

Evaluating the Academic and Behavioral Impact of “School in the Park”

Johanna R. Lacoë*

Gary D. Painter

Danielle Williams

Price School of Public Policy,
University of Southern California

March 9, 2014

Abstract

Arts enrichment programs that leverage existing community cultural institutions such as museums, zoos, and theaters have the potential to improve students’ educational experiences. This paper estimates the impact of School in the Park, a museum-based educational program for low-income students that takes place within the cultural institutions and museums of San Diego’s Balboa Park. This study evaluates the impact of participation in the program on short-term and long-term academic and behavioral outcomes using longitudinal, student-level data since 1996, the natural variation in the timing of program implementation between two elementary schools, and control groups of students from nearby schools that did not receive the program. Findings indicate that participation in the program has positive but limited short-term impacts and more enduring longer-term effects through high school. The results are broadly applicable to cities that seek to leverage existing cultural resources to improve educational outcomes for young people.

Key words: museum education program, academic outcomes, behavioral outcomes, longitudinal analysis, difference-in-difference analysis

*Corresponding author. Contact: Johanna Lacoë, Price School of Public Policy, University of Southern California, lacoë@price.usc.edu.

Introduction

Numerous school-based educational interventions are targeted at closing the achievement gap between students in low-income schools and those in high-income schools (such as universal preschool, full day kindergarten, after school enrichment programs, class size reduction, and charter schools) though none have been deemed a “silver bullet” intervention. Interventions that occur off the school campus, such as experiential enrichment programs, have rarely been studied. Enrichment programs leverage existing community cultural institutions such as museums, zoos, and theaters and have the potential to improve the educational experience of low-income students and bolster academic achievement.

Although evaluations of some programs have found positive associations between program participation and a range of student outcomes in the short run, there is little rigorous evidence to support the assumption that these programs impact student achievement. Further, the previous evaluations tell us little about the long-term outcomes of program participants. Out-of-school interventions may make school administrators wary, as schools are increasingly evaluated on test score gains and curricula are carefully structured to meet these standards. This paper fills the gap in the literature about whether and how arts intervention programs affect academic outcomes by evaluating the impact of School in the Park (SITP), a museum-based educational program in San Diego, on both short and long-term student outcomes.

Literature Review

The body of research on arts and science enrichment programming is limited and very few studies focus specifically on museum education. A meta-analysis of research on the impact of arts programs on academic achievement found only 31 studies to analyze, only one of which was published in a peer reviewed journal (Winner and Cooper, 2000). Within these studies,

there is some indication that high quality arts enrichment programs may increase the school readiness and educational attainment of students in these programs (Brown, Benedett, and Armistead, 2010; Winner and Hetland, 2000; DOJ, 2002; Seidel, 1996). However, these studies suffer from methodological flaws that hamper the authors' ability to isolate program impacts and limit the generalizability of the results. Specifically, most programs operate on a volunteer basis making it difficult for researchers to control for selection bias introduced when some students select to participate, and others do not. SITP presents a unique opportunity to avoid this common problem. Within the two schools where SITP is offered, all students in the grades served participate in the program; therefore selection bias is less of a concern.

Given the recent emphasis on test outcomes students, museums and school educators have made efforts to expand museum-based experiences beyond one-off guided tours and link these experiences more closely to school curricula. Thus far, only a handful of studies have evaluated the impacts of museum enrichment programs for students.¹ These programs vary significantly in terms of their methods and duration. While some involve multiple visits to a museum, the Hands-On Museum's program in Ann Arbor takes place exclusively in the students' school, and programs at the Museum of the City of New York and the Peabody Museum of Natural History consist only of one-time field trips. A science program based out of a Los Angeles County natural history museum only involves a single visit to the museum itself, though the program includes eight additional after-school sessions led by museum educators. Similarly, although the Guggenheim's program includes three museum visits, it is predominantly based in the school with twenty 90 minute sessions led by an artist in residence in the classroom

¹ See Appendix A for program descriptions.

during the course of the school year. There is only one program aside from SITP that involves multiple institutions; the rest of the programs involve only a single institution.

The outcomes measured by evaluations of these programs vary substantially. Some of the studies measure gains in student knowledge of specific content areas while others focus on more qualitative measures of program success, e.g. whether students' attitudes towards specific subjects such as science or history are improved. Although no unifying method has been used to evaluate the programs, it is possible to glean some general information from the existing evaluations. Museum enrichment programs are associated with improved student attitudes towards the subject at hand (Melber, 2003; Paris, Yambor, and Packard, 1998; Randi Korn & Associate, Inc., 2010). They also improve students' content knowledge of the subject matter (Melber, 2003) though single museum visit programs see only modest effects (Randi Korn & Associate, Inc., 2010). Some evaluations report improved student problem solving skills after participation (Paris, Yambor, and Packard, 1998). In most of the evaluations that utilize control groups of similar students, participating students scored higher on critical thinking evaluations than peers who did not participate (Burchenel and Grohe, 2007; Downey et al., 2007). Only one study measures achievement using standardized test scores and although it finds no statistically significant difference between the scores of participants and non-participants, these findings are only included in an unpublished report (Adams, Foutz, Luke, and Stein, 2006). A further weakness of these studies is the fact that they are cross sectional in nature and only provide information on short-term outcomes. A key advantage of the current study is the ability to provide information on longer-term outcomes.

Description of the Program

School in the Park is an educational program for third, fourth, and fifth grade students that takes place among the cultural institutions and museums of San Diego's Balboa Park. Participating institutions include the San Diego Museum of Art, the Museum of Photographic Arts, the History Center, the Natural History Museum, the Museum of Man, the Junior Theater and the Old Globe Theater, the Reuben H. Fleet Science Center, the Air and Space Museum, and the San Diego Zoo.² The program's stated goal is for students to experience "visual, auditory, and kinesthetic information" and to provide students with a "foundation of knowledge and a context in which to place new material" as they progress in school.³ The program has developed an explicit focus on "academic excellence" aimed at helping students excel in school and in the future. The program's implied goal is that students who participate will have better academic outcomes than those who do not participate.

SITP serves students in two inner-city public elementary schools, Rosa Parks and Alexander Hamilton Elementary, who participate in the activities at Balboa Park for up to eight weeks. Both schools are located in the City Heights neighborhood of San Diego, which is a major refugee portal for families from Somalia, Cambodia, Vietnam, Iraq, and Liberia, among other countries, and a destination for immigrants from Latin America. In addition to serving a large refugee and immigrant student population, many of the students who attend these two schools are low-income. In the San Diego Unified School District, 61 percent of school children

² Participating institutions by grade level –3rd Grade: Historical Society, San Diego Zoo, Museum of Art. 4th Grade: Museum of Photographic Arts, Museum of Man, Natural History Museum, Junior Theatre, Fleet Science Center. 5th Grade: Fleet Science Center, Air & Space Museum, Hall of Champions, Natural History Museum.

³ School in the Park website: <http://schoolinthepark.net/>

qualify for free or reduced price lunch⁴, while at the two participating elementary schools over 97% of the students qualify.⁵

SITP was initially developed in response to overcrowding at Rosa Parks in the late 1990s. The lack of space for instruction prompted administrators to think about how to utilize other community spaces to serve all of the City Heights students. Over time, the program has evolved, hiring full-time staff and museum educators to partner with teachers, expanding to Alexander Hamilton Elementary, and developing curricula for the program that align with district academic standards. Today, district administrators are questioning the ongoing utility of the program. Is it best for students to spend valuable class time outside the classroom, when they could be practicing skills that will appear on the annual standardized exams? Are students who are struggling academically worse off after attending the program? Is the district getting the most utility out of teachers who are paid to attend the program even when museum educators are providing instruction? These questions can only be answered with a systematic evaluation of the program.

Research Questions

This paper evaluates the impact of SITP participation on short-term and long-term academic and behavioral outcomes using longitudinal, student-level data, the natural variation in the timing of program implementation between the two schools, and control groups of students from nearby schools that did not receive the program. The study addresses two primary research questions:

⁴ San Diego Unified School District. <http://www.sandi.net/site/Default.aspx?PageID=984>

⁵ California Department of Education, School Fiscal Services Division. October 2010. <http://www.cde.ca.gov/ds/sh/cw/filesafdc.asp>

1. How does participation in the School in the Park program impact academic and behavioral outcomes for students in the years of participation?
2. Do impacts of School in the Park participation extend beyond the year of participation to affect longer-term outcomes in middle school and high school?

Theory

Experiential education programs like SITP aim to utilize non-traditional strategies to promote learning and to broaden students' understanding of the world. A conceptual paper by the American Alliance of Museums describes museum-based educational programs as providing "vital, experiential, multi-modal and trans-disciplinary educational opportunities...more than the ancillary field trip" (Kratz and Merritt, 2012). Work by John Dewey in the early 1900s described museums as contexts for educational experiences. A model of the impact of museum experiences on students influenced by Dewey envisions an educational cycle where students bring previous knowledge from their lives to the initial museum experience, the museum triggers reflection and inquiry, generates new problems and interests, and opens new lines of inquiry, and these new experiences and interests are applied to the student's life beyond the museum (Hein, 2004).

Building from these ideas, the conceptual model in Figure 1 presents the hypothesized impact of participation in SITP on short, medium, and long term student outcomes. Inputs into the program include the characteristics and prior experiences of students and teachers, and the school context in which students and teachers generally operate. Exposure to different cultures, ideas, and fields through time in the Balboa Park institutions, and the curricular units designed to complement these experiences, comprise the outputs of the program. These activities are hypothesized to affect students in the short, medium, and long term. Specifically, in the year of participation in SITP, the process described by Dewey above may result in improved

engagement both during SITP and regular class instruction. Students may also attend school more frequently because they want to participate in the program. As a result, students participating in the program may achieve higher scores on standardized exams in that year, and may be less likely to be retained, suspended, or expelled from school.

However, for students who are already struggling academically, the program may actually have negative impacts on achievement. These students may be better served by spending more time in the classroom and receiving remedial instruction. Therefore the expected direction of the impact of the program on short-term outcomes is ambiguous.

The benefits of exposure to new concepts and cultures may extend beyond the year of participation if students incorporate these new perspectives into their lives. These “broader horizons” might change how students envision their future options, including college attendance and possible careers. One way this may be observable is if the impact of participation in SITP lasts into the high school years. In Figure 1, immediate improvements in engagement, achievement, and behavior are hypothesized to influence academic and behavioral outcomes in middle school and high school and college preparation and enrollment. Even if the program has limited short-term outcomes, the real benefit of participation may appear later on, as students use the motivation and/or skills they learned in the program to persevere through high school and set higher post-graduation goals.

Research Design

To answer the question of how participation in the School in the Park program impacts the academic and behavioral outcomes of students, a standard difference-in-differences framework is employed (See Cannon et al., 2011 for a recent application of this technique by this author) that takes advantage of the fact that the program was implemented in the two schools at

different times. In essence, the research design compares changes in outcomes for students who participate in SITP (before and after participation in the program), to changes in outcomes for students who do not participate. The rich data from the San Diego Unified School District (SDUSD) allows the development of better control groups than past studies and to measure short-term as well as long-term outcomes like graduating from high school.

Equation 1 models short-term academic and behavioral outcomes as a function of participation in SITP, time-varying student characteristics, student fixed effects, and school fixed effects, using individual-level panel data.

$$Y_{isgt} = \alpha + \beta SITP_{sgt} + \gamma ST_{it} + \delta SC_{st} + \eta_i + \mu_s + \zeta_t + w_g + \varepsilon_{isgt} \quad [1]$$

In Equation 1, Y_{isgt} represents the dependent variable for student i in school s at time t ; $SITP_{sgt}$ is a dichotomous variable indicating whether the child attended the School in the Park program; ST_{it} is a vector of time-varying student characteristics; and SC_{st} is a vector of school characteristics for the school the student attended. In addition, the model includes a student fixed effect (η_i), school fixed effect (μ_s), a time fixed effect (ζ_t), a grade fixed effect (w_g), and an error term (ε_{isgt}).

For the long term outcomes, ordinary least squares and linear probability models are estimated, as shown in Equation 2 below:

$$Y_{isg} = \alpha + \beta SITP_{sg} + \gamma ST_i + \delta SC_{st} + \mu_s + \zeta_t + \varepsilon_{isg} \quad [2]$$

The structure of the dataset used in Equation 2 is cross-sectional, with one observation per student. In this equation, Y_{isg} represents the dependent variable for student i in school s in a given grade; $SITP_{sg}$ is a dichotomous variable indicating whether the child ever attended the School in the Park program in third, fourth, or fifth grade; ST_i is a vector of student

characteristics (including 2nd grade achievement); and SC_{st} is a vector of school characteristics for the school the student attended. In addition, the model includes a school fixed effect (μ_s), a series of dummy variables for the year the student was in third grade (ζ_t), and an error term (ε_{isgt}). The inclusion of the school fixed effect and time effects allow us to avoid confounding the impact of SITP with other changes that may be happening in the schools at a particular point in time.

Identification of the effects on student achievement and behavioral outcomes comes from three primary sources (Figure 2). First, we compare students within Rosa Parks and Alexander Hamilton Elementary schools that did not receive the program, with students from those same schools that received the program in later years (within-treatment cohort comparison). Second, we compare students in Rosa Parks to students in Alexander Hamilton Elementary over the period in which students in Rosa Parks received the program and Alexander Hamilton did not (within-treatment school comparison). This second comparison is particularly important because both schools may have received some resources from other Price Charities investments in the community that could also be impacting achievement independent of the School in the Park program. Finally, we compare students in these schools to students in comparable schools in SDUSD that do not receive this enrichment program (i.e., other schools in the Hoover High School system, and the schools in the Crawford High School system) (control school comparison). A remaining necessary assumption for the identification of a causal impact of SITP is that parents with students that may be stronger do not select into Rosa Parks or Alexander Hamilton elementary simply because the program exists at the schools.⁶

⁶ In general, students appear to be more at-risk for academic challenges following the implementation of SITP, and not the reverse. There are slight differences in the average student characteristics before and after the implementation of SITP. The student population at Rosa Parks comprised more black and white students (12 percent

Data

We utilize administrative student records from the SDUSD, which includes student-level demographic information, behavioral records, and measures of academic performance. We also have school-level measures of school size and type, and student poverty.

To conduct the most robust analysis possible, the impact of participation in SITP on individual students is modeled over time, controlling for student demographic characteristics as well as characteristics of their peers, teachers, classrooms, and schools. To do so requires student-level academic records and teacher-level records from 1996 to 2012 for students in the Hoover High School feeder system (which includes the treatment group) and the schools in the adjacent Crawford High School system (to serve as a control group). All students who stay within the SDUSD are followed through high school, regardless of the schools they attend following elementary school. Further, comparing students in the schools that receive the program to their school peers enhances our ability to distinguish program impacts from characteristics of schools. As would be expected, not all students that begin in our study schools will conclude their education in SDUSD. The issue of attrition will only confound the analysis if there are systematic differences between those that have participated in SITP and left the school district and those that did not participate in SITP and left the school district.

Following the primary analysis, several stratified models are estimated to explore the variation in the impact estimates by prior achievement levels, gender, race, ethnicity, special education status, and immigrant status. Further, we estimate dosage models to understand

and 4 percent, respectively) and fewer Hispanics (71 percent) before SITP, compared to after (7 percent black, 2 percent white, 79 percent Hispanic). In the years following SITP implementation, a larger share of students qualifies for English language support and special education. The characteristics of Hamilton students before and after the implementation of SITP are even more similar. There is very little change in the racial/ethnic distribution, although Hispanic students comprise a 4 percent larger share of the student population after SITP, while the Asian population decreases by 3 percent. The share of students in special education increases by 3 percent, and the share of students receiving English language support increases by 6 percent.

whether the length of time a student has the program moderates the effect of the program on student outcomes.

Measures

The various data sets were merged together using unique student identification numbers, school year, and school codes. From this large, longitudinal data set, we created multiple measures for the analysis. There are multiple short-term educational outcomes that we are able to analyze: whether participation in SITP has influenced the probability that students take the standardized exams, the actual scores achieved on those exams, and whether the student performs at or above the standard set by the state.

In the 2001/02 school year, California changed the standardized test used throughout the state from the SAT9 to the California Standards Test (CST). This makes it difficult to compare results for students over time. Therefore, we have created a standardized test score measure that compares student performance to that of his or her grade level peers each year. These z-scores have a mean of zero and a standard deviation of one. During the 2001/02 school year, both the SAT9 and CST exams were conducted. For the students who took both exams, we calculated z-scores and then examined how highly correlated the scores were in that year. We find that the z-scores for the two exams are highly correlated – the ELA z-scores are correlated at 0.894 and the math z-scores are correlated at 0.878. Therefore, in the models that span the period of time where both exams were given, we rely on the z-score measures of performance.

Following the examination of test scores, we consider whether SITP influences grade retention in the 3, 4, or 5th grades, and two short-term behavioral outcome measures – the number of absences in the current year, and the number of suspensions in the current year.

As mentioned earlier, one of the goals of the SITP program is that the impacts extend well beyond the school year in which the student participates in the program. We investigate longer-term outcomes for students as they reach middle school and high school. Three measures explore middle school behavior and achievement – the number of suspensions and expulsions in middle school, and whether the student is retained by the 8th grade. For a subset of students who participated in SITP prior to 2005, we observe progress through high school. For these students, we measure whether they have been retained in high school, whether they have taken an AP course or the SAT test (markers for intended college attendance), SAT scores, whether they pass the California High School Exit Examination (CAHSEE), a requirement for graduation, and whether they graduate with a diploma or earn a GED. We also construct a drop out measure for those students legally old enough to have left school (the state of California mandates compulsory attendance until a student turns 18) from the known dropout records provided by the school district. Students who are present in the drop out file but who have not reached the age of 18 are coded as dropouts. Those who received a GED or diploma were ultimately coded as non-dropouts even if they had previously dropped out.

In addition to these outcome measures, the models include a comprehensive set of control variables at the individual, classroom, and school levels. Individual-level controls include measures of gender, race and ethnicity (black, Hispanic, Asian, white, or other race), English language learner status, whether the student was born outside the United States, special education status, and whether the student moved schools in the current year. Further, for each classroom we include a measure of the teacher's total number of years of teaching experience. Finally, at the school-level, we control for the percent of students eligible for free or reduced

price lunch (a proxy for poverty), total student enrollment, whether or not the school is a charter school, and the years in which a school health clinic existed at Rosa Parks.

Sample

The analytic sample is restricted to students in 3, 4, or 5th grade between the 1995/96 and 2011/12 school years. Schools are included in the comparison group if we observe at least 10 students in the school in a given year. The sample size changes by the outcome measure of interest because there is a smaller sample of students for which we observe high school outcomes. In all models, the school year 1999/00 is omitted from analysis because it was the first year of partial implementation of SITP at Rosa Parks, but only half of the students in the third grade received the program.⁷

Results

Overall, we find consistently evidence that SITP has positive impacts on a range of short and longer-term academic and behavioral outcomes for students who participated in the program.

Descriptive Statistics

Table 1 presents mean descriptive statistics of the students who attended Rosa Parks between the 2000/01 and 2011/12 school years (and received SITP), those who attended Hamilton between 2006/07 and 2011/12 school years (and received SITP), and those students who attended the control schools in the Hoover and Crawford High School systems between 2000/01 and 2011/12 (who never received the program). Both the Rosa Parks and Hamilton SITP samples are majority Hispanic, and while the control group also consists of a Hispanic majority, Hispanic students make up a smaller share of the student population than in either Rosa Parks or Hamilton (64 percent in the control schools, 82 percent at Rosa Parks, and 77 percent at Hamilton). The control group is nearly 15 percent black and 15 percent Asian whereas Rosa

⁷ See Appendix Table B for a timeline of SITP implementation across the two schools.

Parks and Hamilton are 6 and 9 percent black, and 11 and 9 percent Asian, respectively. Students classified as English language learners make up a larger share of the student body at Rosa Parks and Hamilton than the control group average. On all other demographic measures, the treatment schools are comparable to the control schools.

In terms of school characteristics, Rosa Parks and Hamilton both have roughly the same percentage of teachers with over two years of teaching experience as the control group.

However, while the average teaching experience of a Rosa Parks teacher (10.5 years) is similar to the 11 years of average teaching experience amongst the control school teachers, Hamilton's teachers have an average of 15.4 years. Rosa Parks and Hamilton also have student populations that are 100 percent free and reduced lunch eligible whereas the control group population is 94 percent eligible. Rosa Parks is a larger school, serving nearly 1300 students on average, compared to Hamilton's 700 and the control group's 800 students on average.

Test-taking rates across all of the groups are high -92 to 94 percent on average. Rosa Parks and Hamilton SITP students' rate of taking standardized tests are on par with that of the control group. However, students in the control outperform students at Rosa Parks in both the language arts and math z-scores, and Rosa Parks students outperform Hamilton students. Students in both treatment schools perform better on average in terms of absences and suspensions. On average, they experience fewer absences (5.8 and 5.5 absences versus the control group's 6.7), fewer suspensions (0.019 and 0.017 compared to 0.039), and are less likely to be held back than the control group.

During middle school, Rosa and Hamilton SITP students have similar numbers of suspensions and expulsions, which are higher than the control group's averages. Rosa Parks

students have slightly higher average math and language arts z-scores in seventh grade compared to the control group, but both substantially outperform students from Hamilton.

For the high school outcomes, we only compare the students who participate in SITP from Rosa Parks students to the control group, as even the oldest Hamilton students who participated in the first year of the program are not yet old enough to have completed high school. Rosa Parks students have similar rates of high school suspensions, expulsions, grade retention, and dropouts as the control group. Similarly, 17 percent of Rosa Parks students take an Advanced Placement course at some point in high school compared to 16.6 percent of students who attended the control schools as elementary students. The share of students who graduate with a diploma is similar across the treatment and control groups. The one notable dissimilarity is that a higher share of Rosa Parks students take the SAT (34 percent) compared to only the control group (29 percent).

Short term Models

The first set of models in Tables 2 and 3 present the impact of participation in SITP on outcomes in the year of participation. These models utilize panel data and include individual student fixed effects. We find that Rosa Parks students who participate in SITP have a small, but statistically significant increase in the probability of taking the language arts standardized test. Rosa Parks SITP students also score 0.10 standard deviations higher on the standardized math exam, compared to control group students in the same year. SITP participation at Hamilton results in an increase in the likelihood of taking both the language arts and math exams, as well as a 0.07 standard deviation increase in language arts scores. SITP participation at both schools results in a reduction in suspensions.

As noted in the theory section, the expected impact of SITP on short-term outcomes is ambiguous, because while some students may benefit from the experiential learning opportunity, other students who are already struggling may fall farther behind on basic skills such as reading or math. To investigate this issue, Table 4 presents the short-term results stratified by whether or not the student passed the second grade standardized language arts and math exams. We find no impact on Rosa Parks students' ELA scores as a result of SITP – meaning that participation in the program does not appear to hinder the achievement of students struggling in second grade. Further, the program positively impacts these students' math scores for students who passed the second grade standardized math test *and* those who failed it. For Hamilton students, participation in SITP increases ELA test scores among those who had failed their second grade ELA test, with no effect on students who had passed the second grade exam. Not only is there no discernible negative effect of the program on the students who are most academically “at risk,” but we see modest improvements for those students. In addition, students who participate in SITP at Rosa Parks and who passed their second grade math test experience a slight reduction in the number of suspensions, while students who participate at Hamilton and failed their second grade math test also experience a reduction in their number of suspensions.

Long-term models

Data for the long-term models are cross-sectional, with one observation per student and outcomes measured through middle and high school. For the long-term outcomes, we examine academic and behavioral outcomes in both middle school and high school. While the short-term outcomes yield more positive results, the outcomes that occur during middle school are generally negative.⁸ Participation in SITP increases the likelihood of grade retention between sixth and eighth grade and increases the number of suspensions. For students at Rosa Parks, participating

⁸ Tables available from the authors upon request.

in SITP results in a 0.12 standard deviation decrease in seventh grade math scores, on average. It may be the case that beyond those immediate impacts of SITP in the year of participation, the longer term academic and behavioral impacts may take more time to materialize.

SITP results in generally positive outcomes for students in high school. In Tables 5 and 6, Rosa Parks SITP students have fewer high school expulsions, a lower probability of being held back between 9th and 12th grades, and a lower probability of getting a GED, compared to students in the comparison schools. However, there is a small, significant increase in the probability of dropping out of school for SITP students.

Marginal Week Analysis

Changes over the course of the program's existence led to variations in the amount of weeks students spent participating in the program by year. If participating in SITP has positive short and long-term outcomes for students, those students who spent more time in the program may reap larger benefits. Therefore, we analyze the impact of the marginal week of SITP participation. One of the reasons that many other museum enrichment programs analyzed in the literature have no impact may be because they simply do not provide enough exposure to the enrichment activities. In the short-term, the results depicted in Table 7, an additional week of SITP participation for students who attended Hamilton is associated with a small (less than one percent) but significant increase in the likelihood that students took their language arts and math standardized tests. The findings for Rosa for both of these measures were not statistically significant. The marginal week of SITP was also correlated with a decrease in the number of days absent for students both at Rosa and Hamilton (-0.058 days and -0.043 days, respectively). Interestingly, the marginal week was also associated with a less than one percent increase (0.26 percent) in the likelihood of being held back at Rosa.

In high school, Tables 8 and 9, an additional week of SITP participation results in numerous positive outcomes for Rosa Parks SITP students. The marginal week of participation increases tenth grade language arts scores by 0.004 standard deviations and increases both verbal and math SAT scores. There is also a small, but significant, reduction in suspensions and expulsions, as well as a 0.2 percentage point decrease in the likelihood of being retained during high school. An additional week of participation increases the probability of graduating with a diploma by 0.2 percentage points and reduces the probability of earning a GED by 0.1 percentage points. However, the marginal week is also associated with a 0.1 percentage point increase in the likelihood of dropping out amongst Rosa Parks students.

Discussion and Policy Implications

Overall, students who participate in SITP experience small short-term gains in test scores, and mixed effects on behavioral outcomes (decreases in suspensions, but a slight increase in the probability of being held back). Although district officials were concerned about the impact of SITP being potentially negative for students who are struggling academically prior to the program, we find no evidence of a negative effect. In fact, we find that students who previously failed the second grade standardized exam either performed better than or on par with students who passed the second grade exam, after participating in SITP. We find little effect of ever participating in SITP on the long-term academic outcomes of the program, but significant decreases in the expulsions and retention.

When we consider the marginal effect of an additional week of the program however, there are larger long-term effects. In particular, an additional week of the program results in increases in standardized test scores in 10th grade and SAT scores in high school, and the probability of graduating with a diploma, and decreases in suspensions, expulsions, and retention

in high schools. This may reflect differences in the quality of the SITP program over time, as the length of the program varied across years. It may also reflect a dosage effect – that students who received more weeks of the program had better outcomes, than students who received fewer weeks. There are also some small, negative effects of the program, namely increases in the probability that participating students will eventually drop out of high school. However these results may be sensitive to the measurement of drop out, and we intend to test the robustness of the finding with alternative drop out measures.

These findings support several important policy implications. First, exposing students to a new learning environment in lieu of class time does not harm them in terms of educational achievement. In fact we see slight gains for the students who are struggling prior to the program, compared to similar students who do not receive the program. The perceived tradeoff between out-of-classroom time and achievement is not grounded in the evidence from this program.

Second, in the short term, the larger effects are found for the behavioral outcomes. These behavioral changes might translate into improved academic achievement in the longer term. One way to invest in student achievement may be to indirectly build behavioral competencies – such as sitting still, paying attention, speaking in turn – that will facilitate learning for years to come.

Third, the program has lasting effects beyond the school year in which students participate, with positive impacts well into high school. These results indicate that students who receive more of the SITP program are more likely to achieve higher scores on the SAT and graduate with a diploma – both indicators of future college-going. The primary benefits of the program may only materialize when students spend a longer amount of time in the museums.

Fourth, the program utilizes existing community resources that many cities and states nationwide may also be able to access. At first, the program re-purposed museum space as

classroom space, when Rosa Parks was unable to accommodate all of the students in the neighborhood. Over time, the program grew to serve more students in more grades, utilizing multiple cultural institutions at Balboa Park, with teachers partnering with museum educators to provide lessons to the students.

Overall, the program shows impacts in both the short and long-term, well beyond those intended by the program in its design. If we are concerned with affecting student achievement in the long run (despite the current laser focus on end of year test scores), investing in experiential education programs that offer prolonged and structured experiences for students outside the classroom may boost achievement and long-term academic success for even the most at-risk students.

References

- (2005). M-DCPS launches four new magnet schools. *Miami Times*, 7-7B. Retrieved from <http://search.proquest.com/docview/363122829?accountid=14749>
- (2008). "The Museum Magnet Program." The Wolfsonian, Florida International University, 2008. Web. 28 Sept. 2012. <http://www.museummagnet.wolfsonian.org/>.
- Adams, M., Foutz, S., Luke, J., and Stein, J. (2006). *Thinking Through Art* Isabella Stewart Gardner Museum School Partnership Program Year 3 Preliminary Research Results. Edgewater, MD: Institute for Learning Innovation.
- Anderson, D., Piscitelli, B., Weier, K., and Everett, M. (2002). Children's museum experiences: Identifying powerful mediators of learning. *Curator: The Museum Journal* 45(3), 213-231.
- Burchenel, M., and Grohe, M. (2007). Thinking through art: Transforming museum curriculum. *The Journal of Museum Education* 32(2), 111-122.
- Cox-Petersen, A. M., Marsh, D. D., Kisiel, J., and Melber, L. M. (2003). Investigation of guided school tours, student learning, and science reform recommendations at a museum of natural history. *Journal of Research in Science Teaching* 40(2), 200-218.
- Downey, S., Delamatre, J., and Jones, J. (2007). Measuring the impact of museum-school programs: Findings and implications for practice. *The Journal of Museum Education* 32(2), 175-187.
- Fitchburg Public School District. (2012) Fitchburg Arts Academy Pilot School/McKay Elementary School Innovation School Prospectus. http://faa.fitchburgschools.org/files/_cKKdg_/3e45fb172f5c014f3745a49013852ec4/McKay_FAA_Innovation_Plan_May_21_2012.pdf
- Hein, G. E. (2004). John Dewey and museum education. *Curator: The Museum Journal*, 47(4), 413-427.
- Kanevsky, L., Corke, M., and Frangkiser, L. (2008). The academic resilience and psychosocial characteristics of inner-city English learners in a museum-based school program. *Education and Urban Society* 40(4), 452-475.
- Kratz, S., and Merritt, E. (2012). On the Horizon: Museums and the future of education. An American Alliance of Museums conceptual paper. Retrived from <http://www.aamus.org/docs/default-document-library/on-the-horizon-web-version.pdf?sfvrsn=0>.
- Melber, L.M. (2003). Partnerships in science learning: Museum outreach and elementary gifted education. *Gifted Child Quarterly* 47(4), 251-258.
- Paris, S. G., Yambor, K. M., & Packard, B. W. (1998). Hands-on biology: A museum-school-university partnership for enhancing students' interest and learning in science. *The Elementary School Journal* 98(3), 267-288.

Randi Korn & Associates, Inc. (2004). *Dinosaurs and More* school tour: Findings from questionnaires, interviews and observations. Alexandria, VA: Randi Korn & Associates, Inc.

Randi Korn & Associate, Inc. (2007). Program evaluation: School programs of the Frederick A.O. Schwarz Children's Center at the Museum of the City of New York. Alexandria, VA: Randi Korn & Associates, Inc.

Randi Korn & Associates, Inc. (2010). Educational research: Evaluation of Traveling Through Time, a school program of the Museum of the City of New York. Alexandria, VA: Randi Korn & Associates, Inc.

Semon, C. S. (2006, Jan 24). Fitchburg to get Arts-Driven Schools; Boston Model will be used in Pilot Program. *Telegram & Gazette*, pp. 1-A1. Retrieved from <http://search.proquest.com/docview/268954491?accountid=14749>, Innovative step ; art-oriented pilot school looks promising. (2006, Jan 26). *Telegram & Gazette*, pp. 12-A12. Retrieved from <http://search.proquest.com/docview/268950142?accountid=14749>

Figures

Figure 1. Conceptual Model of SITP Impact

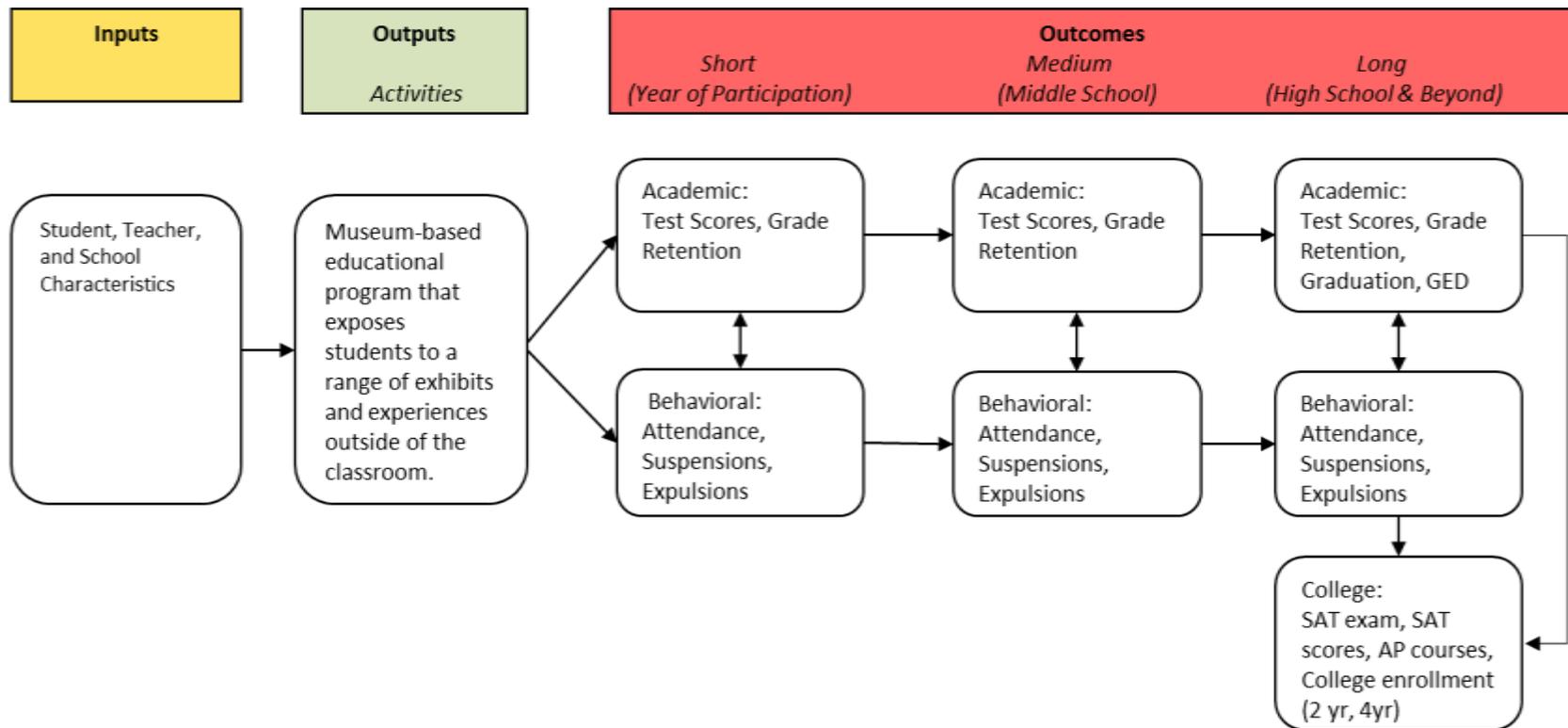
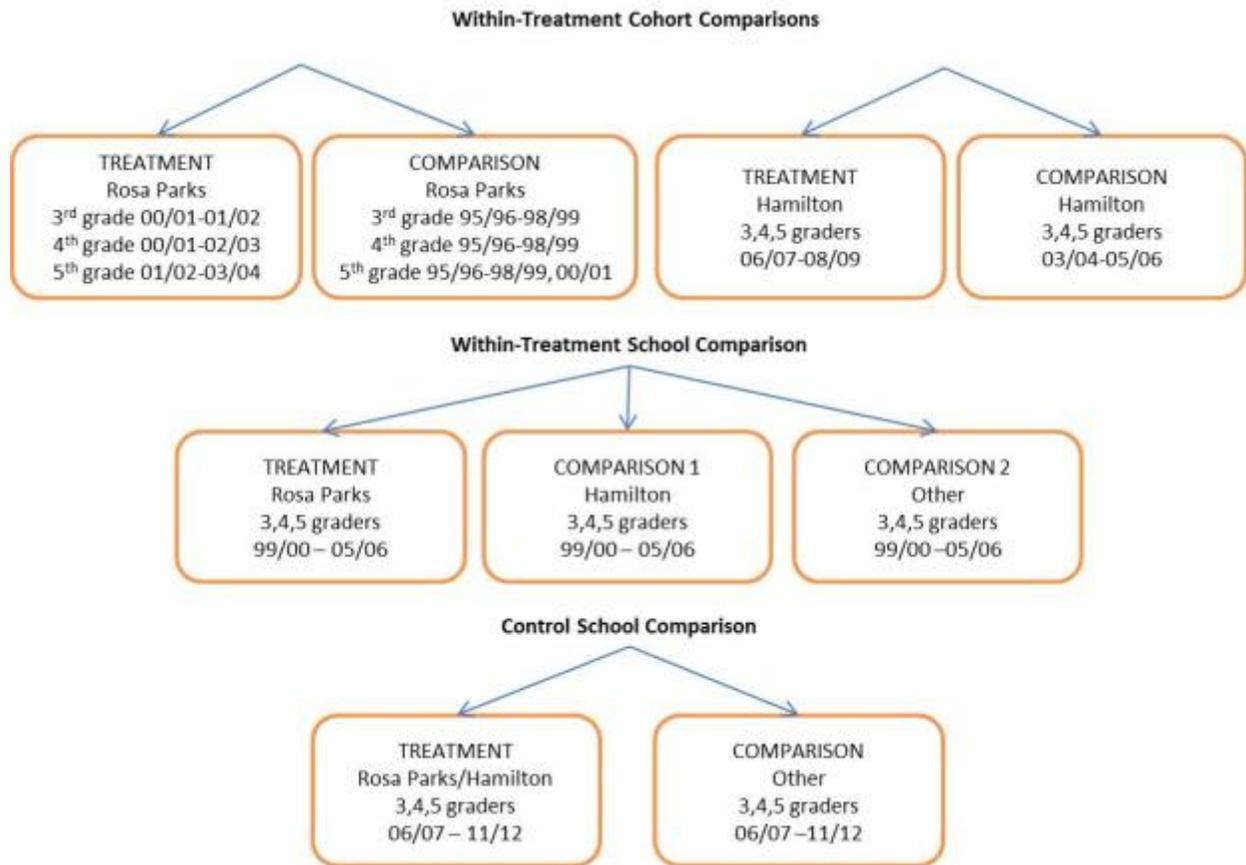


Figure 2. Research Design



Tables

Table 1. Descriptive Statistics, Overall Means for SITP Participants and Full Sample, 2000/01-2011/12

Variable	Rosa Parks 2000/01-2011/12	Alexander Hamilton 2006/07-2011/12	Other Schools ⁹ 2000/01-2011/12
Number of Observations	6,783	1,874	77,819
<i>Student Characteristics</i>			
Male	0.51	0.53	0.51
Black	0.06	0.09	0.15
Hispanic	0.82	0.77	0.64
Asian	0.10	0.10	0.14
White	0.01	0.01	0.05
Other Race	0.01	0.04	0.02
English Learner	0.79	0.77	0.66
Foreign born	0.19	0.17	0.21
Special Education	0.13	0.13	0.21
3 rd Grade	0.35	0.35	0.28
4 th Grade	0.34	0.33	0.23
5 th Grade	0.30	0.32	0.22
<i>Classroom Characteristics</i>			
Teacher Years of Experience	10.8	15.7	11.6
<i>School Characteristics</i>			
% Eligible for Free/Reduced Lunch	100	100	93.05988
Enrollment	1252.2	673.4	757.9
Teacher Average Years of Experience	10.5	15.4	11.1
% Teachers w 2+ Years of Experience	0.97	0.98	0.94
Charter School	0	0	0.034542
<i>Short-term Outcomes</i>			
ELA z-score	0.09	0.04	0.12
Math z-score	0.17	-0.03	0.12
Days absent	5.8	5.5	6.7
Number of suspensions	0.02	0.02	0.04
Retained 3 rd -5 th Grade	0.04	0.02	0.02
<i>Longer-term Outcomes</i>			
Retained 6 th -8 th Grade	0.01	0.02	0.02
Middle school suspensions	0.75	0.77	0.68
Middle school expulsions	0.03	0.02	0.02
Retained 9 th -12 th Grade	0.45	-	0.44
Take AP Course	0.17	-	0.17
Take SAT Exam	0.33	-	0.29
SAT Verbal Score	426.7	-	440.4
SAT Math Score	447.0	-	453.7
HS Diploma	0.43	-	0.40
GED	0.00	-	0.01

⁹ Other schools only include schools in the Crawford or Hoover clusters.

Table 2. Short-Term Academic Outcomes (1996-2012, omit 2000)

VARIABLES	(1) Take ELA	(2) ELA Z Score	(3) Take Math	(4) Math Z Score
SITP*Rosa	0.0122* (0.00564)	0.0425 (0.0297)	0.00791 (0.00607)	0.0984** (0.0310)
Rosa	-0.0491* (0.0200)	0.178 (0.141)	-0.0262 (0.0237)	0.0886 (0.142)
SITP*Hamilton	0.0183** (0.00634)	0.0668* (0.0279)	0.0150* (0.00702)	-0.0239 (0.0302)
Hamilton	-0.0103 (0.0119)	0.223** (0.0696)	0.000790 (0.0135)	0.146 (0.0721)
Special Ed	-0.142** (0.0260)	-0.0981** (0.0192)	-0.116** (0.0210)	-0.0512 (0.0247)
Moved	0.00935 (0.00624)	-0.0459** (0.0150)	0.00665 (0.00660)	-0.0519** (0.0130)
% Free Lunch	-0.000540 (0.000309)	0.00242 (0.00155)	-0.000519 (0.000360)	0.00215 (0.00205)
Enrollment	4.46e-05 (2.31e-05)	-0.000228 (0.000184)	2.31e-05 (3.02e-05)	7.36e-06 (0.000205)
Charter	0.00363 (0.0181)	0.00973 (0.0931)	0.00183 (0.0209)	0.113 (0.104)
Clinic	-0.00168 (0.00851)	-0.0378 (0.0483)	-0.00174 (0.00833)	-0.125** (0.0395)
Constant	-0.248 (0.182)	2.734** (0.316)	-0.126 (0.181)	4.333** (0.302)
Observations	77,819	72,919	77,819	73,010
R-squared	0.519	0.844	0.511	0.823
year FE	Yes	Yes	Yes	Yes
school FE	Yes	Yes	Yes	Yes
student FE	Yes	Yes	Yes	Yes
Clusters	22	22	22	22

Robust standard errors in parentheses

** p<0.01, * p<0.05

Table 3. Short-Term Behavioral Outcomes (1996-2012, omit 2000)

VARIABLES	(1) Days Absent	(2) Suspensions	(3) Held back
SITP*Rosa	-0.00800 (0.0876)	-0.0116* (0.00508)	0.00573* (0.00243)
Rosa	-0.508 (0.428)	-0.0408 (0.0263)	-0.00288 (0.0129)
SITP*Hamilton	0.0575 (0.141)	-0.0203* (0.00845)	0.00575 (0.00303)
Hamilton	-0.281 (0.332)	-0.0262 (0.0241)	-0.00835 (0.00831)
Special Ed	0.380* (0.177)	0.0551** (0.0187)	0.00575 (0.00343)
Moved	1.383** (0.164)	0.0281** (0.00642)	0.00141 (0.00226)
% Free Lunch	0.00449 (0.0113)	0.000341 (0.000470)	-1.38e-05 (0.000248)
Enrollment	0.000266 (0.000666)	7.05e-05* (3.37e-05)	2.20e-05 (1.10e-05)
Charter	0.657 (0.395)	0.0932* (0.0332)	0.00698 (0.0137)
Clinic	0.407* (0.189)	0.0129 (0.0124)	-0.00955** (0.00315)
Constant	5.354 (4.171)	-0.176 (0.0868)	4.659** (0.208)
Observations	77,665	77,819	73,895
R-squared	0.720	0.550	0.647
year FE	Yes	Yes	Yes
school FE	Yes	Yes	Yes
student FE	Yes	Yes	Yes
Clusters	22	22	22

Robust standard errors in parentheses

** p<0.01, * p<0.05

Table 4. Short-Term Outcomes, Stratified by 2nd Grade ELA Pass/Fail (1996-2012, omit 2000)

VARIABLES	ELA Z Score		Math Z Score		Days Absent		Suspension	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
STATIFICATION	2 nd Pass	2 nd Fail	2 nd Pass	2 nd Fail	2 nd Pass	2 nd Fail	2 nd Pass	2 nd Fail
SITP*Rosa	0.0581 (0.0304)	-0.0101 (0.0266)	0.103** (0.0345)	0.108** (0.0317)	0.0457 (0.0970)	-0.154 (0.221)	-0.0145* (0.00583)	-0.00188 (0.00685)
Rosa	0.131 (0.146)	0.156 (0.127)	0.0372 (0.157)	0.129 (0.123)	-0.531 (0.428)	-0.0735 (1.086)	-0.0391 (0.0256)	-0.0287 (0.0793)
SITP*Hamilton	0.00597 (0.0251)	0.113** (0.0302)	-0.0303 (0.0338)	-0.0188 (0.0302)	0.180 (0.107)	-0.482 (0.308)	-0.00940 (0.00820)	-0.0461* (0.0182)
Hamilton	0.212** (0.0743)	0.160* (0.0626)	0.123 (0.0838)	0.154* (0.0694)	-0.372 (0.376)	0.0668 (0.694)	-0.0341 (0.0209)	0.00296 (0.0659)
Constant	2.886** (0.489)	-1.840** (0.294)	4.834** (0.398)	-2.776** (0.310)	0.891 (5.027)	12.15** (1.919)	-0.138 (0.0823)	-0.176 (0.136)
Observations	52,458	20,445	59,606	14,117	62,299	16,049	62,407	16,095
R-squared	0.811	0.676	0.791	0.722	0.717	0.725	0.549	0.558
year FE	Yes	Yes						
school FE	No	No						
school*year FE	No	No						
Student FE	Yes	Yes						
Clusters	22	22	22	22	22	22	22	22

Robust standard errors in parentheses. Models also control for special education status, school moves, share of student body that receives free lunch, total school enrollment, charter school, and presence of school-based health clinic.

** p<0.01, * p<0.05

Table 5. Long-Term Academic Outcomes (1996-2012, omit 2000)

VARIABLES	(1) ELA 10th Z Score	(2) Take AP	(3) Take SAT	(4) SAT Verbal	(5) SAT Math	(6) Grad Diploma	(7) Grad GED	(8) CAHSEE Pass
SITP								
Ever*Rosa	-0.000891 (0.0469)	0.00644 (0.0211)	0.0470 (0.0300)	-5.123 (8.388)	3.773 (10.91)	0.0303 (0.0265)	-0.0190** (0.00535)	-0.00514 (0.0133)
Rosa	0.143 (0.135)	0.119* (0.0422)	0.734** (0.126)	31.25 (56.66)	55.70 (69.05)	0.221* (0.0847)	-0.0105 (0.0204)	0.0462 (0.0623)
Hamilton	0.0175 (0.101)	0.0596** (0.0198)	0.458** (0.0807)	7.968 (39.05)	39.62 (46.33)	0.114 (0.0653)	-0.0134 (0.0122)	0.0108 (0.0445)
Male	-0.113** (0.0121)	-0.0526** (0.00666)	-0.0543** (0.0134)	15.14** (3.578)	41.17** (3.691)	-0.0685** (0.0134)	0.00268 (0.00383)	-0.00683 (0.00732)
Black	-0.220** (0.0757)	-0.0316 (0.0186)	0.0795* (0.0359)	-62.44** (7.441)	-60.16** (9.892)	-0.0487 (0.0353)	-0.0215* (0.00831)	-0.0249 (0.0166)
Hispanic	-0.0663 (0.0821)	0.0183 (0.0202)	0.0843* (0.0353)	-38.78** (8.066)	-14.51 (11.20)	0.00173 (0.0289)	-0.0132 (0.0103)	0.0134 (0.0157)
Asian	0.223** (0.0759)	0.143** (0.0159)	0.257** (0.0385)	-32.79** (5.987)	16.19 (10.62)	0.160** (0.0321)	-0.0202 (0.00987)	0.0542** (0.0157)
Other race	-0.0587 (0.0909)	-0.0124 (0.0307)	0.103 (0.0765)	-92.70** (17.22)	-57.70** (12.33)	-0.0241 (0.0481)	-0.0186* (0.00881)	-0.0117 (0.0176)
El Status	-0.122** (0.0287)	0.0121 (0.0133)	-0.0355 (0.0218)	-18.23* (6.960)	-5.627 (6.762)	0.0248 (0.0188)	-0.00641 (0.00404)	0.000738 (0.0115)
Foreign Born	0.199** (0.0361)	0.0195* (0.0089)	0.00320 (0.0206)	23.44** (7.056)	6.603 (8.261)	0.0105 (0.0225)	0.000477 (0.00521)	0.0501** (0.0099)
Special Ed	-0.295** (0.0365)	-0.0635** (0.0085)	-0.118** (0.00937)	-32.37** (9.604)	-43.76** (9.764)	-0.132** (0.0200)	-0.00891* (0.00392)	-0.130** (0.0193)
Z Ela 2nd	0.515** (0.0155)	0.0883** (0.0055)	0.104** (0.00634)	41.24** (4.258)	44.48** (3.092)	0.102** (0.0070)	0.000135 (0.00117)	0.0484** (0.0061)
Constant	0.0156 (0.386)	0.190 (0.141)	1.303** (0.253)	373.3** (101.0)	441.4** (147.0)	0.976** (0.216)	-0.0126 (0.0382)	1.085** (0.152)
Observations	6,935	10,271	4,067	1,424	1,424	6,366	6,367	5,172
R-squared	0.381	0.195	0.321	0.393	0.405	0.164	0.036	0.343
school FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3rd Grade dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	21	22	18	18	18	21	21	21

Robust standard errors in parentheses. Models also include controls for % of students eligible for free lunch, total enrollment, HS charter school, and ever attended Clark Middle School.

** p<0.01, * p<0.05

Table 6. Long-Term Behavioral Outcomes (1996-2012, omit 2000)

VARIABLES	(1) HS Suspensions	(2) HS Expulsions	(3) Retention 12th	(4) HS Dropout
SITP Ever*Rosa	-0.0888 (0.0552)	-0.0259** (0.00413)	-0.0462* (0.0216)	0.0346** (0.0119)
Rosa	-0.103 (0.0660)	0.0166* (0.00785)	-0.251** (0.0763)	-0.0366 (0.0586)
Hamilton	-0.149** (0.0427)	-0.00662 (0.00633)	-0.145** (0.0468)	-0.000775 (0.0351)
Male	0.403** (0.0295)	0.0199** (0.00295)	0.0745** (0.0120)	0.0138 (0.0103)
Black	0.246* (0.0900)	0.00191 (0.00865)	-0.0145 (0.0342)	0.0160 (0.0226)
Hispanic	0.00945 (0.0725)	-0.00231 (0.00672)	-0.000733 (0.0270)	0.0243 (0.0138)
Asian	-0.239** (0.0708)	-0.0100 (0.00753)	-0.159** (0.0314)	-0.0121 (0.0203)
Other Race	-0.0853 (0.114)	-0.00473 (0.00840)	-0.0288 (0.0321)	0.0316 (0.0175)
El Status	0.0190 (0.0313)	0.000958 (0.00369)	0.0426* (0.0158)	-0.00402 (0.0127)
Foreign Born	-0.0844 (0.0501)	0.00667 (0.00518)	-0.0552** (0.0166)	-0.0204 (0.0108)
Special Ed	0.141* (0.0616)	-0.00188 (0.00370)	0.101** (0.0151)	0.0455** (0.0127)
Z Ela 2 nd	-0.120** (0.0171)	-0.00127 (0.00133)	-0.103** (0.00769)	-0.0249** (0.00465)
Constant	0.0656 (0.239)	-0.0386 (0.0232)	-0.0504 (0.187)	0.0459 (0.116)
Observations	10,271	10,271	6,059	5,503
R-squared	0.065	0.014	0.313	0.046
school FE	Yes	Yes	Yes	Yes
3rd Grade dummies	Yes	Yes	Yes	Yes
Clusters	22	22	21	22

Robust standard errors in parentheses. Models also include controls for % of students eligible for free lunch, total enrollment, HS charter school, and ever attended Clark Middle School.

** p<0.01, * p<0.05

Table 7. Marginal Week Analysis, Short-Term Outcomes (1996-2012, omit 2000)

VARIABLES	(1) Take ELA	(2) ELA Z Score	(3) Take Math	(4) Math Z Score	(5) Days Absent	(6) Suspensions	(7) Held back
Total SITP weeks*Rosa	-0.000173 (0.00129)	-0.000752 (0.00272)	-0.00107 (0.00137)	0.00267 (0.00239)	-0.0548* (0.0215)	0.000258 (0.000922)	0.00262* (0.00114)
Rosa	-0.0357 (0.0221)	0.226 (0.143)	-0.00721 (0.0271)	0.104 (0.136)	0.0359 (0.383)	-0.0553 (0.0275)	-0.0235* (0.0106)
Total SITP weeks*Hamilton	0.00226* (0.000904)	0.00228 (0.00228)	0.00223* (0.000937)	0.000807 (0.00333)	-0.0433* (0.0193)	-0.00216 (0.00106)	0.00111 (0.000689)
Hamilton	-0.0168 (0.0123)	0.234** (0.0734)	-0.00748 (0.0139)	0.137 (0.0770)	-0.0434 (0.390)	-0.0209 (0.0232)	-0.0107 (0.00844)
Constant	-0.245 (0.185)	2.788** (0.310)	-0.125 (0.183)	4.358** (0.293)	5.455 (4.099)	-0.180 (0.0887)	4.660** (0.209)
Observations	77,819	72,919	77,819	73,010	77,665	77,819	73,895
R-squared	0.519	0.844	0.511	0.823	0.720	0.550	0.648
year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
school FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
student FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	22	22	22	22	22	22	22

Robust standard errors in parentheses. Models also control for special education status, moved schools, % of student body eligible for free lunch, total enrollment, charter school, and health clinic.

** p<0.01, * p<0.05

Table 8. Marginal Week Analysis, Long-Term Academic Outcomes (1996-2012, omit 2000)

VARIABLES	(1) ELA 10th Z Score	(2) Take AP	(3) Take SAT	(4) SAT Verbal	(5) SAT Math	(6) Grad Diploma	(7) Grad GED	(8) CAHSEE Pass
Total SITP weeks*Rosa	0.00318** (0.00110)	0.000806 (0.00057)	0.00211 (0.00102)	0.566* (0.251)	0.878* (0.315)	0.00174* (0.00083)	-0.000642** (0.000151)	-0.000346 (0.000355)
Rosa	0.0894 (0.151)	0.110* (0.0398)	0.734** (0.127)	28.13 (56.73)	52.85 (68.60)	0.217* (0.0860)	-0.0133 (0.0213)	0.0477 (0.0642)
Hamilton	0.0193 (0.102)	0.0592** (0.0200)	0.458** (0.0796)	13.32 (38.10)	44.26 (45.04)	0.114 (0.0646)	-0.0125 (0.0123)	0.0107 (0.0442)
Male	-0.112** (0.0122)	-0.0525** (0.00667)	-0.0543** (0.0134)	15.28** (3.600)	41.27** (3.780)	-0.0685** (0.0134)	0.00268 (0.00383)	-0.00685 (0.00731)
Black	-0.220** (0.0758)	-0.0316 (0.0186)	0.0796* (0.0359)	-62.29** (7.467)	-60.08** (9.889)	-0.0487 (0.0353)	-0.0215* (0.00829)	-0.0249 (0.0166)
Hispanic	-0.0658 (0.0820)	0.0183 (0.0202)	0.0840* (0.0352)	-38.42** (7.943)	-14.33 (11.21)	0.00152 (0.0288)	-0.0130 (0.0103)	0.0134 (0.0157)
Asian	0.223** (0.0759)	0.143** (0.0159)	0.257** (0.0385)	-32.54** (5.967)	16.32 (10.66)	0.160** (0.0322)	-0.0201 (0.00983)	0.0542** (0.0157)
Other Race	-0.0582 (0.0908)	-0.0124 (0.0307)	0.103 (0.0765)	-92.73** (17.46)	-57.89** (12.27)	-0.0243 (0.0482)	-0.0185* (0.00879)	-0.0116 (0.0176)
El Status	-0.122** (0.0285)	0.0119 (0.0134)	-0.0353 (0.0216)	-18.84** (6.455)	-6.081 (6.337)	0.0247 (0.0187)	-0.00657 (0.00403)	0.000752 (0.0115)
Foreign Born	0.200** (0.0358)	0.0196* (0.00889)	0.00318 (0.0205)	23.72** (6.994)	6.823 (8.265)	0.0107 (0.0225)	0.000498 (0.00522)	0.0500** (0.00998)
Special Ed	-0.295** (0.0364)	-0.0635** (0.00853)	-0.117** (0.00942)	-32.27** (9.702)	-43.73** (9.768)	-0.131** (0.0199)	-0.00902* (0.00395)	-0.130** (0.0193)
Z Ela 2nd	0.515** (0.0154)	0.0882** (0.00548)	0.104** (0.00634)	41.15** (4.252)	44.38** (3.107)	0.102** (0.00700)	0.000119 (0.00116)	0.0485** (0.00607)
Constant	0.0105 (0.394)	0.188 (0.140)	1.303** (0.255)	370.6** (102.7)	439.3** (148.1)	0.976** (0.215)	-0.0117 (0.0389)	1.085** (0.152)
Observations	6,935	10,271	4,067	1,424	1,424	6,366	6,367	5,172
R-squared	0.381	0.195	0.321	0.393	0.406	0.164	0.036	0.343
school FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3 rd Grade dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	21	22	18	18	18	21	21	21

Robust standard errors in parentheses. Models also include controls for % of students eligible for free lunch, total enrollment, HS charter school, and ever attended Clark Middle School. ** p<0.01, * p<0.05

Table 9. Marginal Week Analysis, Long-Term Outcomes, Behavioral (1996-2012, omit 2000)

VARIABLES	(1) HS Suspensions	(2) HS Expulsions	(3) Retention 12th	(4) HS Dropout
Total SITP weeks*Rosa	-0.00436** (0.00131)	-0.000607** (0.000121)	-0.00168** (0.000586)	0.00126** (0.000289)
Rosa	-0.0999 (0.0674)	0.00502 (0.00803)	-0.257** (0.0743)	-0.0319 (0.0569)
Hamilton	-0.148** (0.0425)	-0.00672 (0.00625)	-0.144** (0.0474)	-0.00172 (0.0352)
Male	0.402** (0.0294)	0.0199** (0.00294)	0.0745** (0.0120)	0.0138 (0.0103)
Black	0.246* (0.0899)	0.00197 (0.00864)	-0.0144 (0.0341)	0.0160 (0.0226)
Hispanic	0.0101 (0.0726)	-0.00198 (0.00680)	-0.000244 (0.0269)	0.0240 (0.0138)
Asian	-0.238** (0.0709)	-0.00978 (0.00756)	-0.159** (0.0315)	-0.0122 (0.0204)
Other Race	-0.0853 (0.114)	-0.00466 (0.00839)	-0.0285 (0.0321)	0.0313 (0.0175)
El Status	0.0190 (0.0313)	0.000723 (0.00369)	0.0424* (0.0157)	-0.00375 (0.0128)
Foreign Born	-0.0847 (0.0503)	0.00670 (0.00517)	-0.0552** (0.0166)	-0.0204 (0.0108)
Special Ed	0.141* (0.0616)	-0.00194 (0.00369)	0.101** (0.0151)	0.0457** (0.0127)
Z Ela 2nd	-0.120** (0.0171)	-0.00127 (0.00133)	-0.103** (0.00769)	-0.0249** (0.00465)
Constant	0.0644 (0.240)	-0.0415 (0.0222)	-0.0505 (0.188)	0.0455 (0.117)
Observations	10,271	10,271	6,059	5,503
R-squared	0.065	0.014	0.313	0.046
school FE	Yes	Yes	Yes	Yes
3rd Grade dummies	Yes	Yes	Yes	Yes
Clusters	22	22	21	22

Robust standard errors in parentheses. Models also include controls for % of students eligible for free lunch, total enrollment, HS charter school, and ever attended Clark Middle School.

** p<0.01, * p<0.05

Appendix

Appendix A. Summary of Existing Museum Education Program Evaluations

<i>Program Name</i>	<i>Location</i>	<i>Duration</i>	<i>Description</i>	<i>Grades</i>	<i>Evaluation Tools</i>	<i>Findings</i>	<i>Title</i>
QUT Museums Collaborative (QUTMC)	Queensland, Australia	10 week	Ten-week, multi-visit museum program	Age 4 -7	Observations, interviews, parental diaries	Evidence that museum experiences "embedded within children's familiar culture and contexts are powerful mediators of memory, enjoyment, and learning in these settings."	Anderson, D., Piscitelli, B., Weier, K., & Everett, M. (2002). Children's museum experiences: Identifying powerful mediators of learning.
Traveling through Time, Leave it to the Beavers, and The Grid	Museum of the City of New York. Alexandria, VA		Traveling through Time: focus on Henry Hudson, the Dutch, and the foundation of New Amsterdam. Leave it to the Beavers: exploration of trade with an emphasis on trade that occurred between the Dutch and the Lenape. The Grid: discussion of urban planning.	2, 4	Observations, interviews with students and teachers.	Observations showed that students were overwhelmingly excited and engaged throughout all the programs. No students were observed to be completely off task or disengaged. Students' participation in the programs was constant, with students answering and asking questions.	Randi Korn & Associate, Inc. (2007). Program evaluation: School programs of the Frederick A.O. Schwarz Children's Center at the Museum of the City of New York. Alexandria, VA: Randi Korn & Associates, Inc.
Learning Through Art (LTA)	The Guggenheim, New York, NY	30 hours/year per class	An artist in residence visits three different classrooms in a public elementary school for ninety minutes each for a total of twenty visits during the school year.	K-5	Measured critical thinking skills among 3rd graders. Control group from similar schools.	Positively impact on five of the six identified skills for both looking at art and interpreting text, and increased student learning in general compared to control group.	Downey, S., Delamatre, J., and J. Jones. Measuring the impact of museum-school programs: Findings and implications for practice. The Journal of Museum Education 32(2), 175-187.
School Partnership Program (SPP)	Isabella Stewart Gardner Museum. Boston, MA		Museum visits tied to the curricula that students were currently already learning in their classrooms, primarily through subject-based connections.	K-8	Recorded student discussions of art in a classroom and in a gallery. Control group from similar schools.	SPP students had statistically significant higher scores than their counterparts on five of the seven criteria: associating, comparing, flexible thinking, observing, and interpreting. They also spoke more about art than the control group and offered more evidence to support their assertions.	Burchenel, M., & Grohe, M. (2007). Thinking through art: Transforming museum curriculum. The Journal of Museum Education 32(2), 111-122.

<i>Program Name</i>	<i>Location</i>	<i>Duration</i>	<i>Description</i>	<i>Grades</i>	<i>Evaluation Tools</i>	<i>Findings</i>	<i>Title</i>
Hands-On Biology	Hands-On Museum, Ann Arbor, MI		Science curriculum for elementary school students based on students' exploration of science through laboratory activities, experiments, and personal projects guided by teachers and docents.	3-5	Student survey of attitudes about science. Assessments and quizzes. Measured students' interest and knowledge of biology.	Students' general attitudes about science became more positive from pretest to posttest. Grade 3 students were more enthusiastic about learning about biology than older students. The program fostered positive attitudes about participating in science activities. Students showed significant learning gains on problem-solving questions.	Paris, S. G., Yambor, K. M., & Packard, B. W. (1998). Hands-on biology: A museum-school-university partnership for enhancing students' interest and learning in science. <i>The Elementary School Journal</i> 98(3), 267-288.
	Unidentified science museum, Los Angeles, CA	8 sessions	Museum science program for academically gifted elementary students. 8 sessions take place at school using museum specimens and artifacts and the final session is at the museum, and students meet scientists and visit areas of the museum that are normally closed to the general public	4-5	Pre and post questionnaire of student attitudes about science, parent questionnaire.	Students indicated a greater understanding of science careers and an increased desire to explore careers in science.	Melber, L.M. (2003). Partnerships in science learning: Museum outreach and elementary gifted education. <i>Gifted Child Quarterly</i> 47(4), 251-258.
Traveling through Time	Museum of the City of New York. Alexandria, VA	90 minute	90-minute program facilitated by a MCNY educator. Students learn how and why New Amsterdam became New York as they tour the Trade and New York Interiors exhibitions. Throughout the program, MCNY educators employ inquiry and object-based education strategies.	4	Pre and post questionnaires, interviews. Measured attitudes about history and achievement on history-related skills. Control group of fourth-grade students who did not attend.	The study revealed that the single visit fieldtrip had positive effects on students' attitudes and ideas about history as well as positive effects on students' achievement of history-related skills, including historical knowledge, historical inquiry, historical perspective, and historical reasoning. Impact on history-related skills overall was modest.	Randi Korn & Associates, Inc. (2010). Educational research: Evaluation of Traveling Through Time, a school program of the Museum of the City of New York.

Appendix B. Participation in SITP, by School, Grade, and Year

<i>School</i>	<i>Grade</i>	School Year												
		<i>99-00</i>	<i>00-01</i>	<i>01-02</i>	<i>02-03</i>	<i>03-04</i>	<i>04-05</i>	<i>05-06</i>	<i>06-07</i>	<i>07-08</i>	<i>08-09</i>	<i>09-10</i>	<i>10-11</i>	<i>11-12</i>
Rosa Parks Elem.	3 rd	50%	Yes											
	4 th	No	Yes											
	5 th	No	No	Yes										
Hamilton Elem.	3 rd	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
	4 th	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
	5 th	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes

Appendix C. Descriptive Statistics, 3rd-5th Grades, 2000/01-2005/06

Variable (means)	Treatment	Comparisons	
	Rosa Parks	Alexander Hamilton	Other Schools ¹⁰
Number of Observations	3,949	3,110	22,315
<i>Student Characteristics</i>			
Male	.5122816	.5138264	.5091642
Black	.0602684	.1459807	.1720367
Hispanic	.804001	.6581994	.5787139
Asian	.11294	.1581994	.156397
White	.0151937	.0324759	.0765404
Other Race	.0075969	.0051447	.0163119
English Learner	.7619651	.6385852	.5821645
Foreign born	.1825779	.2366559	.2326686
Special Education	.1121803	.0942122	.1136903
3 rd Grade	.3694606	.3598071	.3714542
4 th Grade	.3514814	.3305466	.3310329
5 th Grade	.279058	.3096463	.2975129
<i>Classroom Characteristics</i>			
Teacher Years of Experience	7.944096	9.575433	10.44589
<i>School Characteristics</i>			
% Eligible for Free/Reduced Lunch	100	100	88.23254
Enrollment	1444.062	1117.123	716.4938
Teacher Average Years of Experience	8.243254	9.600009	10.17169
% Teachers with 2+ Years of Experience	.9545967	.9284259	.9287573
Charter School	0	0	.0688326
<i>Short-term Outcomes</i>			
ELA z-score	.0872262	.1391391	.1667596
Math z-score	.2123972	.1439511	.1489781
Days absent	6.036258	6.907389	7.082798
Number of suspensions	.0139276	.0334405	.0461125
Retained 3 rd -5 th Grade	.0533837	.0281591	.02011
<i>Longer-term Outcomes</i>			
Retained 6 th -8 th Grade	.0069819	.0197171	.0285697
Middle school suspensions	.8134819	.929233	.7706121
Middle school expulsions	.0319909	.0351982	.0290443
Retained 9 th -12 th Grade	.4323635	.5050188	.4205142
Take AP Course	.1947328	.1633441	.1948017
Take SAT Exam	.3338192	.2588745	.3190622
SAT Verbal Score	426.7478	417.884	444.929
SAT Math Score	447.0354	430.7167	457.6795
HS Diploma	.4279595	.3813758	.4264369
GED	.0042885	.0056232	.0097538

¹⁰ Other schools only include schools in the Crawford or Hoover clusters.

Appendix D. Descriptive Statistics, 3rd-5th Grades, 2006/07-2011/12

Variables (Means)	Treatment		Comparison
	Rosa Parks	Hamilton	Other Schools ¹¹
Total Observations	3,525	2,315	23,475
<i>Student Characteristics</i>			
Male	.5075177	.5265659	.5071778
Black	.0533333	.0941685	.1442386
Hispanic	.8343262	.7572354	.6415761
Asian	.0967376	.1045356	.1434718
White	.0076596	.012959	.0443876
Other Race	.0079433	.0311015	.0263259
English Learner	.830922	.7615551	.7114803
Foreign born	.1997163	.1771058	.1771246
Special Education	.1438298	.1257019	.1422364
3 rd Grade	.3341844	.3473002	.3670714
4 th Grade	.3302128	.3274298	.3461981
5 th Grade	.3356028	.32527	.2867306
<i>Classroom Characteristics</i>			
Teacher Years of Experience	13.872	14.76751	13.13224
<i>School Characteristics</i>			
% Eligible for Free/Reduced Lunch	100	100	93.86069
Enrollment	1062.197	717.5266	559.2076
Teacher Average Years of Experience	12.93596	14.62923	12.80343
% Teachers with 2+ Years of Experience	.995656	.9859116	.9952617
Charter School	0	0	.0230032
<i>Short-term Outcomes</i>			
CST ELA test taken	.9029787	.9218143	.9038126
CST math test taken	.9095035	.924838	.9086688
CST ELA score	332.5994	331.8365	335.9046
CST math score	361.1185	352.0738	363.166
Passed CST ELA test	.7244738	.7197751	.7386058
Passed CST math test	.7932002	.7403083	.7760536
ELA z-score	.0902993	.0720798	.1409247
Math z-score	.1418583	.0155672	.1434088
Days absent	5.53284	5.823377	6.609135
Number of suspensions	.0260993	.0164147	.0550799
Retained 3 rd -5 th Grade	.0332557	.0240209	.0199969
<i>Longer-term Outcomes</i>			
Retained 6 th -8 th Grade	.0045962	.0207877	.016618
Middle school suspensions	.710985	.8495238	.5312144
Middle school expulsions	.024321	.0260317	.0147522

¹¹ Other schools only include schools in the Crawford or Hoover clusters.