Evaluation of Citizen Schools' Expanded Learning Time Model: Final Report

October 21, 2016

Prepared for:
Citizen Schools
308 Congress St., 5th Floor
Boston, MA 02210

Submitted by:
Alyssa Rulf Fountain
Beth Gamse
Melissa Velez
Matthew Hillard
Porsha Cropper
Abt Associates
55 Wheeler Street
Cambridge, MA 02138

Executive Summary

Introduction and Background

Academic achievement in the nation's lowest performing schools is troubling. Despite improvements in students' overall academic achievement over the past few decades, proficiency gaps in reading and mathematics still persist across income, racial, and ethnic groups (National Center for Education Statistics, 2013). One promising strategy for reducing the achievement gaps is the use of a longer school day, often called Expanded Learning Time (ELT); evidence suggests that students' increased access—through ELT programming—to enrichment activities can have positive effects on their academic motivation and social-emotional skill development.

The number of schools across the country implementing ELT continues to grow; by 2013–14, the most recent school year for which data are available, over 2,000 schools were using ELT strategies (National Center for Time and Learning (NCTL)). Nearly half of these schools were charter schools, and the remaining schools include many low-performing regular district schools that adopted ELT strategies in response to federal requirements for the School Improvement Grant (SIG) program. A recent GAO report indicates that approximately 1,800 SIG-funded schools, and 69 schools funded under the 21st Century Community Learning Program are implementing some form of ELT strategies.[1] Current guidance under the SIG program defines increased learning time as "increasing the length of the school day, week, or year to significantly increase the total number of school hours" for instruction in core academic subjects and other subjects, enrichment activities, and teacher collaboration, planning, and professional development (U.S. Department of Education, 2012a, p. 23).

The definition of ELT may seem straightforward, yet its implementation varies considerably in practice. ELT can occur via additional days in the school year, more hours in a school day, or summer programming above and beyond the regular year or day. ELT participation is sometimes mandatory for designated students, grades, or schools, and is sometimes optional. ELT programming is offered by local schools/districts, community-based organizations or non-profit groups, and is delivered by teachers, tutors, community volunteers, or other adults. Further, schools implementing ELT may rely upon the same *structure* (i.e., all after-school or summer or longer day programming), yet may vary substantially in the content of ELT, how additional time is distributed across subjects or extracurricular activities, and in who provides instruction.

Regardless of structure or content, using additional time effectively requires attending to multiple factors or strategies that all affect program implementation. Both the Afterschool Alliance (2012) and NCTL (2012) have articulated some guiding principles as essential for high-quality ELT programming. These include strong school community partnerships; focused learning; family engagement; intentional programming; diverse, prepared staff; participation and access; safety, health, and wellness; and ongoing assessment and improvement.

[1] See http://www.gao.gov/products/GAO-16-141

Key Elements of Successful Expanded Learning Time Approaches

Key elements[a] of expanded learning time initiatives include:

- Increase time in schools for students
- Provide academic support, individualized learning support, and enrichment activities to prepare students for college and career success
- Ensure opportunities for teacher collaboration and professional development to improve the quality of instruction
- Build strong school leadership and support of ELT
- Promote data-driven and evidence-based support to improve student performance
- Cultivate dedicated partnerships with external organizations
- Build a culture of high expectations and accountability

[a] See After School Alliance (January 2012), *Principles of Effective Expanded Learning Programs: A Vision Built on the Afterschool Approach*. Washington, DC: Author. Retrieved from http://www.afterschoolalliance.org/Principles%20of%20Expanded%20Learning%20Programs_Jan_2012(2).pdf; see also Claire Kaplan & Roy Chan (2011), Time Well Spent: Eight Powerful Practices of Successful Expanded-Time Schools, National Center for Time and Learning. Retrieved from http://timeandlearning.org/sites/default/files/resources/timewellspent.pdf

Since 1995, Citizen Schools has developed and implemented its own ELT model. Citizen Schools partners with middle schools serving predominantly low income (90% nationally) racial or ethnic minority (94% nationally), and academically struggling students. The Citizen Schools model relies upon an additional shift of educators and community volunteers to engage middle school students in hands-on apprenticeships, while simultaneously providing individualized supports to ensure academic and future success. Initially, Citizen Schools offered optional out-of-school time (OST) programming to middle school students in Boston, MA; it expanded to other districts in 2002, and further expanded to serve entire grade levels in an ELT framework in 2006. By 2010, Citizen Schools had shifted its focus from opt-in OST to mandatory full-grade ELT programming, and it began its national ELT initiative in 2010. Over the past six years, Citizen Schools ELT programming has expanded to schools in seven states, including California, Illinois, Massachusetts, New Jersey, New Mexico, New York, and Texas.

Abt Associates has been working as an evaluation partner with Citizen Schools since 2010, to learn how schools are implementing and integrating the Citizen Schools ELT model, and whether participation in ELT affects students' shorter-term behavioral and aspirational outcomes as well as students' academic performance. The evaluation was designed to examine implementation and impact from the beginning of Citizen Schools' national ELT expansion in fall 2010. This final report summarizes what the evaluation has learned about ELT implementation and outcomes over five consecutive school years, from 2010–11 through 2014–15.

Background and Findings from Prior Research

The research foundation for identifying the most effective strategies for implementing ELT has not kept pace with the steady expansion of ELT across the country. While there have been numerous studies, relatively few have been both rigorous and multi-faceted enough to assess the diversity of targeted outcomes and the heterogeneity of implementation. On the surface, for example, a study of a summer learning program and a study of ELT may each examine the impact of additional time, yet a summer-based program intervention is quite distinct from an expansion of instructional time during the regular school day. While some research does suggest positive findings about OST or summer

programming, those findings may not be applicable to an integrated ELT program model (McCombs et al., 2011; Redd et al., 2012). Studies that assess whether students attending charter schools (in which longer school days are required) outperform peers in regular district schools have found positive impacts on academic outcomes (Angrist et al., 2010; Tuttle et al., 2013; Tuttle et al., 2015). One of the strongest studies is a large-scale random assignment investigation of the Knowledge is Power Program (KIPP), a multi-faceted school improvement effort with many simultaneously implemented strategies, including ELT; this study found convincing positive effects on academic outcomes, yet cannot disentangle the effects of a longer school day from other key program elements (the seven principles that govern KIPP schools focus on academics, learning, and leadership[2]).

Several recent meta-analyses have examined studies of different strategies for increasing learning time, prioritizing studies based on stronger research designs, and the findings are mixed. For example, some research suggests that expanded learning time improves non-academic student outcomes (e.g., students' attendance, study skills, behavior, social skills, and motivation to learn); however, the evidence is based primarily on non-experimental research, including studies that used simple pre- and post-program comparisons of ELT participants and other quasi-experimental designs to identify correlations between ELT participation and youth outcomes (Zief et al., 2006; Redd et al., 2012). Kidron and Lindsay (2014) synthesized 30 studies, some of which were experimental and some quasi-experimental, and found that OST programs (before- and after-school and weekend programs) had a small, positive, and statistically significant effect on students' academic motivation. Another meta-analysis of the effects of afterschool programs on socio-emotional skill development found some positive and statistically significant effects on socio-emotional skill development, behavior management, school bonding, and positive self-perceptions (Durlak et al., 2010), but because the interventions could occur within and outside the regular school day, the effects of ELT cannot be examined separately from other school improvement strategies.

Overall, the evidence suggests three key findings on the implementation of ELT. First, ELT models vary in focus, structure, and content across school environments (Rocha, 2007). Second, time alone is not sufficient to improve student performance; quality academic learning time matters. Third, schools with successful expanded learning time programs share common features, including bold, visionary leadership; strong community support and partners; ongoing assessment and improvement; engaged students; high attendance and participation; and a culture of high expectation.

Given the large variability in ELT types and elements, in mandatory or voluntary ELT participation, and the concomitant implementation of other school-wide improvement initiatives, it can be difficult to pinpoint the source of any positive effects on students, both academically and non-academically. Further, the observed impacts largely reflect the benefits of instructional time within programs or interventions that are quite distinct from the Citizen Schools ELT model. Describing specific elements of the Citizen Schools' ELT model can help situate findings about both the implementation and impact of Citizen Schools ELT, relative to other research about ELT. The report turns next to a discussion of the Citizen Schools ELT program.

---

[2] See www.kipp.org/our-approach for more detail about key KIPP approaches.

The Citizen Schools ELT Program Model

The Citizen Schools Expanded Learning Time model is defined by three separate components: Apprenticeships, Academic Support, and Explore, which are built into a lengthened school day. The Apprenticeships are the cornerstone of the Citizen Schools ELT program model; they connect students to adult volunteers who teach a skill or content area about which they are passionate. The adult volunteers, called Citizen Teachers, teach about such varied topics as robotics, mock trials, poetry, dance, and numerous other areas. Students learn about possible apprenticeship topics early each semester via an Apprenticeship Fair; they rank their top choices after hearing Citizen Teachers present short pitches. Citizen Schools staff then assign students to apprenticeships based on a combination of student preference and availability. Students take four apprenticeships each year, two each semester. Each apprenticeship consists of 10 90-minute sessions per semester that culminates in a showcase called the WOW! during which students "teach back" to friends, family, and community members what they learned.

Apprenticeships are complemented by two distinct types of academic support. Structured homework time is generally offered for an hour each program day; it includes one-on-one goal setting and tutoring, and is known as AIM, or "aspire, invest, and make the grade." Academic League includes targeted academic support in either math or English/Language Arts (at each school's discretion); it is offered twice a week for between 30 and 90 minutes.

The third program component, called Explore, provides additional enrichment activities through team-building exercises. The Explore block generally reflects each participating school's priorities and circumstances, while establishing connections between students' middle school experiences, knowledge/guidance about college and careers (earlier, this component was known as C3—College to Career Connections), and how students can develop a pathway to future goals. Citizen Schools ELT also includes an eighth grade-specific program element, the Eighth Grade Academy (8GA), a capstone program that supports students as they navigate from middle to high school.

The chief mechanism by which program components are delivered is through an embedded "second shift" of educators, funded by AmeriCorps, hired to support the schools' extended day, either as Teaching Fellows (TFs) or Teaching Associates (TAs). Teaching Fellows generally work full-time and are expected to serve for two years. The requirements have become more explicit over time: TFs are required to have had prior experience working with children, have earned some college credit, and to be U.S. citizens. Teaching Associates, half-time Citizen Schools employees, generally provide Citizen Schools programming to students during the extended portion of the day, while TFs spend their mornings engaged in varied activities (e.g., providing support for the school, preparing for the afternoon programming). ELT programming generally begins around 3:00 PM, and concludes by 6:00 pm.

Illustrative Weekly Schedule for Students

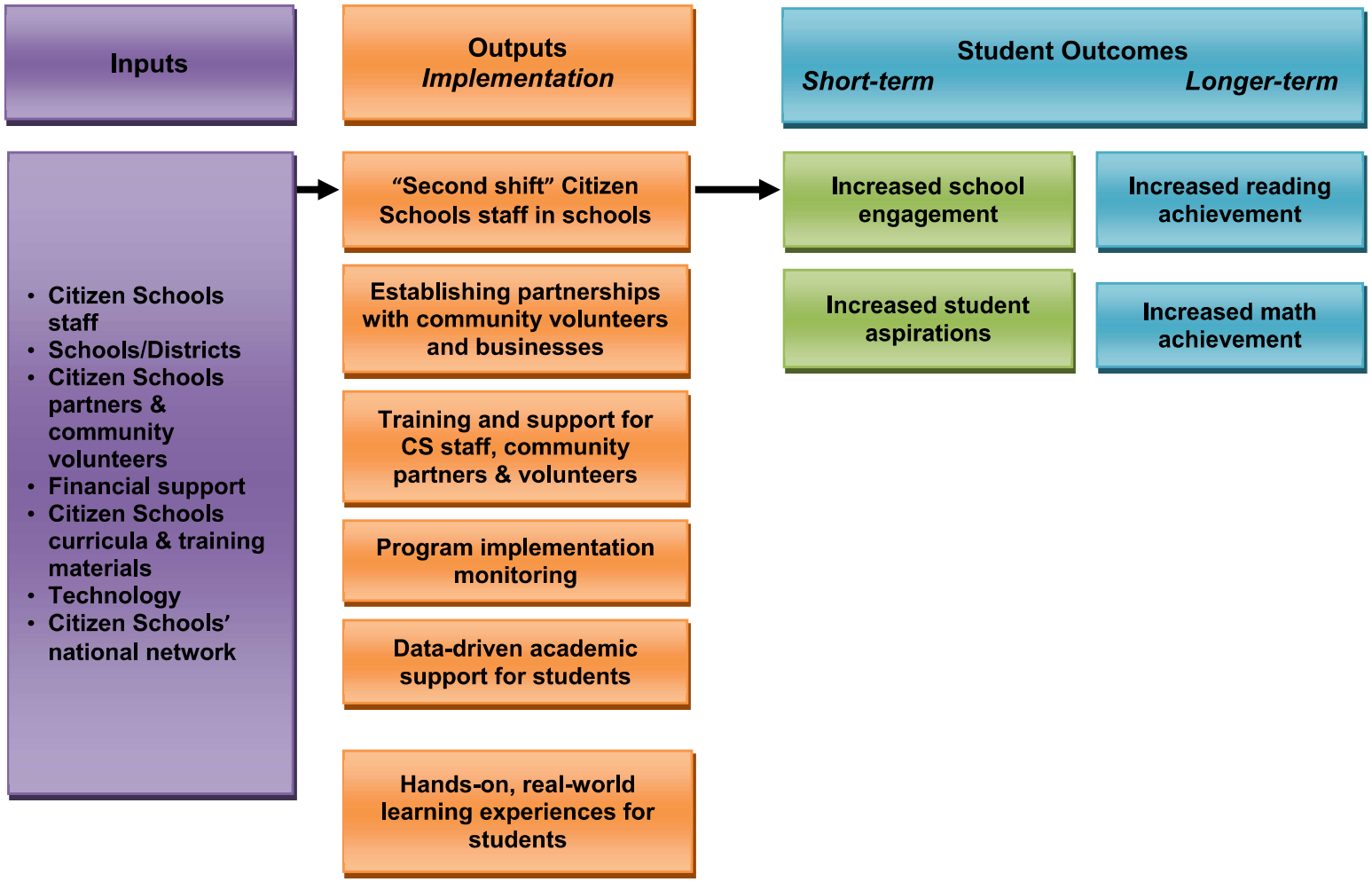
- **Monday:** 60 minutes of homework support, followed by 90-minute Academic League lesson on fractions
- **Tuesday**: 60 minutes of homework support, followed by 90-minute *Robotics* Apprenticeship taught by Citizen Teachers from Google
- **Wednesday:** 60 minutes of homework support, followed by 30 minutes of refresh/review on fractions, then 60 minutes of team-building exercises
- **Thursday**: 60 minutes of homework support, followed by 90-minute Mock Trials Apprenticeship taught by Citizen Teachers from local law firm
- **Friday:** No Citizen Schools programming typically offered

Although partner schools do not necessarily adhere to the same universal schedule network-wide, schools typically schedule activities across all three program components each week.

Through its systematic and loosely structured model, Citizen Schools provides consistency to students and their staff, and takes the burden of additional teaching away from first shift educators in their partner schools. Citizen Schools' ELT model also establishes clear expectations against which it can be evaluated. And, as noted above, the Citizen Schools ELT model is both similar to and different from other programs that include additional instructional time.

The evaluation of the Citizen Schools ELT Model incorporates the program elements described above into a logic model (Exhibit 1). The model illustrates the inputs, outputs, and expected outcomes, moving from left to right. The inputs represent the elements that, in combination, comprise the Citizen Schools ELT model, including the school/district, the Citizen Schools staff, partners and community volunteers, materials (e.g., curricula), supports, and technology. The outputs (e.g., activities/personnel represented in the model) include a second shift of ELT staff, establishing partnerships with community members and organizations, training, supporting, and monitoring Citizen Schools staff, and providing hands-on experiences for students. The underlying hypothesis is that if the model is implemented with fidelity, student engagement in school and student aspirations would increase (short-term outcomes), and correspondingly, student English/Language Arts (ELA) and math achievement, as measured by state standardized test scores, would then increase (longer-term outcomes).

Exhibit 1: Citizen Schools ELT Logic Model



Study Design and Research Questions

The Citizen Schools ELT Evaluation examined how Citizen Schools ELT programming is being implemented across sites nationally, and assesses whether there are impacts of Citizen Schools ELT participation on engagement, aspirational, and academic outcomes.

Key Study Design Features

- ***Multi-site, multi-year study***: staggered cohorts of schools (2010-11 through 2014-15)
  - 35 schools in four cohorts have implemented Citizen Schools ELT for one or more years
  - The cumulative achievement results include up to 27 schools (Cohorts 1, 2, 3, and 4)
- ***Implementation and outcome components***
  - Implementation focuses on Citizen Schools ELT programming
  - Non-academic outcome component focuses on student engagement and aspirational outcomes
  - Academic component uses quasi-experimental design to assess whether Citizen Schools has an impact on student achievement

This final report summary focuses on implementation and student outcomes from 2010-11 through the 2014-15 school years (achievement data are available only through the 2013-14 year). The implementation component of the study incorporates surveys, interviews, and site visits to assess how schools integrate Citizen Schools ELT into their school days, organizationally and instructionally. The data on implementation purposefully draw from multiple sources over different points in time to (1) ensure representation from the various stakeholders engaged in Citizen Schools ELT and (2) attempt to corroborate perspectives across stakeholder groups. Survey analyses assess the student engagement and aspirational outcomes from teacher, Citizen School staff, and student perspectives. Comparative analyses assess differences in these outcomes between Citizen Schools teachers and students and matched comparison individuals. The study also uses these data sources to assess the level of implementation more systematically in an implementation index.

The impact component of the study uses a comparative interrupted time series quasi-experimental (QED) research design; the design heeds the What Works Clearinghouse (WWC) guidance, and its sampling and analytic approaches are intended to meet WWC standards with reservations, which is the highest possible rating for a QED. The comparative interrupted time series is one of the strongest alternatives to a random control trial (RCT) to assess whether there are systematic differences in academic achievement for students in Citizen Schools ELT schools, compared to the same outcomes for their counterparts attending similar schools without Citizen Schools ELT. The study design relies upon carefully matching up to four local comparison schools to each Citizen Schools ELT campus, based on demographic and achievement profiles, to assess differences in academic achievement. Survey analyses are based on results from two of the four comparison schools for each Citizen Schools campus.[3]

The table below illustrates the study's data collection schedule.

[3] If a Citizen Schools ELT campus exits the study, so too do its matched comparison school counterparts.

Exhibit 2. Data Collection Schedule

Data Collection Activity	Study Year / School Year 1 2010–11	2 2011–12	3 2012–13	4 2013–14	*5* *2014–15*
Site visit (CS ELT only)	✓	✓	✓		
Principal Telephone Interview (CS ELT)		✓	✓	✓	✓
Principal Telephone Interview (MC)		✓	✓	✓	
Campus Director Telephone Interview		✓	✓	✓	✓
Survey					
*Student (CS ELT & MC)		✓	✓		
**Teacher (CS ELT & MC)		✓	✓	✓	
CS staff (CS ELT only)		✓	✓	✓	
Extant (test score) data	✓	✓	✓	✓	

*Student surveys were not collected in years 4 and 5.

**Teacher surveys were administered in Citizen Schools ELT schools (not comparison schools) in year 4; no surveys were administered in year 5.

In 2014–15, interviews were conducted with a subset of Citizen Schools campus administrators and campus directors, focusing specifically on schools designated as high implementers in the 2013–14 school year.

The study's guiding research questions include one question about implementation, one about shorter-term non-academic outcomes, and one about longer-term academic outcomes:

- How successfully are schools implementing and integrating expanded learning opportunities into their school days?
- What are the differences between Citizen Schools ELT and MC schools on student engagement and aspirations?
- What are the differences between Citizen Schools ELT and MC schools on student achievement (e.g., on English/Language Arts (ELA) and math achievement test scores)?

Key Findings

The study has found consistent patterns in survey, interview, and achievement findings. Implementation of the Citizen Schools ELT model takes time to establish and operate adequately, yet it also continues to vary substantially across Citizen Schools ELT schools, even as participating schools have demonstrably implemented several core principles of Citizen Schools ELT. Findings about short-term non-academic outcomes indicate some positive findings as well as some negative perceptions.

The data on student achievement in ELA and math indicated no statistically significant impacts of Citizen Schools ELT, on average. Exploratory subgroup analyses suggest a significant positive impact of Citizen Schools ELT on math achievement in the first year of implementation and a marginally significant positive effect of Citizen Schools ELT on 7th grade math achievement. In the context of other research about the impact of ELT on student achievement outcomes, these findings are not surprising. Few studies have found consistent positive outcomes, and the most positive

findings to date, from the most recent KIPP study (Tuttle et al., 2015) reflect implementation of a highly structured program model that emphasizes academic achievement rather than the considerably more flexible Citizen Schools ELT model that emphasizes varied learning opportunities designed to increase student engagement and aspirations.

Below, we summarize key findings about implementation and the implementation index, about student non-academic outcomes, and about impacts on student achievement.

Key Findings about Implementation

Over the course of several years, common patterns in implementation have become more evident. Attention to planning, along with clear and early communication with both school and district partners about roles and responsibilities helped set the stage for smoother program launch and implementation. Once program operations began in earnest, schools faced a number of common challenges, including changes in district or school leadership, and therefore changes in priorities, as well as maintaining program consistency given high staff turnover among the Teaching Fellows.

Because Citizen Schools programming occurs primarily through its staff, recruitment, hiring, preparation and support of staff directly affect how the program operates and is perceived. Participating campuses experienced similar challenges over the study's duration, including training and experience, hiring practices, staff turnover, and staff diversity. Establishing strong working relationships early on was consistently described as essential, as was ensuring that staff had access to training that helped them understand local school needs. Citizens Schools Fellows' capacity to manage student behavior effectively and provide focused instruction was consistently reported as a challenge by Citizen School staff and campus administrators, and staff turnover was a challenge in about half the schools.

Integrating Citizen Schools staff into schools generally occurred through three mechanisms: 1) structured overlap and joint participation between the first and second shift staff; 2) alignment of content, pedagogy, and especially behavior management systems; and 3) establishing and maintaining relationships between stakeholder groups. Developing purposefully structured overlaps between first and second shifts generated benefits, goodwill, and enhanced integration efforts. In about half the schools, Citizen Schools staff had shared access to information about student performance, typically via schools' data systems or joint participation in meetings with first shift teachers. Integration also proceeded more smoothly when there were positive working relationships between Citizen Schools staff and first shift teachers.

Over the past several years, the Citizen Schools Program model has demonstrably changed in response to network schools' continued adaptations of core program elements and the national curricula. The Citizen Schools organization adjusted its approach to developing and using its national curricula to make it more responsive to schools' needs. Even as the national organization has become more flexible, network campuses continued to adapt core program elements to meet their own students' needs.

By the end of the 2014–15 school year, the Citizen Schools network included campuses in their second through fifth years of implementation. The network schools and the Citizen Schools organization have faced persistent questions about how to sustain school/district commitment and participation, particularly for schools relying upon time-limited federal funding. The two most

prevalent concerns include (1) financial sustainability, and (2) campus stability—and therefore programmatic sustainability—given frequent staff transitions.

Implementing the Seven Key Elements of the Citizen Schools ELT Model across the Network and Over Time

Over the course of the Citizen Schools ELT evaluation, it has become and remains evident that implementation varies substantially across participating campuses. The Citizen Schools organization is keenly interested in understanding how the Citizen Schools ELT schools vary in their implementation of the model. The study team, working collaboratively with Citizen Schools, developed a summary metric, called the implementation index, to capture individual campus and network-wide progress on implementation of core Citizen Schools ELT program elements—the seven "non-negotiable" processes and activities that together, represent what successful implementation of the program would require. The implementation index draws from surveys administered in schools' second and subsequent years of implementation as well as annual interviews. The index incorporates information from multiple school-based respondents, including principals, classroom teachers, Citizen Schools CDs and Citizen Schools Teaching Fellows.

The multi-dimensional index includes the following key elements of the Citizen Schools program model:

1) Planning;
2) Leadership;
3) Data collection;
4) Training and professional development;
5) Family/community engagement;
6) Alignment/coordination between partner school and CS; and
7) Perceptions of program quality.

Each school's implementation (in a given year) is characterized as limited, moderate, or full (corresponding to ratings of 1, 2, or 3), for each of the seven core program constructs listed above. School ratings could vary between 7 (representing limited implementation of each construct) and 21 (full implementation across all seven constructs). The index provides information about progress, consistency, and variability—whether for individual campuses, for cohorts of schools that began implementation at specific times, or for the entire network, both at a given point in *time* (e.g., as of the 2012–13 academic year) as well as at a given point in *implementation* duration (e.g., as of the second year of program implementation).

Overall index scores remain clustered in the moderate range, with average scores hovering around 17 (out of 21) across multiple cohorts and up to four implementation years. There are few patterns evident in schools' levels of implementation over time. While schools with lower index scores tended to exit the network over time, continuing schools' implementation levels do not change substantially from one year to the next. The specific constructs schools were most consistently able to implement fully *across* implementation years include planning, leadership, perceived quality, and data collection. For example, 16 schools (of 19 for which index scores are available) in their second year, 9 (of 11) in their third year, and 3 (of 3) in their fourth year scored a "3" for the planning construct. Fewer schools implemented either the family and community engagement or alignment of partner school and Citizen Schools constructs fully, again, *across* implementation years.

How Does Citizen Schools ELT Affect Student and Teacher Perceptions?

Survey findings help to describe teacher and student perceptions about the Citizen Schools ELT model, and provide some insight into whether and how the program is progressing toward the shorter-term outcomes depicted in the logic model. The results are consistent with implementation findings that some program elements are positively perceived (exposure to learning about colleges and careers), and that there are some ongoing puzzles and challenges (less positive perceptions of Citizen Schools staff).

Overall, significantly more students in Citizen Schools ELT campuses than the counterfactual reported positive behavioral engagement of peers, that the ELT program helped their self-esteem and pro-social behaviors, and that they participated in activities to help them learn about college and careers. At the same time, however, significantly more Citizen Schools students also reported concerns with Citizen Schools staff and the Citizen Schools ELT program than those comparison students who reported that they attended after-school programming.

Student and faculty staff perceptions differed on some dimensions. Specifically, significantly more Citizen Schools' ELT students indicated that they were positively engaged and had high aspirations, while more school teachers/Citizen Schools staff reported student engagement and student aspirations were a problem than occurred in schools without ELT.

How Does Citizen Schools ELT Affect Student Achievement?

Following the recommendations of the U.S. Department of Education's What Works Clearinghouse (WWC) standards for rigorous designs, the student achievement analyses were divided into confirmatory (central hypotheses to the evaluation) and exploratory analyses (hypotheses that are important yet are not central and may not be well powered).[4] The data on student achievement in ELA and math indicated no statistically significant impacts of Citizen Schools ELT, on average. Exploratory subgroup analyses suggest a significant positive impact of Citizen Schools ELT on math achievement in the first year of implementation roughly equivalent to just over three months of math growth in a school year and a marginally significant positive effect of Citizen Schools ELT on 7th grade math achievement roughly equivalent to about three-and-a -half months of math growth in a school year. In the context of other research about the impact of ELT on student achievement outcomes, these findings are not surprising. Few studies have found consistent positive outcomes, and the most positive findings to date, from the most recent KIPP study (Tuttle et al., 2015) reflect implementation of a highly structured program model that emphasizes academic achievement rather than the considerably more flexible Citizen Schools ELT model that emphasizes varied learning opportunities designed to increase student engagement and aspirations.

Discussion

This final report provides a comprehensive overview of the Citizen Schools ELT model after five years of implementation. Over that time, the Citizen Schools ELT network has matured and expanded as new schools joined each year, and schools exited after one, two, three, or four years. The ELT program model has become more flexible as partner schools have refined and adapted the program

[4] http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_draft_standards_handbook.pdf

model to meet their needs. Over the past five years, the study has examined both implementation and outcomes in detail.

The implementation-focused findings presented in this report center on implementation variability. Overall, the ELT schools are clearly committed to implementing Citizen Schools' ELT programming with fidelity, and at the same time, Citizen Schools has continued to recognize the need for flexibility with its campus partners. As a result, there is considerable variation across individual campuses in how the model is incorporated into their respective school contexts; indeed, that variability seems to be essential for the model to be adaptable across such diverse contexts. The implementation findings also highlight some of the challenges associated with launching a multi-faceted model in dynamic settings, coupled with built-in staffing changes.

The outcome-focused findings described in this report are mixed. The study finds both positive effects on student engagement and aspirations and negative perceptions about students' Citizen Schools ELT experiences. The confirmatory findings indicate no overall significant impact on student performance, as measured by standardized achievement test scores in ELA and math; exploratory findings indicated a significant positive impact of Citizen Schools ELT on math achievement in the first year of implementation and suggest a marginally significant positive effect of Citizen Schools ELT on 7th grade math achievement. Statically significant impacts on student achievement have proven persistently elusive and it is unclear whether this is a function of variability in implementation, model variation across campuses, statistical power, the fit between the intervention and the outcome measures, or some combination of the above.

Prior research on other interventions that include expanded learning time has yielded mixed evidence; few studies found positive achievement impacts, most studies found no achievement impacts, and some found negative effects. Perhaps unsurprisingly, interventions designed to improve academic achievement, such as KIPP, or Higher Achievement, were more likely to affect academic outcomes than were more holistic interventions designed to broaden students' enrichment, socio-emotional, or other non-academic experiences. One of the distinctive features of the Citizen Schools ELT model is its emphasis on non-academic learning opportunities that are hypothesized to be necessary precursors to improved achievement. Perhaps student engagement and aspirational short-term outcomes are more appropriate outcomes on which to focus, given Citizen Schools' emphasis on team-building and exposing students to novel, hands-on, real-world experiences through apprenticeships.

Limitations

The study was designed to answer key questions about implementation and impact using the strongest possible approaches, yet the study's analytic approaches have some limitations. For example, data on implementation relied upon semi-structured interviews in which respondents could and did answer questions in varying detail, and, because campuses differed in the number (and availability) of stakeholders participating in interviews, the study was not necessarily able to interview staff members with comparable positions at each site. Additionally, interviews and surveys were completed after the ELT program had been introduced, and while student achievement data could be obtained retrospectively—dating back to before the program was implemented—data about school climate and short-term outcomes reflect experiences *after* schools had begun to implement Citizen Schools ELT, and may or may not reflect *changes* as a result of ELT.

Two other limitations are worth heeding: one, the study examines Citizen Schools' specific approach to ELT, and two, it uses a purposefully constructed sample of schools. The Citizen Schools ELT model is distinctive, and lessons derived from this study apply to this specific approach to ELT, and reflect the idiosyncrasies of the study sample. The study sample has experienced considerable fluctuation over its duration, and while some attrition may occur because either schools or the Citizen Schools organization recognize the "fit" (or lack thereof) for the ELT model, the fluidity of the sample raises a concern about external validity and the sample's representativeness. The study describes implementation progress and challenges for all 35 schools for at least one year of implementation—but the program is not currently designed as a one-year or even a two-year intervention. The cumulative results (both implementation and outcomes) are therefore based on a subset of schools and may not apply to all schools that had ever participated in the Citizen Schools network.

Finally, one other consideration is noteworthy. Performance on standardized achievement tests may not be the most sensitive measure of ELT's potential impact on student learning. Perhaps achievement outcomes are too distal a measure for Citizen Schools' ELT model, given its emphasis on such non-academic outcomes as student motivation, engagement, and aspiration, as well as on better understanding the high school application processes and what comes after secondary school (e.g., college and careers). Performance on core academic subject assessments is clearly important, but it may not be the right outcome for the ELT model.

Future Steps

At the conclusion of this five-year evaluation, the Citizen Schools ELT initiative continues to anticipate its school partners' needs, as new campuses join the network and some current school partners exit. The Citizen Schools' ELT model has undergone some fundamental changes as well, responding to school partners and secular shifts in both academic and non-academic priorities. Citizen Schools' programming will also likely adjust to changes in the education landscape resulting from the recently authorized Every Student Succeeds Act (ESSA), as the legislation simultaneously reduces federal reach into local education decisions and provides states and districts more control over teacher evaluation, standards, school turnarounds, and accountability. Given that Citizen Schools has focused chiefly on working with turnaround schools, the new legislation will undoubtedly influence not only how the organization engages new district and school partners, but also how it will adjust its programming and staffing. Hopefully the findings described in this report can provide useful information as the Citizen Schools organization plans for the future.

1. Introduction

Academic achievement in the nation's lowest performing schools is a persistent concern and object of scrutiny. Despite improvements in students' overall academic achievement over the past few decades, proficiency gaps in reading and mathematics still persist across income, racial, and ethnic groups (National Center for Education Statistics, 2013). One promising strategy for reducing the achievement gaps is the use of a longer school day, often called Expanded Learning Time (ELT). There is also evidence that ELT programs that increase students' access to enrichment activities can have positive effects on students' academic motivation social-emotional skill development.

In recent years, the number of schools across the country implementing ELT has increased. According to the National Center for Time and Learning (NCTL), by 2013–14 (the most recent school year for which data are available), over 2,000 schools were using ELT strategies. Nearly half of these schools were charter schools, while many low-performing regular district schools have adopted ELT strategies in response to federal requirements for the School Improvement Grant (SIG) program. According to a recent GAO report, approximately 1,800 SIG-funded schools, and 69 schools funded under the 21st Century Community Learning Program are implementing some form of ELT strategies.[5] Current guidance under the SIG program defines increased learning time as "increasing the length of the school day, week, or year to significantly increase the total number of school hours" for instruction in core academic subjects and other subjects, enrichment activities, and teacher collaboration, planning, and professional development (U.S. Department of Education, 2012a, p. 23).

The definition of ELT may seem straightforward, yet ELT implementation varies considerably in practice: it can take the form of additional days in the school year, or more hours in a school day, or can occur in summer programming above and beyond the regular year or day. Participation in ELT can also be mandatory for designated students, grades, or schools, or can permit students to opt in or out. Moreover, responsibility for organizing and delivering ELT programming might be in districts' or schools' hands, or may be overseen by community-based organizations such as Boys and Girls Clubs, or other non-profit organizations, regardless of whether teachers, tutors, community volunteers, or other adults provide actual instruction. And even when schools indicate they are implementing ELT using an identical *structure* (i.e., all after-school or summer or longer day programming), there may well be meaningful differences in the content of ELT, or in how additional time is distributed across subject areas, extracurricular activities, and in responsibility for instruction. The definition of ELT, therefore, includes a broad array of approaches and programs that encompass extensions of the school day or year, and that include both programs situated within and outside K–12 school settings.

Citizen Schools works primarily with middle schools serving largely low-income (90% nationally), racial or ethnic minority (94% nationally), and academically struggling students. Citizen Schools relies upon an additional shift of educators and community volunteers to engage middle school students in hands-on apprenticeships, while simultaneously providing individualized supports to ensure academic and future success. Since 1995, Citizen Schools has developed and implemented its

[5] See http://www.gao.gov/products/GAO-16-141

own approaches to expanding the school day. Initially, Citizen Schools offered optional out-of-school time (OST) programming to middle school students in Boston, MA; it expanded to other districts in 2002, and further expanded to serve entire grade levels in an ELT framework in 2006. By 2010, Citizen Schools had shifted its focus from opt-in OST to mandatory full-grade ELT programming, and it began its national ELT initiative in 2010. Over the past six years, Citizen Schools has purposefully expanded to serve increased numbers of schools and students in districts across seven states.

Abt Associates has been working with Citizen Schools since 2010, as an evaluation partner, to learn how schools are implementing and integrating the Citizen Schools ELT model, and whether participation in ELT affects students' shorter term behavioral, and aspirational outcomes as well as students' academic performance . The evaluation was purposefully designed to examine implementation and impact from the beginning of Citizen School's national ELT expansion in fall 2010. This final report summarizes what the evaluation has learned about ELT implementation and outcomes over five consecutive school years, 2010–11 through 2014–15.

1.1 Approaches to ELT

The research foundation for identifying the most effective strategies for implementing ELT has not kept pace with the steady expansion of ELT across the country. While there have been numerous studies, relatively few have been both rigorous and multi-faceted enough to assess the diversity of targeted outcomes and the heterogeneity of implementation. Summarizing the research evidence about effective ELT implementation and impact, therefore, requires examining multiple studies that individually may highlight isolated contributions of different facets of ELT programming. Studies that examine OST or summer programming may point to strategies that contribute to positive outcomes for students, yet the nature of such programming is qualitatively distinct from an integrated ELT program model (McCombs et al., 2011; Redd et al., 2012). Other studies have examined multi-faceted school improvement efforts, of which ELT is one of many simultaneously implemented strategies; for example, a large-scale random assignment investigation of the Knowledge is Power Program (KIPP) found convincing positive effects, yet did not disentangle the effects of a longer school day from other key program elements (the seven principles that govern KIPP schools focus on academics, learning, and leadership[6]).

Districts and schools have implemented multiple ELT strategies which vary in focus, structure, and content (Rocha, 2007). Recent research highlights several findings, including:

- ***Increasing Time in School.*** ELT implementation can mean longer academic years, longer school days, summer programming, or mandatory or voluntary after-school programming, any of which can be staffed by teachers and/or community-based organizations (GAO, 2015; Kolbe et al., 2012). While School Improvement Grants (SIGs) and state policies establish minimum thresholds, districts and schools have considerable flexibility in how they build time into school schedules. Kolbe and colleagues found that urban schools, including schools with large percentages of minority students and higher enrollments, were more likely to extend the school year beyond the average 180 days a year. Moreover, traditional public schools that lengthened

[6] See www.kipp.org/our-approach for more detail about key KIPP approaches.

their school day (7 or more hours) were more likely to serve students at risk for academic failure (Kolbe et al., 2012).

- ***Prioritizing Time for Student Learning***. Even within schools that are implementing ELT, variation in how additional time is allocated across different subject areas can vary dramatically. Some research suggests that students in schools with an extended day receive more instruction in core academic courses (ELA, mathematics, social studies, science) than do their peers in schools without an extended day (Kolbe et al., 2012). An earlier study of after-school programs focused on English/Language Arts and math found initial positive effects for math and no effects for reading (Black et al., 2009). In its publication, *Time Well Spent*, NCTL found that high-performing schools with ELT increased students' access to more rigorous academic instruction (NCTL, 2012). Most recently, the Center on Education Policy found that schools may add instructional time for students by extending the beginning or end time of the regular school day, decreasing non-instructional time in the day, or by including additional days to the school year (McMurrer et al., 2015).
- ***Improving the Quality of Instruction***. There is increasing evidence that the impact of additional time is influenced by the quality of the instruction students receive during that time, as students may be participating in additional instruction while their regular classroom teachers are participating in professional development (Berliner, 1990; Aronson et al., 1999; Hammond et al., 2009; Yoon et al., 2007). A meta-analysis of over 1,300 studies found that programs demonstrating "a positive and significant effect" on student achievement averaged almost 50 hours in professional development a year (Yoon et al., 2007). It is not surprisingly, therefore, that some ELT schools have allocated more time for teachers to collaborate on instructional planning and participate in professional development opportunities (After School Alliance, 2012; Checkoway et al., 2012; McMurrer et al., 2015).
- ***Increasing Focus on Academic Support and Enrichment Activities***. Many schools use additional time to include enrichment and academic activities, and rely on community-based organizations or third-party providers to offer these services to students. In some cases, these activities are designed to address expanded learning program goals in academics, remediation, or enrichment. For example, some schools engage youth by providing hands-on learning experience that focus on STEM, sports, music or the arts. Additional opportunities may include apprenticeships, tutoring, or college and career exploration (Afterschool Alliance, 2012; Rulf Fountain et al., 2013; McMurrer et al., 2015).

Making additional time for learning effective for students means heeding multiple strategies about ELT simultaneously. For example, the Afterschool Alliance (2012) identified eight research-based principles it believes characterize strong ELT programming, including strong school community partnerships; engaged learning; family engagement; intentional programming; diverse, prepared staff; participation and access; safety, health, and wellness; and ongoing assessment and improvement.

Key Elements of Successful Expanded Learning Time Approaches

Key elements[a] of expanded learning time initiatives include:

- Increase time in schools for students
- Provide academic support, individualized learning support, and enrichment activities to prepare students for college and career success
- Ensure opportunities for teacher collaboration and professional development to improve the quality of instruction
- Build strong school leadership and support of ELT
- Promote data-driven and evidence based support to improve student performance
- Cultivate dedicated partnerships with external organizations
- Build a culture of high expectations and accountability

[a] See, for example: After School Alliance (January 2012), *Principles of Effective Expanded Learning Programs: A Vision Built on the Afterschool Approach*. Washington, DC: Author. Retrieved from http://www.afterschoolalliance.org/Principles%20of%20Expanded%20Learning%20Programs_Jan_2012(2).pdf; see also Claire Kaplan & Roy Chan (2011), Time Well Spent: Eight Powerful Practices of Successful Expanded-Time Schools, National Center for Time and Learning. Retrieved from http://timeandlearning.org/sites/default/files/resources/timewellspent.pdf

Leadership also plays a key role in successful reform efforts that include ELT. For example, ECONorthwest researchers suggested that successful ELT programs share multiple characteristics: strong school leadership; teacher commitment and leadership; use of evidence-based and data-driven decision-making; active engagement of parents, partners, and the community to help support ELT; and clear focus on core academics and enrichment activities aligned with other goals and reforms (2008).

Several recent studies of charter schools have identified both promising practices for ELT, as well as positive associations between ELT and improved student achievement. NCTL examined 30 high-performing expanded time schools, and identified eight practices common among them: maximizing instructional time; prioritizing instruction according to an instructional focus; providing individualized academic support; establishing a culture of high expectations; providing a well-rounded education including enrichment activities; readying students for college and careers; working to strengthen instruction; and using data to direct instruction (Kaplan & Chan, 2011). Fryer and Dobbie (2010) conducted a study of New York City charter schools (for which longer school days were fundamental), and found that improved student achievement was positively associated with frequent teacher feedback, the use of data to guide instruction, and high-dosage tutoring.

1.1.1 Holistic Education through ELT

Schools are widely perceived to help students develop not only academic skills, but also a broader set of knowledge, abilities, and experiences designed to help them achieve both academic and life success (Bushaw & Lopez, 2013). Indeed, there is evidence that a well-rounded, holistic education, which includes increased access to arts, music, and other enrichment activities, has positive and significant impacts on student learning and growth (Burton et al., 1999; Hetland et al., 2007; Kisiel et al., 2006). For example, students who participated in at least three years of arts instruction reported more positive attitudes toward school and learning than peers with less access to arts education, using student self-reports (Burton et al., 1999). Similarly, "culturally enriching" field trips, which have a long history in American education, can play an important role in child/youth development; one recent study compared students who participated in a program that included tours of a local art museum (Arkansas' Crystal Bridges Museum of American Art) to students who did not attend such

tours, and program students demonstrated higher critical thinking skills than comparison students. These gains were more pronounced among poorer and rural students (Greene et al., 2014).

While the research draws positive correlations between enrichment programming and student learning outcomes, many schools, particularly lower performing schools, find it challenging to provide these opportunities within the confines of the standard American school schedule (Farbman, 2015; Government Accountability Office, 2009; Rabkin & Hedberg, 2008). Unfortunately, enrichment activities and classes can take second place when academic performance is prioritized. For many schools, ELT is seen as one promising approach for providing a holistic education and improving academic achievement.

NCTL's *Time Well Spent* report found that schools allocate about an hour per day for enrichment activities in 30 academically high-performing schools with longer school days and/or years (Kaplan & Chan, 2011; Farbman, 2015). Similarly, data from the U.S. Department of Education's (ED's) Schools and Staffing Survey (SASS) indicated that schools with more time in their days were able to provide students with a broad range of instructional programs in art, foreign language, and physical education (Kolbe et al., 2012). Through partnerships with community-based organizations, expanded learning time schools can offer an extensive array of learning opportunities that may include apprenticeships with local businesses, STEM classes (e.g., robotics, engineering, astronomy), drama, dance, and martial arts (Checkoway et al., 2012; Kolbe et al., 2012).

1.1.2 The Impact of ELT on Non-Academic Outcomes

While some research suggests that expanded learning time improves non-academic outcomes for youth, the evidence is based primarily on non-experimental research, including studies that used simple pre and post-program comparisons of ELT participants and other quasi-experimental designs to identify correlations between ELT participation and youth outcomes (Zief et al., 2006; Redd et al., 2012). The vast majority of studies include such non-academic outcomes as attendance, students' study skills, behavior, social skills, and motivation to learn. Moreover, research on these outcomes often includes ELT as well as OST or other programs that occur outside the regular school day. For example, Kidron and Lindsay (2014) synthesized 30 studies, some of which were experimental and some quasi-experimental, and found that OST programs (before- and afterschool and weekend programs) had a small, positive, and statistically significant effect on students' academic motivation (0.04 standard deviations), and a negligible effect on social-emotional skill development (0.03 standard deviations).[7]

Yet other research is mixed. Another meta-analysis of the effects of afterschool programs on socio-emotional skills development found examples of afterschool programs in which there were positive and statistically significant effects on socio-emotional skill development, behavior management, school bonding, and positive self-perceptions (effect sizes of 0.19, 0.14, and 0.34, respectively) (Durlak et al., 2010). Unlike Kidron and Lindsay (2014), these other syntheses included both one-on-

[7] Kidron and Lindsay (2014) used an effect size of 0.25 as a benchmark of "educational significance," as recommended by Hill et al. (2008).

one and group-based interventions, which can occur within and outside the regular school day, and therefore could not disentangle ELT from other school improvement strategies.[8]

The evidence from experimental studies about the impact of OST participation on non-academic outcomes is mixed; Zief and colleagues conducted a meta-analysis of six experiments, and found one study with positive and significant effects on social and emotional outcomes (Lauver, 2002), and five studies with null effects of OST participation (Zief et al., 2006). Redd and colleagues (2012) reviewed 22 studies with either experimental or quasi-experimental designs and found equivocal results; four experimental studies showed positive effects on attendance and engagement and five studies showed no effects on such outcomes. The most recent review, from Kidron and Lindsay (2014), included 30 studies, and found some statistically significant yet modest effects on students' academic motivation. On balance, then, the evidence about the positive effects of ELT or OST on non-academic outcomes is further tempered by potential confounding of the effects of ELT and the effects of the larger interventions of which ELT is one element.

1.1.3 The Impact of ELT on Academic Outcomes

The most rigorous evidence about the impact of ELT on students' academic performance is mixed; Kidron and Lindsay (2014) found no statistically significant effects of out-of-school programs on literacy and math achievement, echoing the 2006 results from the earlier Zief et al. meta-analysis. However, while some current research has suggested a positive relationship between ELT and improved academic achievement, the positive findings could not disentangle the specific contributions of ELT from other components of multi-faceted interventions.

Most recently, Tuttle and colleagues found a positive impact of attending KIPP public charter schools on student achievement in both elementary and middle schools. Using both a lottery-based and a quasi-experimental, matched comparison group design, the authors suggested that KIPP middle school attendance significantly increases a student's average math score (0.24 and 0.18 standard deviation units, respectively, for one and two years) (Tuttle et al., 2015). The results are consistent with previous quasi-experimental studies of KIPP schools, which reported positive and statistically significant impacts in math and English/language arts among KIPP students compared to similar non-KIPP students over a 10-year period (Tuttle et al., 2015; Angrist et al., 2010; Doran et al., 2002; Economic Policy Institute, 2005; Gallagher & Ross, 2005; MacIver & Farley-Ripple, 2007; McDonald et al., 2008; Musher et al., 2005; Ross et al., 2007; Tuttle et al., 2010; Woodworth et al., 2008). Yet prior research, including the research on KIPP schools, has been unable to isolate the effects of ELT programs from other comprehensive school models to measure student achievement (Redd et al., 2012). Most disadvantaged schools adopt ELT activities and structures as part of a broader school improvement plan, which typically includes multiple school reform components. Because these varied program components are implemented simultaneously, it is difficult to determine whether reported academic gains derive from the programming offered during the regular or the extended school day, from other aspects of school improvement models, and/or the combination of these elements.

[8] The studies included in the Kidron and Lindsay review do not overlap with those included in the earlier syntheses.

Recent charter school research links charter school characteristics with improvements in student achievement. Fryer and Dobbie (2011) found that instructional time of 300 or more hours than the standard district calendar was a strong predictor of academic achievement (including tutoring, feedback to teachers, use of data, and high expectations). Hoxby and colleagues (2009) reported that charter school enrollment in New York City had a significant effect on student achievement; an additional 10 instructional days was associated with an increase of 0.2 standard deviation units in annual achievement gains. Angrist, Pathak, and Walters (2011) showed that urban charter schools were more effective at raising test scores than non-urban charter schools. Other research has indicated that additional instructional time leads to statistically significant and substantially important positive effects on the literacy achievement among at-risk students and students performing below standards (Kidron & Lindsay, 2014; Redd et al., 2012; Tuttle et al., 2010; Patal et al., 2010; Wheeler, 1987). For example, expanded learning time provided particular benefits to students at increased risk of academic failure and dropout, including members of minority groups, students with low performance records for standardized testing, and students eligible for free- or reduced-price lunch (Redd et al., 2012). However, Redd and colleagues also found that when programs were not implemented well, expanded learning time was associated with some negative impacts, and that it may have had adverse consequences for higher achieving students.

Overall, the evidence suggests three key findings on the implementation of ELT. First, ELT models vary in focus, structure, and content across school environments (Rocha, 2007). Second, time alone is not sufficient to improve student performance; quality academic learning time matters. Third, schools with successful expanded learning time programs share common features, including bold, visionary leadership; strong community support and partners; ongoing assessment and improvement; engaged students; high attendance and participation; and a culture of high expectation.

Given the large variability in ELT types and elements, in mandatory or voluntary ELT participation, and the concomitant implementation of other school-wide improvement initiatives, it can be difficult to pinpoint the source of any positive effects on students, both academically and non-academically. Further, the observed impacts largely reflect the benefits of instructional time within programs or interventions that are quite distinct from the Citizen Schools ELT model. One way to examine both the implementation and impact of Citizen Schools ELT, therefore, is to describe the specific elements of Citizen Schools' ELT model. The report turns next to a discussion of the program model.

1.2 The Citizen Schools ELT Program Model

The Citizen Schools Expanded Learning Time model is defined by three separate components: Apprenticeships, Academic Support, and Explore, which are built into a lengthened school day. The Apprenticeships are the cornerstone of the Citizen Schools ELT program model; they connect students to adult volunteers interested in teaching a skill or content area about which they are passionate and have clear practice knowledge. The adult volunteers, called Citizen Teachers, reflect a wide range of backgrounds, experiences, and content knowledge, and teach about such varied topics as robotics, mock trials, poetry, dance, and numerous other areas. Students learn about possible apprenticeship topics at the beginning of each semester via an Apprenticeship Fair, where Citizen Teachers present short pitches, and students rank their top choices. Citizen Schools staff then assign students based on a combination of student preference and apprenticeship availability. Students take four apprenticeships each year, two each semester. Each apprenticeship consists of 10 90-minute

sessions during a semester that culminates in a showcase called the WOW! during which students "teach back" to friends, family, and community members what they learned.

Apprenticeships are complemented by two distinct types of academic support. Structured homework time is generally offered for an hour each program day; it includes one-on-one goal setting and tutoring, and is known as AIM, or "aspire, invest, and make the grade." Academic League includes targeted academic support in either math or English/Language Arts (at each school's discretion); it is offered twice a week for between 30 and 90 minutes.

The third program component, called Explore, provides additional enrichment activities through team-building exercises. The Explore block generally reflects each participating school's priorities and circumstances, while establishing connections between students' middle school experiences, knowledge/guidance about college and careers (earlier, this component was known as C3—College to Career Connections), and how students can develop a pathway to future goals. Citizen Schools ELT also includes an eighth grade-specific program element, the Eighth Grade Academy (8GA), a capstone program that supports students as they navigate from middle to high school.

The chief mechanism by which program components are delivered is through a "second shift" of educators hired to support the schools' extended day, either as Teaching Fellows (TFs) or Teaching Associates (TAs) who are embedded with the schools. Teaching Fellows generally work full-time, are funded by AmeriCorps, and are expected to serve for two years. The requirements have become more explicit over time: TFs are required to have had prior experience working with children, have earned some college credit, and to be U.S. citizens. In practice, most have a college degree. Teaching Associates, half-time Citizen Schools employees, generally provide Citizen Schools programming to students during the extended portion of the day, while TFs spend their mornings engaged in varied activities (e.g., providing support for the school, preparing for the afternoon programming). ELT programming generally begins around 3:00 PM, and concludes by 6:00 pm. Although partner schools do not necessarily adhere to the same universal schedule network-wide, schools typically schedule activities across all three program components each week.

Illustrative Weekly Schedule for Students

- **Monday:** 60 minutes of homework support, followed by 90-minute Academic League lesson on fractions
- **Tuesday**: 60 minutes of homework support, followed by 90-minute *Robotics* Apprenticeship taught by Citizen Teachers from Google
- **Wednesday:** 60 minutes of homework support, followed by 30 minutes of refresh/review on fractions, then 60 minutes of team-building exercises
- **Thursday**: 60 minutes of homework support, followed by 90-minute Mock Trials Apprenticeship taught by Citizen Teachers from local law firm
- **Friday:** No Citizen Schools programming typically offered

Through its systematic and loosely structured model, Citizen Schools provides consistency to students and their staff, and takes the burden of additional teaching away from first shift educators in their partner schools. Citizen Schools' ELT model also establishes clear expectations against which it can be evaluated. And, as noted above, the Citizen Schools ELT model is both similar to and different from other programs that include additional instructional time.

1.2.1 Organization of the Report

The remainder of this report presents study findings on the implementation and outcomes of the evaluation. Chapter 2 describes the study, including its design, the research questions, the study sample, and data collection and analysis. Chapter 3 summarizes key implementation themes to assess how successfully Citizen Schools ELT schools are implementing and integrating expanded learning opportunities into their school days. Chapter 4 examines how participation in Citizen Schools programming affects student academic and non-academic outcomes. Finally, Chapter 5 summarizes the findings, including lessons learned over this multi-year study.

2. Study Design and Methods

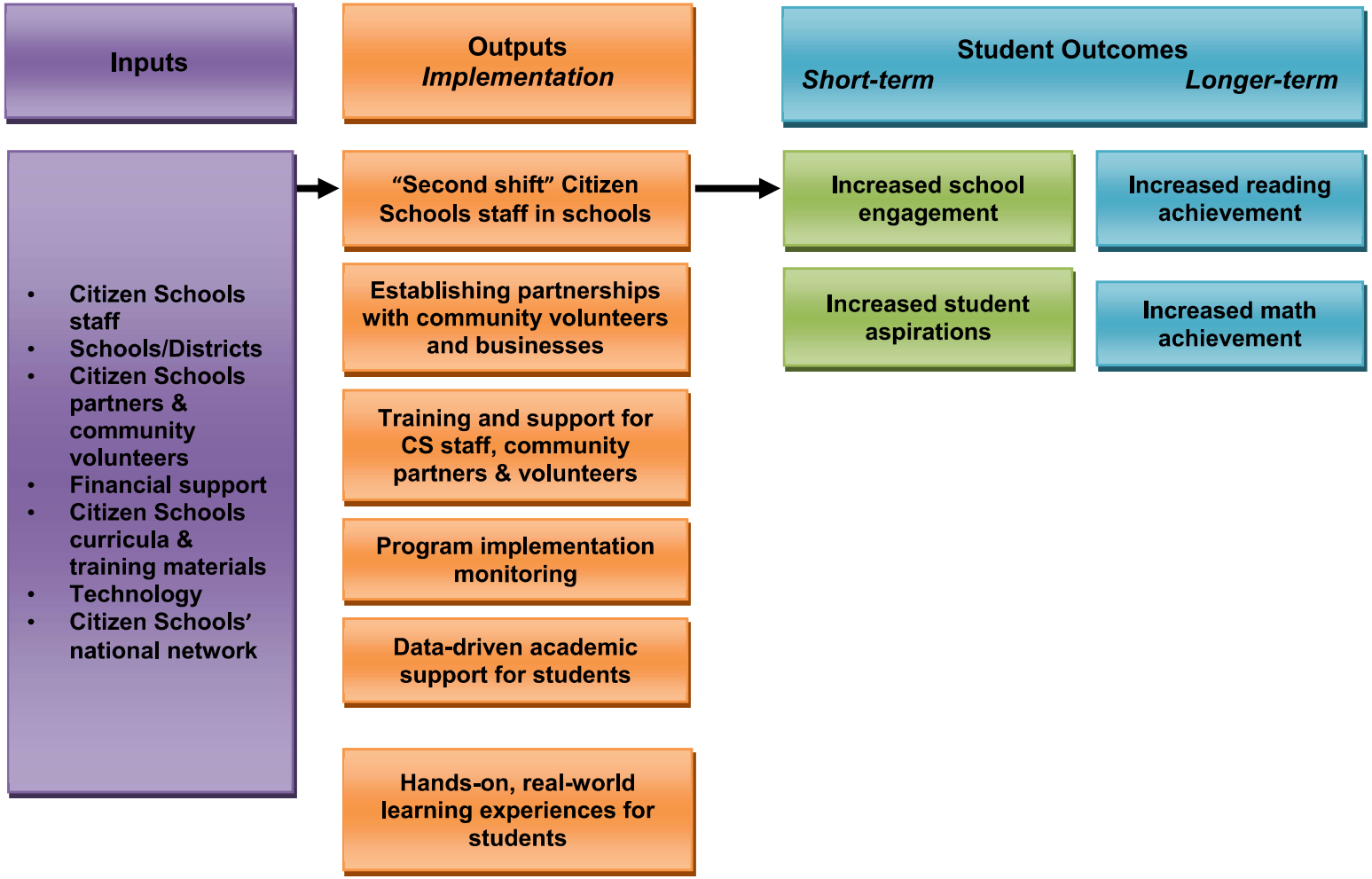
The Citizen Schools Expanded Learning Time (Citizen Schools ELT) Evaluation has a dual purpose: 1) to collect information about how Citizen Schools ELT programming has been implemented across sites, and 2) to determine whether Citizen Schools ELT participation affects key outcomes, including student engagement, aspirations, and achievement, for middle school students. Taken together, both purposes help understand the observed effects of expanding the learning day. The implementation data contextualizes variation in fidelity to the Citizen Schools ELT model, and may help explain findings about the impact of participation. This chapter describes the study's logic model, design, research questions, study sample, data collection measures and administration, and the data analysis.

2.1 Logic Model, Study Design, and Research Questions

2.1.1 Logic Model

The Citizen Schools ELT logic model illustrates the inputs, outputs, and expected outcomes, moving from left to right in Exhibit 2.1. The inputs represent the elements that together comprise the Citizen Schools ELT model, including the school/district, the Citizen Schools staff, partners and community volunteers, materials (e.g., curricula), supports, and technology. The outputs (activities/personnel encompassed in the model) include a second shift of ELT staff, establishing partnerships with community members and organizations, training, supporting, and monitoring Citizen Schools staff, and hands-on experiences for students. If the model is implemented with fidelity, the theorized short-term outcomes are an increase in student engagement in school and an increase in student aspirations. The theorized longer-term outcomes are increases in student English/Language Arts (ELA) and math achievement as captured by state standardized test scores.

Exhibit 2.1: Citizen Schools ELT Logic Model



2.1.2 Study Design

The National Evaluation of Citizen Schools' ELT Model is a multi-year study that evaluates both how the model is operationalized (outputs) and its effects on middle school students across the country (short and longer-term outcomes). The implementation component of the study (outputs on the logic model), incorporates several modes of data (surveys, site visits, and interviews) with multiple respondents within each school to obtain stakeholder perspectives; this component is designed to assess how schools integrate Citizen Schools ELT into their school days, organizationally and instructionally. Over the five years of the study, Citizen Schools expanded; as planned, it brought on new cohorts of schools in successive school years while continuing to serve schools from previous cohorts. For the purposes of this study, schools that began implementing Citizen Schools ELT in 2010–11 are called Cohort 1 schools, those beginning in 2011–12 are called Cohort 2 schools, and Cohorts 3 and 4 began in 2012–13 and 2013–14, respectively (see Section 2.2 for additional detail on the study sample). Exhibit 2.2 illustrates the data collection schedule over the study period.

Exhibit 2.2: Data Collection Schedule

	Study Year School Year				
Data Collection Activity	**1 2010–11**	**2 2011–12**	**3 2012–13**	**4 2013–14**	***5 2014–15***
Site visit (CS ELT only)	✓	✓	✓		
Principal Telephone Interview (CS ELT)		✓	✓	✓	✓
Principal Telephone Interview (MC)		✓	✓	✓	
Campus Director Telephone Interview		✓	✓	✓	✓
Survey					
*Student (CS ELT & MC)		✓	✓		
**Teacher (CS ELT & MC)		✓	✓	✓	
CS staff (CS ELT only)		✓	✓	✓	
Extant (test score) data	✓	✓	✓	✓	

*Student surveys were not collected in years 4 and 5.
**Teacher surveys were administered in Citizen Schools ELT schools (not comparison schools) in year 4; no surveys were administered in year 5.
In 2014–15, interviews were conducted with a subset of Citizen Schools campus administrators and campus directors, focusing specifically on schools designated as high implementers in the 2013–14 school year.

The impact component of the study is designed to assess short and longer-term student outcomes. To estimate the effect of Citizen Schools ELT on those outcomes, the study relies on well-established methods to control for as many reasonable alternative hypotheses as possible, the most common of which are: 1) changes in general education policy that might cause changes in outcomes across all schools; and 2) pre-existing, persistent characteristics of schools that affect outcomes. Analyses of short-term and longer-term outcomes address the first category of alternative hypotheses using a group of comparison schools—matched on key pre-intervention observable characteristics. Matched comparison schools help control for changes in federal, state, and district education direction that could affect outcomes across some or all schools. Although pre-ELT measures were not available for student non-academic outcomes, the use of matched comparison schools matched on pre-intervention characteristics help account for baseline differences in these outcomes.

Because longitudinal data are available to examine academic outcomes, the study builds upon the matched comparison design described above by using a comparative short interrupted time series (C-SITS) which models the implementation of Citizen Schools ELT as an "interruption" in what would otherwise be assumed to be somewhat stable levels of a particular outcome. Secular changes in policy are addressed through the inclusion of comparison schools, as well as by estimating year fixed effects across both Citizen Schools ELT and matched comparison schools. Assuming that any policies or programs implemented in all of the schools would affect matched comparison and Citizen Schools ELT schools within a particular grade similarly, including year fixed effects would control for their effects, allowing the estimating of the effect of Citizen Schools ELT over and above the effects of the other factors.

In addition, models use school fixed effects to address the second category of alternative hypotheses mentioned above: pre-existing, *persistent* characteristics of schools that affect outcomes. For example, within a given school, students' parental motivation presumably affects pre- as well as post-ELT student achievement. Including school fixed effects in the models allows the estimation of the effect of Citizen Schools ELT over and above what might be expected given the effects of stable school characteristics.

The C-SITS design represents one of the strongest non-experimental designs available to obtain credible estimates of the effect(s) of Citizen Schools ELT on student outcomes (Somers et al., 2013; St. Clair et al., 2014). Indeed, by using both matched comparison schools and controlling for baseline measures, the study effectively controls for baseline outcome measures twice, yielding "doubly-robust" results designed to decrease bias in estimating impacts (Bang & Robins, 2005). Thus, the design yields impact estimates that provide both robust and credible evidence of the impacts of Citizen Schools ELT.

2.1.3 Research Questions

The evaluation's research questions, presented below, follow the left-to-right pathway depicted by the logic model.

Outputs: Implementation

The primary research question for implementation focuses on how effectively the Citizen Schools ELT outputs are being carried out in schools nationally. Additionally, the question seeks to answer how well Citizen Schools ELT is being integrated into partner schools' cultures. The implementation research question is stated below.

How successfully are schools implementing and integrating expanded learning opportunities into their school days?

Student Outcomes

Research questions for the short and longer-term outcomes (non-academic and academic) portion of the study were delineated as confirmatory and exploratory prior to analysis. Following Schochet (2008), the study's confirmatory analyses test central research questions and hypotheses of the study; they are driven by the theory of change and are limited in number. Exploratory analyses help to generate hypotheses, but may be limited by less statistical power or a lack of pre-ELT measures for outcomes.

Short-Term Non-Academic Outcomes

All analyses conducted with student non-academic outcomes are considered **exploratory** because pre-ELT versions of these measures are not available. Consequently, effects can only be assessed based using post-ELT data and matched comparison schools, which is a strong, but less rigorous design than the C-SITS model, as it limits the study's capacity to control for pre-ELT factors. Research questions for teacher and student analyses are listed below.

Student-Level

*What is the relationship between Citizen Schools ELT and **student engagement** and **aspirations**?*

Teacher-Level

*What is the relationship between Citizen Schools ELT and teacher perceptions of **student engagement** and **aspirations**?*

Longer-Term Academic Outcomes

The impact of Citizen Schools ELT on academic outcomes is assessed using both **confirmatory** and **exploratory** analyses. Cohort 1 schools have data for up to four implementation years, Cohort 2 schools have data for three years, Cohort 3 schools have data for up to two years, and Cohort 4 schools have data for one year (Cohort 5 schools are not included in the analyses as their data were not available in time for processing). The **confirmatory** research questions test the overall impact of Citizen Schools' ELT model on students' ELA and math state achievement scores (the study team and Citizen Schools determined that the confirmatory tests should be limited to the overall effect of the Citizen Schools ELT model rather than by subgroups) and the **exploratory** analyses test the impact of the Citizen Schools' ELT model for subgroups of schools and/or students (e.g., by grade, implementation year, dosage, and implementation levels):

Confirmatory:

What is the effect of the Citizen Schools ELT initiative for all middle grade students who participated in the initiative in the 2010–11 - 2013–14 school years on their scores from the state ELA and math tests?

Exploratory[9]:

- *What is the effect of the Citizen Schools ELT initiative for all middle grade students who participated in the initiative in the 2010–11 - 2013–14 school years on their scores from the state ELA and math tests, **by implementation year**?*
- *What is the effect of the Citizen Schools ELT initiative for all middle grade students who participated in the initiative in the 2010–11 - 2013–14 school years on their scores from the state ELA and math tests, estimated separately **by grade (6, 7, and 8)**?*

[9] The study was not able to conduct exploratory analyses by gender or race/ethnicity as the required data were not available for three of seven states. Additionally, because the large majority of participating schools were racially/ethnically homogenous, dividing the sample into subgroups would have been problematic.

- *What is the effect of the Citizen Schools ELT initiative for all middle grade students who participated in the initiative in the 2010–11 - 2013–14 school years on their scores from the state ELA and math tests,* ***by dosage****?*
- *What is the effect of the Citizen Schools ELT initiative for all middle grade students who participated in the initiative in the 2010–11 - 2013–14 school years on their scores from the state ELA and math tests,* ***by implementation level****?*

The research questions that divide the sample by implementation year, grade, dosage, and implementation level are considered exploratory because statistical power for these analyses is limited. While all schools in the sample have been implementing Citizen Schools ELT for at least one year, only those schools continuing in Cohorts 1, 2, and 3 have at least two years of data available, and although three Cohort 1 schools continued in 2013–14 (see Exhibit 2.3), the analytic sample (for that year) includes only the two schools for which data are available (data were not released for the third school for the 2013–14 school year). Similarly, grade-specific and dosage effect estimates on students' test scores are considered exploratory because they hinge on grade distributions within schools and the Citizen Schools ELT initiative did not serve all middle grades in each campus (see Exhibit 2.3 below for specific grades served in each school). Analyses by implementation level, like the other exploratory analyses, also divide the sample into subgroups that decrease statistical power.

2.2 Study Sample

Citizen Schools uses an iterative, multi-step selection process for Citizen Schools ELT campuses, utilizing "goodness of fit" measures and intensive, ongoing communication with district and school administrative staff. In Fall 2010, an initial cohort of eight schools began to implement Citizen Schools ELT, followed by 10 schools in the Fall of 2011, nine additional schools in the Fall of 2012, and eight schools in the Fall of 2013. Schools joined the network serving one, two, or three grade levels, and nine (of 35) schools expanded to serve additional grades over the course of the study.

Despite considerable attention to site selection, a number of schools within each cohort exited the network. In Cohort 1, for example, two schools exited after their second year, a third after its third year, and two after their fourth year. Over the study's duration, 22 schools exited the network. Exhibit 2.3 illustrates the staggered entry of schools, by cohort, and indicates which grades were served in a given school year for each school. The exhibit also shows when schools exited the network (depicted by shaded cells).

Exhibit 2.3. Grades Served by Citizen Schools ELT Schools

Cohort/ELT School	2010–11	2011–12	2012–13	2013–14	2014–15
Cohort 1					
School B	6, 7	6, 7	6, 7	5, 6, 7, 8	6, 7, 8
School H	6, 7	6, 7	6, 7, 8	6, 7, 8	6, 7, 8
School L	6	6, 7	6, 7	6, 7	6, 7
School C	6	6	6		
School G	6	6	6, 7		
School K	6, 7, 8	6, 7, 8[a]			
School P	6, 7, 8				
School R	6				
Cohort 2					
School E		6	6, 7	6, 7, 8	6, 7, 8
School J		6	6, 7	6, 7	6, 7
School N		6	6	6	6, 7, 8
School A		6, 7, 8	6, 7, 8	6, 7, 8	
School D		6	6	6	
School I		6	6	6	
School F		7	7		
School M		6, 7, 8	6, 7, 8		
School Q		6, 7, 8			
School T		6, 7			
Cohort 3					
School X			6	6	6
School Y			6	6, 7	6, 7, 8
School O			6	6	
School S			6, 7, 8	6, 7, 8	
School U			6	5, 6, 7	
School Z			6	6, 7	
School V			6, 7, 8		
School AA			6, 7, 8		
School W			6, 7		
Cohort 4					
School AG				6	6
School AB				6, 7, 8	5, 6, 7, 8
School AC				5, 6, 7, 8	5, 6, 7, 8
School AJ				6	5, 6
School AI				6	6
School AD				6	
School AE				6, 7, 8	
School AH				6	

EXHIBIT READS: School B in Cohort 1 served 6th and 7th graders in each school year from 2010 to 2015.

Notes:

School IDs (letters) have been randomly assigned in order to present school-level data anonymously. Schools are organized by cohort, then by the year they exited from the study (if applicable), as indicated by school names shaded in gray. After a school has exited a study, subsequent school years are shaded in red.

[a] This school was closed by the district after the 2011-2012 school year.

Source: Citizen Schools website (www.citizenschools.org), interviews and other communications with Citizen Schools sites.

Exhibit 2.4 illustrates the number and home states of Citizen Schools ELT campuses between 2010 and 2015.

Exhibit 2.4 Citizen Schools ELT Campuses by State

2.2.1 Citizen Schools ELT School Sample

It is important to note that the full sample of Citizen Schools ELT campuses did not necessarily participate in all forms of the study's data collection and analysis activities. As summarized in Exhibit 2.3 above, the study's data collection activities were not identical in each of the study years. Only schools that had demonstrably committed to the Citizen Schools ELT model, and were continuing to meet minimal implementation thresholds were included in survey data collection activities beginning in schools' second year of implementation. Additionally, the study reduced data collection activities to minimize burden on schools in the 2012–13 and 2013–14 school years. Exhibit 2.5 below illustrates the number of schools included in each type of data analysis. The study generally obtained student achievement data from state education databases in the fall following spring test administration, and schools without the requisite data were (of necessity) excluded from some analyses.[10] Consequently, the number of schools in a given analytic sample varies depending on the specific research question (i.e., each research question has an associated analytic sample). At a broad level, the study team conducted interviews in each of the study's five years, fielded student surveys in two years, teacher and Citizen Schools staff surveys in three years (student and teacher surveys were fielded in matched comparison schools for two years), and obtained student achievement data for four school years.

[10] Eight of the 35 schools in the study could not be included in the achievement analytic sample due to various data issues: one school was so unique that no viable matched comparison schools could be found; four schools had no baseline data for the period prior to implementing Citizen Schools ELT; two schools were in a state that did not release achievement data during their first year of implementation; and one school had implemented a variant of ELT during its baseline period.

Exhibit 2.5. Number of Citizen Schools Campuses Included in Each Type of Data Analysis

Analyses Based on…	Number of Schools Included				
	2010–11	2011–12	2012–13	2013–14	2014–15
Interviews	8	15	22	22	9
Student surveys	n/a	5	12	n/a	n/a
Teacher surveys	n/a	5	13	15	n/a
Citizen Schools staff surveys	n/a	6	13	15	n/a
Student Achievement data	8	17	22	27	n/a

2.2.2 Matched Comparison Schools

By including a set of carefully selected and quite similar schools that are not implementing Citizen Schools ELT, the study can better support attribution of observed changes to Citizen Schools, since external factors would presumably affect both Citizen Schools ELT and matched comparison schools in roughly similar ways. By holding such secular effects constant, one can then more confidently attribute observed changes to the Citizen Schools ELT program.

Each Citizen Schools ELT school was matched to multiple comparison schools with similar demographic and achievement profiles within the same district[11] by using a combination of calipers[12] and Mahalanobis distance matching methods (see Appendix A.2[13] for a more detailed description of the matching process).[14] Thus, matches were both focal and local, two features that research on rigorous quasi-experimental designs has found to minimize bias (Bifulco, 2012; Cook et al., 2008; Fortsonet al., 2015; Steiner et al., 2010). Further, school and district staff were able to provide specific feedback about the suitability of the prospective matched comparisons regarding important contextual factors such as school status (e.g., slated for closure or restructuring) or impending

[11] In districts with more than one Citizen Schools ELT school, matching was conducted with replacement, i.e., a comparison school could be matched with more than one ELT school. Enough within-district matches for one Citizen Schools ELT school could not be found due to the size of the district, so one of the matched comparison schools was drawn from a demographically similar district.

[12] Calipers (ranges of appropriate values) for student enrollment and scores on subject tests narrowed down the number of appropriate match comparison schools. See also Rosenbaum and Rubin (1985) for a discussion of how calipers can be used in conjunction with Mahalanobis matching techniques.

[13] Note: Appendices A-C are in Appendix Volume I.

[14] The Mahalanobis distance matching method ranks the overall similarity of the Citizen Schools ELT school to different non-ELT (potential comparison) schools based on various demographic and testing characteristics. The smaller the differences between the potential comparison and the Citizen Schools ELT school across matching characteristics, the higher the potential comparison school is ranked, and the more likely it is to be chosen as a comparison school. The four schools with the highest ranks are chosen as primary matches, and the next two schools are chosen as alternate matches. One advantage of the Mahalanobis distance metric over other metrics, such as Euclidian distance, is that it takes into account that variables have different scales, and that some variables are related to one another (and thus does not overstate variables with large scales or variables that measure similar characteristics).

implementation of other, similar expanded learning time initiatives. Following this, the top four ranked schools became the Citizen School ELT school's set of matches.

The study team assessed the comparability of the Citizen Schools ELT and matched comparison schools on baseline test scores to confirm that the matched comparison schools were, in fact, similar enough to the Citizen Schools ELT schools to represent a plausible counterfactual condition—that is, to be sufficiently comparable that one could comfortably believe the comparisons are credible. Specifically, the study compared the average standardized effect sizes of test scores during the baseline period for the Citizen Schools ELT and comparison schools. Any baseline differences larger than .25 standard deviations would be considered meaningful enough that they might affect the analysis (Ho et al., 2007). Exhibit 2.6 presents the average baseline differences between the Citizen Schools ELT and matched comparison schools in the final sample; it indicates that after restricting the sample, baseline differences between Citizen Schools ELT and matched comparison schools were about a tenth of a standard deviation. This approach is consistent with guidelines outlined by the U.S. Department of Education's What Works Clearinghouse (U.S. Department of Education, 2011).

Exhibit 2.6. Baseline Equivalence of Citizen Schools ELT and Matched Comparison Schools

	Difference (in effect size units)	
	ELA	Math
Overall	0.10	0.10

EXHIBIT READS: The average baseline difference (in effect size units) between Citizen Schools ELT and matched comparison schools' test scores was .10 for ELA and math.

Source: Publicly available school-level test scores data downloaded from the websites of schools in study states.

Sample: Twenty-seven Citizen Schools ELT schools and 95 matched comparison schools in the ELA sample and 87 matched comparison schools in the math sample.

2.3 Data Collection Measures and Administration

Data collection for Citizen Schools ELT Study began in the 2010-2011 school year, and followed a staggered schedule (refer to Exhibit 2.2 above). The text below describes each data collection element (data collection instruments can be found in Appendix Volume II).

2.3.1 Interviews and Site Visits

Citizen Schools ELT Schools. The research team conducted site visits to Citizen Schools ELT campuses in their first year of ELT implementation, and conducted either site visits or telephone interviews in subsequent years. The research team interviewed Citizen Schools regional and program directors, school principals, Citizen Schools Campus Directors, as well as other school or Citizen Schools personnel with program implementation leadership roles. Stakeholder interview protocols asked about sites' background and motivation to implement Citizen Schools ELT, goals and expectations for Citizen Schools ELT, decision-making, communication, and responses to Citizen Schools ELT, Citizen Schools ELT operations, and historical/contextual information (see Volume II Appendix for interview protocols). Follow up interviews (whether on site or by telephone) with principals and campus directors addressed the topics mentioned above, as well as program growth and evolution, and any changes made over each additional year of implementation. Interviews in Years 4 and 5 were conducted by telephone only.

Matched Comparison Schools. The research team conducted telephone interviews with the principals of matched comparison (MC) schools in each of the first two years of the study to learn about schools' backgrounds and contexts, including student demographics, neighborhood characteristics, and about any extra-curricular programming in which students participate. The interviews provided contextual information for the findings from the impact study.

2.3.2 Surveys

Citizen Schools staff, teacher, and student surveys were administered in Spring 2012 and 2013 (see the Volume II Appendix for copies of the surveys). Because of both cost and logistical difficulties associated with recruiting non-ELT schools to participate in all study data collection activities, the study team administered surveys to two of the four selected matched comparison schools. The study team targeted the top two best matches among the four matched comparison schools for data collection, and in cases where matched comparison schools refused to participate, surveys were administered to the non-ELT schools with the next highest ranks.[15] In Spring 2014, the study team, in consultation with Citizen Schools, elected not to administer surveys to matched comparison schools' students or teachers, as findings had remained stable over prior administrations, and survey data collection required substantial resources. Instead, surveys were administered solely to Citizen Schools ELT teachers and staff, both to ensure that the study team could continue to monitor ongoing implementation generally and to ensure that data were available for the study's index of implementation (described in greater detail below and in Chapter 3). Exhibit 2.7 presents the staff, teacher, and student response rates in 2012, 2013, and 2014. For more information about survey administration, see Appendix A.1.

[15] In one case, surveys were only administered to one MC school, rather than two, because of difficulty recruiting other matched schools.

Exhibit 2.7: Survey Response Rates, Spring 2012, 2013 and 2014

	2012		2013		2014
	Citizen Schools ELT	MC	Citizen Schools ELT	MC	Citizen Schools ELT
Citizen Schools Staff					
N	68	--	187	--	195
Response Rate	89%	--	89%	--	88%
Teachers					
N	57	68	239	196	158
Response Rate	84%	82%	77%	83%	88%
Students					
N	145	154	390	614	--
Response Rate	71%	67%	65%	61%	--

EXHIBIT READS: Sixty-eight Citizen Schools staff completed a survey in 2012. The average response rate across schools was 89 percent.

Notes:

[1]Recall that matched comparison teachers and students were not administered surveys in 2014.

Source: Abt Associates' Surveys of Citizen Schools staff, Spring 2012, 2013, and 2014; Abt Associates' Surveys of Citizen Schools ELT and Matched Comparison School Teachers, Spring 2012, 2013 and 2014; and Abt Associates' Surveys of Citizen Schools ELT and Matched Comparison School Students, Spring 2012 and 2013.

2.3.3 Extant Data

The study team obtained publicly available ELA and math school-level achievement data[16] for both the Citizen Schools ELT and the matched comparison schools for up to five years prior to the start of the ELT initiative, and up to four post-ELT years. Data from each state were processed individually before they were combined for analyses, so that key variables could be standardized across states. Given that test scores were obtained from multiple states using different achievement tests, the study team converted them into a standardized z-score metric, which allowed for the pooling of data across states.[17] Additional school-level characteristics such as the percent of students in schools that were male or received free or reduced price lunch were coded identically across states.

2.4 Implementation Data Analysis

Evaluation staff recorded notes during site visit interviews with Citizen Schools and campus staff. All interview data were synthesized, cleaned and standardized before being uploaded into NVivo 10, a qualitative data analysis software package. Coding categories were constructed in NVivo at a macro

[16] Two states provide publically available, de-identified student-level datasets. These student-level data were aggregated to the school level so that they'd be comparable to the data from other states.

[17] For each test subject, the z-score is calculated by subtracting the average scaled score for all students in that grade and year in the state from the school-level mean score, and dividing by the standard deviation in that grade and year. For example, if students' average ELA score in a given school is identical to the statewide average for that year, the ELA z-score for that school would be zero. Positive and negative scores would indicate that the school's students scored above and below the state average, respectively.

level and refined in preparation for the cross-site analysis of Citizen Schools ELT implementation. Once broader thematic categories had been determined (e.g., planning, staffing, integration, adaptation, sustainability), more detailed subcategories were created based on emerging patterns within category. The qualitative data were synthesized by macro and micro categories in cross-site implementation findings presented in Chapter 3.

The study team developed an index keyed to core principles of effective ELT implementation, as articulated by the Citizen Schools ELT model. The purpose of the index is two-fold: one, to integrate information from multiple data sources into one summary measure that describes variation in schools' implementation efforts, both for individual schools and for the network as a whole; and two, to create a measure that could potentially be used to explore relationships between level of implementation and student achievement. Additional information and results from the implementation index are presented in Chapter 3.

2.5 Non-Academic and Academic Outcomes Data Analysis

As described above, addressing the key research questions for the outcomes study involves estimating the difference between outcomes for students in Citizen Schools ELT and matched comparison schools such that differences between the two groups can be attributed to Citizen Schools ELT. The specific approaches used to model the effects of Citizen Schools ELT on various outcomes are described below.

2.5.1 Modeling the Effects of Citizen Schools ELT on Short-Term Non-Academic Outcomes

To address questions about the effect of Citizen Schools ELT on short-term, non-academic outcomes, the study uses three-level hierarchical linear models wherein students are nested in classrooms, and students and teachers are clustered in schools (see Appendix A.3 for formal model specifications). Prior to analysis, the study team identified the items from each survey that corresponded with the Citizen Schools ELT outcomes of interest (i.e., student engagement and aspirations). The team used a technique called factor analysis to identify items that could reasonably be grouped into substantively linked constructs. All survey items that addressed a non-academic outcome of interest were averaged together in a single construct.[18] See Appendix C for comprehensive information about all individual survey items as well as which survey items were combined into the constructs presented in Chapter 4.

2.5.2 Modeling the Impacts of Citizen Schools ELT on Longer-Term Academic Outcomes

The C-SITS models that test for impacts of Citizen Schools ELT on academic outcomes represent the most rigorous models used in the study (see Appendix A.4 for formal model specifications). As mentioned in section 2.1.2, by including school and year fixed effects, the statistical models can account for secular and time-varying factors, and are therefore more robust model than those testing non-academic outcomes.[19] Further, the models appropriately account for clustering of grades within

---

[18] In order for a construct to be calculated, respondents had to answer at least two thirds of the survey items included in the construct. For example, if 10 items comprised a particular construct, respondents had to answer at least seven of them for the items to be averaged together to create the construct.

[19] The inclusion of year fixed effects precludes the ability to establish a particular trend line (or curve); rather, any systematic trends will be captured by including these variables in the models. This approach does assume, however, that the pre-ELT trends are the same in Citizen Schools ELT and matched comparison

schools and schools over time, using the cluster-robust variance estimator (also known as the "sandwich" standard errors; White, 1984 and Liang and Zeger, 1986). Models are flexible and allow for the entry of new cohorts over time, as well as the exit of any schools from the Citizen Schools network. Although data from schools that exit the network are included in the analytic sample during the years they participated in the program, no additional data were collected once they exited.[20]

The impact of Citizen Schools ELT on school-level test scores is modeled using an indicator for Citizen Schools ELT school,[21] and provides an estimate of the impacts of Citizen Schools ELT across cohorts and implementation years. The models can detect differences over and above any changes in the outcomes that might be expected given outcomes prior to ELT and in schools not implementing Citizen Schools ELT. The models include school-level covariates to control for their impacts on outcomes and increase the precision of the estimates, in addition to school and year fixed effects variables. These covariates included race/ethnicity, limited English proficiency status, low income status, and gender.[22]

Because the study was designed to have sufficient statistical power to detect the impact of Citizen Schools ELT as a whole across its network, and not to detect impacts in certain subgroups,[23] **confirmatory** analyses pool data across implementation years. Confirmatory analyses alone, however, may mask potentially meaningful differences in outcomes by, for example, grade or implementation year (e.g., it may be the case that the program has differential effects for students in 6th vs. 7th grade), so the study also conducted **exploratory** analyses of Citizen Schools ELT impacts

---

schools. Sensitivity analyses conducted for the Year 1 analyses did not find systematic trends across all schools in the sample. Given that the inclusion of trend lines into the model is predicated on an assumption that the trends both exist and will continue to exist after the intervention (strong assumptions), these were not included in the achievement models. Another consideration in making this decision is that some states only had three years of baseline data, too few to draw conclusions about future trends.

[20] See section 4.2.4 in Chapter 4 for additional information about the implications of this analytic decision. Sensitivity analyses suggest that removing these schools from the analytic sample once they exited the network does not bias the results.

[21] For the exploratory analyses that run these models by implementation year, grade, dosage, and implementation levels, treatment indicators are created for each subgroup of interest (e.g., an indicator for 6th graders in Citizen Schools ELT schools or schools in their first year of Citizen Schools ELT implementation).

[22] Because these were obtained at the school level, there was virtually no missing data on these covariates. In the rare instance that a school-level covariate was missing in a given year, it was imputed using the average of that covariate across the previous three years. This was deemed preferable to imputing the data by replacing missings with data from the Common Core of Data, as it did not introduce a new data source into the state's data. There is one exception: a state in which there was so much suppressed data for certain covariates that data for those covariates were obtained solely from the Common Core of Data.

[23] For example, as of 2013–14 (the most recent school year for which student achievement data are available), four cohorts of schools had implemented Citizen Schools ELT for one or more years, yielding a reasonable sample on which to base estimates of impacts of the program after at least one year. However, the sample size decreases for each additional year of implementation due to the staggered nature of cohorts entering the network: 15 schools had implemented Citizen Schools ELT for two years, 10 schools for three years, and three schools for four years (again, as of the 2013–14 school year).

on academic outcomes by implementation year, grade, years of exposure to the program ("dosage"), and implementation levels. Note that because the confirmatory outcomes are in different domains (ELA and math), the study team did not need to adjust for multiple comparisons. Similarly, no adjustments were made to the exploratory analyses given that these are preliminary. Finally, as detailed in Chapter 4, the study team conducted a variety of specification and sensitivity tests to assess the robustness of the student achievement outcomes. These additional tests indicate that the results are stable and robust.

2.5.3 Interpreting the Results

The goal of these analyses is to estimate the impacts (e.g., test scores for students or perceptions from teachers) in Citizen Schools ELT schools and to test whether these outcomes differ from what would be expected in the absence of Citizen Schools ELT, often referred to as the counterfactual. Given that the counterfactual cannot be observed, the study uses a statistical technique to estimate the counterfactual using data from students and teachers in schools that did not implement the program (i.e., the matched comparison schools). Once the models have been estimated, the responses are no longer referred to as averages for matched comparison schools, since they do not accurately reflect the real outcomes in matched comparison schools. Rather, they are regression-adjusted estimates of *what would have happened* in Citizen Schools ELT schools had the initiative not been implemented. Therefore, all findings are described in terms of the actual outcomes measured in Citizen Schools ELT schools relative to the estimated outcomes in the counterfactual. If differences are observed, there is evidence that those differences can be attributed to the effects of the implementation of Citizen Schools ELT.

2.6 Study Limitations

2.6.1 Study Sample

The study sample experienced considerable fluidity, as nearly two-thirds of the participating schools exited the network during the study period. This raises a concern about the representativeness of the study sample as well about the stability of the intervention for the study schools. Unlike experimental studies, attrition in a quasi-experimental framework does not necessarily threaten internal validity; on the other hand, it may influence external validity simply because the final sample does not accurately represent the full set of schools originally in the evaluation. The study can document implementation progress and challenges for all 35 schools for at least their first year, and it can obtain unbiased achievement results for all schools with requisite data after one year of implementation. The information from those two study components (based on the first year alone) can yield useful information for the program. Yet the program is not currently designed as a one-year or even a two-year intervention, but as a program that could potentially be integrated into a school for an indefinite period. The cumulative results (both implementation and outcomes) are based on a subset of schools and may therefore be less informative; the results may not even apply to the full set of 35 schools. While the study design was flexible enough to accommodate modest attrition, the fluidity of the sample was substantial. Nonetheless, the study's purposeful triangulation of data across multiple sources and years (for all analyses), and use of robustness and sensitivity checks (for outcome analyses) bolster confidence in the study's internal validity.

2.6.2 Implementation

The primary limitations of the implementation data include the variable number of interview respondents available across sites and consistency of questions asked of each respondent type, as the evaluation team was not able to systematically interview staff members with comparable positions at each site. When and where possible, the study team interviewed multiple stakeholders with different positions and perspectives, which then allowed triangulation across respondents, and also generated a rich base of information. However, in some schools, evaluation staff interviewed fewer stakeholders, thus limiting the amount of information obtained. Similarly, because the interviews were semi-structured, not all interviewees were consistently asked or responded to all questions.

2.6.3 Outcomes

Although analyses modeling the relationship between Citizen Schools ELT and student outcomes include numerous controls to increase their rigor, there are several possible limitations to the study's analytic approaches that should be noted prior to presenting findings. For example, although models that assess survey outcomes include the matched comparison group, they are cross-sectional, and thus do not include the school and year fixed effects employed by the C-SITS design in the longitudinal analyses. As a result, these models control for pre-existing differences in characteristics used to match Citizen Schools ELT schools to the comparison group, but cannot control for any remaining differences that are not accounted for in the matching process. In addition, the representativeness of these results is subject to survey response rates, which, while high, may arguably not generalize to all teachers or students in a particular school.

In addition, while analyses that use multiple years of data to assess the impacts of Citizen Schools ELT on academic outcomes are more rigorous than those used to model survey outcomes, they may still omit unobservable characteristics that affect outcomes. For example, if secular changes in a given academic year affect Citizen Schools ELT and matched comparison schools differently in the post-ELT period, then the year fixed effects included in the models will not be sufficient to control for these differences. That said, by using a combination of multiple sources of implementation and impact data, using appropriate statistical modeling, and conducting extensive sensitivity analyses, the study team has confidence in the results described in the report.

3. Implementation of Citizen Schools ELT

How schools implement the Citizen Schools Expanded Learning Time (ELT) program model is critical to its success. Over the past several years, the study team has examined program implementation as the network has expanded into additional schools and districts, and as the program model itself has undergone important refinements. Before describing implementation in detail, it may be helpful to step back to consider the study, and the network, as a whole. The study purposefully builds on the networks' gradual expansion to include additional schools, both to reflect the wider scope of Citizen Schools' reach to larger numbers of schools in different districts and regions, and to ensure that the study would have sufficient statistical power to be able to detect differences in student outcomes that may be attributable to the Citizen Schools ELT model. Exhibit 3.1 illustrates; the study began with eight program schools in Cohort 1, and added new schools in Cohorts 2, 3, and 4 in the following years. Over the study's duration, 35 schools joined the network. Each year, however, some schools also exited the network, due to changes in district/school leadership and priorities, or lack of fit between the ELT program and the campuses. By the end of 2014–15, 22 schools, representing two-thirds of the network, had exited. The discussion of program implementation that follows, therefore, is situated in the context of a fluid network of participating schools.[24]

Exhibit 3.1 Number of Participating Schools, By Cohort and School Year

Cohort	Number of Schools, by Cohort and School Year				
	2010–11	2011–12	2012–13	2013–14	2014–15
Cohort 1	8	6	5	3	3
Cohort 2		10	8	6	3
Cohort 3			9	6	2
Cohort 4				8	5
Total	**8**	**16**	**22**	**23**	**13**

This chapter summarizes findings about the major building blocks of implementation, including program launch, staffing, and staff integration, program adaptation, and sustainability, and describes commonly experienced accomplishments and challenges across the network as a whole. It draws from interviews and surveys of school and Citizen Schools staff as well as survey data collection from educators and students in selected comparison schools (see Chapter 2 for additional information about data collection). The chapter also describes a summary measure of implementation, the implementation index, which synthesizes information about participating schools in their second, third, and fourth years; note that the index incorporates information gathered during interviews with schools during the first year into the index scores for the schools' second year.

3.1 Planning and Roll-out for Expanded Learning Time

Successful launching of new Citizen Schools campuses sets the stage for much of the day-to-day program operations. Observations from four successive cohorts were quite consistent: how Citizen

[24] The fluidity of the study sample does not materially change the findings on program implementation, although that same fluidity may limit the applicability of findings to schools that remain in the sample.

Schools selected appropriate partner schools, planned and communicated about ELT with stakeholders, and anticipated logistical and programmatic challenges to the model demonstrably influenced the ELT implementation.

3.1.1 Key Findings

- Criteria for selecting partner schools for Citizen Schools ELT were similar across the network.
- Ensuring the adequacy of local (both school and district) resources mattered. The program launched more smoothly when school/district resources were aligned: schools had the financial and decision-making wherewithal to commit to ELT, and school and district leaders were knowledgeable and supportive of Citizen Schools ELT.
- Clear and timely communication to stakeholders during the planning process facilitated smooth program launch, and late or incomplete communication efforts translated into missed opportunities to engage local school community(ies) from the beginning.

3.1.2 Site Selection

Choosing which schools to implement Citizen Schools ELT has clear consequences, reflecting the importance of positive school/district partnerships and shared commitment to program goals. In some cases, schools made autonomous decisions, whereas in others, districts identified sometimes willing, sometimes reluctant school partners. Citizen Schools staff described explicit criteria for strong site selection, including: 1) demonstrated academic and socioeconomic need; 2) access to funding (federal, state, grant, or school budget) sufficient to sustain ELT programming; 3) principal capacity to prioritize collaborative relationships and support ELT as integral to the school culture; and 4) functionality, including a reasonably established and effective behavior and support system.

Yet choosing new partner schools remains challenging, as few schools fully met *all* the desired selection criteria. Volatility in turnaround school leadership, district support and commitment, and school autonomy continued to pose challenges for site selection. Particularly in those several instances where districts, not schools, made decisions about school participation, principals had insufficient time to cultivate support before decisions were made, and consequently had to develop support after-the-fact. Even as the Citizen Schools organization has recognized that explicit criteria and careful vetting are important for site selection, there are participation decisions over which it has no control.

3.1.3 Clear Communication about Program Launch

Having sufficient time to communicate with stakeholders before program launch meant that schools had time to conduct outreach, communication, and planning activities, and absent sufficient lead time, lack of communication contributed to chaotic start-up, lack of clear role definition, and misconceptions about Citizen Schools ELT. Establishing joint committees and/or conducting summer workshops involving administrators, teachers, and Citizen Schools staff, or other efforts to help foster positive working relationship among school staff and Citizen Schools staff and/or inform families about the purposes of adopting an expanded learning time model represented strategies some schools used to promote greater integration between first and second shifts; however, such efforts were more likely in schools where the decisions to launch had occurred well in advance. Campuses in which decisions had been made in the summer (or later) had much less time to communicate, negotiate, and plan effectively, and school staff, parents, and Citizen Schools respondents reported friction and

challenges related to misconceptions about Citizen Schools programming and Citizen Schools' staff capacity, roles, and responsibilities.

School administration and Citizen Schools staff used multiple approaches to inform families and communities about ELT, including community forums, constituent focus groups, visits to feeder grades and schools, and open city meetings. These activities were designed to disseminate information and solicit support for the partnership and the extended day programming, and such activities were perceived as more effective when the majority of families learned about and understood what would happen with Citizen Schools ELT soon enough to make their own decisions and plans. In particular, communication that informed families of the extended day programming (and its mandatory attendance requirements) well in advance of school lottery and other enrollment deadlines was critically important, as that allowed families to plan. With sufficient time, school administrators could (and did) organize afterschool care for younger siblings of students participating in Citizen Schools ELT, which helped alleviate the child care issues faced by ELT families. Absent such advance planning, schools were not able to assist families with childcare needs for younger siblings, and in a few instances, principals deliberately delayed Citizen Schools ELT programming for a month or even a full year to establish adequate communication with stakeholders and identify solutions to issues such as care for younger siblings. Program launch tended to be better received for schools in which principals had planned in advance to take care of parent concerns such as child care for younger siblings and transportation.

3.1.4 Common Challenges at Program Launch

The importance of advance planning was also reflected in how campuses managed the logistical and administrative challenges that can affect any new initiative. The most common obstacles included transportation, physical space/resources, participation in sports and extra-curricular activities, and determining how best to serve students with special needs. When campuses had enough time to prepare for program launch, they were more likely to be able to anticipate how to navigate challenges specific to their own student populations. For example, some campuses worked with Citizen Schools leadership to identify resources to support transportation, whether via finding funds to pay for additional buses, or by deploying Citizen Schools staff to accompany students home or wait until families arrived.

Determining whether and how to accommodate students' athletic and other extracurricular commitments meant that in some cases, students were simply not able to participate in Citizen Schools ELT programming during sports or other extra-curricular time, whereas in other cases, campuses were able to build sports and other activities into apprenticeships and choice time. Another common issue focused on how to address the needs of students with disabilities (SWDs) or English language learners (ELLs) during the afternoon shift, as most Citizen Schools staff have not been trained to provide necessary accommodations for such special populations. Schools used a variety of adaptations, ranging from providing higher staff-to-student ratios for SWD students, organizing training about how to work with SWDs for Citizen Schools staff, or shifting the content focus from math to English/Language Arts to accommodate large numbers of ELL students. However, in some schools, SWD or bilingual students had early dismissal and could opt out, especially when the CS program lacked adequate resources to support them after school.

Access to sufficient space was an issue in some schools, particularly those schools with multiple community partners, each of whom needed space to work with students. Space limitations could

constrain the capacity to provide lower student-staff ratios, particularly when coupled with programming that required greater numbers of classrooms. Access to schools' technology systems and administrative records was another challenge that could impede smooth program launch, disrupt programming, and require creative problem-solving well after the school year had begun.

3.2 Program Staffing

Citizen Schools programming occurs primarily through its staff, and consequently, recruitment, hiring, preparation and support of staff have clear ramifications for how the program operates and is perceived. Data from several consecutive school years indicate that participating campuses continue to experience some of the same challenges, including training and experience, hiring practices, staff turnover, and staff diversity.

3.2.1 Key Findings

- School and Citizen Schools staff alike consistently emphasized the importance of establishing and maintaining strong relationships with first shift educators, and emphasized the significance of these relationships for implementing large scale systemic changes.
- In earlier years, Citizen Schools staff expressed concern about the match between the content of national training and local school needs, although fewer campus staff reported this concern for later cohorts, which may reflect the addition of region-specific training.
- Citizens Schools Teaching Fellows' capacity to manage student behavior effectively and provide focused instruction was consistently reported as a challenge by Citizen School staff and campus administrators.
- Staff turnover remained a persistent issue, although not all schools experienced the same levels of staff transitions.

3.2.2 Teaching Fellows' and Campus Directors' Training and Experience

Citizen Schools Teaching Fellows are primarily recent college graduates hired as Citizen Schools AmeriCorps members. The Teaching Fellowship is intended to be a two-year commitment. For many Teaching Fellows, working in Citizen Schools campuses is their first post-college position. Campus directors (CDs) generally have some additional experience, although they, too, are often new to their roles and to the specific schools in which they work.

Citizen Schools has used a combination of national, regional, and campus-based training to prepare incoming staff. Earlier cohorts participated in national training, and more recently, Citizen Schools has moved toward the use of regional trainings. While national trainings were characterized as beneficial, particularly in terms of team building, they were perceived as insufficient for preparing fellows quickly and effectively to assume teaching—along with classroom and behavior management—responsibilities in middle school contexts. Teaching Fellows generally described campus-based trainings as more beneficial, because insights could immediately be applied to their specific school contexts.

> A lot of them experienced a disconnect between the training they received in Boston and the reality of the work on the ground. There were definitely differences in expectations for the workload. Once they got their hands dirty, some just couldn't handle it.
>
> –CS Campus Director

> More training is necessary to ensure CS staff have experience working with youth. Some CS staff just don't have the instincts to deal with problematic students. They require additional training in management, curriculum building, and instruction.
>
> –Principal

Some of the same concerns about national training were raised by regional directors—that not all topics necessarily applied across the region. In some cases, this led to redesigning in-service trainings to reflect a purposeful mix of regional and campus-based trainings on such topics as data training, differentiated instruction, behavior management techniques, classroom management, and lesson planning. Such flexibility and attention to local needs were reported to contribute to a more effective ELT program.

Teaching Fellows also benefited from on-going training and professional development designed to develop their classroom management skills provided throughout the school year. Campus directors and deputy campus directors consistently described offering diverse training, coaching, and support activities to fellows, ranging from one-on-one coaching, to conducting classroom observations and providing feedback, to participation in school-provided training and professional development, to participation in regional sessions attended by staff from multiple Citizen Schools campuses. Campuses at which Teaching Fellows participated in school-wide professional development also contributed to greater integration of first and second shift teachers.

> Unfortunately, the fellows don't necessarily have the content background in the identified area of need, nor the pedagogy experiences. This results in a pretty difficult situation—the fellows are expected to deliver content support, following a curriculum developed by CS, that is somewhat aligned with the state standards—but they don't really come equipped with the necessary tools to pull it off well.
>
> –CS Campus Director

Despite the shift toward more regional and/or local training for first-year Teaching Fellows, however, inadequate classroom and behavior management and content knowledge continued to be perceived by administrators and CDs as major challenges. At the same time, administrators consistently acknowledged the benefits of having Teaching Fellows in their schools, as they brought enthusiasm and excitement to the program, were invested in the program's success, and often became positive role models for the students. And some administrators observed that second-year Teaching Fellows did not face the same challenges, having benefited from a year of experience and on the job training that helped to develop their skills, confidence, and competence.

CDs generally had considerably more relevant experience, whether through Teach for America, charter school, or international teaching experience, which allowed them to provide professional development, coaching and mentoring to their Teaching Fellows. Yet in a few instances, school administrators described CDs with underdeveloped management skills or leadership experience as an area of concern, although they also described observing tangible improvement. Opportunities for CDs to meet together within a region provided them access to peers and resources from more experienced colleagues.

3.2.3 Hiring, Retention, and Turnover

The national hiring policy for Teaching Fellows is generally similar across the network; Citizen Schools regional staff typically recruit from college and university campuses, which means that CDs are not part of the process. A small proportion of CDs and administrators consistently lamented the absence of campus-specific involvement in hiring decisions, and noted that individuals already familiar with the local context are more likely to understand a given school's specific needs and likely "fit" with potential candidates. Hiring in and of itself was not perceived as a major issue.

Turnover, however, posed consistent challenges, in two ways: one, the two-year service period guarantees some turnover within a typical school's three-plus planned years of commitment to the ELT program model, and two, some Teaching Fellows exit earlier than planned. The first is a structural feature of the AmeriCorps program outside Citizen Schools' purview, yet Citizen Schools staff and administrators consistently articulated concerns about the built-in disruption and loss of hard-earned experience. The factors that contribute to unanticipated departures include Teaching Fellows' lack of teaching experience and training for working in often challenging middle school environments. Clearly, high turnover posed substantial challenges for staff and for students. When Citizen Schools ELT campuses were short-staffed, CDs (and deputy CDs) themselves often covered staffing shortfalls, which hindered their capacity to oversee Citizen Schools ELT programming efficiently, support the remaining Teaching Fellows, and maintain high esprit de corps. Staff changes also had unintended negative consequences for students who had developed relationships with Teaching Fellows and then were faced with different teaching and leadership styles when replacement staff were hired.

Turnover also affected campus leadership; many campuses experienced transitions among the Campus Directors (Exhibit 3.2). In some cases, CDs left to take other positions within the Citizen Schools network, and their departures created opportunities for Deputy CDs or qualified fellows to assume greater responsibility. Nonetheless, the combination of staff and campus leader turnover meant that nearly every school experienced some disruption to programming.

Exhibit 3.2 Number of Schools with Leadership Changes from Year to Year, by Cohort

Cohort	Number of Schools									
	2010-2011		2011-2012		2012-2013		2013-2014		2014-2015	
	P	CD	P	CD	P	CD	P	CD	P	CD
1	0	0	0	1	1	2	1	1	1	2
2			0	0	2	1	0	5	0	2
3					0	1	2	2	1	1
4							0	2	2	1

P=Principal, CD=Campus Director

3.2.4 Diversity

Differences between the racial and socioeconomic profiles of Citizen Schools staff and those of the students represented another staffing challenge. While noted explicitly only in a small proportion of schools in any given year, ensuring that fellows are prepared to work in campuses attended primarily by low-income students of color is important—network-wide. This includes training and preparation for communicating with students and families from diverse backgrounds, who may speak different languages, and whose cultural norms may well differ from those of the fellows. Examples of school-specific efforts to address this challenge occurred through purposeful hiring of Teaching Fellows and teaching assistants from the local community with pre-existing ties to the school, and reliance on Spanish-speaking fellows to help communicate with families of students attending the school.

3.3 Staff Integration

Citizen Schools enters into partnerships with schools that function best when the expanded day staff who lead the second shift are meaningfully integrated into the life of partner schools. Insights from several years of studying implementation suggest three primary mechanisms for integrating Citizen Schools staff into schools: 1) structured overlap and joint participation between the first and second shift staff; 2) alignment of content, pedagogy and behavior systems; and 3) establishing and maintaining relationships between stakeholder groups.

3.3.1 Key Findings

- Developing purposeful structured overlaps between first and second shifts generated benefits, goodwill, and greater integration.
- Shared access to some data about student performance was common practice in about half the schools, as was participation by Citizen Schools staff in grade-level or subject meetings with first shift teachers.
- Managing student behavior during the second shift remained a persistent challenge for the majority of schools.

3.3.2 Structured Staffing Overlaps

Schools relied on several different strategies to ensure continuity between first and second shift learning for students, including purposeful overlaps with both first and second shift educators present, involving Citizen Schools staff in first shift activities, joint participation in campus-specific planning and professional development activities, and shared teaching responsibilities for designated academic

blocks. For example, integration was generally characterized as stronger where Citizen Schools fellows co-taught or led small-group instruction, observed first-shift teachers, and shared playground or cafeteria duty, which allowed them to learn firsthand from first shift educators and helped strengthen connections.

A majority of schools scheduled overlaps of 15 to 90 minutes during the transition between the first and second shifts. This meant that students observed first and second shift faculty working side-by-side, and smoothed the transition of responsibility from one shift to the next. In some cases, the transition block allowed ELT programming to begin earlier, which allayed parents' concerns about children walking home after dark. During the overlap block, school teachers had time for "real differentiated instruction" and fellows were able to learn from the school teachers. Generally, teaching fellows were paired with school faculty to conduct small-group, pull-out instruction, or individual tutoring, and in some sites, the transition block was an advisory or instructional period, or used for homework support.

> Just having more adults in the school or asking CS staff to check in with kids who are having a problem is helpful to teachers.
>
> –CS Deputy Campus Director

> The integration between the first and the second shift is so strong [at one school]. CS staff are in classrooms in the morning. Students and parents there don't know who's who.
>
> –CS Managing Director

Having purposeful overlap did not necessarily eliminate all potential frictions; determining who was responsible for managing student behavior when both first and second shift educators were working with students posed an occasional challenge, yet that challenge seemed less daunting than the challenges faced in schools where the two shifts operated as separate and uncoordinated portions of the day. Examples of the consequences included lack of clarity about who was responsible for making sure that participating students went to ELT programming instead of leaving the building for the day, chaotic conditions in hallways during the transition; and diminished participation as students simply skipped out.

3.3.3 Alignment of Content, Pedagogy, and Behavior Management

Effective Citizen Schools ELT implementation requires tight alignment of content, pedagogy, and behavior expectations across both the first and second shifts. Common approaches across the network included joint (first and second shift) participation in planning and professional development meetings, ensuring access to shared data systems, and developing common expectations and consequences for student behavior, although campuses varied in the specific strategies used.

In about half of ELT campuses, Citizen Schools Teaching Fellows were included in faculty and/or staff meetings, planning, and professional development sessions about instruction, content, student needs, and behavior management. Participation in these meetings facilitated greater consistency in lesson delivery and behavior management between first and second shifts, and provided Citizen Schools staff insights into how their work complemented the school day lessons. Citizen Schools staff also participated in professional development sessions designed primarily for

> Put yourself in the shoes of a student. The nine- hour day should seem seamless. Math strategy should not be different. Behavior management should not be different. CS staff need to be seen as faculty. If they are not, it's strange
>
> –CS Regional Director

first shift educators, although not as consistently; scheduling constraints sometimes meant that professional development workshops occurred precisely during the second shift, or in the summer before Teaching Fellows typically are placed at schools. In a few instances, Teaching Fellows benefited from supportive first shift educators who spontaneously served as mentors or coaches.

Citizen Schools staff could access some form of student performance data (student grades, state test scores, report cards, and progress reports) in about half the schools. Not all first shift educators were comfortable sharing student performance data although information sharing become more commonplace as Citizen Schools and school staff became more familiar with one another. When Citizen Schools staff served in such specific roles as Academic Program Lead or Data Lead, data sharing occurred more smoothly, as designated Citizen Schools representatives were responsible for obtaining and sharing performance data and/or homework assignments or expectations. However, fewer schools had formal processes for sharing information about student homework; the majority relied upon informal efforts on the part of Citizen Schools staff.

> All the daytime teachers have a panic button in their classrooms and when they press it an administrator comes over the loud speaker like God (or classroom intercom), "How can I help you Ms. So-and-so?" It's very formal. In the ELT program, our staff members text us. But this looks unprofessional since they have their phones out and they're not using the same system as the school. So there is no degree of severity [for students] for breaking the rules
>
> –CS Deputy Campus Director

Different approaches to managing student behavior problems in first and second shifts were problematic for the majority of CS schools. Generally, Citizen Schools staff have had substantially less training or experience in how to manage behavior problems than first shift educators, and perhaps as a result, there were often inconsistencies between approaches to managing student behavior, Many schools attempted to address this inconsistency through more integrated referral processes, explicit alignment of disciplinary policies across shifts, offering official training, mentoring and co-teaching, although these approaches were ad hoc and varied across the network—and even these efforts faced additional pressures when unanticipated turnover among Citizen Schools staff occurred.

3.3.4 Relationships

Having strong, trusting, and positive working relationships between Citizen Schools staff and school administration, and school faculty, respectively, was important network-wide. The presence of good relationships both signaled the value of the Citizen Schools contribution and helped mitigate challenges that arose; this is consistent with earlier research documenting the importance of "relational trust" in Chicago schools (Bryk and Schneider, 2003). The absence of such relationships, not surprisingly, seemed to be a factor in schools that struggled with integration of staff, programming, and effective supports for student learning.

> For the ELT partnership, it's the administrative relationship that has to be the cornerstone. If it's not strong, it's hard to implement the CS program. If they are a resource to school, then they must use it as a calculated resource. They need to make sure the school has a plan for how CS programming fits into the school.
>
> –ELT School Principal

Integration efforts indicate how the partner schools welcomed and accommodated Citizen Schools ELT; adaptation describes how effectively Citizen Schools altered program implementation based on schools' specific needs and structures. The next section describes how school and Citizen Schools

staff the types of adapting that occurred both by school staff and Citizen Schools to best serve their students.

3.4 Adaptations to the Model

3.4.1 Key Findings

- Comments from the majority of schools in earlier cohorts about the capacity of the Citizen Schools national curriculum to engage students, align with and complement student learning in the first shift, and function effectively for students with disabilities contributed to changes in how Citizen Schools developed, reviewed, and used its national curricula.
- Schools consistently implemented their own creative approaches to the primary academic component of the CS ELT model (Academic League) to meet their own students' needs.
- Campuses' experiences implementing—and adapting—the Citizen Schools ELT model, particularly the College to Careers Connection (C3) program component, led Citizen Schools to adjust its earlier expectations that participating schools would implement the core components as initially designed. The Citizen Schools ELT program model continued to reflect the same four emphases, although schools were able to adapt those elements to accommodate their own students' needs and circumstances.

Citizen Schools initially developed and refined its National English/Language Arts and Math Curriculum in order to provide resources for program launch in diverse settings. Yet locally, the National Curriculum was not perceived as sufficiently engaging or flexible. Citizen Schools CDs and Teaching Fellows (in earlier cohorts) consistently observed mismatches between the national curriculum and school needs, in terms of engaging students effectively when they had already spent six or seven hours in school, or in terms of content coverage (e.g., skills and topics emphasized in a given school's instruction were omitted completely in the Citizen Schools curriculum).

> What was happening was that students had all this exciting technological learning during the day, and most of our teachers were in their 7th year teaching and have proven track records—and then you have a recent college grad teaching a pen and paper program at the end of the day. It was flopping.
>
> –ELT School Principal

At the organizational level, Citizen Schools has focused substantial attention on tightening its curricular offerings, and at the individual campus level, ELT campuses have implemented modifications designed to accommodate students' needs. These adaptations include restructuring various components of the Citizen Schools ELT model to ensure student engagement across all middle school grade levels. Overall, adaptive strategies have provided school leaders and Citizen Schools staff greater flexibility and voice in ELT planning and implementation, and have also enhanced students' access to academic and enrichment opportunities within the school environment.

3.4.2 Adaptations to Core Program Activities

Over several consecutive years, the organization's reliance on and use of a centrally developed national curriculum continued to evolve. What worked for one school did not necessarily work for another; similarly, what worked in the first year of Citizen Schools ELT was not necessarily appropriate two years later. Schools varied in their efforts to adapt Academic Leagues, for example, by augmenting the national curriculum with alternative strategies (e.g., using word generation

curriculum tied to hip-hop or junk food—categories designed to interest middle school students specifically), or using additional assessments to inform gaps in student learning missed within the regular school day (e.g., using Achievement Network-generated data to help first-shift educators identify weaknesses in students' knowledge and understanding.

> As we have closed the gap between our kids and proficiency, we have had to change what we are teaching. What we had to teach two years ago is no longer relevant. Everyone is adapting to that, even Citizen Schools. We change according to the data we observe.
>
> –ELT School Principal

Persistent questions about why the C3 some program elements were not integrated in the C3 program element was separate from Academic League, homework assistance, or apprenticeships contributed to Citizen Schools' recognition that individual campuses should be able to incorporate C3 into other program activities as appropriate to their circumstances. Schools varied in their adaptations; some redesigned the C3 curriculum to focus specifically on 8th graders, others broadened its content to apply to all middle school grades, and some integrated strands of C3 into other Citizen Schools activities. For example, in a few schools whose districts have competitive high school application processes (e.g., pilot or exam high schools), Citizen Schools staff tailored the C3 curriculum to help students prepare for the admissions process, and made the connections to subsequent college admissions explicit. Other strategies included workshops on time management, note-taking, and leadership skills to make college relevant for lower-middle school students; taking students on visits to college campuses; helping students think about what they need to do to get to college; or developing "College 101" sessions to introduce students to what college is about.

> The campus is moving in this direction because C3 feels like 'one and done'. Integrating C3 concepts into apprenticeships makes the material more engaging for students.
>
> –CS Campus Director

Adaptations to the ELT model, both at the organization and individual campus levels, helped the Citizen Schools ELT program develop stronger local partnerships. Sustaining those partnerships is the focus of the next section, below.

3.5 Sustainability

By the end of the 2014–15 school year, the Citizen Schools network included campuses in their second through fifth years of implementation. Questions about sustainability have been raised by schools at different points in their Citizen Schools' timelines, and internally by the Citizen Schools organization, about how to sustain school/district commitment and participation, particularly for schools relying upon time-limited federal funding. While the specific questions vary somewhat as a function of schools' own implementation progress and regional contexts, the two most prevalent concerns include (1) financial sustainability, and (2) campus stability—and therefore programmatic sustainability—given frequent staff transitions.

3.5.1 Key Findings

- Future funding was an issue raised by the majority of schools across implementation years.
- Maintaining programmatic continuity in the face of staff transitions was a concern for about half of schools across implementation years.

- Expansion of Citizen Schools to additional grades within current schools and to other districts/regions raised questions about the balance between organizational growth and programmatic sustainability.

3.5.2 Funding

> Now that SIG funds are going away, the school isn't mandated to offer extended day… The funding issue was the deciding factor in ELT leaving. It was one quarter to one third of the SIG budget. If money was not an issue (for extended day) [I] would keep CS.
>
> –ELT School Principal

Schools accessed various sources to pay for Citizen Schools ELT programming, including federal Title I, School Improvement grants, and 21st Century Community Learning Center funding, as well as host school districts' use of their own discretionary funding and local philanthropic donations. On one hand, the diversity of sources means that the Citizen Schools ELT model is not dependent on a single funding stream, yet on the other hand, it means continuous pressure to seek and obtain grant support from varied sources. It also means that philanthropic donors may be approached at both individual school and district or regional levels—whether under school/district or Citizen Schools' auspices. In rare instances, Citizen Schools supported ELT funding with its own resources, reflecting an especially compelling partnership it hoped could subsequently obtain other funding support. Because local educational priorities could—and did—change, the funding for ELT has often been subject to shifting budget allocations, and that instability adversely affected CS' program sustainability.

How schools could continue to support future Citizen Schools ELT programming was an issue for the large majority of schools nearing the end of three-year grant cycles (see text box). Schools that used their own discretionary resources to fund some ELT costs could exert greater control over the future of ELT programming, simply because they were not fully dependent on grant or district funding. Principals observed that the overall program cost was generally more manageable for larger schools, since year-to-year (or even within-year) enrollment fluctuations could be absorbed more easily than at smaller schools.

3.5.3 Staff Consistency

The Citizen Schools staffing model was an issue raised by over a third of schools. The built-in turnover of Teaching Fellows, as well as the transitions (both planned and unanticipated) when campus directors and deputy directors exit—even to take positions elsewhere in the network—posed challenges to the continuity and sustainability of relationships between first and second shift staff. Teaching Fellows typically exit just when they have gained expertise and knowledge of the school context, and the schools then inherit a new cohort of inexperienced staff. By contrast, some campuses credited their Citizen Schools ELT programmatic stability and success to staff longevity, because the CDs had been on campus for several years.

> [She worries] … about the talent pipeline as so much of the program's success rests on the teaching fellows. It is important that the sites are fully staffed and that the fellows are stable.
>
> –CS Regional Program Director

3.5.4 Growth

Citizen Schools' growth trajectory also affected the sustainability of ELT funding and staffing, both because schools (or districts) could not necessarily expand to other grades or campuses without obtaining additional financial support, and because expanding the number of students served also meant corresponding cost increases due to higher numbers of Citizen Schools staff to be trained, mentored, and supervised. One regional director reported that the expansion of sites and customization of program implementation across regions makes both campus oversight and fundraising more challenging. "People don't want to fund unicorns," she explained. "People want to fund things that are sustainable."

3.6 Implementation index

Over the course of the Citizen Schools ELT evaluation, it has become and remains evident that implementation varies substantially across participating campuses. The Citizen Schools organization is keenly interested in understanding how the Citizen Schools ELT schools vary in their implementation of the model. To that end, the Abt study team, working collaboratively with Citizen Schools, developed a summary metric, called the implementation index, to capture individual campus and network-wide progress on implementation of core Citizen Schools ELT program elements—the seven "non-negotiable" processes and activities that together, represent what successful implementation of the program would require (see Exhibit 3.3 below). The implementation index draws from surveys administered in schools' second and subsequent years of implementation as well as annual interviews. (Specific survey outcomes will be discussed in Chapter 4.) The index incorporates information from multiple school-based respondents, including principals, classroom teachers, Citizen Schools CDs and Citizen Schools Teaching Fellows.

The index provides information about progress, consistency, and variability—whether for individual campuses, for cohorts of schools that began implementation at specific times, or for the entire network, both at a given point in *time* (e.g., as of the 2012–13 academic year) as well as at a given point in *implementation* duration (e.g., as of the second year of program implementation).

The multi-dimensional index includes seven key elements of the Citizen Schools program model (Appendix B.1 describes how the index was created):

1) Planning;
2) Leadership;
3) Data collection;
4) Training and professional development;
5) Family/community engagement;
6) Alignment/coordination between partner school and CS; and
7) Perceptions of program quality.

The level of implementation fidelity for each program element is measured by a number of related indicators that draw from survey and interview data. The levels reflect increasing and corroborated evidence that the core element has been implemented as intended. Scores are based upon the presence or absence of the structural elements of the implementation index.

Exhibit 3.3. Implementation Index Constructs and Indicators

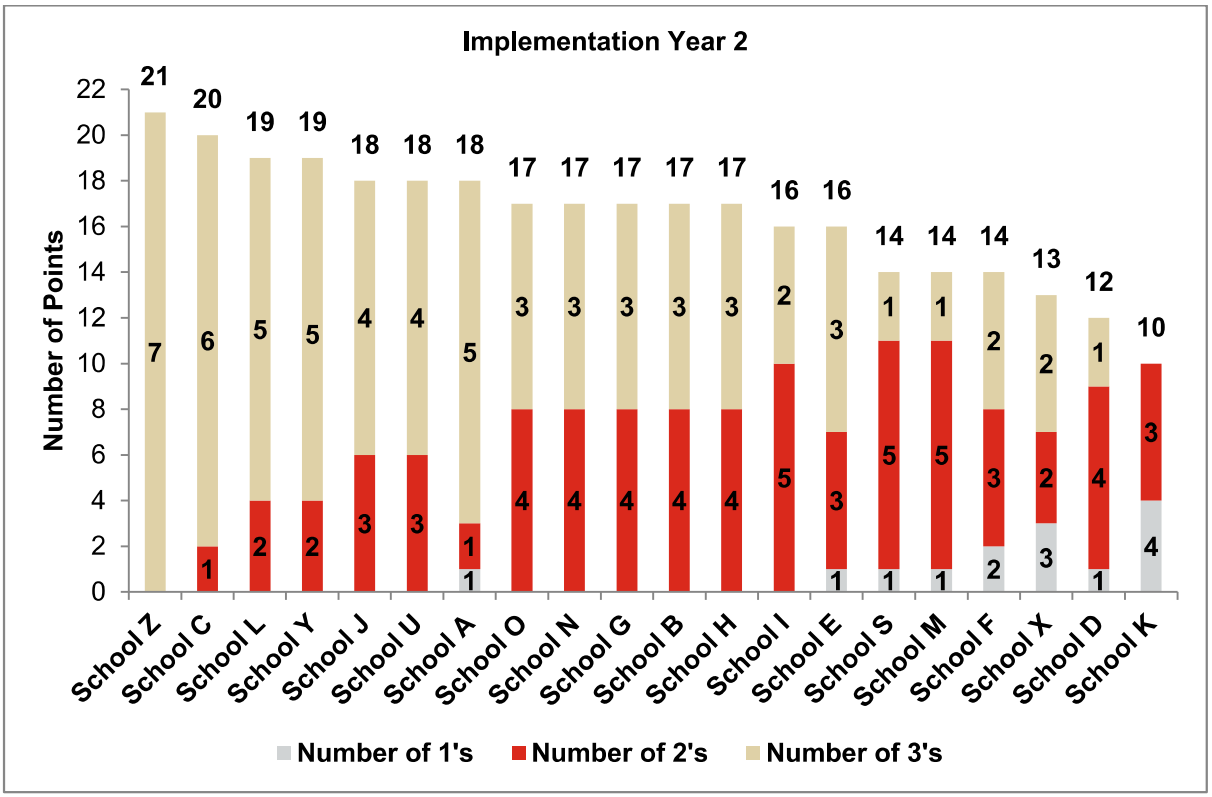
Construct Fidelity of implementation reflects …	Indicators
1. Planning	a. Ongoing planning between CS and School
2. Leadership	a. Engaged leadership
	b. Support for ELT
3. Data Collection	a. Data sharing between 1st and 2nd
	b. CS Aligned with data results
4. Training and PD	a. Training for CS staff
	b. Increased responsibility for fellows
5. Family and Community	a. Families are engaged and informed
6. Alignment and Coordination between Partner School and CS	a. Alignment of content and behavior standards
	b. Nature and frequency of communication
	c. Nature and frequency of integration
7. Program Quality	a. High-quality programming

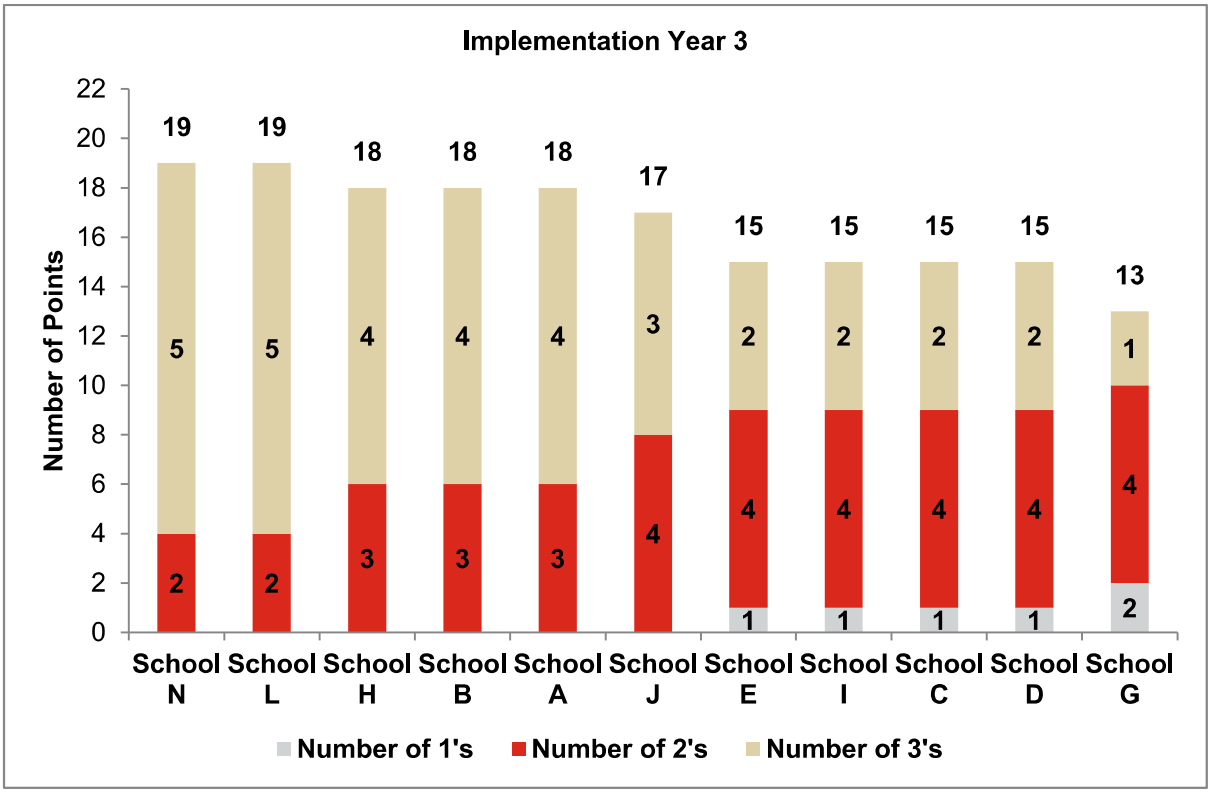
EXHIBIT READS: For Construct 1, Planning, fidelity of implementation reflects ongoing planning between Citizen Schools and School Partners, and there is one indicator that contributes to the overall score for this construct.

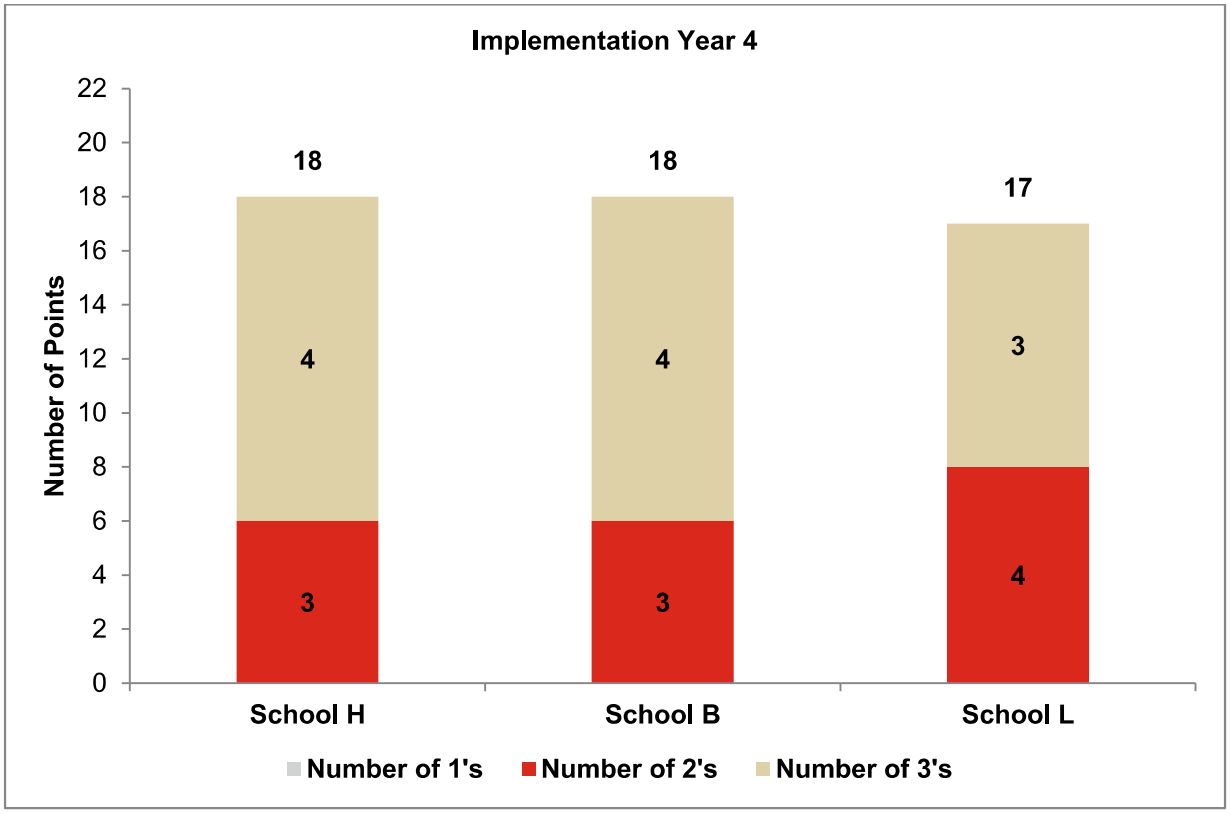
Each school's implementation (in a given year) is characterized as limited, moderate, or full (corresponding to ratings of 1, 2, or 3), for each of the seven core program constructs described in Exhibit 3.3 above. School ratings could vary between 7 (representing limited implementation of each construct) and 21 (full implementation across all seven constructs). In the exhibits below, schools are not named, but are identified as School A, School B, School C, etc., as they were in Exhibit 2.3 in chapter 2.

Exhibit 3.4 below displays schools' overall scores for each cohort and for each implementation year, on the implementation index. [Note that because surveys were not fielded until schools' second year of program implementation, that represents the first year for which index data are available.] While schools are implementing several program elements consistently, implementation continues to vary substantially, both within individual schools over time, and for the network writ large. The exhibit illustrates such variation; it presents information for schools that had completed their second, third, and/or fourth years, respectively. Because the number of schools continuing to implement ELT decreases for each subsequent year or implementation, the number of schools also decreases in each of the three panels below. In their second year of implementation, 20 schools' ratings ranged from 10 to 21 (with 10 representing more limited implementation, and 21 representing full implementation), and the average was about 17, suggesting, on average, moderate to some full implementation. Ratings for the 11 schools in their third year ranged from 13 to 19, and the three fourth-year schools' ratings ranged from 17 to 18. Interestingly, most campuses' ratings remained in the "moderate" category, and instances of "limited" ratings generally faded as implementation matured—and as the schools that had struggled with various programmatic elements exited the network.

Exhibit 3.4A-C Overall School Scores, by Cohort and Implementation Year






Implementation Year 4
Number of Points
22
20
18
16
14
12
10
8
6
4
2
0
18
18
17
4
4
3
3
3
4
School H
School B
School L
Number of 1's
Number of 2's
Number of 3's

Key findings about implementation based on application of the implementation index include:

- Overall index scores remain clustered in the moderate range, with average scores hovering around 17 (out of 21) across cohorts and implementation years.
- There are few patterns evident in schools' levels of implementation over time, although schools with lower index scores tended to exit the network over time.
- Overall, continuing schools' implementation levels do not improve from one year to the next.
- The specific constructs schools were most consistently able to implement fully *across* implementation years include planning, leadership, perceived quality, and data collection. For example, 16 schools (of 19 for which index scores are available) in their second year, 9 (of 11) in their third year, and 3 (of 3) in their fourth year scored a "3" for the planning construct.
- Fewer schools implemented the family and community engagement and alignment of partner school and Citizen Schools constructs fully, again, *across* implementation years.

The study team created pie charts displaying the evidence scores for Cohorts 1, 2, 3, and 4 for 2013–14. Each pie has seven parts, one for each program element (see Exhibit 3.6). The exhibit below illustrates each school's level of evidence for each element. Exhibit 3.6 presents data for the remaining Cohort 1 schools, in their fourth year of implementation; Cohort 2 schools, for their third year; Cohort 3, for their second year; and Cohort 4 schools, for their first year of implementation. Recall that the index is based on a combination of survey and interview data, with one exception: Cohort 4 schools' scores are based solely on interview data.[25] (Similar pie charts for earlier years, albeit with fewer schools, are included in Appendix B.4.)

Looking at the set of individual school charts, it is clear that the majority of schools were able to implement Citizen Schools ELT with moderate levels of fidelity. Few schools have multiple slices (constructs) with low scores, and the large majority was implementing at least two constructs at a moderate level. Examining the schools by cohort suggests a slightly different pattern. In their second year of implementation, one Cohort 1 school, five Cohort 2 schools, and one Cohort 3 school had scores of "1" (white slices). Some—but not all—of the schools with low scores exited the network. From the second to the third implementation year, most continuing schools earned very similar overall scores, although none did so by scoring the same way as in the prior year. The fact that schools appeared to be implementing particular program elements fully in one year did not guarantee a repeat performance in the subsequent year. To the contrary, not one school had the same profile from one year to the next, regardless of that school's overall score.

The most striking finding, whether focusing on individual schools, cohorts, or academic year, is that schools' implementation of the program varies substantially. The discussions earlier in this chapter offer some insights about why schools with demonstrated early implementation successes struggle to sustain those successes.

---

[25] The study team explored whether index scores changed as a function of including survey and interview data separately or in combination, and determined that the results were generally indistinguishable. This exploration meant that information from Cohort 4 school interviews in 2013–14 could be used to generate index scores.

Exhibit 3.5. Pie Chart Key

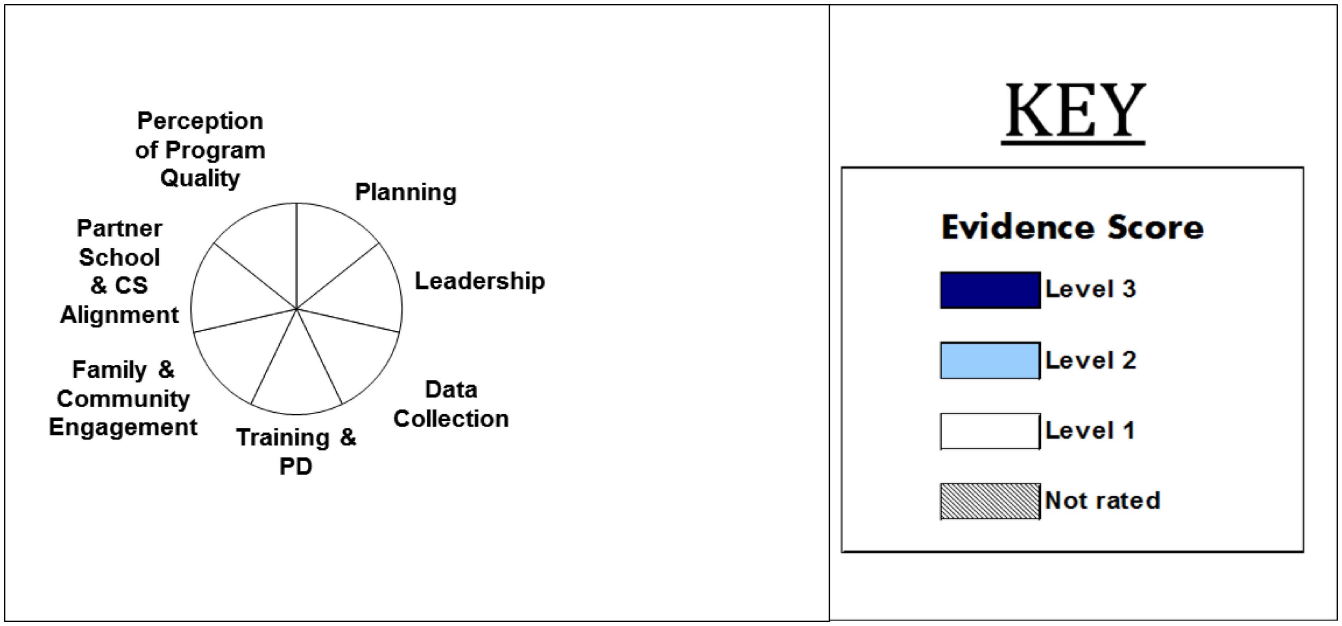
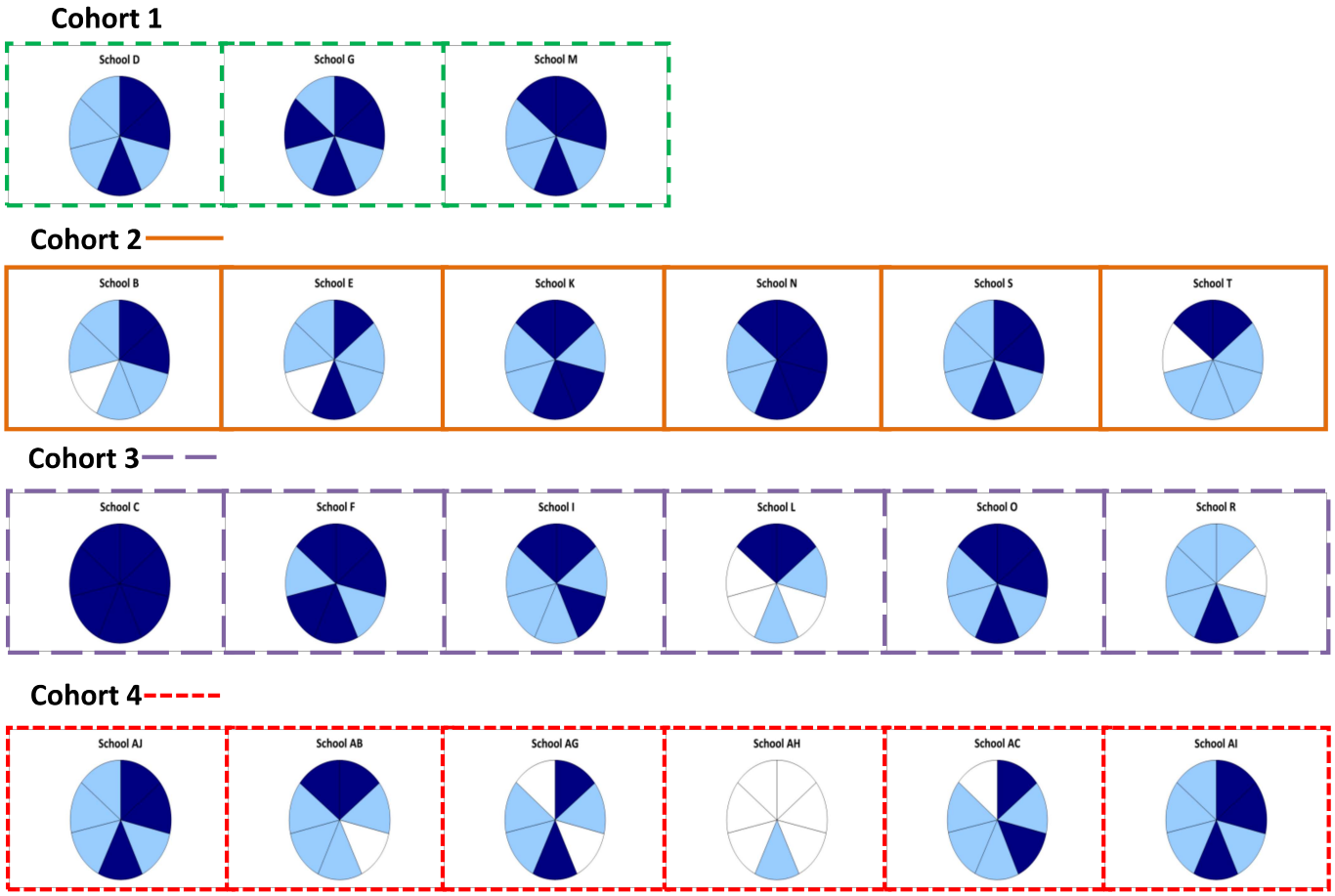


Exhibit 3.6. Implementation Index Scores for Citizen Schools ELT Schools 2013–14, by Cohort



4. Impact of Citizen Schools ELT on Student Outcomes

This chapter examines the relationship between participation in Citizen Schools ELT and student outcomes. As described in Chapter 1, prior research suggests that participation in expanded learning opportunities can improve students' academic skills, school work, and educational expectations in school and into their life trajectories (Redd et al., 2012). Citizen Schools ELT programming is designed to provide students additional learning time as well as enrichment experiences that broaden students' awareness of options in high school, college, and careers. This chapter first examines short-term outcomes (depicted in the logic model in Chapter 2), including such non-academic outcomes as student engagement and aspirations. The chapter then progresses through the logic model to explore the hypothesized distal or longer-term outcomes—the impact of Citizen Schools' ELT programming on academic achievement.

4.1 Student Short-Term Outcomes

The Citizen Schools ELT program model hypothesizes that participation will positively affect short-term student outcomes, including increased school engagement and student aspirations. The study team uses data from surveys to explore these hypothesized proximal student outcomes in the Citizen Schools model,[26] and draws from surveys administered to Citizen Schools ELT and matched comparison students and teachers, along with Citizen Schools' ELT staff, during the 2011–12, 2012–13, and 2013–14 academic years.[27] Evidence about the effects of participation in ELT and similar programs suggests that positive school experiences develop students' academic capabilities, as well as broader life skills, some of which may in turn influence affect academic outcomes. By providing holistic, enriching activities, Citizen Schools ELT programming hopes students become more engaged with the larger academic community, and develop a set of tools they can use to pursue academic and other goals.

For the ease of describing survey items that each relate to a given category (e.g., academic engagement), the team aggregated individual survey items into constructs which are presented as percentages (e.g., percentage of individuals participating in an activity, agreeing with a concept, etc.). For example, if there were 10 items corresponding to student engagement, these were averaged to create one overarching construct.[28] Appendix C provides additional information about individual

[26] Ideally, these outcomes would have been measured using pre and post Citizen Schools student-level data; however, surveys could not be administered at baseline as the evaluation began concurrently with the entrance of the first cohort in the study.

[27] Recall that no surveys were administered during the first year of Citizen Schools ELT implementation (2010–11), student surveys were discontinued after the 2012–13 school year, and teacher and staff surveys were administered only to Citizen Schools campuses in the 2013–14 academic year to inform the implementation index.

[28] As discussed in Chapter 2, the team used a technique called factor analysis to identify items that could reasonably be grouped into substantively linked constructs. All survey items that addressed a non-academic outcome of interest were averaged together in a single construct. In order for a construct to be calculated, respondents had to answer at least two thirds of the survey items included in the construct. For example, if 10 items comprised a particular construct, respondents had to answer at least seven of them for the items to be averaged together to create the construct.

items within various constructs along with results of all survey scales across years.[29] Three other notes about the results described below include the following: a) because schools joined (and exited) the Citizen Schools network over time, each round of survey administration reflects a different combination of schools in each cohort; b) the term "some evidence" refers to findings that were not consistently statistically significant across years; and c) comparison means are regression-adjusted estimates of what *would* have happened in Citizen Schools ELT schools without Citizen Schools ELT, and therefore are described as "what would be expected in the absence of Citizen Schools ELT," as the estimated non-ELT mean, or as the counterfactual.

4.1.1 Key Findings

Key findings related to outcomes are summarized below. Note that the findings include both cross-sectional outcomes as well as results from comparative analyses, and are noted as such below. We also highlight instances where comparative results were based on a *subset* of the comparison group. In particular, survey questions that asked comparison students about after-school activities were necessarily subset only to those comparison students who had participated in after-school programming (and thus may be a highly selected group of individuals):

Guide to Quantitative Descriptors

- Most: 75 – 100% (of survey respondents)
- Majority: 50 – 74%
- Substantial minority: 25 – 49%
- Few: < 24%

- Citizen Schools students reported high levels of social engagement, feelings of belonging, and perceptions that teachers care and encourage students. Citizen Schools staff and teachers in Citizen Schools campuses, however, reported that student engagement was a problem area in their schools (cross-sectional).
- Students in Citizen Schools ELT schools reported significantly more positive behavioral engagement of their peers than the counterfactual, but there was some evidence that significantly more Citizen School ELT students engaged in behavioral misconduct than the estimated non-ELT mean (e.g., skipping school; comparative).
- A majority of Citizen Schools students expressed positive feelings of belonging in their after-school programming, engagement with peers, and positive relationships with Citizen Schools staff (cross-sectional).
- There was some evidence that a lower percentage of Citizen Schools ELT students had positive feelings of belonging after-school and satisfaction with their personal agency during after-school time than the estimated non- ELT mean (comparative, subset).
- A significantly higher percentage of students in Citizen Schools ELT campuses expressed negative feelings about Citizen Schools staff and programming than the counterfactual. However,

[29] As described in Chapter 2 and Appendix A, each item or construct is analyzed using multi-level models that adjust standard errors for the clustering of students within classrooms or teachers/staff within schools, and include a CS ELT indicator, dummy variables for each pair of matched schools, and individual-level demographics. Results in tables present actual CS ELT means for a particular outcome, along with the estimated non-ELT mean, the estimated *absolute* difference (presenting the magnitude of the difference without indicating if the difference is positive or negative) and the statistical significance of the estimated difference.

significantly more Citizen Schools ELT students also reported that the program helped their self-esteem and pro-social behaviors, and there was some evidence that a higher percentage of Citizen Schools ELT students reported that the program helped their performance in core academic subjects than would be expected in the absence of Citizen Schools ELT (comparative, subset)

- Most Citizen Schools students aspired to finish college. Yet, generally, Citizen Schools staff and teachers in Citizen Schools campuses did not report high levels of student aspirations in their schools (cross-sectional).
- There was some evidence that Citizen Schools teachers perceived student expectations for academic success were significantly higher than the estimated non-ELT mean, and a significantly higher percentage of Citizen Schools students reported participating in activities to learn about colleges and careers during the school day than the estimated non- ELT mean (comparative).
- There was also some evidence that a higher percentage of Citizen Schools students felt that Citizen Schools helped them learn about college, careers, and job skills than the estimated non-ELT mean (comparative, subset).

4.1.2 Student Engagement

General Student Engagement

A key expectation for the Citizen Schools model is that increased access to various activities and enrichment opportunities make learning more relevant and meaningful in students' daily lives. These changes to the school day, in turn, are expected to increase students' social, academic, and behavioral engagement. Exhibit 4.1 presents various engagement metrics that draw from student, teacher, and Citizen Schools staff perspectives. Over 70 percent of students reported positive levels of social engagement and feelings of belonging in their schools, and over 80 percent perceived that teachers care about and encourage students (neither is significantly different from the estimated non-ELT mean). However, students in Citizen Schools ELT schools reported significantly more positive behavioral engagement of their peers than the counterfactual. Levels of behavioral misconduct were relatively low; less than 30 percent of students across both Citizen Schools and comparison campuses reported having engaged in negative behaviors such as skipping school during the school year, although there was some evidence that Citizen Schools ELT students had participated in somewhat more negative behaviors than their counterfactual counterparts.

The overall measure of student engagement, which captures information from teachers' and staff perspectives, indicates that student engagement is an area of concern across all schools: fewer than half of teachers and less than 30 percent of Citizen Schools staff reported positive student engagement. It is also interesting to note that teachers, who generally may be more experienced educators than Citizen Schools staff, with broader reference points, tended to report more positive student engagement than did Citizen Schools staff.

Exhibit 4.1. Relationship between Citizen Schools ELT and Student Engagement, by Academic Year

	2011–12				2012–13				2013–14
	Actual ELT Mean (%)	Estimated Non-ELT Mean (%)	Estimated Difference (%)	Statistical Significance (p-value)	Actual ELT Mean (%)	Estimated Non-ELT Mean (%)	Estimated Difference (%)	Statistical Significance (p-value)	Actual ELT Mean (%)
Student Items									
Social engagement and feelings of belonging[1]	81	73	7.3	0.310	75	72	2.8	0.497	--
Perceptions that teachers care and encourage students[2]	87	80	7.5	0.106	80	81	0.9	0.802	--
Behavioral engagement of peers[3]	61	30	31.0*	0.030	43	31	12.1+	0.098	--
Behavioral misconduct[4]	29	27	2.4	0.613	29	23	5.9+	0.057	--
School Teacher Items									
Student engagement[5]	44	29	14.6	0.184	31	29	2.3	0.624	39
Citizen Schools Staff Items									
Student engagement[6]	26	--	--	--	21	--	--	--	27

EXHIBIT READS: Eighty-one percent of students in CS ELT schools reported positive feelings of belonging in the 2011–12 academic year, compared to an estimated 73 percent in the absence of ELT.

Notes:

[1]Items 12a–12e.
[2]Items 13d, 13e, 13i, 13j, 13k, 13m, 13p, 13q, 13t.
[3]Items 13g, 13u.
[4]Items 18b, 18d, 18f, 18l, 18n.
[5]Items 10a, 10b, 10c, 10d, 10e, 10h. This construct was reverse coded so that higher numbers indicate positive responses.
[6]Items 9a, 9b, 9c, 9d, 9j. This construct was reverse coded so that higher numbers indicate positive responses.

\+ marginally significant at $p < .10$ level
* statistically significant at $p < .05$ level
** statistically significant at $p < .01$ level
*** statistically significant at $p < .001$ level

Source: Abt Associates' Surveys of ELT and Matched Comparison School Students, Spring 2012 and 2013 and Teachers/Citizen Schools Staff, Spring 2012, 2013, and 2014.

Sample: 238 students in 2012 and 837 students in 2013. Nonresponse rates across ELT and matched comparison student survey items ranged from 6.7 to 31.3. 90 teachers in 2012, 369 teachers in 2013, and 158 teachers in 2014. Nonresponse rates across ELT and matched comparison teachers ranged from 1.1 to 3.5 percent . 36 Citizen Schools staff in 2012, 135 Citizen Schools staff in 2013, and 195 Citizen Schools staff in 2014. Nonresponse rates ranged from 0.0 to 1.5 percent.

After-School Engagement

Additional opportunities for enrichment, instruction, and interaction for Citizen Schools ELT students are expected to improve engagement not only during the regular school day, but also during ELT programming. Citizen Schools ELT and matched comparison students (limited to those who participated in after-school programs) reported on aspects of engagement during their after-school time.[30] Similar to the results reported above, the majority of Citizen Schools students expressed positive feelings of belonging in their after-school programming, engagement with peers, and positive relationships with Citizen Schools staff, although there was some evidence that a lower percentage of Citizen Schools ELT students had positive feelings of belonging after-school than the estimated non-ELT mean (Exhibit 4.2). Further, a significantly higher percentage of students in Citizen Schools ELT campuses expressed negative feelings about Citizen Schools staff and Citizen Schools programming than would be expected in the absence of ELT; there is some evidence that a smaller percentage of Citizen Schools students were satisfied with their personal agency during after-school time.

These latter findings are worth considering further, for two reasons, related to the comparison school student sample and the nature of their after-school options. First, only the 50 percent of comparison school students who indicated they participated in school-based after-school programming were included in the analysis. Second, matched comparison students' after-school programming may reflect considerably broader ranges of optional participation as well as content, structure, and location, whereas Citizen Schools programming likely reflects more consistency and structured programming. Successfully implementing the Citizen Schools ELT model requires Citizen Schools staff to enforce rules, set boundaries, and manage student behavior. And while the Citizen Schools ELT model offers students choices about selecting apprenticeships, student autonomy may well be greater for students in comparison schools, who can choose all of their activities during the after-school period. As a result, Citizen Schools students may well have qualitatively different (i.e., stricter, more structured) experiences after-school than comparison students.

[30] It is important to note that these comparisons are based on a subset of students from matched comparison schools, specifically those who reported participating in an after-school program. As such, these comparison students are self-selected, and they may (or may not) be participating in after-school programs similar to the Citizen Schools ELT model. Note also that Citizen Schools ELT surveys asked specifically about Citizen Schools programming, while matched comparison surveys asked more generally about after-school programming. For brevity, these are characterized as Citizen Schools programming (instead of Citizen Schools/after-school programming).

Exhibit 4.2. Relationship between Citizen Schools ELT and After-School Engagement, by Academic Year

	2011–12				2012–13			
	Actual ELT Mean (%)	**Estimated Non-ELT Mean** (%)	**Estimated Difference** (%)	**Statistical Significance (p-value)**	**Actual ELT Mean** (%)	**Estimated Non-ELT Mean** (%)	**Estimated Difference** (%)	**Statistical Significance (p-value)**
Feelings of belonging after school[1]	72	74	2.2	0.724	61	70	9.3+	0.066
Social engagement after school with peers[2]	78	70	8.1	0.155	66	70	3.2	0.407
Satisfaction with student agency[3]	47	48	1.2	0.851	32	45	12.7**	0.008
Positive relationships with Citizen Schools staff[4]	77	79	2.6	0.703	69	75	5.7	0.152
Negative perceptions of Citizen Schools staff[5]	62	45	16.9*	0.035	55	44	11.7*	0.032
General dissatisfaction with the Citizen Schools program[6]	63	39	24.0**	0.008	59	41	18.3***	<0.001

EXHIBIT READS: Seventy-two percent of students in CS ELT schools in the 2011–12 academic year reported positive feeling of belonging after school, compared to an estimated 74 percent in the absence of ELT.

Notes:

[1]Items 14a_14, 14a_15, 14a_16, 14a_18, 14a_19, 14c_8, 14c_9.
[2]Items 14a_1, 14a_3, 14a_6, 14a_9, 14a_13.
[3]Items 14a_8, 14a_10, 14a_11, 14a_12, 14c_12.
[4]Items 14c_1, 14c_2, 14c_4, 14c_7, 14c_10, 14c_11.
[5]Items 14c_3, 14c_5, 14c_6, 14c_13.
[6]Items 14a_2, Q14a_4, Q14a_7.

\+ marginally significant at p < .10 level
* statistically significant at p < .05 level
** statistically significant at p < .01 level
*** statistically significant at p < .001 level

Source: Abt Associates' Surveys of ELT and Matched Comparison School Students, Spring 2012 and 2013.

Sample: 210 students in 2012 and 634 students in 2013 (matched comparison students in both years who indicated they did not participate in after-school programming were excluded from the sample). Nonresponse rates across ELT and matched comparison student survey items ranged from 19.3 to 33.4 percent.

Perceived Impact of Citizen Schools on Student Engagement

The study team also asked whether students perceive that their after-school programs made a difference in their school (and after-school) engagement. Overall, a higher percentage (and, indeed, most) of Citizen Schools ELT students agreed that Citizen Schools helped their social, academic, and behavioral engagement. A significantly higher percentage of Citizen Schools ELT students reported that Citizen Schools helped their self-esteem and pro-social behaviors. The specific outcomes within this construct include whether Citizen Schools helped students learn about assisting others in their school and community, learn skills to work with and lead peers, and feel better about themselves. There is also some suggestive (marginally significant) evidence that a higher percentage of Citizen Schools ELT students indicated that the program helped their performance in core academic subjects.

Exhibit 4.3. Perceived Impacts of Citizen Schools ELT on Student Engagement, by Academic Year

	2011–12				2012–13			
	Actual ELT Mean (%)	**Estimated Non-ELT Mean** (%)	**Estimated Difference** (%)	**Statistical Significance (p-value)**	**Actual ELT Mean** (%)	**Estimated Non-ELT Mean** (%)	**Estimated Difference** (%)	**Statistical Significance (p-value)**
Self-esteem and pro-social behaviors[1]	85	69	16.0*	0.038	77	71	6.6+	0.093
Performance in core subjects (e.g., math)[2]	80	69	11.0	0.231	77	69	8.0+	0.064
Work habits[3]	83	71	11.9	0.270	81	75	5.7	0.141

EXHIBIT READS: Eighty-five percent of students in CS ELT schools in the 2011–12 academic year agreed that CS ELT programming impacted their social engagement, compared to an estimated 69 percent in the absence of ELT. The difference was statistically significant.

Notes:

[1]Items 17g, 17h, 17i, 17j, 17l, 17m, 17n.
[2]Items 17o, 17p, and 17q.
[3]Items 17c, 17d, 17e, and 17f.

\+ marginally significant at p < .10 level
* statistically significant at p < .05 level
** statistically significant at p < .01 level
*** statistically significant at p < .001 level

Source: Abt Associates' Surveys of ELT and Matched Comparison School Students, Spring 2012 and 2013.

Sample: 210 students in 2012 and 634 students in 2013 (matched comparison students in both years who indicated they did not participate in after-school programming were excluded from the sample). Nonresponse rates across ELT and matched comparison student survey items ranged from 13.5 to 23.0 percent.

4.1.3 Student Aspirations

The underlying theoretical model for ELT hypothesizes that increased time for learning improves student engagement, and also improves student aspirations. Exhibit 4.4 indicates that across all schools, most students planned to finish college. Interestingly, however, fewer teachers and staff reported positive levels of student aspirations, although there was some evidence that Citizen Schools teachers perceived student expectations for academic success to be significantly higher than the counterfactual.

Exhibit 4.4. Relationship between Citizen Schools ELT and Student Aspirations, by Academic Year

	2011–12				2012–13				2013–14
	Actual ELT Mean (%)	Estimated Non-ELT Mean (%)	Estimated Difference (%)	Statistical Significance (p-value)	Actual ELT Mean (%)	Estimated Non-ELT Mean (%)	Estimated Difference (%)	Statistical Significance (p-value)	Actual ELT Mean (%)
Student Items									
Plans to finish college[1]	79	87	7.9	0.299	83	83	0.1	0.965	--
School Teacher Items									
Student aspirations for college and/or career[2]	57	47	9.9	0.523	49	43	6.2	0.276	62
Student expectations for academic success[3]	68	33	34.7*	0.039	44	40	3.9	0.523	49
Citizen Schools Staff Items									
Student aspirations for college and/or career[4]	47	--	--	--	42	--	--	--	41
Student expectations for academic success[5]	42	--	--	--	46	--	--	--	38

EXHIBIT READS: Seventy-nine percent of students in CS ELT schools in the 2011–12 academic year reported that they planned to finish college, compared to an estimated 87percent in the absence of ELT.

Notes:

[1]Item 16.
[2]Item 10f. This item was reverse coded so that higher numbers indicate positive responses.
[3]Item 10o. This item was reverse coded so that higher numbers indicate positive responses.
[4]Item 9f. This item was reverse coded so that higher numbers indicate positive responses.
[5]Item 9p. This item was reverse coded so that higher numbers indicate positive responses.

\+ marginally significant at p < .10 level
* statistically significant at p < .05 level
** statistically significant at p < .01 level
*** statistically significant at p < .001 level

Source: Abt Associates' Surveys of ELT and Matched Comparison School Students, Spring 2012 and 2013 and Teachers/Citizen Schools Staff, Spring 2012, 2013, and 2014.

Sample: 238 students in 2012 and 837 students in 2013. Nonresponse rates across ELT and matched comparison student survey items ranged from 21.4 to 24.5 percent. 90 teachers in 2012, 369 teachers in 2013, and 158 teachers in 2014. Nonresponse rates across ELT and matched comparison teachers ranged from 2.5 to 7.0 percent. 36 Citizen Schools staff in 2012, 135 Citizen Schools staff in 2013, and 195 Citizen Schools staff in 2014. Nonresponse rates ranged from 0.0 to 4.6 percent.

The Citizen Schools logic model posits that increasing aspirations occurs through pathways to college and careers, using such mechanisms as enrichment opportunities to promote future career interest and success. Exhibit 4.5 summarizes Citizen Schools students' reports about their exposure to resources for the future. A significantly higher percentage of students in Citizen Schools ELT schools reported participating in activities to learn about college and careers during the school day. In addition, there is some evidence that students felt that Citizen Schools helped them learn about college, careers, and job skills.

Exhibit 4.5. Relationship between ELT and College and Career Connections, by Academic Year

	2011–12				2012–13			
	Actual ELT Mean (%)	**Estimated Non-ELT Mean** (%)	**Estimated Difference** (%)	**Statistical Significance (p-value)**	**Actual ELT Mean** (%)	**Estimated Non-ELT Mean** (%)	**Estimated Difference** (%)	**Statistical Significance (p-value)**
Activities to learn about college and careers during the school day[1]	86	36	49.8***	<.001	68	38	29.8***	<.001
Learning about college, careers, and job skills[2]	89	74	14.7	0.135	87	74	13.8**	0.001

EXHIBIT READS: Eighty-six percent of students in CS ELT schools in the 2011–12 academic year participated in activities to learn about college and career during the day, compared to an estimated 36 percent in the absence of ELT. The difference was statistically significant.

Notes:

[1]Items 10d, 10e.
[2]Items 17o, 17p, and 17q.

\+ marginally significant at p < .10 level
* statistically significant at p < .05 level
** statistically significant at p < .01 level
*** statistically significant at p < .001 level

Source: Abt Associates' Surveys of ELT and Matched Comparison School Students, Spring 2012 and 2013.

Sample: For the first construct, 238 students in 2012 and 837 students in 2013. For the second construct, 210 students in 2012 and 634 students in 2013 (matched comparison students in both years who indicated they did not participate in after-school programming were excluded from the sample). Nonresponse rates across ELT and matched comparison student survey items ranged from 4.1 to 19.4 percent.

Although mixed, the findings discussed are consistent with prior research about the outcomes of ELT; there are examples of positive—as well as neutral or even negative—impacts on non-academic domains, including student engagement and student aspirations for the future. Citizen Schools students reported high levels of engagement both during and after-school, along with high aspirations for future educational plans, yet Citizen Schools teachers and staff consistently reported less positive perceptions about those same outcomes. At the same time, Citizen Schools students expressed significantly more discontent about Citizen Schools staff, Citizen Schools programming, and individual agency than would be expected in the absence of ELT. This disconnect may reflect inherent structural differences between the Citizen Schools models and other, typical after-school programs. Nonetheless, Citizen Schools students reported a number of perceived benefits of the program, particularly in helping their self-esteem and pro-social behaviors. These benefits may represent the outcomes that the Citizen Schools ELT model can plausibly affect, through structured opportunities for holistic education and enrichment activities that target real life skills.

4.2 Student Longer-Term Outcomes

The chapter now turns to an examination of the impact of Citizen Schools ELT implementation on more distal outcomes—student achievement, as measured by student scores on English/Language Arts (ELA) and mathematics tests administered by states. First, this section presents results from confirmatory analyses, using a measure that pools results across grade levels and implementation years. Next, the section describes findings from exploratory research questions about whether and how the impact of Citizen Schools ELT varies as a function of implementation year, grade level, program dosage, and implementation level.[31] These results are followed by a discussion about selection bias and the chapter's summary. Recall that because the confirmatory outcomes are in different domains (ELA and math), the study team did not need to adjust for multiple comparisons. Similarly, no adjustments were made to the exploratory analyses given that these are preliminary.

The impact models for the study leverage (1) pre-ELT data to control for observable and unobservable, stable characteristics of schools that affect outcome measures, and (2) data from matched comparison schools to control for year-to-year variation that affect outcomes similarly across schools. The statistical models estimate the impact of Citizen Schools ELT on student achievement by comparing the observed outcomes for ELT schools and the counterfactual outcomes (e.g., regression-adjusted estimates of what would have occurred in the absence of ELT). These models control for many alternative hypotheses that might explain observed differences between Citizen Schools ELT and comparison schools, and, therefore represent strong, quasi-experimental analyses.

Data for the impact analysis include student achievement data from seven different states, each of which has its own ELA and math achievement tests and associated scores. For ease of interpretation and analysis, the impact estimates are presented as z-scores which have been scaled to student-level state means and standard deviations for a given grade, subject, and year, and can thus be interpreted

[31] The confirmatory analyses represent the most rigorous tests of the study's primary research questions and hypotheses; exploratory analyses examine variation in outcomes across subgroups of interest, and also serve to generate hypotheses to be tested more rigorously in the future.

as the average score for a school relative to the student statewide average.[32] Using z-scores as the analysis metric also allows for the estimated impact of Citizen Schools ELT to be interpreted as an effect size, or a proportion of the student-level standard deviation.

4.2.1 Key Findings

- Overall impacts of Citizen Schools ELT on student achievement outcomes were not statistically significant.
- Exploratory analyses suggest a significant positive impact of Citizen Schools ELT on math achievement in the ***first year of implementation,*** and a marginally significant positive effect of Citizen Schools ELT on ***7th grade*** math achievement.

4.2.2 Impact of Citizen Schools ELT on Student Achievement

Exhibit 4.6 presents the pooled estimated impact of Citizen Schools ELT on student achievement across grade levels and for up to four years of implementation (see Appendix Exhibit C.4.1 for sample sizes associated with each model; Appendix Exhibit C.4.2 presents standard errors, p-values, and 95% confidence interval for all estimated impacts).[33] The y-axes indicate the size of the impact of Citizen Schools ELT, presented in effect size metrics, or standard deviation units. The bars above the line labeled "0" represent positive impacts of Citizen Schools ELT, indicating that students in ELT schools scored higher than would have been predicted in the absence of ELT on those tests. Bars below the "0" line indicate negative impacts of Citizen Schools ELT, such that the scores for ELT students are lower than what would have been predicted in the absence of ELT. The confidence interval above and below the bars represent the range in which the values for the true effect of Citizen School ELT can be found. Overall results are presented by subject and indicate that there were no statistically significant overall impacts of Citizen Schools ELT on student achievement scores.[34]

---

[32] For each subject, the z-score is calculated by subtracting the average scaled score for all students in that grade and year in the state from the school-level mean score, and dividing by the standard deviation in that grade and year. For example, if students' average ELA score in a given school is identical to the statewide average for that year, the ELA z-score for that school would be zero. Positive and negative scores would indicate that the school's students scored above and below the state average, respectively.

[33] The analytic models used to generate these estimates compare data from Citizen Schools ELT and matched comparison schools prior to and after the implementation of ELT, and the models use school-level covariates and school, year, and grade fixed effects to control for a number of plausible alternative explanations for the estimated impacts. Further, the models account for clustering at the school level using the cluster-robust variance estimator (also known as the "sandwich" standard errors; White, 1984 and Liang and Zeger, 1986). In school-level models, these are intended to adjust for the correlation of grades within schools and schools over time.

[34] Had impact estimates been significant, they would have been roughly equivalent to nearly two months of growth in reading in a school year, and slightly over two months of growth in math in a school year. Growth in reading and math are calculated using effect-size benchmarks aggregated from various national standardized tests (Bloom et al., 2008; p. 16). Average yearly gains in reading and math for middle school students were divided by nine to estimate how much growth occurs in each month of school. These monthly estimates were converted to the number of months of growth in the present study.

Exhibit 4.6. Overall Impact of Citizen Schools ELT on ELA and Math Test Scores

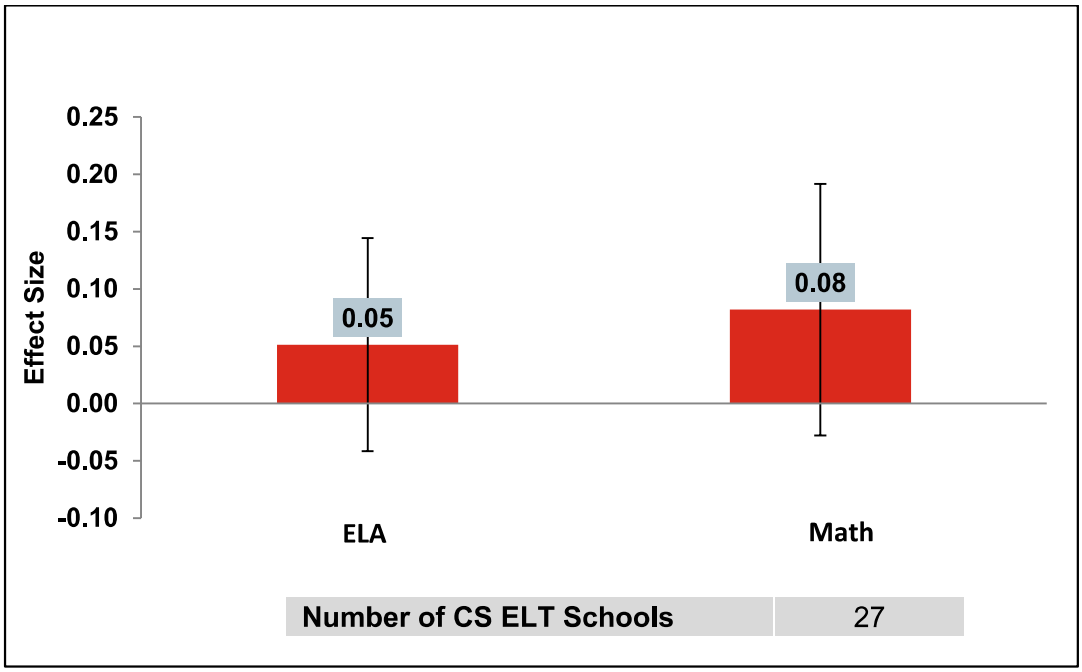


EXHIBIT READS: Students in Citizen Schools ELT schools, on average, scored 0.05 standard deviations higher on their state ELA test than would have been expected in the absence of ELT. This difference (estimated impact of ELT) was not statistically significant. The upper limit of the 95% confidence interval for the impact is 0.14, the lower limit is -0.04.

Notes:

\+ marginally significant at p < .10 level
* statistically significant at p < .05 level
** statistically significant at p < .01 level
*** statistically significant at p < .001 level

Source: Publically available school-level achievement data from states with schools in the Citizen Schools ELT Network.

Sample: Standardized ELA test scores from 27 Citizen Schools ELT and 95 matched comparison schools; standardized math test scores from 27 Citizen Schools ELT and 87 matched comparison schools.

Sensitivity Analyses

The study team conducted a large number of sensitivity analyses to assess the robustness of the results of the student achievement analyses, i.e., whether similar results would have been obtained if different analytic decisions had been made (see Appendix C.4.2). The alternative model specifications included in the sensitivity analyses include:

- A model that adjusts standard errors for clustering at the group-level (groups are matched sets consisting of a Citizen Schools ELT school and its matched comparison schools) instead of at the school level.
- The cluster-robust variance estimator was used to account for clustering; however this method is known to introduce bias when there are a small number of clusters (Angrist & Pischke, 2008). Running the models using clustering at the group level diminishes the number of clusters. Therefore, the percentile-t bootstrap method was used to obtain p-values, as simulation studies by Cameron, Gelbach, and Miller (2008) provide evidence of the greater accuracy of this method.

- Although using the cluster-robust variance estimator when schools are the unit of clustering should provide a sufficient number of schools to cluster at this level, the team also calculated percentile-t bootstrapped p-values for the main models as guidance varies regarding how many clusters are considered sufficient to use the cluster-robust variance estimator (Cameron & Miller, 2015).

Results of these alternatively specified models do not change the conclusions drawn from the confirmatory analytic model, which suggests that the results are robust.

4.2.3 Exploratory Analyses of the Impact of Citizen Schools ELT on Student Achievement by Implementation Year, Grade, Dosage, and Implementation Level

The exploratory analyses examine questions about whether and how the impact of Citizen Schools ELT might vary according to implementation year, grade level, dosage, and implementation level.[35] Because these subgroup analyses divide the sample into smaller groups of schools, and statistical power hinges on the number of schools in a subgroup, they generally have less power than the confirmatory analyses. Additionally, because the number of schools in subgroups could vary, power was not equal across all analyses, and this, in turn, could lead to impact estimates that appear large, but are severely underpowered. Appendix Exhibit C.4.3 includes the standard errors, p-value, and 95% confidence intervals for all estimated impacts, as well as the bootstrapped p-value for each impact, and the text below notes any differences in results based on variance estimation strategy. As described above, using the percentile-t bootstrap method of calculating p-values may be more accurate when the number of clusters is small, as is the case for certain subgroup analyses presented below.

Impact of Citizen Schools ELT on Student Achievement by Implementation Year

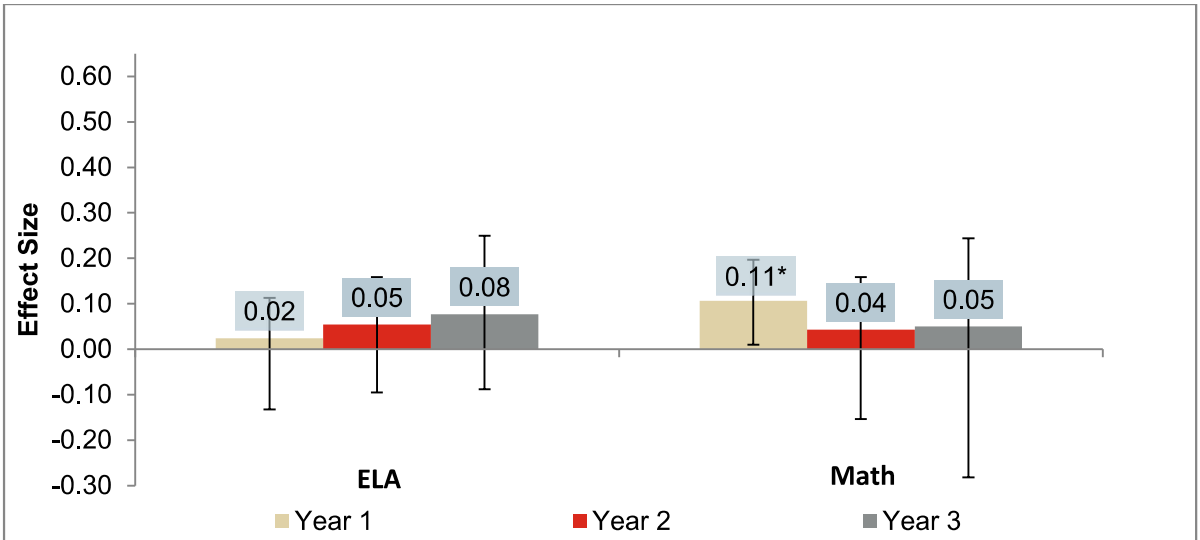
Given that schools and the Citizen Schools network can mature over time, Exhibit 4.7 presents the estimated impact of ELT on student achievement scores by implementation year. These analyses seek to understand whether the length of a school's Citizen Schools ELT implementation (one, two, or three years) is associated with student outcomes. Note that this is different than examining impacts by program exposure or dosage, which will be explored in a subsequent section. Tan bars correspond to impacts after the first year of implementation, red bars correspond to the impact after two years of implementation, and gray bars correspond to the impact after three years of implementation. As such, the Year 1 impacts include results from all cohorts, Year 2 impacts only include Cohorts 1, 2, and 3, and Year 3 impacts include Cohorts 1 and 2.[36]

---

[35] The exploratory analyses presented below (similar to confirmatory analyses described above) compare data from Citizen Schools ELT and matched comparison schools prior to and after the implementation of ELT, use school-level covariates and school, year, and grade fixed effects to control for plausible alternative explanations for the estimated impacts, and account for clustering at the school-level.

[36] Due to small sample sizes, results from Year 4 (with only two Citizen Schools ELT schools) are not shown in the charts (see Appendix C.4.3).

Exhibit 4.7. Impact of Citizen Schools ELT on ELA and Math Test Scores, by Implementation Year



	Year 1	Year 2	Year 3
Number of CS ELT Schools	27	15	ELA: 10 Math: 9

EXHIBIT READS: Students in schools that have been implementing Citizen Schools ELT for one year, on average, scored 0.02 standard deviations higher on their state ELA test than would have been expected in the absence of Citizen Schools ELT. The estimated impact was not statistically significant. The upper confidence limit of the impact is 0.11, the lower confidence limit is -0.13.

Notes:

\+ marginally significant at p < .10 level
* statistically significant at p < .05 level
** statistically significant at p < .01 level
*** statistically significant at p < .001 level

Source: Publically available school-level achievement data from states with schools in the Citizen Schools ELT Network.

Sample: Standardized ELA test scores from 27 schools participating in Citizen Schools ELT for one year and their 95 matched comparison schools, 15 ELT schools participating in Citizen Schools ELT for two years and their 45 matched comparison schools, and 10 ELT schools participating in Citizen Schools ELT for three years and their 28 matched comparison schools. Standardized Math test scores from 27 schools participating in Citizen Schools ELT for one year and their 87 matched comparison schools, 15 ELT schools participating in Citizen Schools ELT for two years and their 44 matched comparison schools, and 9 ELT schools participating in Citizen Schools ELT for three years and their 22 matched comparison schools.

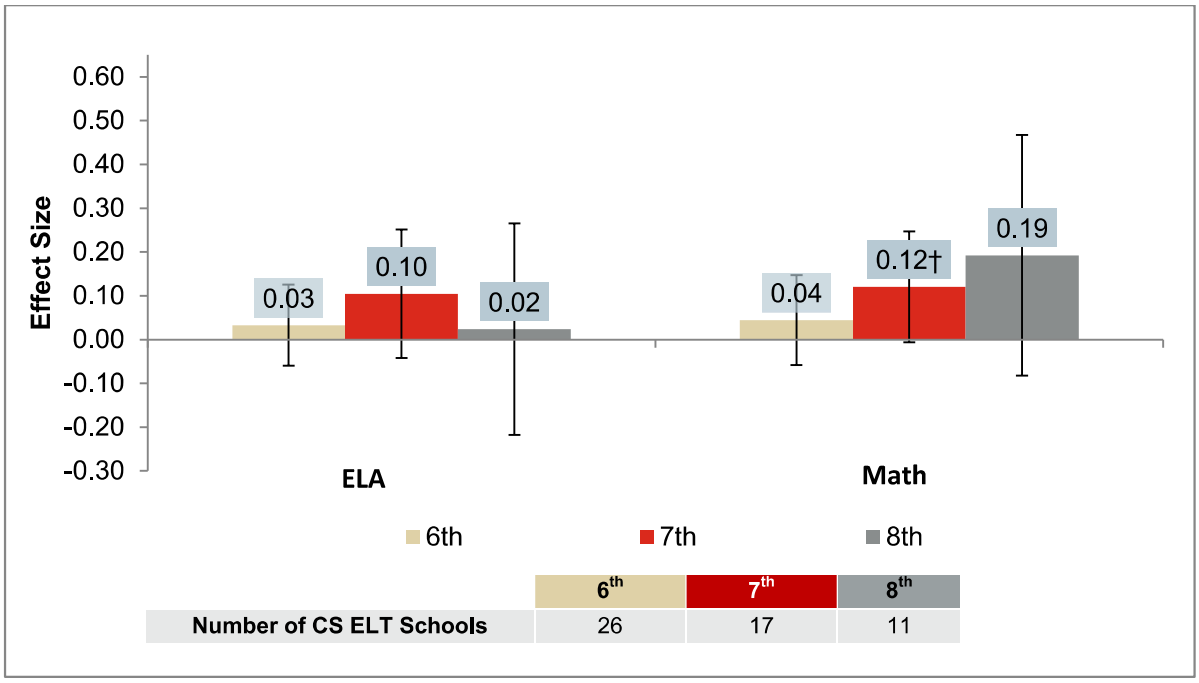
The analysis suggests that impacts of Citizen Schools ELT on reading scores did not vary as a function of implementation year; however, students in Citizen Schools ELT schools in their first year of implementation score, on average, scored 0.11 standard deviations higher on their state math achievement tests than would be expected in the absence of ELT. This effect size is roughly equivalent to just over three months of growth in math.

Impact of Citizen Schools ELT on Student Achievement by Grade

Exhibit 4.8 presents the estimated impact of Citizen Schools ELT on student achievement scores by grade. Tan bars correspond to 6th grade scores, red bars correspond to 7th grade scores, and gray bars

correspond to 8th grade scores. The analysis suggests that Citizen Schools ELT had a marginally significant impact on math test scores for 7th graders, an effect roughly equivalent to about three-and-a -half months of math growth in a school year.[37] Effect sizes for the other grades/subjects were not statistically significant, although they were considerable in size; .10 standard deviations higher for the 7th grade state ELA achievement test and .19 standard deviations higher for the 8th grade state math tests, respectively, than would be expected in the absence of ELT.

Exhibit 4.8. Impact of Citizen Schools ELT on ELA and Math Test Scores, by Grade



	6th	7th	8th
Number of CS ELT Schools	26	17	11

READS: Sixth grade students in Citizen Schools ELT schools, on average, scored 0.03 standard deviations higher on their state ELA test than would have been expected in the absence of ELT. The estimated impact was not statistically significant. The upper confidence limit of the impact is 0.12, the lower confidence limit is -0.06.

Notes:

\+ marginally significant at p < .10 level
* statistically significant at p < .05 level
** statistically significant at p < .01 level
*** statistically significant at p < .001 level

Source: Publically available school-level achievement data from states with schools in the Citizen Schools ELT Network.

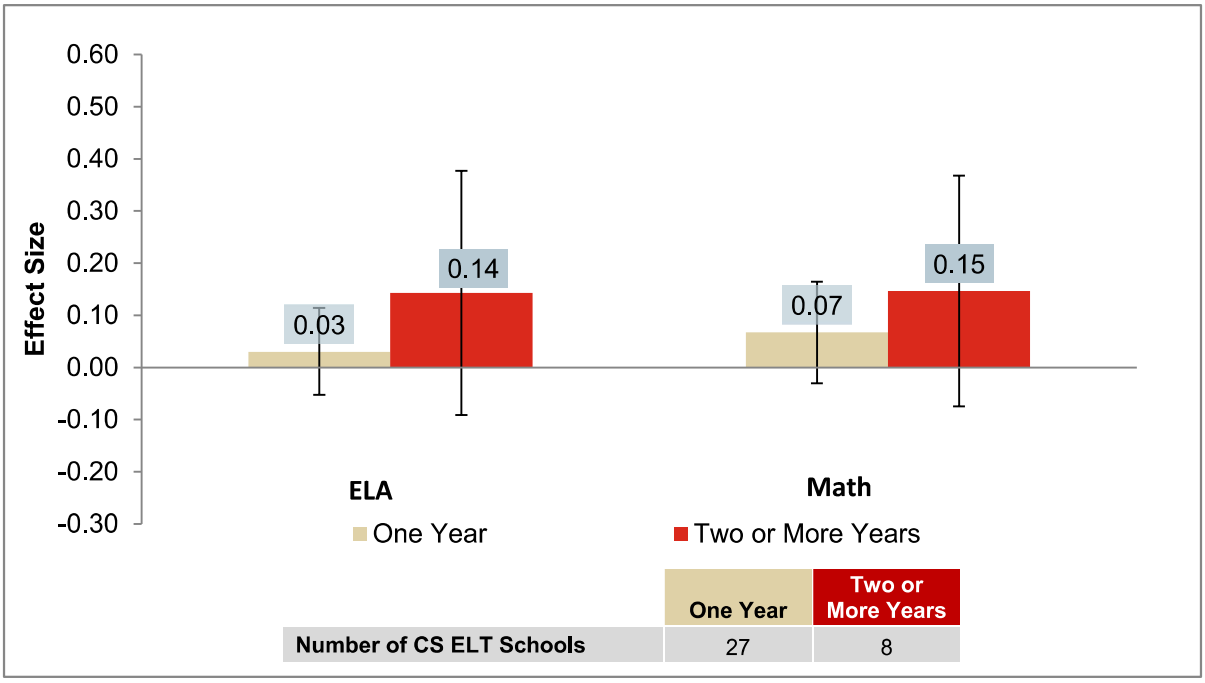
Sample: Standardized ELA test scores from 26 ELT schools serving 6th graders, 17 ELT schools serving 7th graders, and 11 ELT schools searching 8th graders; standardized ELA test scores from 76 matched comparison schools with 6th graders, 56 matched comparison schools with 7th graders, and 37 matched comparison schools with 8th graders. Standardized math test scores from 26 ELT schools serving 6th graders, 17 ELT schools serving 7th graders, and 11 ELT schools searching 8th graders; standardized ELA test scores from 75 matched comparison schools with 6th graders, 49 matched comparison schools with 7th graders, and 33 matched comparison schools with 8th graders.

[37] Given that this is a marginally significant effect, it is particularly sensitive to any changes in estimation strategy. Calculating p-values using the percentile-t method (Appendix Exhibit C.4.3) shifts the p-value over the marginally significant threshold.

Impact of Citizen Schools ELT on Student Achievement by Dosage

Schools that offer Citizen Schools ELT programming to more than one grade can serve students over multiple years, which may result in cumulative achievement impacts. This possibility is explored in Exhibit 4.9, which indicates that across both outcomes, although the impact estimates were not statistically significant, effects sizes were larger for students who had the potential to have two or more years of Citizen Schools ELT exposure, compared to those with the potential for a single year of dosage.[38]

Exhibit 4.9. Impact of Citizen Schools ELT on ELA and Math Test Scores, by Dosage



	One Year	Two or More Years
Number of CS ELT Schools	27	8

EXHIBIT READS: After one year of exposure to the Citizen Schools ELT program, students in Citizen Schools ELT schools, on average, scored 0.03 standard deviations higher on their state ELA test than would have been expected in the absence of ELT. The estimated impact was not statistically significant. The upper confidence limit of the impact is 0.11, the lower confidence limit is -0.05.

Notes:

\+ marginally significant at p < .10 level
* statistically significant at p < .05 level
** statistically significant at p < .01 level
*** statistically significant at p < .001 level

Source: Publically available school-level achievement data from states with schools in the Citizen Schools ELT Network.

Sample: Standardized ELA test scores from students in 27 ELT schools who received one year of Citizen Schools ELT and their 94 matched comparison schools; standardized ELA test scores from students in 8 ELT schools who received two years or more years of Citizen Schools ELT and their 27 matched comparison schools. Standardized math test scores from students in 27 ELT schools who received one year of Citizen

[38] Because analyses were conducted using school-level data, we are assuming that most of the students that are present in one grade remain in subsequent grades. Mobility data were not available consistently across states to test this assumption or adjust impacts for mobility rates.

Schools ELT and their 87 matched comparison schools; standardized math test scores from students in 8 ELT schools who received up to two years of Citizen Schools ELT and their 23 matched comparison schools.

Impact of Citizen Schools ELT on Student Achievement by Implementation Level

As discussed in Chapter 3, schools implement the Citizen School model with varying degrees of fidelity to the intended program. The study also explored whether this variability affects student outcomes by using the implementation index described in the previous chapter to identify "higher" and "lower" implementing Citizen School campuses.[39] Unlike the bar charts presented earlier in the chapter, the bars in Exhibit 4.10 below represent the *difference* in impact between higher and lower implementing schools. Although not statistically significant, students in higher implementing schools scored .07 and .05 standard deviations higher on their ELA and math achievement tests respectively than lower implementing schools.

Exhibit 4.10. Difference in Impact of Citizen Schools ELT on ELA and Math Test Scores of Higher and Lower Implementing Schools

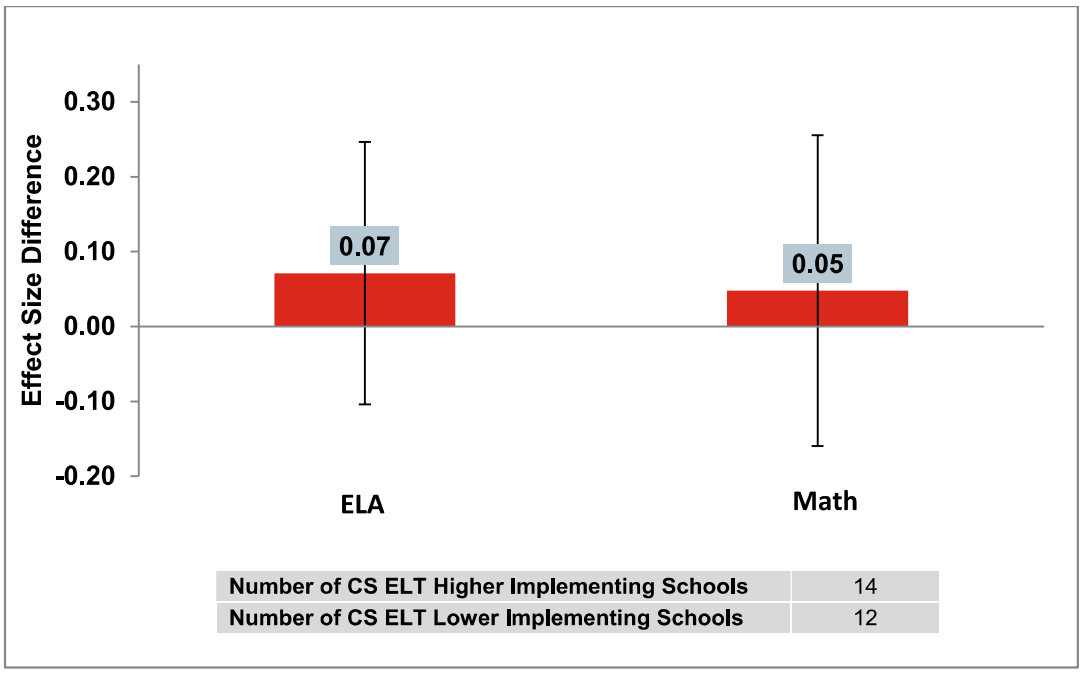


EXHIBIT READS: The impact on students in higher implementing Citizen Schools ELT schools on their state ELA test is, on average, 0.07 standard deviations higher than the impact of students in lower implementing Citizen Schools ELT schools. The difference between the high and low implementing group was not statistically significant. The upper confidence limit of the difference of the impact is 0.25, the lower confidence limit is -0.10.

Notes:

[39] For the purposes of these analyses, the study team relied solely on interview data, which were collected in schools' first, second, third, and fourth years of implementation (where applicable), whereas survey data were obtained only in schools' second, third and fourth years. Schools were categorized as high or low implementing schools at the median value of the implementation index total score. Sensitivity analyses assessed whether "high" or "low" designation would change if the index included both interview and survey/quantitative data, and determined that the interview data alone did not yield materially different implementation level designations than the combined index scores.

+ marginally significant at p < .10 level
* statistically significant at p < .05 level
** statistically significant at p < .01 level
*** statistically significant at p < .001 level

Source: Publically available school-level achievement data from states with schools in the Citizen Schools ELT Network.

Sample: Citizen Schools ELT schools with qualitative implementation index data are included in the analysis. Standardized ELA test scores from students in 12 lower implementing schools and their 43 matched comparison schools; standardized ELA test scores from students in 14 higher implementing schools and their 46 matched comparison schools. Standardized math test scores from students in 12 lower implementing schools and their 41 matched comparison schools; standardized math test scores from students in 14 higher implementing schools and their 40 matched comparison schools. One Citizen Schools ELT school is omitted from the analyses as it did not have implementation index data available.

4.2.4 Selection Bias

As highlighted in Chapter 3, schools included in the Citizen Schools network experience a certain degree of fluidity for many reasons, ranging from issues with implementation to staffing turnover to changes in district priorities. The models presented in this chapter are flexible to these changes, allowing for new cohorts of schools to enter the sample each year, and also for the exit of schools no longer in the network. In particular, after a school leaves the network, the study team ceased to collect that school's data, and the school is included in the analytic sample only for the years it participated in Citizen Schools ELT. As a result, although the models accurately reflect the state of the network at any given time, it is possible that schools that remain in the sample are different on unobservable characteristics that affect their likelihood of staying in (or leaving) the program and their students' achievement scores. If this were the case, results presented above might not be representative of the full sample of schools that have participated in Citizen Schools programming.

The study team addressed this possibility in several ways. From a theoretical perspective, one concern is that the continuing schools were those implementing with the most fidelity, and thus those with the greatest potential for detecting significant impacts. However, of the ten schools that exited the network (and therefore the analytic sample), three were among the highest implementing schools. Further, as indicated in Exhibit 4.10, the study found no significant differences in achievement outcomes between higher and lower implementing schools. Finally, as Chapter 3 highlighted, a number of schools left the network because of funding difficulties, an issue which would not necessarily be hypothesized to be related to achievement outcomes.

To explore selection bias empirically, the study also conducted additional sensitivity analyses to assess whether confirmatory and exploratory outcomes for schools that remained in the network

("stayers") differed from those that exited "leavers").[40] Results indicate that there were no significant differences in impacts (confirmatory or exploratory) between stayers and leavers,[41] suggesting that the reasons for leaving the network were unrelated to achievement outcomes (see Appendix Exhibit C.4.4. See also Appendix Exhibit C.4.5 which indicates that there were no significant differences between stayers and leavers' baseline achievement scores).[42]

4.3 Summary

As the Citizen Schools logic model illustrates, the program is intended to affect short-term outcomes first, through providing students with experiences and presumably a skillset that they can use in and outside of the academic arena, and that may also influence academic achievement more distally. And, indeed, although students in Citizen Schools campuses expressed some dissatisfaction about the more structured nature of this after-school model, they also recognized benefits to social, academic, and behavioral outcomes, and most dramatically, college and career connections. For these students, increasing their self-esteem, empowering them to help their communities and lead their peers, or having a clearer understanding of how to attain their future educational and career goals, are all powerful effects that set the stage for improved outcomes in the longer-term.

If the hypothesized short-term outcomes are achieved, one of the more distal outcomes the Citizen Schools models may impact is academic achievement. To explore this possibility, the study examined the impact of Citizen Schools ELT on ELA and math achievement outcomes. The results were similar to prior quasi-experimental research on the relationship between after-school ELT models and academic achievement that have not found consistently positive, significant impacts of ELT on achievement. The most positive findings to date, from the recent KIPP study (Tuttle et al., 2015) reflect implementation of a highly structured program model rather than the considerably more flexible Citizen Schools ELT model. Exploratory analyses revealed suggestive, although not consistently statistically significant, results. As Chapter 1 points out, some of volatility in results may be attributable to differences in ELT models—while some models increase the length of the school day or year (e.g., KIPP), others, such as Citizen Schools, complement it. Focusing specifically on those programs that complement the existing school day, Redd et al. (2012) concluded that such

---

[40] Another way to explore this idea would have been to keep all schools in the sample regardless of whether they ever left the program, and compare the results to the original model. The study team considered, but ultimately did not pursue this approach. Given that Citizen Schools targets underperforming schools, it is likely that schools that left Citizen Schools took up another intervention. If this was the case, and there were significant impacts using the full sample, the study would not be able to tease out how much of the effect was due to keeping those schools in the sample and how much was due to those schools taking up a different intervention. In fact, an exploration about the nine "exiters" from the analytic sample that had remained open after exiting the Citizen Schools network indicates that most took up another intervention after Citizen Schools, some of which were mandatory for the entire student body.

[41] There was a marginally significant impact of being a "leaver" for the Year 1 impact estimate, but this would contradict a priori hypotheses that those that stay in the network are more likely to have positive achievement outcomes.

[42] There were small, significant differences in some of the baseline demographic characteristics between stayers and leavers (e.g., percent of students with limited English proficiency), but all differences were less than 10 percentages points and are thus unlikely to be substantively meaningful.

programs have not consistently produced positive and sustained impacts on academic achievement, and some programs have produced negative impacts. The review also noted that it is impossible to disentangle positive findings from self-selection in analyses of programs that permit students to opt in (thus, positive impacts may be overstated). Given such differences in the models and foci, it is not surprising that the body of literature is mixed about the relationship between ELT and achievement outcomes. And yet, the Citizen Schools analyses revealed some promising and statistically significant effects in exploratory analyses, particularly for the math outcomes. Further, these impacts were detected despite the fact that, as a mandatory program, self-selection that could upwardly bias estimates is minimized.

5. Conclusion

This final report provides a comprehensive overview of the Citizen Schools ELT model after five years of implementation. Over that time, the Citizen Schools ELT network has matured and expanded as new schools joined each year, and schools exited after one, two, three, or four years. The Citizen Schools ELT program model has become more flexible as partner schools have refined and adapted the program model to meet their needs. Over the past five years, the study has examined both implementation and outcomes, and the prior chapters describe findings about each in detail.

The implementation-focused findings presented in this report center on implementation variability. Overall, the ELT schools are clearly committed to implementing Citizen Schools' ELT programming with fidelity, and at the same time, Citizen Schools has continued to recognize the need for flexibility with its campus partners. As a result, there is considerable variation across individual campuses in how the model is incorporated into their respective school contexts; indeed, that variability seems to be essential for the model to be adaptable across such diverse contexts. The implementation findings also highlight some of the challenges associated with launching a multi-faceted model in dynamic settings, coupled with built-in staffing changes.

The outcome-focused findings described in this report are mixed. The study finds both positive effects on student engagement and aspirations and negative perceptions about students' personal agency and Citizen Schools ELT experiences. The confirmatory findings indicate no overall significant impact on student performance, as measured by standardized achievement test scores in English/Language Arts and math; exploratory findings suggest significant positive impact of Citizen Schools ELT on math achievement in the first year of implementation and a marginally significant positive effect of Citizen Schools ELT on 7th grade math achievement. Statistically significant impacts on student achievement have proven persistently elusive and it is unclear whether this is a function of variability in implementation, model variation across campuses, statistical power, the fit between the intervention and the outcome measures, or some combination of the above.

Prior research on other interventions that include expanded learning time has yielded mixed evidence; few studies found positive achievement impacts, most studies found no achievement impacts, and some found negative effects. Perhaps unsurprisingly, interventions designed to improve academic achievement, such as KIPP, or Higher Achievement, were more likely to affect academic outcomes than were interventions designed to broaden students' enrichment, socio-emotional, or other non-academic experiences. One of the distinctive features of the Citizen Schools ELT model is its emphasis on non-academic learning opportunities that are hypothesized to be necessary precursors to improved achievement. Perhaps student engagement and aspirational short-term outcomes are more appropriate outcomes on which to focus, given Citizen Schools' emphasis on team-building and exposing students to novel, hands-on, real-world experiences through apprenticeships.

Key lessons about implementation, outcomes, and impact are summarized below.

5.1 Implementation

There are clearly many complexities inherent in implementing the Citizen Schools ELT program model, which requires initial site selection and continued planning; communication with varied stakeholders and program rollout; staffing (and hiring, training, and ongoing support activities);

adaptation; and attention to sustainability. The cumulative evidence over five successive school years suggests that full implementation continues to be a moving target—not only because schools are dynamic organizations, but also because the Citizen Schools staffing model is structurally limited to two-year terms, and because funding is typically limited as well. Nonetheless, the ELT schools have demonstrated their capacity to be nimble in the face of challenges and have been quite resilient and creative in how they operate their respective ELT programs. The persistence of local adaptation means that Citizen Schools must balance the trade-offs between adherence to core principles and flexibility in local context.

The overarching implementation findings remained consistent across years. Once Citizen Schools solidified the initial criteria for selecting school partners, those criteria remained relatively stable over time, and program rollout was generally smoother, reflecting timely stakeholder communication and support from school and district leaders. Programming benefitted from context-specific training for Citizen Schools staff and strong relationships between the first and second shifts (e.g., structured overlaps between teaching shifts, shared access to student data, joint participation in grade-level meetings). However, student behavior management, consistent instructional delivery, and staff turnover were persistent challenges for Citizen Schools staff. Schools reliably implemented creative approaches to adapt Citizen Schools' ELT model to meet their needs, and the most prevalent threats to program sustainability continued to be funding and school staff transitions.

The consistency of implementation successes and challenges reported by stakeholders over the years did not translate into systematic patterns in implementation reflected in the implementation index. By contrast, the index findings suggest that continuing schools were consistently at moderate and occasionally high levels of implementation, that struggling schools tended to exit, and that no schools were able to sustain the identical levels of implementation from year to year. Although implementation of some constructs (e.g., planning, leadership, data collection, and perceived program quality) was executed with the highest degree of fidelity, there was generally no evidence of implementation improvement across years.

5.2 Survey Outcomes

In general, survey results were somewhat mixed. On the one hand, a significantly higher percentage of Citizen Schools ELT students reported that Citizen Schools helped their self-esteem and pro-social behaviors than their counterparts, and most Citizen Schools students indicated having positive feelings of belonging in their after-school programming, engagement with peers, and positive relationships with Citizen Schools staff. Similarly, a significantly higher percentage of Citizen Schools students reported participating in activities to learn about colleges and careers during the school day than their peers. On the other hand, students in Citizen Schools ELT programming reported being less satisfied with the amount of autonomy in their respective after-school settings, and reported more negative feelings about Citizen Schools staff and programming than their counterparts; these findings may well reflect the structured nature of the Citizen Schools model relative to the voluntary participation of their peers. At the same time, however, a higher percentage of Citizen Schools ELT students reported the program helped their performance in core academic subjects, students in Citizen Schools' ELT programming felt that Citizen Schools helped them learn about college, careers, and job skills, and the majority of these students aspired to finish college.

These findings highlight a persistent tension between providing middle school students with opportunities to explore and develop interests in novel, substantive areas within a highly structured, longer day and allowing them the autonomy they desire. Despite the inconsistent findings, evidence clearly suggests that Citizen Schools' ELT model is influencing student engagement and aspirations.

5.3 Student Achievement

The student achievement results employ a comparative short interrupted time series (C-SITS) design, one of the strongest quasi-experimental designs possible, and the analyses leverage both the matched comparison group design and the inclusion of fixed effects to account for persistent characteristics of schools and secular changes. The results indicate no statistically significant overall impacts of ELT on students' ELA and math test scores. Exploratory analyses by grade and implementation year indicate a significant positive impact of Citizen Schools ELT on math achievement in the first year of implementation (roughly equivalent to just over three months of growth in math) and a marginally significant positive effect of Citizen Schools ELT on 7th grade math achievement (an effect roughly equivalent to about three and a half months of math growth in a school year).

5.4 Limitations

The study was designed to answer the key impact questions using the strongest possible approaches, given available data and the challenges of obtaining approvals to collect data from participating districts/schools. However, the study's overall analytic approaches have some limitations.

The interview data were collected using a standardized semi-structured protocol which allows for flexibility in responses. Consequently, respondents did not address all questions with equivalent detail. Additionally, answers to interview questions represented a snapshot in time and did not capture respondents' perceptions or experiences over the entire school year. Lastly, the evaluation team was not able to systematically interview staff members with comparable positions at each site. For example, at some sites, the team had the benefit of speaking with numerous stakeholders with different positions and perspectives, providing the benefit of a rich base of information and triangulation of the information. At other sites, evaluation staff were only able to speak with one or two stakeholders, limiting the conclusions the team was able to draw.

The study began survey and interview data collection after the ELT program had been introduced, and while student achievement data could be obtained retrospectively, survey data could only be obtained after schools had begun to implement Citizen Schools ELT. Therefore, any patterns in responses may or may not reflect *changes* due to the introduction of ELT, as the same patterns might have been present in prior years. Data from comparison schools provide useful contemporaneous contextual information that can help to mitigate that limitation, however. The survey data collection has two other limitations: one, limitation is that survey data are all self-reported, and respondents may have recorded socially acceptable rather than accurate responses, although the fact that both positive and negative perceptions were reported suggests that social acceptability may not have played an influential role. Another limitation is that the study surveyed a subset of students in each ELT grade within study schools to minimize burden on schools, and to use limited study resources as efficiently as possible; consequently, the results cannot be assumed to generalize across years or to represent all students within the grades served by ELT.

Two other limitations are worth heeding: one, the study examines Citizen Schools' specific approach to ELT, and two, it uses a purposefully constructed sample of schools. The Citizen Schools ELT model is distinctive, and lessons derived from this study apply to this specific approach to ELT, and further, reflect the idiosyncrasies of the study sample. The study sample has proven to be a fluid one, with considerable fluctuation over the study's five years of data collection. Some fluctuation can be explained by mutual recognition of the "fit" (or lack thereof) between schools and the Citizen Schools ELT model, yet there may be other factors that contribute to schools' decisions to exit the network (e.g., schools' ability to raise the funds necessary to continue partnership with Citizen Schools, new school leadership committed to a different school reform strategy). The sample volatility raises a concern about external validity and the representativeness of the study sample. And although the study can document implementation progress and challenges for all 35 schools for at least their first year of implementation, the program is not currently designed as a one-year or even a two-year intervention. The cumulative results (both implementation and outcomes) are therefore based on a subset of schools and may be less informative.

The student achievement analyses are more rigorous than those used to model survey outcomes both because it was possible to collect baseline data on these outcomes and because they rely upon externally measured assessments, not self-reported measures. Yet performance on ELA and mathematics standardized tests may not be the most sensitive measures of ELT's potential impact on student learning. It may well be the case that achievement outcomes are too distal a measure for Citizen Schools' ELT model. The Citizen Schools ELT model puts significant emphasis on non-academic outcomes such as student motivation, engagement, and aspiration, as well as on better understanding the high school application processes and what comes after secondary school (e.g., college and careers). The academic components of Citizen Schools' model also rely on AmeriCorps members, typically young college graduates who have little formal teaching experience and often have limited prior experience working in urban school environments with diverse student populations. Nonetheless, performance on core academic subject assessments represents the most commonly accepted metric for testing the impact of educational programs on students. Also, despite the fact that the impact analyses use a strong QED, and the statistical models control for many observable factors, there may be unobservable characteristics for which the models do not account.

5.5 Future Steps

This final report summarizes *implementation* data collected over five school years (2010–11 through 2014–15), and *impact* data collected in both baseline years (up to five years before participating schools first launch ELT) and up to four post-ELT years (2010–11 through 2013–14). Over the coming months and years, Citizen Schools may draw from the findings in this report to inform programming and make decisions about future directions. Reflection about the model's ability to effect change in student achievement as measured by standardized test scores over a relatively short time period is likely to be one topic for future consideration. Unlike programs that are focused primarily on academic achievement (e.g., KIPP), Citizen Schools' model delivers a holistic approach to ELT, attending to enrichment and team-building activities more likely aligned with improving students' school engagement and future aspirations. Assessing which outcomes are most appropriate to Citizen Schools' model is likely to be important.

At the conclusion of this five-year evaluation, the Citizen Schools ELT initiative continues to anticipate the needs of its school partners, as new campuses join the network and some current school

partners exit. It is worth noting that Citizen Schools' ELT model has endured some fundamental changes, reflecting a national shift toward the Common Core State Standards in 2013–14 and accompanying assessments (although the future of such standards and assessments is now less certain). Citizen Schools' programming will also need to accommodate upcoming changes in the education landscape resulting from the recently authorized Every Student Succeeds Act (ESSA). While ESSA seems likely to reduce federal reach into local education decisions, it will also provide states and districts more control over teacher evaluation, standards, school turnarounds, and accountability. Given that Citizen Schools has focused chiefly on working with turnaround schools, the new legislation will undoubtedly influence not only how the organization engages new district and school partners, but also how it will adjust its programming and staffing. Although the landscape for instruction and assessment will change as a result of the legislation, hopefully the findings described in this report can provide useful information as Citizen Schools plans for the future.

References

AfterSchool Alliance. (2012). *Principles of effective expanded learning programs: A vision built on the afterschool approach*. Washington, DC: Author.

Angrist, J. D., & Pischke, J.S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Angrist, J. D., Dynarski, S. M., Kane, T. J., Pathak, P. A., & Walters, C. R. (2010). *Who benefits from KIPP?* (NBER Working Paper 15740). Cambridge, MA: National Bureau of Economic Research.

Aronson, J., Zimmerman, J. & Carlos, L. (1999). *Improving student achievement by extending school: Is it just a matter of time.* San Fransisco: WestEd.

Bang, H., & Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics,* 61, 962–972.

Berliner, D. (1990). What's all the fuss about instructional time? In M. Ben-Peretz & R. Bromme, *The nature of time in schools theoretical concepts, practitioner perceptions*. New York: Teachers College Press, 1990.

Bifulco, R. (2012). Can nonrandomized estimates replicate estimates based on random assignment in evaluations of school choice? A within-study comparison. *Journal of Policy Analysis and Management*, 31(3), 729–751.

Black, A. R., Somers, M.-A., Doolittle, F., Unterman, R., and Grossman, J. B. (2009). *The Evaluation of Enhanced Academic Instruction in After-School Programs: Final Report* (NCEE 2009-4077). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, 1, 289–328.

Bryk, A.S., & Schneider, B. (2003). Creating Caring Schools: Trust in Schools: A Core Resource for School Reform. *Educational Leadership* 60(6), 40-45.

Burton, J., Horowitz, R., & Abeles, H. (1999) Learning in and through the arts: Curriculum implications." In Fiske, E. B., ed., *Champions of change: The impact of the arts on learning*. Washington, DC: President's Committee on the Arts and the Humanities; Arts Education Partnership.

Bushaw, W., & Lopez, S. J. (2013). The 45th annual PDK/Gallup Poll of the Public's Attitudes Toward the Public Schools. Available at http://pdkintl.org/noindex/2013_PDKGallup.pdf

Cameron, A. C. and & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2), 317–373.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3), 414–427.

Checkoway, A., Gamse, B., Velez, M., Caven, M., de la Cruz, R., Donoghue, N.,… & Woodford, M. (2012). *Evaluation of the Massachusetts Expanded Learning Time (ELT) Initiative, year five final report: 2010-2011.* Cambridge, MA: Abt Associates.

Cook, T. D., Shadish, W. R., & Wong, V. C. (2008). Three conditions under which experiments and observational studies often produce comparable causal estimates: New findings from within-study comparisons. *Journal of Policy Analysis and Management,* 27(4), 724–750.

Doran, H. C., & Drury, D. W. (2002). *Evaluating success: KIPP educational program evaluation.* Alexandria, VA: New American Schools.

Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of afterschool programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology, 45*(3/4), 294–309.

ECONorthwest. (2008). *A review of research on extended learning time in K–12 schools.* Eugene, OR.

Educational Policy Institute. (2005). *Focus on results: An academic impact analysis of the Knowledge is Power Program (KIPP)*. Virginia Beach, VA: Author.

Farbman, D. (2012, April). *The case for improving and expanding time in school: A review of key research and practice.* Boston, MA: National Center on Time and Learning.

Fortson, K., Verbitsky-Savitz, N., Kopa, E., & Gleason, P. (2015). Horseshoes, hand grenades, and treatment effects? Reassessing whether nonexperimental estimators are biased. *Economics of Education Review,* 44, 100–113.

Fountain, A. R., Velez, M., Gamse, B., Sahni, S., Caven, M., Roy, R., Cropper, P., Dadisman, K., & Lamothe, H. (2013). *Evaluation of Citizens Schools' Expanded Learning Time Model: Year 3 Interim Report.* Cambridge, MA: Abt Associates.

Fryer, R. G., & Dobbie, W. (2011). *Getting beneath the veil of effective schools: Evidence from New York City* . Available at http://www.economics.harvard.edu/faculty/fryer

Gallagher, B. M., & Ross, S. M. (2005). *Analysis of year 2 (2003-2004) student achievement outcomes for the Memphis KIPP DIAMOND Academy*. Memphis, TN: Center for Research in Educational Policy.

Greene, J.P., Kisida, B., & Bowen, D.H. (2014). The educational value of field trips. *Education Next,* 14(1). Available at http://educationnext.org/the-educational-value-of-field-trips/

Hammond, D., Wei, L., Chung, R., Althea, A., Nikole, R., & Stelios, O. (2009). *Professional learning in the learning profession: A Status report on teacher development in the United States and abroad*. Prepared by the National Staff Development Council.

Hetland, L., Winner, E., Veenema, S., & Sheridan, K. (2007). *Studio thinking: The real benefits of visual arts education.* New York: Teachers College Press.

Ho, D., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Policy Analysis,* 15(3), 199–236.

Hoxby, C. M., Murarka, S., & Kang, J. (2009). *How New York City's charter schools affect student achievement* . Stanford, CA: The New York City Charter Schools Evaluation Project.

Kaplan, C., & Chan, R. (2011). *Time well spent: Eight powerful practices of successful, expanded-time schools.* Boston, MA: National Center on Time and Learning .

Kidron, Y, & Lindsay, J. (2014). *The effects of increased learning time on student academic and nonacademic outcomes: Findings from a meta-analytic review.* Prepared by American Institutes for Research for the Institute for Education Sciences, US Department of Education.

Kisiel, C., Blaustein, M., Spinazzola, J., Schmidt, C. Zucker, M., & Kolk, B. (2006). Evaluation of a theater- based youth violence prevention program for elementary school children. *Journal of School Violence*, 5(2), 19–36.

Kolbe, T., Partridge, M., & O'Reilly, F. (2011). *Time and learning in schools: A national profile.* Boston: National Center on Time and Learning and The Center for Eduacation Policy Analysis, University of Connecticut.

Liang, K., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika,* 73(1), 13–22.

MacIver, M. A., & Farley-Ripple, E. (2007). *The Baltimore KIPP Ujima Village Academy*. Baltimore, MD: Center for Social Organization of Schools, Johns Hopkins University.

McCombs, J., Augustine, J., Schwartz, H., Bodilly, S., McInnis, B., Lichter, D., & Cross, A. (2011). *Making summer count: How summer programs can boost children's learning*. Prepared by RAND for the Wallace Foundation.

McDonald, A. J., Ross, S. M., Abney, J., & Zoblotsky, T. (2008). *Urban school reform: Year 4 outcomes for the Knowledge is Power Program in an urban middle school*. Memphis, TN: Center for Research in Educational Policy, University of Memphis.

McMurrer, J., Frizzell, M. Yoshioka, N., Scott, C., and Ostler, N. (2015). *Expanded learning time: A summary of findings from case studies in four states.* Prepared by the Center on Education Policy.

Musher, K. K., Musher, D. M., Graviss, E. A., & Strudler, R. M. (2005). Can an intense educational experience improve performance on objective tests?: Results from one charter school. *The Educational Forum,* 69(4), 352–366.

National Center for Education Statistics. (2013). *The nation's report card: Trends in academic progress 2012* (NCES 2013–456). Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

National Center on Time and Learning (2012). *Mapping the field: A report on expanded-time schools in America*. Boston, MA National Center on Time and Learning.

National Center on Time and Learning (2015). *The Growing Field of Expanded-Time Schools.* Boston, MA National Center on Time and Learning.

Patall, E., Cooper, H., & Allen, A. (2010). Extending the school day or school Year: A systematic review of research (1985-2009). *Review of Educational Research,* 80, 401–436.

Redd, A., Boccanfuso, C., Walker, K., Princiotta, D., Knewstub, D., & Moore, K. (2012). *Expanding time for learning both inside and outside the classroom: A Review of the evidence bases*. Prepared by Child Trends for the Wallace Foundation.

Rocha, E. (2007). *Choosing more time for students: The what, why, and how of expanded learning*. Washington, D.C.: Center for American Progress.

Rosenbaum, P.R. & Rubin, D. B. (1985). Constructing a control group by multivariate matched sampling methods that incorporate the propensity score. *The American Statistician,* 39(1), 33–38.

Ross, S. M., McDonald, A. J., Alberg, M., & McSparrin-Gallagher, B. (2007). Achievement and climate outcomes for the Knowledge is Power Program in an inner-city middle school. *Journal of Education for Students Placed at Risk,* 12(2), 137–165.

Schochet, P. (2008). Statistical power for random assignment evaluations of education programs. *Journal of Educational and Behavioral Statistics*, 33(1), 62–87.

Somers, M., Zhu, P., Jacob, R., & Bloom, H. (2013). *The validity and precision of the comparative interrupted time series design and the difference-in-difference design in educational evaluation.* MDRC Working Paper. New York: MDRC.

St. Clair, T., Cook, T. D. & Hallberg, K. (2014). Examining the internal validity and statistical precision of the comparative interrupted times series design by comparison with a randomized experiment. *American Journal of Evaluation*, 35(3), 311–327.

Steiner, P. M., Cook; T. D., Shadish, W. R., & Clark M. H. (2010). The importance of covariate selection in controlling for selection bias in observational studies. *Psychological Methods*. 15(3), 250–267.

Tuttle, C. C., Teh, B.-R., Nichols-Barrer, I., Gill, B. P., & Gleason, P. (2010). *Student characteristics and achievement in 22 KIPP middle schools*. Washington, DC: Mathematica Policy Research Inc.

Tuttle, C.C., Gill, B., Gleason, P., Knechtel, V., Nichols-Barrer, I., Resch, A. (2013.) KIPP Middle Schools: Impacts on Achievement and Other Outcomes: Final Report. Mathematica Policy Research. Princeton NJ.

Tuttle, C.C., Gleason, P., Knechtel, V., Nichols-Barrer, I., Booker, K., Chojnacki, G., Coen, T., & Goble, L. (2015). *Understanding the effect of KIPP as it scales: Volume I, impacts on achievement and other outcomes.* Washington, DC: Mathematica Policy Research Inc.

U.S. Department of Education. (2012a, March 1). *Guidance on fiscal year 2010 school improvement grants under section 1003(g) of the Elementary and Secondary Education Act of 1965.* Available at http://www2.ed.gov/programs/sif/legislation.html.

U.S. Department of Education, What Works Clearinghouse. (2011). *Procedures and standards handbook (version 2.1)*. Washington, DC: US Department of Education, Institute for Education Sciences. Available at http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=19 (accessed September 30, 2013).

Wheeler, P. (1987). The relationship between grade six test scores and the length of the school day. *Educational Research Quarterly,* 11(3), 10.

White, H. (1984). *Asymptotic theory for econometricians*. Orlando, FL: Academic Press.

Woodworth, K. R., David, J. L., Guha, R., Wang, H., & Lopez-Torkos, A. (2008). *San Francisco Bay Area KIPP schools: A study of early implementation and achievement final report*. Menlo Park, CA: SRI International.

Yoon, K., Duncan, T., Lee, S., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement.* Prepared by REL Southwest for the Institute of Education Sciences, US Department of Education.

Zief, S. G., Lauver, S., & Maynard, R. (2006). The impacts of afterschool programs on student outcomes. *Campbell Systematic Reviews,* 3.