatment groups spent their out-of-school time during the summer, with a special focus on academically enriching activities. Data for this aspect of the evaluation come from the survey of parents of both the treatment and control group children.

A common concern about many evaluation studies is that the research questions or methods used are changed based on the findings. This can invalidate statistical tests if care is not taken. For this reason it is important to note that the research questions and methods used here

are very similar to those proposed when we applied for funding for this project. Minor changes that were made are described in Appendix A.

Background Variables

Our sample population appears to be fairly disadvantaged educationally based on a number of measures (see table 1). The average income is below $30,000 per year, only about 40 percent have fathers who have attended college, less than 40 percent live with their fathers, and over 90 percent are minorities (Black or Hispanic). On the other hand around 60 percent have mothers who have attended college and over 90 percent live with their mothers.

These data suggest no statistically significant differences between the treatment and control groups based on observed characteristics. However, there were significant differences in the degree to which the treatment and control groups provided data. The information on family background (income, education, and who the child lives with) was all collected from the parents after random assignment. Consequently it was possible for these differences to arise. Data on the child's race and gender were collected from BELL also after random assignment and thus could also be affected, though there were no statistically significant differences in reporting of race or gender by treatment status. Since all of these variables could have been affected by the treatment status and since statistically significant differences were found for many of them, we do not use these variables as controls in most of our regressions, though we do estimate a few models (reported in Appendix C) controlling for these variables and find similar results.

Table 1
Mean Characteristics by if in Treatment Group

Characteristic	Predicted Means		Difference		
	Controls	Treatment	Value	P-Value	
Annual Family Income	$28,950	$26,467	-$2,483	0.18	
Mother has Post-Secondary Ed	0.59	0.61	0.02	0.69	
Father has Post-Secondary Ed	0.43	0.39	-0.04	0.36	
Lives with Mother	0.93	0.92	-0.01	0.69	
Lives with Father	0.35	0.39	0.04	0.36	
Black	0.51	0.45	-0.06	0.10	
Hispanic	0.38	0.42	0.04	0.26	
Asian, White, or Other	0.06	0.07	0.01	0.62	
Male	0.48	0.51	0.03	0.45	
Missing Value Dummies					
Family Income	0.11	0.18	0.074	0.01	* * *
Mother's Education	0.17	0.27	0.099	0.01	* * *
Father's Education	0.34	0.36	0.020	0.62	
Living with Mother	0.07	0.17	0.104	0.01	* * *
Living with Father	0.07	0.17	0.104	0.01	* * *
Race	0.14	0.13	-0.011	0.72	
Gender	0.08	0.09	0.013	0.67	

Notes:

Weights control for differential randomization probabilities and non-response.

Predicted mean is actual mean for control group.

Predicted mean for treatment group equals mean for control group plus estimated impact of treatment.

All models include dummy variables for randomization pools.

* means statistically significant at the .10 level, ** at .05 and *** at .01.

The treatment and control groups were also very similar based on their grade levels, sites, and rounds of randomization (see table 2). They are fairly evenly distributed across grade levels with slightly more students in grades 2 and 3 than in the other grades and somewhat fewer in grade 6.[9] Just over half of the sample is at one site with the remainder distributed across the two remaining sites. Most of the sample comes from the last round of randomization and less than 10 percent from the first round.

Table 2
Grade, Site, and Round
by Treatment Status

	Fraction with Characteristic	
Characteristic	Controls	Treatment
Grade 1	0.17	0.18
Grade 2	0.21	0.21
Grade 3	0.21	0.22
Grade 4	0.15	0.14
Grade 5	0.15	0.15
Grade 6	0.11	0.10
Site 1	0.26	0.24
Site 2	0.21	0.21
Site 3	0.53	0.54
Round 1	0.07	0.08
Round 2	0.38	0.37
Round 3	0.54	0.55

Weighted to sample in study.

Numbers do not always sum to 100 because of rounding.

None of these differences are statistically significant.

The similarities between the treatment and control groups based on these characteristics are virtually guaranteed because randomization was done within the grade, site, and randomization round categories. These variables (and all possible interactions between them) are

9. The grade levels are the grades for testing. Students entering grades 6 and 7 both took the grade 6 test.

used in all of the models presented below to help reduce the standard errors of the resulting estimates. Without these controls the standard errors in table 5 would be about twice as large. The grade level dummies by themselves explain over 70 percent of the variation in test scores.

Results

Table 3 presents the estimated impacts of being in the treatment group compared to the control group on summer activities. It includes predicted means for the treatment and control groups, the estimated impacts of being in the treatment group, the effect size, the p-value, the statistical significance levels of the estimated impacts, and the sample sizes. The predicted mean for the control group is their actual mean. The predicted mean for the treatment group is the mean for the control group plus the estimated impact. The effect size is the estimated impact divided by the standard deviation for the control group. More detailed results are presented in Appendix D, tables D1 to D4.

Table 3
Estimated Impacts of being in the Treatment Group on Summer Activities

Outcome	Predicted Means Treatment	Control	Std Dev	Estimated Impact	Standard Error	Effect Size	P Value			N
Fraction participated in BELL program	0.65	0.06	0.26	0.59	(0.04)	2.27	0.00	*	* *	847
Days in BELL Program	15.18	1.40	5.99	13.78	(0.76)	2.30	0.00	*	* *	840
Hours of Academics per Week in July	18.14	11.76	15.58	6.38	(1.39)	0.41	0.00	*	* *	769
Books Read in July	10.51	6.64	7.75	3.87	(1.44)	0.50	0.01		* *	677
Fraction of children doing activity during a typical week in July of 2005 between 8 a.m. and 5 p.m.										
TV/Computer Games	0.18	0.51	0.50	-0.33	(0.035)	-0.66	0.00	*	* *	762
Chores	0.10	0.32	0.47	-0.22	(0.031)	-0.47	0.00	*	* *	762
Camp	0.06	0.24	0.43	-0.18	(0.029)	-0.42	0.00	*	* *	762
Other academic activity	0.08	0.24	0.43	-0.16	(0.026)	-0.37	0.00	*	* *	762
Hanging out with friends	0.06	0.22	0.41	-0.16	(0.026)	-0.40	0.00	*	* *	762
Internet/Computer	0.06	0.19	0.39	-0.13	(0.026)	-0.33	0.00	*	* *	762
Museum/Cultural activity	0.10	0.20	0.40	-0.10	(0.025)	-0.26	0.00	*	* *	762
School	0.10	0.17	0.37	-0.07	(0.024)	-0.19	0.00	*	* *	765
At Another Academic Program	0.05	0.19	0.40	-0.14	(0.028)	-0.35	0.00	*	* *	765
At Home	0.18	0.43	0.49	-0.25	(0.035)	-0.51	0.00	*	* *	765
With Parent	0.22	0.46	0.50	-0.24	(0.036)	-0.48	0.00	*	* *	760
With Program Staff	0.76	0.35	0.48	0.41	(0.035)	0.86	0.00	*	* *	760
Time Child was without anyone over the age of 12 during July of 2005										
Hours per week	0.07	0.08	0.30	-0.01	(0.022)	-0.03	0.73			672
Fraction of parents doing an activity during a typical week in July of 2005 between 8 a.m. and 5 p.m.										
Adult literacy classes	0.03	0.05	0.25	-0.02	(0.018)	-0.09	0.14			742
Computer classes	0.19	0.10	0.34	0.09	(0.040)	0.28	0.02	*	*	742

Notes:

Weights control for differential randomization probabilities and non-response.

Effect sizes are equal to impact of treatment over standard deviation for control group.

Predicted mean is actual mean for control group.

Predicted mean for treatment group equals mean for control group plus estimated impact of treatment.

All models include dummy variables for randomization pools.

* means statistically significant at the .10 level, ** at .05 and *** at .01.

Summer Activities: The random assignment lottery had a large impact on participation in the BELL program. Most of the treatment group attended BELL (65 percent) and those who attended did so for most of the time. In contrast and as expected, the control group had very low participation rates in BELL. The first line of table 3 shows that being in the treatment group increased participation rates in BELL by 59 percentage points.[10] This translates to approximately 14 additional days in attendance at the BELL program.

While the large impacts of being in the treatment group on participation in BELL are important it should be noted that application for the BELL program may signify that a family has a great deal of interest in promoting the academic achievement of their children. Consequently it is not surprising that even among control group members (all of whom applied to the BELL program), participation in academic activities during the month of July 2005 was substantial. We focus on July 2005 because the program operated during all of that month, starting just after July 4th.

On average parents in the control group reported that their children spent almost 12 hours per week in academic activities and read 6.6 books in July. That said, being in the treatment group did increase both of these numbers by about 50 percent; hours per week in academic activities by 6.4 hours per week and books read by 3.9. Some parents reported that their children read as many as 50 books during the month. This is possible especially given that some books

10. Both participation and days in the program are based on data obtained from the BELL program. Parents were also asked four questions about participation in the BELL program and their answers agreed with those of the BELL program over 90 percent of the time for three of the four questions. Some differences would be expected for parents whose children only participated briefly in BELL. In addition, it appears that the parents may have misunderstood the fourth question as the agreement rate for that question dropped to around 75 percent. This last question asked, "…did your child attend one or more summer learning programs (including BELL)…" It is possible that some parents assumed they should say no unless their child attended two or more programs.

are very short. However, to test the robustness of these results we also estimated a model setting the maximum number of books to 10. The impact of being accepted into the treatment group was still moderately large (at 1.10 books) and statistically significant at the 1 percent level.

The impact on hours of academic activities is largest during the hours when the BELL program provides academic activities (weekday mornings). Positive impacts were also found for weekday afternoons and evenings, though not for weekends (see table D1 in Appendix D). Since the program goes for six weeks, an addition of 6.4 hours per week translates to about 40 additional hours of academic activities during the summer. This is about the same as 10 additional days in school (assuming four hours of academic activities per day in school).

In order to participate in BELL members of the treatment group had to reduce their time in other activities. The largest impacts were found on time spent watching TV and playing computer games, doing chores, attending camp, other academic activities, hanging out with friends, Internet/Computer use, and cultural activities. Small reductions were found for participation in other academic programs; sports, music, or arts; volunteer work; and religious activities (see table D3). No impacts were found for summer school or caring for other children. One site administrator told us that the BELL program counted as summer school. This would explain the lack of a negative impact on participation in this activity. The negative impact of being in the treatment group on participation in academic programs and activities other than BELL might be expected to partly offset the positive impacts of BELL on the treatment group.

Participation in the BELL program also had impacts on where the members of the treatment group spent their time and whom they were with. As table 3 shows, treatment group members were less likely to be at home, in school, or in another program, though 19 percent of

the control group was in another program and 17 percent was in summer school.[11] The treatment group was also less likely to be with their parents and more likely to be with program staff. No impacts were found on hours per week spent without someone over the age of 12. This last result also holds during the weekday hours when the BELL program was in operation (see table D2).

Parents were also asked a number of questions about their adult education activities during the summer (see bottom of table 3 and top of table D4 for details). Estimated impacts were generally statistically insignificant with the exception that treatment group members had a 9 percent point higher probability of taking computer classes suggesting that the BELL program may have facilitated this type of educational activity, perhaps by serving as a source of childcare for the parents. This result should be interpreted with caution, however, as the estimated impact on a summary measure of parent adult education activity variables was not statistically significant.

Academic Self-Concept: The BELL program aims to improve academic performance in part by increasing student academic self-concept. To measure this construct students were given the API and PASS measures for grades 1–2 and 3–6, respectively. As shown in table 4, the estimated impacts on these two measures of academic self-concept were statistically insignificant and fairly small based on their effect sizes.

11. The program variable includes both academic and non-academic programs as it comes from Question 3 of the parent survey.

Table 4
Estimated Impacts of being in the Treatment Group on Post-Program Outcomes

Outcome	Predicted Means Treatment	Control	Estimated Impact	Standard Error	Effect Size	P Value	N
Data from Students							
Academic Self Concept							
Grades 1-2 (-36 is lowest possible, + 36 is highest)	26.70	25.08	1.62	(1.356)	0.14	0.23	326
Grades 3-6, English (0 is lowest possible, 1 is highest)	0.69	0.71	-0.02	(0.019)	-0.11	0.25	517
Data from Parents							
Parent Activities							
Encourages child to read (1=never, 4=daily)	2.59	2.48	0.11	(0.052)	0.15	0.04 * *	725
Reads books to child (1=never, 4=daily)	1.95	1.78	0.17	(0.075)	0.21	0.02 * *	580
Limits TV watching (1=never, 4=daily)	2.19	2.14	0.05	(0.095)	0.05	0.63	660
Child Behaviors							
Believes can succeed in school (1=never, 3=often)	1.76	1.73	0.03	(0.035)	0.06	0.42	750
Tries to solve math problems (1=never, 3=often)	1.64	1.61	0.03	(0.041)	0.05	0.54	723
Gets along with others (1=strongly disagree, 4=strongly agree)	3.36	3.43	-0.07	(0.061)	-0.11	0.23	754
Is happy (1=strongly disagree, 4=strongly agree)	3.52	3.49	0.03	(0.055)	0.05	0.56	756

Notes:

Weights control for differential randomization probabilities and non-response.

Effect sizes are equal to impact of treatment over standard deviation for control group.

Predicted mean is actual mean for control group.

Predicted mean for treatment group equals mean for control group plus estimated impact of treatment.

All models include dummy variables for randomization pools.

* means statistically significant at the .10 level, ** at .05 and *** at .01.

Other Behaviors: The BELL program does not only aim to improve test scores and academic self-concept. It also has the goal of improving parent involvement and other child behaviors. These concepts were measured using parent surveys. Parents were asked a number of questions about their own behaviors and those of their children. Table 4 presents a small subset of those findings. The only statistically significant impacts were for the parents encouraging their children to read and actually reading to their children. Table D4 includes a number of summary measures of child behaviors. Estimated impacts on the summary measures are also statistically insignificant. The overall statistical significance of the results for parents encouraging reading at the time of data collection aligns well with the large impacts on books read during the summer found above. For more details see tables D3 and D4.

Test Scores: The major outcome for this analysis is reading test scores. As noted earlier, there was a small but important difference in the average number of school days between when the BELL program ended and student testing between the treatment and control groups. On average the control group had 16 days more of regular school. While this difference is small relative to the total period of data collection, this is a large number of days compared to the 14-day difference in participation in the BELL program between the treatment and control groups (see table 3). Consequently, a straight comparison of test scores is, in effect, comparing the effect of 14 days of BELL to the effect of 16 days in school. In addition, as discussed above there was a difference in non-BELL academic activities between the treatment and control groups. In total, the difference between the treatment and control groups in hours of academic activities during the summer was similar to just ten days in school.[12] Indeed, if the BELL program had no impacts

12. The treatment/control group difference was around 6.4 hours per week, or around 38 hours during the 6-week summer program. The members of the treatment group spent around 18 hours per

one might expect a negative effect of being in the treatment group since the control group received 16 additional days in school before being tested. Instead, as shown in the first three lines of table 5, the treatment and control groups received similar scores on the reading tests.

One interpretation of this result is that 14 days in BELL produces about the same amount of change in test scores as 16 days of regular school. An alternative interpretation is that our estimates are not precise enough to detect a difference. To test for this latter possibility we corrected the test scores for days in school so that the results would give us an estimate of the impact of BELL given no difference between the treatment and control groups in days in school. As shown in table 5, once this correction is made, the estimated impact on test scores is positive and statistically significant at around 5.2 points on the extended scale score of the Gates-MacGinitie reading test.

---

week in academic activities. If we assume that they would spend a similar amount of time in academic activities during the summer then two weeks of school (10 days) translates to around 36 hours of academic activities.

Table 5
Estimated Impacts of being in the Treatment Group on Reading Test Scale Scores

Outcome	Predicted Means Treatment	Predicted Means Control	Standard Deviation	Estimated Impact	Standard Error	Effect Size	P Value		N
Compared to Control Group with 16 Additional Days of School									
Total Score	423	422	65	1.05	(2.42)	0.02	0.67		847
Vocabulary	420	422	67	-1.89	(2.54)	-0.03	-0.73		847
Comprehension	418	417	73	1.46	(3.22)	0.02	0.45		847
Compared to Control Group with Same Level of Schooling									
Total Score	418	413	69	5.16	(2.46)	0.08	0.04	* *	835
Vocabulary	416	413	68	2.53	(2.51)	0.04	0.31		831
Comprehension	414	408	76	6.06	(3.19)	0.08	0.06	*	835

Notes:

Sample sizes vary because of missing values on days in school since being in BELL program or norming information.

Weights control for differential randomization probabilities and non-response.

Effect sizes are equal to impact of treatment over standard deviation for control group.

Predicted mean is actual mean for control group.

Predicted mean for treatment group equals mean for control group plus estimated impact of treatment.

Standard Deviation is for the control group.

All models include dummy variables for randomization pools.

* means statistically significant at the .10 level, ** at .05 and *** at .01.

The estimated effects shown in the second panel of table 5 are of the 14 additional days of BELL attendance that the treatment group got compared to the control group. It does not differ statistically from the 3.4 point gain we would expect from 14 additional days in regular school.[13] Thus, we find no evidence that the impacts of participation in the BELL program differ from those of attending a similar number of days of regular school. This is in spite of the fact that control group members were more likely than the treatment group to participate in academic activities outside of the BELL program. This also implies that the results in the first three lines of table 5 were precise enough to show a negative impact, if BELL had not improved student learning.[14]

The effect size in the fourth line of table 5 is around 0.08. While this appears small it should be kept in mind that this represents the impact of only a 0.59 shift in participation rates. The implied effect of participation in a single summer in the BELL program for participants is 0.14 standard deviations or around 8.8 points. The impact of participation can also be estimated using an instrumental variables model with being accepted into the program serving as the instrumental variable for participation (Angrist and Rubens 1996). Similar results were found.[15] In comparison Borman and Dowling (2006) estimate that participation in the Baltimore summer

13. As noted in Appendix B, the average growth rate is 43 points per year and the school year is 178 days. 3.4 = (14/178)*43.

14. The sample sizes using the adjusted test scores in Table 5 are smaller than those using the unadjusted scores because of missing data on when the test was administered. When the unadjusted model is run without those cases the results are very similar (statistically insignificant results with coefficients between –1 and +1).

15. The coefficient on participation was 8.79 with a standard error of 3.71 meaning that the result is statistically significant at the 5 percent level. This model was estimated using the same controls as used in table 5 and the standard errors were adjusted for clustering. In order to estimate the impact of participation, as opposed to the impact of being in the treatment group, this model makes somewhat stronger assumptions than the models used in the rest of this paper (Angrist and Ruebens 1996).

program for at least two years increased test scores by around 0.30 standard deviations, just over twice the impact of a single summer of BELL.

It should also be noted here that the effect sizes for test scores might have been larger had we been estimating impacts for a smaller range of grades. The effect sizes reported in table 5 are based on standard deviations of test scores for grades 1 through 6 of the control group combined. The standard deviation within a grade-level is on average less than half as large at around 32 as compared to 69 for the overall reading test adjusted for days in school.

Interactions: We found no evidence of differential impacts of being in the treatment group on total reading test scores by grade, site, race, or gender (estimated separately). The estimated impact for grade 3 was positive and statistically significant at the 1 percent level but the estimated impacts for grades 2 and 5 were negative, though not statistically significant. The estimated impact for one of the sites was large, positive, and statistically significant at the 10 percent level. However, statistical tests of the interactions between being accepted and these sets of variables had p-values of 0.27, 0.28, 0.97 and 0.74 respectively meaning that none of the interactions were statistically significant. At the same time it should be kept in mind that these tests for interactions are not very precise because the study was not designed to have enough data to estimate such interactions.

Robustness Tests: The test score results are also robust to a number of alternative model specifications. Models that allow for larger test score growth show larger impacts of BELL while those that allow for less test score growth show smaller impacts. We note that our preferred method of estimating test score growth, used in the second panel of table 5, is conservative in that it is likely to underestimate growth and, consequently, to underestimate impacts of BELL (see end of Appendix B for a detailed discussion of this).

Our results are somewhat weaker with alternative test score metrics, but none has all of the benefits of the metric used here (scale scores). The results generally get stronger, as expected, when the sample is limited to subsets with higher differences between the treatment and control groups in participation in the BELL program. In addition, the results remain similar when alternative weights and sets of control variables are used and when impacts are estimated using a hierarchical linear model (see Appendix C for details).

Discussion

The BELL program staff agreed to participate in a random assignment evaluation. Random assignment was conducted and the results suggest that the evaluation itself was successful in the sense that the treatment and control groups did not differ based on baseline characteristics but the treatment group participated in BELL at a much higher rate than the control group.

Child and parent behaviors during the summer were measured and large changes were observed between the treatment and control groups that are consistent with the high participation rate differences. For example, in order to participate in BELL students must give up other types of activities. We found that students in the BELL treatment group spent more time reading and participating in academic activities during the summer and less time with their parents and in non-academic activities than the control group.

Parents of students in the BELL program are encouraged to read to their children and to sign logs indicating how much their children have read during the summer program (T. Cooper 2003). We found positive impacts on how much parents encouraged their children to read after the program was over and no clear impacts on other measures of parent involvement that were not directly targeted by the program, such as the degree to which they limit TV watching or eat

with their children. An increase in parent involvement in reading aligns well with the program goals and is an important outcome, though the fact that it is self-reported means that this result should be interpreted with some caution.

The BELL program also included components designed to improve non-academic student behaviors. On the other hand, participation in BELL would likely increase time spent interacting with other children and decrease the amount of time spent with parents. This could have a negative impact on measured child behaviors. These counteracting effects may explain why we found no impacts on non-academic behaviors such as a summary of bad behaviors and whether the children were happy, sad, depressed, able to get along with others, self-confident, or creative.

We found no impacts of the BELL program on academic self-perceptions. While this might be viewed as cause for concern it should be noted that, with the exception of the students in grade 1, on average these children score well below grade-level based on the Gates-MacGinitie tests (see table B2 for details). Participation in the program may have provided the children with important information about their academic problems. While this may reduce their academic self-concept, it might also inspire them to work harder to improve these skills for a negligible net impact on their academic self-concept. In addition, the API survey, administered to the students in grades 1 and 2, had a high level of top coding, around 22 percent, which may have reduced the potential for large impacts.

We also note that while we think the PASS is a valid and reliable instrument, BELL staff told us that they have used this instrument in the past and found no changes during the summer months for the summer participants, even though test scores were rising. In addition, the PASS manual states that changes should not be expected over such a short period of time. Larger changes might be found for an instrument that measures the somewhat narrower concept of

academic self-efficacy, as opposed to the more global measure of academic self-concept, which is captured in PASS.[16] In addition, there are data to suggest that academic achievement outcomes are perhaps more closely related to academic self-efficacy (Bandura et al. 1999). Unfortunately, we were not able to identify an appropriate instrument to measure academic self-efficacy.

Our results suggest that the treatment group learned about one more month worth of reading skills than the control group. This result depends on how we control for growth in student test scores over time but our assumptions about growth were conservative suggesting that our estimates of the impacts of BELL are also likely to be low. The test score results were robust to a number of alternative model specifications (with and without additional control variables and using a variety of weights) and, as expected, became stronger when subsets of the data were analyzed that had higher treatment/control group differences in BELL participation rates.

When analyzed based on cost-effectiveness BELL appears to be similar to regular school. The BELL summer program costs around $1,500 per student to implement (parents pay far less because the program receives external funding). Our estimates suggest that participants gain around 8.7 points on the reading scale test during the 6-week program for a cost of $170 per point. We also estimate that these applicants to the BELL program would gain about 43 points per year on the reading test used here in the absence of the BELL program (see Appendix B for details). This would cost around $7,310 to produce based on the cost per point ratio of BELL, well within the annual per pupil cost of the regular public schools.[17]

---

16. Self-efficacy refers to the sense that one can accomplish certain tasks (Bandura 1997) whereas self-concept is a broader construct that includes feelings and assumptions about ones self that can vary independently of ones ability to accomplish specific tasks.

17. Indeed this amount is lower than the public school average but it should be noted that BELL is not serving special education students.

T. Cooper (2002, 2003) reports a few months of growth in reading scores among participants in the BELL summer learning program. This is consistent with the findings reported here. We estimate that the treatment group only experienced about one more month of test score growth than the control group but not all members of the treatment group participated in BELL and some members of the control group did. When we adjust for these factors (using the instrumental variables method) we estimate that participating in the BELL summer learning program improves test scores by around two months. In addition, the control group spent two thirds as much time in academic activities during the summer as the treatment group, suggesting that the control group also probably experienced test score growth.

The BELL summer program had a strong academic focus, used a reading curriculum developed by a well-regarded national firm, and had highly qualified teachers. In addition, as noted above, parents were encouraged to become engaged in their children's learning. On the other hand, a noticeable fraction of the treatment group in our study did not participate in the BELL summer program and the control group children, all of whom had also applied for the BELL program, were also engaged in a substantial amount of academic activity during the summer. In spite of these factors, strong impacts were found on reading test scores. In future work we hope to test to see if these impacts can also be detected on longer-term outcomes.

These results are also of interest in light of the lack of impacts found for the 21st Century Community Learning Centers after school programs (James-Burdumy et al. 2005). While the BELL after-school program is probably very different from the typical 21st CCLC program, it is also possible that the marginal benefits of additional academic activities during the summer may be higher than during the school year. Testing the impacts of a summer learning program

compared to an after-school program with similar costs is another area that could benefit from additional future research.

Conclusion

This paper reports estimated impacts of a summer learning program on a variety of outcomes. The estimates suggest that the program has important benefits on academic outcomes including reading test scores, hours of academic activities, reading during the summer and the degree to which parents encouraged reading at the time the data were collected. No impacts are found on academic self-concept or other child behaviors. The impact of the program on test scores appears similar to that of a similar amount of school and is precise enough that it is unlikely that it was caused by chance. If the true impact were 0 there would only be a 4 percent chance of observing an impact of this size. Based on cost BELL appears to be about as cost-effective as regular public schools.

In summary these estimates suggest that the BELL program, as it was implemented in the summer of 2005, had important impacts on summer learning activities, parent involvement in reading, and on reading test scores. These results are of particular importance given the public policy focus on raising achievement levels of low-income students and since they were based on a rigorous random assignment study. Programs that work are hard to find. Our results suggest that the BELL program has positive and substantively important impacts. These findings may also be relevant for other summer learning programs like BELL and for questions related to extending the regular school year.

References

Advisory Committee on Head Start Research and Evaluation. 1999. "October 1999 Report Evaluating Head Start: A Recommended Framework for Studying the Impact of the Head Start Program." Washington, DC: U.S. Department of Health and Human Services.

Agodini, Robert, and Mark Dynarski. 2001. "Are Experiments the Only Option? A Look at Dropout Prevention Programs." Washington, DC: Mathematica Policy Research Inc..

Angrist, Joshua, Guido Imbens, and Don Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association***,** Applications Invited Paper.

Bandura, A. 1997. *Self-Efficacy: The Exercise of Control.* New York: Freeman.

Bandura, A., Pastorelli, C., Barbaranelli, C. and Caprara, G.V. 1999. "Self Efficacy Pathways to Childhood Depression." *Journal of Personality and Social Psychology* 76(2): 258–69.

BELL. 2004. "Overview of the BELL Accelerated Learning Summer Program." Building Educated Leaders for Life program document.

Bloom, Howard S., Charles Michalopoulos, Carolyn J. Hill, and Ying Lei. 2002. "Can Nonexperimental Comparison Group Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs?" New York City: Manpower Demonstration Research Corporation Working Paper on Research Methodology.

Borman, Geoffrey, and N. Maritza Dowling. 2006. "Longitudinal Achievement Effects of Multiyear Summer School: Evidence from the Teach Baltimore Randomized Field Trial." *Educational Evaluation and Policy Analysis* 28(1): 25–48.

Capizzano, Jeffrey, Kendra Bischoff and Duncan Chaplin. 2005. "Reversing the Summer Slide: Experimental Evidence: Early Findings from a Process Study of the Building Educated Leaders for Life (BELL) Accelerated Learning Summer Program." Washington, DC: The Urban Institute.

Capizzano, Jeffrey, Kendra Bischoff, Duncan Chaplin, and Nicola Woodroffe. 2006. "Reversing the Summer Slide: Experimental Evidence, A Process Study of the Building Educated Leaders for Life (BELL) Accelerated Learning Summer Program." Washington, DC: The Urban Institute.

Chaplin, Duncan. 2006. "Generalizing Results from an Evaluation of the BELL Summer Learning Program," Mathematica Policy Research, Prepared for the annual meeting of the American Education Finance Association, Denver, Colorado, March 24th.

Chaplin, Duncan, and Maggie Stanislawski. 2005. "Spring 2005 Progress Report on the Evaluation of the BELL Summer Learning Program." Washington, DC: The Urban Institute.

Cooper, Tiffany. 2002. "BELL Accelerated Learning Summer Program 2001: Evaluation Report." Dorchester, MA. Building Educated Leaders for Life (BELL).

———. 2003. BELL Accelerated Learning Summer Program 2002: National Evaluation Report. Dorchester, MA. Building Educated Leaders for Life (BELL).

———. 2004. "BELL After School Program: 2003–2004 National Program Outcomes." Dorchester, MA. Building Educated Leaders for Life (BELL).

Cooper, Harris, Barbara Nye, Kelly Charlton, James Lindsay, and Scott Greathouse. 1996. "The Effects of Summer Vacation on Achievement Test Scores: A Narrative and Meta-Analytic Review." *Review of Educational Research* 66(3): 227–68.

Cooper, Harris, Kelly Charlton, Jeff C. Valentine, and Laura Muhlenbruck. 2000. "Making the Most of Summer School: A Meta-Analytic and Narrative Review." *Monographs of the Society for Research in Child Development* 1(201): 1–109.

Dynarski, M., Moore, M., Mullens, J., Gleason, P., James-Burdumy, S., Rosenberg, L., et al. 2003. "When Schools Stay Open Late: The National Evaluation of the 21st-Century Community Learning Centers Program." Washington, DC: Mathematica Policy Research.

Entwisle, Doris R., and Karl L. Alexander. 1992. "Summer Setback: Race, Poverty, School Composition, and Mathematics Achievement in the First Two Years of School." *American Sociological Review* 57:72-84.

Glazerman, S., Levy, D., and Myers, D. 2003. "Nonexperimental Replications of Social Experiments: A Systematic Review." Princeton, NJ: Mathematica Policy Research, Inc.

Hay, I., Ashman, A.F., and van Kraayenoord, C.E. (1997) "Investigating the Influence of Achievement on Self-Concept Using an Intra-Class Design and a Comparison of the PASS and SDQ-1 Self-Concept Tests." *British Journal of Educational Psychology* 67: 311–21.

Jacob, Brian A., and Lars Lefgren. 2002. "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis." National Bureau of Economic Research Working Paper 8918, May.

James-Burdumy, Susanne, Mark Dynarski, and John Deke. 2005. "When Elementary Schools Stay Open Late: Results from The National Evaluation of the 21st Century Community Learning Centers Program." Princeton, NJ: Mathematica Policy Research, Inc.

Kessler, Carol E. 2000. "Review of the Academic Perceptions Inventory [2000 Revision]." Buros Institute of Mental Measurements, Test Reviews Online. http://www.unl.edu/buros.

Krueger, Alan, and Diane Whitmore. 2002. "Would Smaller Classes Help Close the Black-White Achievement Gap?" In *Bridging the Achievement Gap*, edited by J. Chubb and T. Loveless. Washington, DC: Brookings Institute Press.

Logerfo, Laura, Austin Nichols, and Sean Reardon. 2006. "Achievement Gains in Elementary and High School." Washington, DC: The Urban Institute.

Imbens, Guido, and Joshua Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2): 467–75.

Michalopoulos, Charles. 2005. "Precedents and Prospects for Randomized Experiments." In *Learning More from Social Experiments, Evolving Analytic Approaches,* edited by Howard Bloom. New York: Russell Sage Foundation.

Murnane, Richard, and Richard Nelson. 2005. "Improving the Performance of the Education Sector: The Valuable, Challenging, and Limited Role of Random Assignment Evaluations." , Working Paper no. 11846. Cambridge, MA: National Bureau of Economic Research.

National Research Council and Institute of Medicine. 2002. *Community Programs to Promote Youth Development.* Washington, DC: National Academy Press.

Patterson, J., and T. Czajkowski. 1979. "Implementation: Neglected Phase in Curriculum Change." *Educational Leadership* 37: 204–6.

Sunmonu, Kola, John Larson, Yolanda Van Horn, Elizabeth Cooper-Martin, and Jennifer Nielsen. 2002. *Evaluation of the Extended Learning Opportunities Summer Program.* Rockville, MD: Office of Shared Accountability, Montgomery County Public Schools.

U.S. Department of Education, National Center for Education Statistics. 1996. "Assessing the Social Competence, Adaptive Behaviors, and Approaches to Learning with Young Children." Working Paper no. 96-18, by Samuel J. Meisels, Sally Atkins-Bunett, and Julie Nicholson. Jerry West, project officer. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Wilde, Elizabeth Ty, and Robinson Hollister. 2002. "How Close Is Close Enough? Testing Nonexperimental Estimates of Impact against Experimental Estimates of Impact with Education Test Scores as Outcomes." Discussion Paper no. 1242-02. Madison, Wisconsin. Institute for Research on Poverty.

Appendix A

Changes from Original Analysis Plans

The analysis methods used in this paper are very similar to those described in the research grant applications for this project submitted to potential funders in 2003 and 2004. There were a few changes caused by unforeseen circumstances. In this appendix we describe the additions and subtractions from the original analysis plan.

Additions

Adjusting Test Scores for Time in School: In the original proposal we were going to control for time in school using time as a covariate. As explained in Appendix B we now feel that it is more appropriate to adjust test scores based on the norming data provided by the test developer.

Race: We did not expect to have data on race. Since those data were available for a large fraction of our sample we have included them in some analyses.

Robustness Tests: We did not originally plan to conduct the robustness tests given in Appendix C. However, we believe that they are important given the apparent sensitivity of our findings.

Subtractions

Pre-Participation Outcomes: The proposal mentioned controlling for pre-participation outcomes. Lacking such data this could not be done. However, we still hope to collect administrative data from the Boston and New York City school systems. If those data are obtained they will be used to control for the pre-participation test scores in future analyses.

Time between Pre and Post-Test: The proposal mentioned controlling for the time between the pre- and post-tests. Lacking a pre-test the analyses presented here control for the time between when the summer program ended and the post-test.

2004 Data: The original plan was to use data on applicants to both the 2004 and 2005 BELL summer programs. Random assignment and data collection was conducted in 2004 but because of low response rates on the student tests in 2004 only the 2005 data are analyzed here (see Chaplin and Stanislawski [2005] for more details on the 2004 data collection). If school district administrative data can be obtained then the 2004 data will be used in later analyses.

Math Tests: The original plan mentioned administering math and reading tests. Due to cost considerations only reading tests were administered. If administrative records with math tests can be obtained those will be analyzed in future analyses.

Clustering by Classroom: The original proposal mentioned controlling for clustering by classroom. This adjustment would likely result in larger standard errors and was not done in this paper. This affects the interpretation of the findings somewhat. It means that in order to generalize to a larger set of classrooms or staff we need to assume that the current staff and classrooms BELL had in place in 2005 at the sites covered by this study were fairly typical.

Dropping Data with Missing Control Variables: The original proposal mentioned dropping cases that were missing "any of the primary control variables." These variables have been excluded from the main analysis because of concerns about endogeneity of missing values, as discussed above. When they are included missing values are replaced with zeros and a dummy variable is used to indicate these cases.

Fully Implemented Model: The original plans called for using measures of implementation by classroom to estimate the impact of a fully implemented model. Given cost

considerations we were only able to visit a subset of classrooms and only visited each classroom once. Consequently we do not have adequate information to measure program implementation precisely. Based on the data we did have our overall assessment was that the program was well implemented (Capizzano et al. 2005, 2006). In addition we find no evidence of differential impacts across sites.

Grades and Attendance: In the original proposal we mentioned estimating models with grades and school attendance as outcomes. We will estimate such models in the future if we are able to obtain those data from the New York and Boston school districts.

Appendix B

Adjusting for Time in School before Test

As explained in the text, the control group was in school on average for about 16 days more than the treatment group. Consequently their test scores were likely increased by this additional education. In order to adjust for this we had planned to include days in school as a regressor in a multivariate regression model. After further thought we determined that this might produce biased results if test scores are positively impacted by time in school but the time when data are provided is negatively correlated with unobserved factors that improve test scores. For example, parents who do not answer the phone often and who do not keep appointments will be harder to collect data from and may also not provide their children with as much academic support. Thus, the coefficient on days in school will capture both the positive impacts of time in school and any negative impacts of these unobserved characteristics. The net coefficient will be smaller than it would have been if it were only capturing the impacts of schooling. For this reason using time in school as a covariate would underestimate the benefits of regular school obtained by the control group.

The test developer recommends a method for adjusting for time in school based on quarter-months rather than on days in school. This would not be precise enough for the purposes of this study given the need to adjust for only 16 days in school, on average. Fortunately, the test developers also provide information on the grade equivalent of each test score value. This enabled us to estimate expected growth per year based on the current test score of each student. To do this we first smoothed the data that matched test scores to grade equivalents[18] and then

18. These data were provided for each test level by number of test questions answered correctly. Within each test level the number of questions answered correctly mapped uniquely to the scale score but

calculated an estimate of the annual growth rate for each tenth of a grade using data on test scores one tenth of a grade level above and below that point. Using this information, and the days in school since the beginning of the school year (calculated using the school calendars for the New York City and Boston school districts), we estimated their score at the beginning of the year—just after the BELL program ended. More precisely we calculated

Adjusted Test Score=

Original Test Score – (Days in school/178)*Expected Growth

This adjustment factor might be too large if these students obtain lower test score growth than the average student in the population used to create the norms reported by the test developers. For this reason we adjusted the growth rates down by the ratio of their actual grade equivalent score divided by the grade-level they were entering. Thus,

Expected Growth=

Expected Growth Given Score * (Grade Equivalent Score/Current Grade)

Table B1 below shows how the expected growth in test scores varies between the treatment and control groups and by subset. The first three rows show expected growth not adjusted for the grades the students are in. These rates are higher for the control group than the treatment group. However, once they have been adjusted for their grade-equivalent performance, their expected growth rates are similar, as shown in the last three rows of table B1.

Being in the treatment group could impact the grade level of the student in the 2005–2006 school year and, thus, the level of the test taken. We estimated models to test for this, with

---

not to the grade equivalent score since the latter were only in tenths of grades. Thus, one grade-equivalent score generally mapped onto many possible scale test scores both within grades and across test levels. We combined all of these norming data and took the average scale score reported for each grade-equivalent score. This gave us one scale score for each grade-equivalent score (measured in tenths of grades).

and without controls for the randomization pool, and found no evidence of impacts on the level of the test taken.

In our main models we did not use days in school as a control variable because we were concerned that this method would underestimate the impacts of days in school and consequently underestimate the impacts of the BELL program. To test for this possibility we estimated a model based on the scale score not adjusted for days in school but using days in school as a control variable in the regression. The estimated impact of being in the treatment group on the total scale score unadjusted for days in school is smaller than in the original model and statistically not significant. The reason is that the coefficient on days in school is only 0.13 scale score points implying an average annual growth rate of only 23 points. In contrast expected annual growth is almost twice as large, at 43 points. As discussed earlier, the smaller coefficient on days in school probably occurs because this variable is capturing both the impact of days in school and the impacts of any unobserved characteristics that affect both when data are provided and test scores. These unobserved factors might bias the estimated effect of days in school downwards if, for example, parents who were harder to reach during the data collection period were also providing their children with less academic support.

The adjustments used here are likely to underestimate the impacts of the BELL program for a number of reasons. First, we are calculating estimated test score growth based on the current test score and not based on the test score average during the period in question which was likely somewhat lower. Adjusting for this would yield a higher test score growth estimate and, consequently, a higher estimated impact of BELL. Second, we assumed that test score growth occurred only during school. There was also a substantial treatment/control group difference in days out of school between the end of the BELL program ended and when testing occurred. This

difference also favored the control group. Finally, we assumed that everyone enters first grade with grade equivalent scores. The numbers in table B2 supports this assumption. If, in fact, many of the children in the higher grades entered first grade already below grade level then their test score growth rates were probably somewhat higher then we are calculating here, again resulting in larger estimated impacts of the BELL program.

Table B1
Predicted Annual Test Score Growth

Type/Subject	Control	Treatment	All
Unadjusted			
Vocabulary	52	49	50
Comprehension	58	57	57
Total	56	51	54
Adjusted for Grade Level			
Vocabulary	37	37	37
Comprehension	43	42	42
Total	43	44	43

Notes:

Expected annual test score growth given individual score.

Adjusted for Grade Level based on growth per grade to date.

See text for details.

Weights control for differential randomization probabilities and non-response.

Table B2
Tenths of Year ahead of Grade Level
Based on Total Reading Test Score

Grade	Grade Level Differences
1	0.00
2	-0.26
3	-0.79
4	-0.91
5	-1.41
6	-1.02

Notes:

Weights control for differential randomization probabilities and non-response.

Appendix C

Robustness Tests

In this Appendix we discuss alternative ways of adjusting the test scores for days in school, alternative test score metrics, and alternative models specifications.

Alternative Ways of Adjusting Test scores for Days in School

The estimated impacts on test scores are somewhat sensitive to how we adjust for when the data were collected. In order to adjust for the fact that the control group spent more time in school before being tested than the treatment group, the results presented in the second panel of table 5 are based on reading test scores adjusted for days in school using expected annual growth rates based on each student's test score and grade level (see Appendix B for details). The estimated impact of the BELL program goes up if we assume higher growth during school and down if we assume lower growth. This is shown table C1. The second line is the method used in table 5, panel 2, which is the estimate of the impact of BELL compared to a control group with the same number of days of regular school. It uses information on the translation between scale scores and grade equivalents provided by the test developer to estimate typical growth in a year and adjusts this estimate of growth down to account for the fact that most of these children are below grade-level. This method suggests that on average these students gain about 43 points per year. However, if they were on-grade level their growth would be closer to 50 points per year. The first line of table C1 shows that, as expected, if their test score growth were 50 points per year during the regular school year, the estimated impact of the BELL program would also be somewhat larger. In contrast, however, the third line of table C1 shows that if expected growth were substantially lower, at only 30 points per year, the estimated effect of BELL would be smaller and not quite statistically significant at the 10 percent level.

Table C1
Using Alternative Adjustments for Days in School
Estimated Impacts of being in Treatment Group
on Total Reading Scale Score

Adjustment Method	Assumed Annual Test Score Growth	Estimated Impact	p-Value	
Not adjusting for grade in school	50	5.73	0.02	* *
Adjusting for Grade in School (same as Table 5, Part 2)	43	5.16	0.04	* *
Lower Growth	30	3.95	0.11	
No Growth (same as Table 5, Part 1)	0	1.05	0.67	

Notes:
First two lines use expected test score growth given score without and with adjustments for grade.
Weights control for differential randomization probabilities and non-response.
All models include dummy variables for randomization pools.
* means statistically significant at the .10 level, ** at .05 and *** at .01.

Alternative Test Score Metrics

The test developer provided two metrics that can easily be adjusted for days in school—the scale score used in table 5 and grade equivalents. The grade equivalent scores are measured in tenths of grade levels, from 0.0 to 12.0 and are adjusted for days in school by subtracting an estimate of their expected growth rate during this period. This is estimated using the ratio of days in school over 178 (the average annual days in school) multiplied by the grade equivalent test score divided by the test level (i.e. (days in school/178)*(grade equivalent unit/test level)). The second ratio adjusts for the fact that students behind in grade level are expected to be learning at a slower rate.

Table C2 presents estimated impacts on test scores in grade level equivalent units adjusted for days in school. The estimated impact on total scores is not significant but the standard error is large. The estimated impact on comprehension scores measured using grade equivalent units is statistically significant at the 10 percent level.

Table C2
Using Grade Equivalent Scores adjusted for days in school
Estimated Impacts of being in Treatment Group
on Reading Test Scores

Variable	Estimated Impact	
Grade Equivalent Score[a]		
Total	0.07	
Vocabulary	-0.07	
Comprehension	0.15	*

Notes:
Weights control for differential randomization probabilities and non-response.
All models include dummy variables for randomization pools.
* means statistically significant at the .10 level, ** at .05 and *** at .01.

The scale scores used in tables 5 and C1 and the grade equivalent units presented in table C2 can be easily adjusted for days in school. This is not the case for the other test score metrics available from the test developer (National Curve Equivalents and Percentiles). In addition, the test developer does not recommend using any of these metrics in regressions with multiple grade levels because the scale score is the only metric that is designed to have units that can be compared both within and across grade levels (the NCE scores can also be compared within grade levels, but not across). For these reasons we focus our analyses on the scale scores.

Alternative Model Specifications

The estimated impacts on test scores also held up well to a number of other model specifications. First, we estimated models based on subsets of the data with relatively high participation rates in the BELL program for the treatment group. Second, we estimated impacts controlling for additional background factors. Third, we estimated models using a variety of alternative weights. Finally, we tried controlling for the level of the test that was taken by the student.

Table C3
Using Subsamples by Randomization Pools with Largest Differences in Participation Rates between Treatment and Control Groups
Estimated Impacts of being in the Treatment Group on Reading Scale Score

Subsample	Participation Rate Difference	Estimated Impact	Std Error	P-Value	
Total (N=835)	0.59	5.16	2.46	0.04	* *
628	0.63	2.87	2.87	0.32	
440	0.65	7.78	3.42	0.02	* *
185	0.71	13.79	5.72	0.02	* *
79	0.78	12.33	7.42	0.10	

Notes:
First row presents results for full sample, sample model as in Table 2.
Subsequent rows present subsets of data taking randomization pools
with largest differences in BELL participation rates between the treatment and controls.
Reading scale score is adjusted for days in school.
Weights control for differential randomization probabilities and non-response.
All models include dummy variables for randomization pools.
* means statistically significant at the .10 level, ** at .05 and *** at .01.
N means sample size.

Table C3 shows that the estimated impacts generally increase when subsets of the data are analyzed with higher differences in the participation rates in BELL between the treatment and control groups. To estimate these models we calculated participation rates in the treatment and control groups by site and grade and then took the grade/site combinations with the highest participation rate differences. We made a number of cuts giving participation rate differences ranging from 59 percent (for the full dataset) up to 78 percent, for a subset with only 79 cases. As shown in table C3, the estimated impacts of the program generally increase with the participation rate differences and remain statistically significant even when the subset with only 185 cases is analyzed. While this is encouraging it should be noted that this method might yield biased results because the resulting subsets of the data are endogenously chosen. The same

factors that generated the large participation rate differences between the treatment and control groups may have also generated differences in the outcomes, independent of the impacts of the BELL program.

In table C4 we show that the estimated impacts on the total reading test score hold up to controlling for background factors (at the 10 percent level), to estimating the model with a variety of weights, and to controlling for the level of the test taken. In addition, estimated impacts on the comprehension sub-score are often statistically significant. The first three lines of table C4 give the preferred specification (from table 5, panel 2), without controls for the background variables from table 1 (gender, race, parent education, family structure, and family income). Age was not included because we had data on that variable for less than half of the sample. As shown in table 1, the background variables are missing differentially for the treatment and control groups. Thus, they are endogenous. Not surprisingly, controlling for these variables weakens the estimated impact of the program slightly, though it remains statistically significant at the 5 percent level. If the control for family income alone is dropped the estimated impact is reduced slightly further but remains statistically significant at the 10 percent level.

The results above are all weighted to the full sample of applicants who agreed to be in the study in each randomization pool. If we instead weight to the full set of applicants, including those not in the study, the point estimate decreases slightly but the results remain statistically significant, as shown in rows seven to nine of table C4.

Unweighted results are unbiased because dummy variables for each randomization pool are included in the regression model. However, the unweighted results apply to a population that differs from the weighted results. This is because the weights vary across observations because of differences in the probability of being accepted into the program and in non-response rates. As

shown in table C4, the unweighted results are very similar to the weighted results. In addition, the estimated impact on the comprehension sub-test moves from being statistically significant at the 10 percent level to being statistically significant at the 5 percent level.

Table C4
Varying weights and control variables
Estimated Impacts of being in Treatment Group
On Reading Scale Scores and Subtests

Model/Outcome	Estimated Impact	P-Value	
Baseline Model (weighted)			
Total	5.16	0.04	* *
Vocabulary Sub-Test	2.53	0.31	
Comprehension Sub-Test	6.06	0.06	*
Baseline Model **Controlling for Background Factors**			
Total	4.90	0.04	* *
Vocabulary Sub-Test	2.23	0.36	
Comprehension Sub-Test	5.69	0.07	*
Baseline Model **weighted to all applicants** randomized			
Total	4.94	0.05	* *
Vocabulary Sub-Test	2.51	0.32	
Comprehension Sub-Test	5.32	0.10	
Baseline Model **Unweighted**			
Total	5.40	0.03	* *
Vocabulary Sub-Test	2.58	0.33	
Comprehension Sub-Test	6.50	0.04	* *
Baseline Model Estimated using **HLM** Method			
Total	5.16	0.02	* *
Vocabulary Sub-Test	2.53	0.28	
Comprehension Sub-Test	6.06	0.03	* *
Baseline Model **controlling for Level of Test**			
Total	5.04	0.04	* *
Vocabulary Sub-Test	2.37	0.34	
Comprehension Sub-Test	5.92	0.06	*

Notes:

Weights control for differential randomization probabilities and non-response.

All models include dummy variables for randomization pools.

* means statistically significant at the .10 level, ** at .05 and *** at .01.

These models are estimated using ordinary least squares regressions with standard errors adjusted for the variation in weights and clustering within families. The information on clustering within families can be used to re-weight the data and thereby improve the precision of the estimates. This can be done using Hierarchical Linear Models (HLM) also known as random effects models in the economics literature when only the intercept varies. The estimated impacts of being in the treatment group on test scores become slightly more statistically significant when estimated using an HLM model because the standard errors are reduced by around 10 percent. The HLM model was estimated using Proc Mixed in SAS. This reduction in standard errors is somewhat surprising since the coefficient estimates did not change. The lack of change in the coefficient estimates suggests that the re-weighting did not matter given the control variables in the model. The reduction in the standard errors may be related to differences in how Proc SurveyReg (used in the rest of the models) and Proc Mixed deal with weights.

All of the regressions control for the randomization pool which is determined by the grade the student completed in the previous school year. Since most students progress on to the next grade level this also determines the level of the test they took. However, in a very small number of cases students took a different test. In most cases this was because they had been held back in school. Since the scale scores used here are designed to be comparable this control should not matter and, indeed, the estimated impact remains about the same and statistically significant when we include test level indicator variables as controls (see table C4) or a continuous test level variable.

Appendix D

Supplementary Tables

This appendix has supplementary tables that contain additional details and outcomes not provided in the tables in the main body of the report. The sample sizes in these tables vary mostly because of item non-response in the parent survey. In addition, we could not calculate an expected test score growth for four students in vocabulary because their scores were higher than the range for which the test developer provided precise grade equivalent information. Thus, their cases are missing when using scale scores in vocabulary adjusted for days in school. We were able to calculate expected test score growth for these students for their total test score. We were also missing test scores adjusted for days in school for 10 cases because the time of data collection was missing.

Table D1
Hours of Academic Activities and Academic Self Concept for Grades 3-6
Estimated Impacts of Being in BELL Treatment Group

Outcome	Control Group Mean	Control Group Std Dev	Estimated Impact	Standard Error	Effect Size	P-Value	N
Hours of Academics Per Week (7/05)							
Total	11.76	15.58	6.38	(1.385)	0.41	0.00 * * *	769
Weekday Mornings	5.05	8.80	2.94	(0.840)	0.33	0.00 * * *	743
Weekday Afternoons	3.00	6.85	1.82	(0.600)	0.27	0.00 * * *	756
Weekday Nights	2.80	6.15	1.22	(0.515)	0.20	0.00 * * *	740
Weekends	1.26	3.74	0.94	(0.935)	0.25	0.20	761
Academic Self Concept (Grades 3-6)							
English	0.71	0.18	-0.02	(0.019)	-0.11	0.25	517
General	0.76	0.23	0.01	(0.021)	0.04	0.51	515
Math	0.69	0.20	0.00	(0.024)	0.00	0.94	517

Notes:

Weights control for differential randomization probabilities and non-response.

All models include dummy variables for randomization pools.

* means statistically significant at the .10 level, ** at .05 and *** at .01.

Table D2
Where Child Was and Who Child was With
Estimated Impacts of Being in BELL Treatment Group

Survey Question/Response (check all that apply)	Control Group Mean	Control Group Std	Estimated Impact	P-Value	Effect Size	Sample Size
Where was your child between 8 a.m. and 5 p.m. during a typical week in July 2005?						
Own home	0.43	0.49	-0.25	0.00 * * *	-0.51	765
Home of someone else	0.25	0.43	-0.19	0.00 * * *	-0.44	765
BELL	0.08	0.28	0.65	0.00 * * *	2.32	765
School	0.17	0.37	-0.07	0.00 * * *	-0.19	765
Another Program	0.19	0.40	-0.14	0.00 * * *	-0.35	765
Somewhere else hanging out	0.06	0.23	-0.07	0.00 * * *	-0.30	765
At another place	0.16	0.37	-0.12	0.00 * * *	-0.32	765
Who was your child with?						
Parent	0.46	0.50	-0.02	0.00 * * *	-0.05	760
Program staff	0.35	0.48	0.41	0.00 * * *	0.85	760
Other adults	0.22	0.41	-0.15	0.00 * * *	-0.37	760
Siblings 13-18 years old	0.09	0.29	-0.04	0.00 * * *	-0.14	760
Siblings under age 13	0.11	0.31	-0.08	0.00 * * *	-0.26	760
Other youth 13-18 years old	0.05	0.22	-0.04	0.00 * * *	-0.18	760
Other kids under age 13	0.12	0.32	-0.07	0.00 * * *	-0.22	760
Alone	0.01	0.09	0.00	0.70	0.00	760
Someone else	0.14	0.35	-0.09	0.00 * * *	-0.26	760
How many hours per day did your child spend without anyone over 12 during July of 2005?						
During weekdays?	0.53	3.84	-0.14	0.75	-0.04	668
During weeknights?	0.30	3.10	0.15	0.62	0.05	670
During weekends?	0.60	6.09	-0.06	0.83	-0.01	673
For how many weeks was your child on vacation during July of 2005?	0.45	0.97	-0.08	0.26	-0.08	742

Notes:

Weights control for differential randomization probabilities and non-response.

All models include dummy variables for randomization pools.

* means statistically significant at the .10 level, ** at .05 and *** at .01.

Table D3
What Child was Doing and Parent Involvement
Estimated Impacts of Being in BELL Treatment Group

Survey Question/Response (check all that apply)	Control Group Mean	Control Group Std	Estimated Impact	P-Value	Effect Size	Sample Size
What was your child doing during a typical week from 8 a.m.-5 p.m. during July of 2005?						
Summer school	0.08	0.27	-0.02	0.26	-0.07	762
BELL	0.08	0.28	0.64	0.00 ***	2.29	762
Other academic program	0.12	0.33	-0.08	0.00 ***	-0.24	762
Camp	0.24	0.43	-0.18	0.00 ***	-0.42	762
Other academic activity	0.24	0.43	-0.16	0.00 ***	-0.37	762
Sports, music, arts	0.11	0.31	-0.03	0.00 ***	-0.10	762
Museum/Cultural activity	0.20	0.40	-0.10	0.00 ***	-0.25	762
TV/Computer games	0.51	0.50	-0.33	0.00 ***	-0.66	762
Internet/Computer	0.19	0.39	-0.13	0.00 ***	-0.33	762
Hanging out with friends	0.22	0.41	-0.16	0.00 ***	-0.39	762
Volunteer work	0.03	0.17	-0.02	0.02 **	-0.12	762
Religious activities	0.13	0.33	-0.05	0.05 **	-0.15	762
Chores	0.32	0.47	-0.22	0.00 ***	-0.47	762
Caring for children	0.03	0.16	0.00	0.77	0.00	762
Doing something else	0.11	0.31	-0.06	0.00 ***	-0.19	762
How often did you do the following last week? (1=never, 4=daily)						
Eat dinner with child	2.59	0.65	0.03	0.58	0.05	719
Encourage child to read	2.48	0.74	0.11	0.04 **	0.15	725
Read book to child	1.78	0.82	0.17	0.02 **	0.21	580
Visit library with child or check out books for child	1.40	0.63	-0.04	0.62	-0.06	383
Limit amount of TV child can watch	2.14	0.88	0.05	0.63	0.06	660
Tell child what to watch on TV	2.58	0.76	0.09	0.19	0.12	642
Mean of Parent Involvement (6 variables above)	2.21	0.50	0.10	0.00 ***	0.20	759
Mean of book related behaviors	1.99	0.62	0.14	0.00 ***	0.23	738

Notes:

Weights control for differential randomization probabilities and non-response.

All models include dummy variables for randomization pools.

* means statistically significant at the .10 level, ** at .05 and *** at .01.

Table B4

Parent Educational Activities and Child Behaviors
Estimated Impacts of Being in BELL Treatment Group

Survey Question/Response (check all that apply)	Control Group Mean	Std	Estimated Impact	P-Value	Effect Size	Sample Size
Did you participate in the following activities in July of 2005? (Yes/No)						
Adult literacy classes	0.05	0.25	-0.02	0.14	-0.08	742
Parenting classes	0.05	0.24	0.01	0.60	0.04	742
Employment counseling, training, and placement	0.14	0.39	0.02	0.66	0.05	742
Computer classes	0.10	0.34	0.09	0.02 * *	0.26	742
Other adult development	0.08	0.31	0.01	0.67	0.04	742
Mean Response	0.08	0.19	0.02	0.19	0.12	742
How often does your child do any of the following things? (1=never, 3=often)[a]						
Acts sad or depressed	1.11	0.31	-0.03	0.37	-0.10	406
Gets angry easily	1.19	0.39	-0.06	0.16	-0.15	459
Reads books on own	1.48	0.50	-0.01	0.86	-0.02	511
Doesn't listen to what others say	1.16	0.37	0.03	0.51	0.08	697
Tries to solve math problems independently	1.61	0.49	0.03	0.54	0.06	723
Believes he/she can succeed in school	1.73	0.45	0.03	0.42	0.07	750
Mean of Bad Behaviors (based on 6 question responses)	1.11	0.24	0.00	0.99	0.00	715
Mean of Good Behaviors (based on 21 question responses)	1.48	0.25	0.01	0.76	0.04	768
Mean of Academic Behaviors	1.61	0.36	0.02	0.55	0.06	759
How much do you agree with the following statements about your child? (1=strongly disagree, 4=strongly agree)						
Gets along with others	3.43	0.70	-0.07	0.23	-0.10	754
Likes school	3.42	0.67	-0.03	0.61	-0.04	754
Works hard at school	3.34	0.73	0.00	0.97	0.00	756
Is self-confident	3.23	0.75	0.03	0.65	0.04	747
Is creative	3.41	0.67	-0.03	0.65	-0.04	747
Is happy	3.49	0.65	0.03	0.56	0.05	756
Mean of Good Behaviors	3.35	0.57	-0.01	0.83	-0.02	765
Mean of Academic Behaviors	3.38	0.65	-0.01	0.87	-0.02	762

Notes:

Weights control for differential randomization probabilities and non-response.

All models include dummy variables for randomization pools.

* means statistically significant at the .10 level, ** at .05 and *** at .01.

a Only 6 of the responses to this question are reported here. The remainder are proprietary questions from the Social Skills Rating System survey. The SSRS responses are included in the means.

Appendix E

BELL 2005 Parent Survey

Building Educated Leaders for Life (BELL) STUDY
PARENT/GUARDIAN SURVEY 2005

THE URBAN INSTITUTE

ATTACH LABEL HERE
PARENT/Guardian OF Student Name
Student ID
Parent/Guardian ID

This survey covers the activities of children who applied for the 2005 BELL summer learning program as well as some background information about their schools and families. The questions on this survey are about one of your children. You may receive more than one of these surveys if you have more than one child selected for our study. Please fill out this survey for the child listed on the label above. We ask that you fill out a different survey for each child.

This survey is part of a study about the loss of knowledge that children experience over the summer and how the BELL summer learning program can help to prevent this summer learning loss and improve child outcomes more generally. The study is being conducted by the Urban Institute, a nationally recognized research organization in Washington, D.C. with a long history of work in the area of out-of-school time activities for youth. Results from the study will be used to improve summer programs for children across the country.

The Urban Institute wants to protect the privacy of individuals who participate in surveys. Your answers will be combined with other surveys to describe the BELL experience. Your answers will be confidential and no one else will know how you answered the questions.

If you have any questions about the study, you may contact Dr. Duncan Chaplin, the Study Director, at 1-866-450-6651 or write to him at The Urban Institute, 2100 M Street, NW, Washington, DC 20037 or at Dchaplin@ui.urban.org.

As part of the study, we are testing/surveying all children/parents who have applied for the BELL program. We are asking for your consent for your child to be tested and for you and your child to be surveyed. Please read the following statement and sign if you agree.

I consent to participate in the survey and to allow my child to be tested and surveyed. I understand that my participation is voluntary and will not affect my child's participation in the BELL program in 2005. I understand that I can stop participating at any time or refuse to answer questions on the survey.

Please print your name: ☐ Mr. ☐ Ms. _____ _____
First Name Last Name

_____ _____
Signature Date

Thank you very much for helping us to learn more about children and their summer experiences.

Instructions

When answering the survey questions you may use either a pen or pencil. Please mark your response or write in the requested information for each question. This survey will be hand-coded, so you may simply check the boxes and neatly write-in your responses as appropriate; however, we would greatly appreciate every attempt to make your answers as clear as possible.

BELL Study Parent Survey

Your Child

1. What is your relationship to the child named on the cover label? Are you his or her:
MARK ONE ANSWER

- ☐ Mother
- ☐ Stepmother
- ☐ Grandmother
- ☐ Other female relative
- ☐ Foster mother or other female guardian
- ☐ Father
- ☐ Stepfather
- ☐ Grandfather
- ☐ Other male relative
- ☐ Foster father or other male guardian

2. Did you submit an application for this child to participate in the 2005 BELL summer program?

- ☐ Yes
- ☐ No

* * *

THIS SECTION ASKS YOU TO TELL US ABOUT YOUR CHILD'S ACTIVITIES DURING THIS SUMMER. PLEASE THINK ABOUT A TYPICAL WEEK IN JULY WHEN ANSWERING THESE QUESTIONS.

Summer Activities 2005

The next question asks where your child was between 8 a.m. and 5 p.m. during a typical week in **July 2005**. ***MARK ALL THAT APPLY.***

3. Where was your child?

Between 8 a.m. and 5 p.m., my child was:	**Monday- Friday**
a. At home	☐
b. At the home of someone else (like a friend, relative, or sitter)	☐
c. At the BELL Summer Learning Program	☐
d. At a program at school	☐
e. At another program someplace else	☐
f. Someplace else to "hang out"	☐
f. Other: **Where?** _____	☐

The next two questions ask who your child was with, and what s/he was doing between 8 a.m. and 5 p.m. during a typical week in **July 2005**. ***MARK ALL THAT APPLY.***

4. Who was your child with?

Between 8 a.m. and 5 p.m., my child was with the following people:	**Monday- Friday**
a. A parent or guardian	☐
b. Staff at a program (i.e. YMCA, Summer camp, BELL, etc.)	☐
c. Other adults over 18 who are not the child's parent or guardian	☐
d. Brothers or sisters age 13-18	☐
e. Brothers or sisters 12 or younger	☐
f. Other youth age 13-18.	☐
g. Other children 12 or younger.	☐
h. My child was alone	☐
i. Someone else: **Who?** _____	☐

5. What was your child doing?

Between 8 a.m. and 5 p.m., my child:	**Monday-Friday**
a. Attended *required* summer school	☐
b. Attended the BELL summer program	☐
c. Attended another program or went to tutoring to improve academic skills	☐
d. Attended a summer camp	☐
e. Did independent academic activities that were not for school or a program	☐
f. Participated in activities such as sports, music, art, drama, dance, or clubs such as the Boys & Girls Club.	☐
g. Visited museums or other cultural sites	☐
h. Watched TV/videos or played video games	☐
i. Surfed the Internet or did other things on a computer	☐
j. Hung out with friends	☐
k. Did volunteer work or community service	☐
l. Did religious activities	☐
m. Did chores around the house	☐
n. Took care of a sister or brother or other children	☐
o. Other: **What?** _____	☐

6. Did your child take a vacation away from home in July 2005?

☐ Yes
☐ No….. **SKIP NEXT QUESTION**

7. IF YES, for how many weeks?

1 week…...☐ 3 weeks…☐
2 weeks…..☐ 4 weeks…☐

8. Did your child do any doing learning activities (reading, writing, homework, classes) in July of 2005? ☐ Yes ☐ No

IF YES PLEASE FILL IN THE APPROXIMATE HOURS PER DAY. ENTER 0 IF NONE.

a. Weekday mornings before noon ☐☐ Hours/day
b. Weekday afternoons from noon to 5:00 p.m. ☐☐ Hours/day
c. Weekday evenings after 5:00 p.m. ☐☐ Hours/day
d. Weekend days ☐☐ Hours/day

9. Did your child read any books in July of 2005? ☐ Yes ☐ No
IF YES PLEASE FILL IN THE NUMBER.

☐☐ Books

10. Did your child spend any time at home without anyone over the age of 12 during July of 2005? ☐ Yes ☐ No
IF YES PLEASE *FILL IN THE NUMBER OF HOURS PER DAY. ENTER 0 IF NONE*.

a. Weekdays Before 5:00 p.m. ☐☐ Hours/day
b. Weekdays after 5:00 p.m. ☐☐ Hours/day
c. Weekend days ☐☐ Hours/day

* * *

THIS SECTION ASKS YOU TO TELL US ABOUT YOUR CHILD'S PARTICIPATION IN LEARNING PROGRAMS THIS SUMMER. A "SUMMER LEARNING PROGRAM" IS ANY PROGRAM WITH AN ACADEMIC COMPONENT (I.E. WRITING, READING, MATH).

Summer Learning Programs 2005

11. During this summer (2005), did your child attend one or more summer learning programs (including BELL)?
MARK ONE ANSWER

- ☐ Yes
- ☐ No….**SKIP TO QUESTION 20**

12. IF YES, in which weeks in July and August of 2005 did your child attend the program(s)?
MARK ALL THAT APPLY

1st Week: July 5-9…… ☐
2nd Week: July 12-16…… ☐
3rd Week: July 19-26……☐
4th Week: August 2-6….. ☐
5th Week: August 9-13….☐
6th Week: August 16-20...☐

13. How much time did your child usually spend attending those summer-learning programs or centers each day? ***FILL IN APPROXIMATE HOURS PER DAY. ENTER 0 IF NONE.***

a. Monday-Thursday ☐☐ Hours/day
b. Friday ☐☐ Hours/day
c. Weekend days ☐☐ Hours/day

ANSWER THE FOLLOWING SIX QUESTIONS FOR THE PROGRAM THAT YOUR CHILD ATTENDED FOR THE LONGEST PERIOD OF TIME DURING THE SUMMER OF 2005.

14. How did you find out about the summer learning program your child attended for the longest period of time during the summer of 2005? ***MARK ALL THAT APPLY***

- ☐ From the school
- ☐ From another parent or guardian
- ☐ From a community organization
- ☐ From somewhere else: **Where?** _____

15. What is the name of this summer learning program?

Please circle the category/name and write in the name of the specific program/site.

Category/Name	Program/Site
General	
BELL	_____
Boys and/or Girls Club	_____
Church program	_____
Girl and/or Boy Scouts	_____
Police program	_____
Summer School	_____
Other School program	_____
YMCA/YWCA	_____
Other	_____
Boston Only	
ACCESS	_____
Alerta/Verano	_____
Blackstone Community Camp	_____
Boston Catholic Chinese Community	_____
Boston Centers for Youth and Families (BCYF)	_____
Brookview House	_____
Camp Unity	_____
Derby	_____
Ella J. Baker House Academic Sports Camp	_____
Federated Dorchester Neighborhood Houses	_____
ID Tech Camp	_____
IDIIL Learning Center	_____
Kennedy Memorial School	_____
New England (Summer) Scores	_____
Newton Summer Success Program	_____
Harvard (Phillips Brooks House Association)	_____
R.I.S.E. Summer Program	_____
Roxbury Youth	_____
Stepping Stone Academy	_____
Tenacity	_____
New York Only	
Alianza Dominicana Inc.	_____
Beacon	_____
ChaRosa Foundation Corporation	_____
Frederick Douglas Center	_____
Harlem Zone	_____
New York City Housing Authority (NYCHA)	_____
Pathways for Youth	_____
Reading Institute	_____
South Brooklyn Youth Consortium	_____
Summer Freedom Day Camp	_____
Sunshine Neo-Elite Day Camp	_____

16. Why was your child in this summer learning program?
CHECK THE <u>THREE (3)</u> MOST IMPORTANT REASONS

- ☐ My child wanted to go
- ☐ School staff suggested that my child enroll
- ☐ It provides affordable care
- ☐ It provides dependable and safe care
- ☐ It will help my child do better in school
- ☐ It will help my child stay out of trouble
- ☐ Other: **What reason?** _____

17. How much do you agree or disagree with the following statements about this summer learning program? ***MARK <u>ONE</u> ANSWER IN EACH LINE***

	Strongly Disagree	Disagree	Agree	Strongly Agree
a. The program meets my child's individual needs	☐	☐	☐	☐
b. I am pleased with the number and variety of activities offered	☐	☐	☐	☐
c. My child usually likes going to the program	☐	☐	☐	☐
d. I am pleased with the quality of program staff	☐	☐	☐	☐
e. The program helps my child with reading, writing, and math	☐	☐	☐	☐

18. Did you participate in the following activities at this summer learning program?
MARK <u>ONE</u> ANSWER IN EACH LINE

	Yes	No	Not Offered
a. Attended parent meetings	☐	☐	☐
b. Parent-teacher conferences	☐	☐	☐
c. Field trips	☐	☐	☐
d. Helping out at the children's program	☐	☐	☐
e. Attended adult classes (i.e. literacy, parenting, employment, counseling, computing, etc.)	☐	☐	☐

19. What <u>overall grade</u> would you give this summer learning program?
MARK <u>ONE</u> ANSWER.

- ☐ A . . . Excellent
- ☐ B . . . Good
- ☐ C . . . Fair
- ☐ D . . .Unsatisfactory
- ☐ F . . . Failing

*** * ***

THIS SECTION NOW ASKS YOU TO TELL US ABOUT YOUR CHILD'S PARTICIPATION IN LEARNING PROGRAMS LAST SUMMER (2004).

Summer Learning Programs 2004

20. Did your child attend a summer learning program last summer (2004)?
MARK ONE ANSWER

- ☐ Yes
- ☐ No…**SKIP TO QUESTION 22**

21. How much of your child's time in this program was spent on academic activities (activities such as reading, writing or math) during the summer of 2004? ***MARK ONE ANSWER***

- ☐ Most of the time
- ☐ Some of the time
- ☐ Not much of the time
- ☐ None of the time
- ☐ Don't Know

* * *

THIS SECTION ASKS YOU TO TELL US ABOUT YOUR ACTIVITIES THIS SUMMER.

Parent/Guardian Activities 2005

22. How often did you do the following last week?
MARK ONE ANSWER IN EACH LINE

	Not at all	1 or 2 times	3 or 4 times	Daily
a. Eat dinner with your child	☐	☐	☐	☐
b. Encourage your child to read a book	☐	☐	☐	☐
c. Read a book to your child	☐	☐	☐	☐
d. Visit the library with your child or check out books for your child	☐	☐	☐	☐
e. Limit the amount of time your child could watch TV	☐	☐	☐	☐
f. Tell your child which shows he or she could or couldn't watch	☐	☐	☐	☐

23. Did you participate in the following activities during July of 2005?
MARK ONE ANSWER IN EACH LINE

	Yes	No
a. Literacy classes	☐	☐
b. Parenting classes	☐	☐
c. Employment counseling, training, and placement	☐	☐
d. Computer classes	☐	☐
e. Other adult development……..	☐	☐

Describe:_____

* * *

THIS SECTION ASKS YOU TO TELL US ABOUT YOUR CHILD'S ACTIONS AND EMOTIONS.

Child Behavior

24. How often does your child do any of the following things?
MARK ONE ANSWER IN EACH LINE

	Never	Sometimes	Often
a-u. <<Questions from the Social Skills Rating Scale, Elementary School Parent Form PE>>			
v. Acts sad or depressed	☐	☐	☐
w. Gets angry easily	☐	☐	☐
x. Doesn't listen to what others say	☐	☐	☐
y. Reads books on own.	☐	☐	☐
z. Tries to solve math problems independently	☐	☐	☐
aa. Believes he/she can succeed in school	☐	☐	☐

25. How much do you agree or disagree with the following statements about your child?
MARK ONE ANSWER IN EACH LINE

	Strongly Disagree	Disagree	Agree	Strongly Agree
My Child:				
a. Gets along with others	☐	☐	☐	☐
b. Likes school	☐	☐	☐	☐
c. Works hard at school	☐	☐	☐	☐
d. Is self-confident	☐	☐	☐	☐
e. Is creative	☐	☐	☐	☐
f. Is happy	☐	☐	☐	☐

* * *

THIS SECTION ASKS YOU TO TELL US A LITTLE BIT ABOUT YOUR FAMILY AND YOUR CHILD'S LIFE AT HOME.

About Your Family

26. Who does your child live with most of the time? ***MARK ALL THAT APPLY***

Adults (over age 18)
☐Mother
☐Other female Guardian
☐Father
☐Other male guardian
☐Grandparents
☐Other relatives
☐Other non-relatives

Children (age 18 and below)
☐Sisters (including stepsisters and half sisters)
☐Brothers (including stepbrothers and half brothers)
☐Other relatives
☐Other non-relatives

27. What is the main language you speak with your child?
MARK ONE ANSWER

- ☐ English
- ☐ Spanish
- ☐ Another language: **What language?_____**

28. How far in school do you expect your child will get? ***MARK ONE ANSWER***

- ☐ Less than a high school graduate
- ☐ Graduate from high school
- ☐ Graduate from a two-year college
- ☐ Graduate from college
- ☐ Go on after college and get a master's, doctorate, law, medical, or other advanced degree

29. Has your child moved to a new home in the last year? ☐ **Yes** ☐ **No**
IF YES PLEASE MARK ONE ANSWER

- ☐ 1 move
- ☐ 2 moves
- ☐ 3 or more moves

*** * ***

YOU ONLY NEED TO FILL OUT THIS PORTION OF THE SURVEY ONCE! IF YOU HAVE ALREADY FILLED OUT A SURVEY FOR ONE OF YOUR OTHER CHILDREN YOU ARE DONE WITH THE SURVEY FOR THIS CHILD. THANK YOU FOR YOUR HELP.

ABOUT THE MOTHER OR FEMALE GUARDIAN (THIS MAY BE YOU)

30. Does this child have a mother or female guardian?

Yes…☐
No….☐ **SKIP TO QUESTION 35**

31. What is the HIGHEST level of education the child's mother (or female guardian) has completed? ***MARK ONE ANSWER***

- ☐ 8th grade or less
- ☐ Some high school (did not graduate)
- ☐ High school equivalency (GED)
- ☐ High school graduate
- ☐ Vocational, trade, or business school after completing or leaving high school
- ☐ Some college (did not graduate)
- ☐ Graduated from a 2-year college (junior college or community college)
- ☐ 4-year college degree or other advanced degree

32. In what year was the child's mother (or female guardian) born? ***FILL IN THE YEAR***

19 ☐☐

33. Is the child's mother (or female guardian) employed? ☐ **Yes** ☐ **No**
If Yes are they working ☐ Full-time or ☐ Part-time ?
If No please mark one of the following.

- ☐ Not working now, but looking for work
- ☐ Not working now, and not looking for work
- ☐ Retired
- ☐ Disabled/unable to work

} *Skip next question if not working.*

34. How many hours/week does the child's mother or female guardian normally work for pay? ***FILL IN NUMBER OF HOURS. ENTER 0 IF NONE.***

a. 6 a.m.- 6 p.m. Monday-Friday? ..... ☐☐ Hours/week
b. Before 6 a.m. or after 6 p.m. Monday-Friday? ... ☐☐ Hours/week
c. Saturdays/Sundays? ..... ☐☐ Hours/week

ABOUT THE FATHER OR MALE GUARDIAN (THIS MAY BE YOU)

35. Does this child have a father or male guardian?

Yes…☐
No….☐ **SKIP TO QUESTION 40**

36. In what year was the child's father (or male guardian) born? ***FILL IN THE YEAR***

19 ☐☐

37. What is the HIGHEST level of education the child's father (or male guardian) has completed? ***MARK ONE ANSWER***

- ☐ 8th grade or less
- ☐ Some high school (did not graduate)
- ☐ High school equivalency (GED)
- ☐ High school graduate
- ☐ Vocational, trade, or business school after completing or leaving high school
- ☐ Some college (did not graduate)
- ☐ Graduated from a 2-year college (junior college or community college)
- ☐ 4-year college degree or other advanced degree

38. Is the child's father (or male guardian) employed? ☐ **Yes** ☐ **No**

If Yes are they working ☐ Full-time or ☐ Part-Time?

If No Please Mark one of the following:

- ☐ Not working now, but looking for work
- ☐ Not working now, and not looking for work
- ☐ Retired
- ☐ Disabled/unable to work

} Skip next question if not working.

39. How many hours/week does the child's father or male guardian normally work for pay? ***FILL IN THE NUMBER OF HOURS. ENTER 0 IF NONE.***

a. 6 a.m.- 6 p.m. Monday-Friday? ..... ☐☐ Hours/week
b. Before 6 a.m. or after 6 p.m. Monday-Friday? .... ☐☐ Hours/week
c. Saturdays/Sundays? ..... ☐☐ Hours/week

*** * ***

THIS SECTION ASKS YOU TO TELL US ABOUT YOUR HOUSEHOLD INCOME. IF YOU ARE UNSURE OF THE EXACT NUMBERS, PLEASE TRY TO ESTIMATE TO THE BEST OF YOUR ABILITY.

ABOUT THE HOUSEHOLD

40. What is your best guess of the yearly income of your household before taxes? Include all earnings from work (before taxes are deducted), pensions, and child support. Also include government assistance, such as cash welfare benefits, food stamps, Medicaid, or public housing. ***MARK ONE ANSWER***

☐	Less than $5,000	☐	$15,000 - $19,999	☐	$35,000 - $39,999
☐	$5,000 - $7,999	☐	$20,000 - $24,999	☐	$40,000 - $49,999
☐	$8,000 - $10,999	☐	$25,000 - $29,999	☐	$50,000 - $59,000
☐	$11,000 - $14,999	☐	$30,000 - $34,999	☐	$60,000 or more

Please enter your monthly income here if you do not know your annual income. $_____

41. Does anyone in this household receive any government assistance? ☐ Yes ☐ No

IF YES PLEASE MARK ONE ANSWER IN EACH LINE

	Yes	No
a. Food stamps	☐	☐
b. Welfare (also known as TANF or public assistance)	☐	☐
c. Social Security	☐	☐
d. Medicaid	☐	☐
e. Supplemental Security Income (SSI)	☐	☐
f. Subsidized government housing	☐	☐
g. Subsidized private housing	☐	☐

YOU ARE DONE WITH THIS CHILD'S SURVEY. IF YOU HAVE OTHER CHILDREN IN THE PROGRAM PLEASE FILL OUT ONE FORM FOR EACH ADDITIONAL CHILD (QUESTIONS 1-29 ONLY).

THANK YOU!!