



---

# **Chicago Public Schools Community Schools Initiative Evaluation**

## **2011–12 Impact Report**

**Neil Naftzger**

**Ryan Williams, Ph.D.**

**Feng Liu, Ph.D.**

American Institutes for Research

November 2014



# **Chicago Public Schools Community Schools Initiative Evaluation: 2011–12 Impact Report**

**November 2014**

**Neil Naftzger  
Ryan Williams, Ph.D.  
Feng Liu, Ph.D.**



**AIR<sup>®</sup>**

AMERICAN INSTITUTES FOR RESEARCH<sup>®</sup>

1120 East Diehl Road, Suite 200  
Naperville, IL 60563-1486  
800-356-2735 • 630-649-6500

**[www.air.org](http://www.air.org)**

Copyright © 2014 American Institutes for Research. All rights reserved.

3907\_11/14





# Contents

	<b>Page</b>
Executive Summary .....	1
Introduction.....	6
CSI Implementation Framework.....	8
Measuring CSI Implementation .....	12
Refining the List of Implementation Elements.....	16
Using Attendance Metrics to Explore the Quality of Implementation .....	18
Assigning Schools to Higher and Lower Implementing Groups .....	20
Assessing the Impact of Participation in a Higher Implementing Community School .....	24
Youth Outcomes Examined .....	24
Analytic Approach .....	26
Results—Higher Implementing Versus Lower Implementing Schools.....	27
Summary of Findings—Higher Implementing Versus Lower Implementing Schools .....	33
Results—Higher Implementing Versus Non-CSI Schools.....	34
Summary of Findings—Higher Implementing Versus Non-CSI Schools .....	40
Conclusion and Recommendations.....	42
References.....	45



## Executive Summary

This report outlines key findings from an impact evaluation completed by American Institutes for Research (AIR) of the Community Schools Initiative (CSI) administered by Chicago Public Schools (CPS). Two primary steps were taken in undertaking the impact evaluation:

1. Steps were taken to classify a sample of community schools providing programming and services during the 2011–12 school year into higher and lower implementing groups based on the criteria outlined in the CSI Implementation Framework.
2. Analyses were then conducted to assess if youth enrolled in programming in higher implementing schools demonstrated better functioning on a variety of academic, behavioral, and learning experience outcomes relative to:
  - Youth enrolled in programming in lower implementing schools
  - Similar youth attending non-CSI schools

It was expected that the effects of community schooling would be tied to the quality with which the school has implemented practices and approaches described by the CSI Implementation Framework, with schools classified as higher implementing demonstrating greater impact on the domain of youth outcomes examined.

### **Classifying Schools and Practices Related to the Value of Leveraged Partnerships**

A variety of measures were employed to assign a sample of 33 schools to either the higher or lower implementing group:

- *Measures obtained from the 2010 leveraged partnership survey.* Leveraged partnerships represent new providers that the resource coordinator recruits to contribute to the community schooling effort based on results from the needs and resources assessment to provide additional, in-kind services to the school. The goal is that leveraged partnerships result in substantive new resources flowing into the school to meet the needs of enrolled students and their families, which go beyond the funds provided to the lead partner agency to run 21st Century Community Learning Centers (21st CCLC)–funded afterschool and summer learning programs. Two measures from the leveraged partnership survey were used in classification efforts:
  - Value of programming provided through leveraged partnerships
  - Value of other services and supports provided through leveraged partnerships
- *Measures obtained from YouthServices.net for program attendance during the 2011–12 school year.* All schools enrolled in CSI are required to track youth participation in programming, including what activities youth were enrolled in and whether or not they attended that activity on a given day that it was provided. Data on program attendance are key to both exploring how afterschool and summer learning activities impact youth and the relative quality of programming. The following attendance-related measures derived from YouthServices.net were used to support the classification process:

- Percentage of youth participating in two consecutive years of programming (2010–11 and 2011–12)
  - Mean number of minutes of programming attended during the 2011–12 school year
  - Mean duration of participation (in days) during the 2011–12 school year
  - Percentage of school population enrolled in programming during the 2011–12 school year
- *Measures of CSI Framework Implementation derived from the 2012 key stakeholder surveys.* During the course of 2011 and into the early part of 2012, the evaluation team set about the task of constructing a series of surveys that were oriented at measuring how well a sample of community schools were implementing the criteria represented in the CSI Implementation Framework. The goal of this approach was to obtain perceptions from those key school staff and stakeholders who are critical to implementation of the community school strategy at a given school. Survey elements found to be related to the value of programming derived from leveraged partnerships were combined into a single implementation index on a 0 to 100 scale, with higher values indicative of greater implementation of these practices. Practices referenced by how they appear in the CSI Self-Assessment Quality Improvement Rubric especially correlated with the value of leveraged partnerships are outlined as follows:
- Element 1A: Shared Goals and Expectations
  - Element 1B: Needs and Resources Assessment
  - Element 1C: Resource Coordinator Received Enough Job Training
  - Element 1C: Resource Coordinator Receives Support in Performing His or Her Job
  - Element 1D: Advisory Committee Role in Decision Making
  - Element 2A–B: Adoption of Practices Supporting Youth Development
  - Element 2E: Absence of Staffing Challenges
  - Element 3C–D: Number of Professional Development Events Focused on Adoption of the Community School Framework
  - Element 3C–D: Presence of Professional Learning Community

Efforts to classify CSI schools into higher and lower implementing schools resulted in 13 schools being assigned to each group.

## **Youth Outcomes Examined and Analytic Approach**

A variety of academic, behavioral, and learning experience outcomes were examined in order to assess how enrollment in programming provided at higher implementing schools was impacting youth outcomes. All outcomes examined pertained to the 2011–12 school year. Given that the domain of outcomes varied somewhat by grade level, separate analyses were done for youth in Grades K–3, 4–8, and 9–12. Behavioral outcomes like school-day absences, school-day misconducts, and school-day suspensions were examined for all grade levels. In addition, for students in Grades 4–8, Illinois Standards Achievement Test (ISAT) reading and mathematics

scores were examined, and for students in Grades 6–8, various scales obtained from the *My Voice, My School* survey were also examined pertaining to both youth experiences in school and their general well-being. For youth in Grades 9–12, academic outcomes like cumulative weighted grade point average (GPA) and the number of credits earned during the year were taken into consideration, as well as the same domain of scales from the *My Voice, My School* survey utilized in conducting impact analyses for youth in Grades 6–8. Some analyses at this grade level also involved ACT and Prairie State Achievement Examination (PSAE) assessment scores.

In any evaluation of a program where participants are not randomly assigned to treatment or control groups, the problem of selection is paramount. Youth from higher implementing schools may have differed from those from lower implementing schools or schools not enrolled in CSI in a variety of ways. These differences could bias estimates of impact. If we were simply to compare youth from higher implementing schools to those from lower implementing schools or schools not enrolled in the initiative, we would not be able to disentangle the effect of the program implementation level from the preexisting difference between these two groups of youth.

To address these potential confounders of estimating the effect of attending a higher implementing school, the research team employed a propensity score stratification approach. Propensity score stratification is a statistical method that allows researchers to more closely estimate the causal effect of interest by creating a comparison group that looks like the treatment group on all observable characteristics. Propensity score models were run when assessing the impact of being enrolled in programming at higher implementing schools for all youth outcomes, except for student in Grades 9–12 when comparing outcomes for higher and lower implementing schools. In this case, the student populations were too different for us to employ a propensity score approach. Instead, multilevel modeling approaches were employed where prior achievement was controlled for in undertaking the analyses.

## **Results**

### **Higher Versus Lower Implementing Community Schools**

Even though a wide variety of outcomes were examined, almost all findings demonstrated no significant relationship between participating in higher implementing schools and better functioning on the youth examined relative to lower implementing schools. The only finding where the hypothesized relationship was found was in relation to youth in Grades 9–12 who demonstrated significantly fewer unexcused absences than their peers in lower implementing programs. Although this effect would be deemed large (an effect size of  $-.79$  or 59 percent fewer absences), these analyses were the least robust conducted in the report given an inability to employ a propensity score matching approach as noted in the previous section.

### **Higher Implementing Community Schools Versus Non-CSI Schools**

Analyses assessing the effect of higher implementing community schools relative to similar youth enrolled in non-CSI school were more positive.

Youth in Grades K–3 enrolled in higher implementing community schools had:

- Fewer misconducts relative to similar youth in non-CSI schools (moderate effect, -.31 or 55 percent fewer misconducts)
- Fewer suspensions relative to similar youth in non-CSI schools (moderate effect, -.29 or 53 percent fewer suspensions)

Youth in Grades 4–8 enrolled in higher implementing community schools had:

- Higher scores on the Emotional Health scale of the *My Voice, My School* survey relative to similar youth in non-CSI schools (moderate effect, .43)

Youth in Grades 9–12 enrolled in higher implementing community schools had:

- Fewer unexcused absences relative to similar youth in non-CSI schools (large effect, -.68 or 61 percent fewer unexcused absences)
- Higher scores on a series of *My Voice, My School* survey scales relative to similar youth in non-CSI schools:
  - Psychological Sense of School Membership (moderate effect, .42)
  - Inquiry-Based Science Instruction (moderate effect, .41)
  - Student-Teacher Trust (small effect, .25)
  - Academic Personalism (small effect, .24)

Many of the findings related to youth outcomes suggest the benefits associated with participating in a higher implementing community school (based on how implementation has been measured to date) are more likely to be witnessed on youth behaviors (like absences, misconducts, and suspensions) and experiences in school as opposed to substantial effects on academic achievement, at least in the one to two year period of participation in CSI programming examined in this report. Increasingly, however, attention is being paid to the role that youth behavior, school bonding, and social and emotional functioning play in ensuring youth success, both in school and in life more generally, so the positive effects noted here are important and relevant.

It is important to note that the analyses conducted in this study were limited to a relatively small number of students who attended higher implementing schools in the 2011–12 academic year. This small sample limits the statistical power of the study to detect effects. As in any study where random assignment to treatment and control conditions is not feasible, it is possible that there were unobservable differences between students who attended higher implementing schools and comparison schools that we were unable to control for. This is especially true where propensity score matching was not possible. To the extent that unobserved differences are related to student outcomes, these differences would bias the estimated differences between higher implementing schools and schools in the comparison groups. To that end, these analyses provide initial evidence about the impact of attending higher implementing afterschool programs but should not necessarily be considered equivalent to experimental studies that have stronger internal validity.

Finally, when key stakeholders associated with the initiative articulate what they have witnessed in terms of how community schools have a positive impact on the lives of enrolled youth and their families, what is often referenced are a series of youth development outcomes (e.g., self-efficacy, self-management and regulation, social skills, positive mind-sets, etc.) that are not represented in the school-related youth outcome data or the school climate data explored in this report. There is a substantial need to better measure and document how CSI is impacting these important components of youth development. Only when this is done well and carefully will the full impact of the Community Schools Initiative in Chicago truly be understood.

## Introduction

Since 2002, Chicago Public Schools (CPS) has looked to community schooling as a strategy to support students, their families, and the broader school community. Through the creation of partnerships between district schools and community-based organizations and providers, community schooling has resulted in additional programming and services being provided in CPS schools. A key component of the CPS Community Schools Initiative (CSI) is ensuring that the needs of the school community are identified and that high-quality programming and services are provided that address these needs, particularly the academic, social, and emotional needs of students enrolled in initiative schools. It is expected that the investment that CPS has made in community schools will result in a wide variety of possible positive outcomes for participating youth, including improvements in academic achievement, the development of behaviors important to school success, and better health and well-being among enrolled youth.

Past efforts to evaluate the impact of programming and services provided at CSI schools on student outcomes have demonstrated mixed results (McCrary & Hoffman, 2010; Whalen, 2007; Whalen, Fujimoto, & Xiong, 2008). The most recent evaluation of student outcomes was conducted by WestEd in which the impact of the initiative was examined for youth that participated in programming during the 2007–08 school year (McCrary & Hoffman, 2010). Key findings from this evaluation suggested that the initiative had no meaningful impact on reading and mathematics achievement and mixed results in terms of having a positive impact on student attitudes toward school. However, the 2007–08 evaluation did demonstrate that program participation did lead to reduced school-day absences across a wide array of student groups.

In light of mixed results from the 2007–08 evaluation, the internal CSI evaluation team set about the task of identifying those characteristics of implementation that seemed critical for a viable community schooling effort to take root in a given school. After extensive key informant interviews in a sample of schools, the CSI Implementation Framework was created (Zander, Burnside, & Poff, 2010). The framework articulates the core features of community schooling in Chicago and the steps and processes that schools need to go through and adopt in order to fully implement the strategy.

Starting in 2011, CPS began working with its internal evaluation team, American Institutes for Research (AIR) and Diehl Evaluation and Consulting Services, Inc., to utilize the CSI Implementation Framework as the platform for designing a local evaluation effort oriented at achieving two primary goals:

1. Assess the relationship between how well a community school has implemented elements represented in the CSI Implementation Framework and the types of youth outcomes expected from schools enrolled in the initiative.
2. The development of a quality improvement tool and process predicated on the CSI Implementation Framework that schools enrolled in the initiative can use to promote strategies for successful community schooling.

The underlying premise driving these goals is the expectation that the effects of community schooling on student outcomes will be tied to the quality with which a school has implemented practices and approaches described by the CSI Implementation Framework.

The purpose of this report is to outline the findings from a series of analyses conducted by AIR:

1. To classify a sample of community schools providing programming and services during the 2011–12 school year into higher and lower implementing groups based on the elements for effective implementation codified in the CSI Implementation Framework.
2. Assess whether youth enrolled in programming in higher implementing schools demonstrated better functioning on a variety of academic, behavioral, and learning experience outcomes relative to (a) youth enrolled in programming in lower implementing schools and (b) similar youth attending non-CSI schools.

In the sections that follow, steps will first be taken to further describe the CSI Implementation Framework and how steps were taken by the evaluation team to use various measures and indicators to sort a sample of community schools into higher and lower implementing groups based on framework elements. Next, the analytic approach used to assess the impact of higher implementation of framework elements will be described, results from these analyses will be provided and summarized, and conclusions will be outlined relative to what these results mean for the initiative moving ahead.

## CSI Implementation Framework

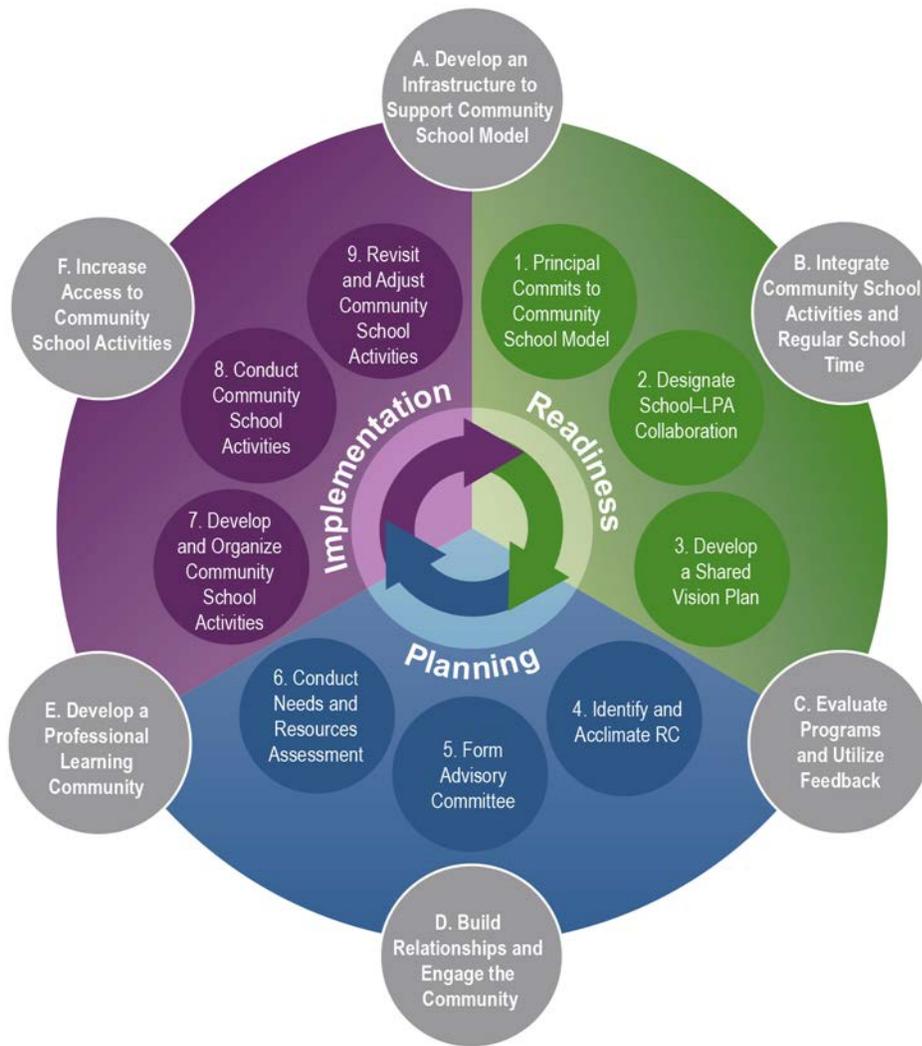
The CSI Implementation Framework is organized around two key processes:

1. The steps a new community school needs to take in order to implement the strategy
2. The factors an established community school needs to be attentive to in order to preserve the health and vitality of the community school strategy as implemented in their school

As shown in Figure 1, the framework includes nine steps related to implementation, with these steps addressing *readiness* (principal commitment to the community school, a designated partnership between the school and a lead partner agency (LPA), and the development of a shared vision); *preparation* (hiring and preparing a resource coordinator (RC) who manages the community school strategy, forming an advisory committee, and conducting a needs and resources assessment); and *implementation* (developing, organizing, and conducting community school activities and revisiting these activities with the goal of program improvement). The framework conveys to new CSI schools the steps they should follow in transforming themselves into a community school and also reminds existing CSI schools about the need to evaluate and adapt programming throughout the lifetime of the community school.

The six sustainability factors are closely related to the implementation steps—emphasizing infrastructure, community relationships, evaluation, professional learning communities (PLCs), program expansion as needed, and continued and close links to the school. Represented in the outer ring in Figure 1, the sustainability factors are essential for maintaining and developing the community school, particularly as certain components—such as funding sources and partnerships—change.

**Figure 1. Community Schools Initiative Implementation Framework**



Starting in 2011, the evaluation team at AIR and Diehl Evaluation worked to convert the CSI Implementation Framework into a viable self-assessment tool by describing what constitutes *exemplary*, *proficient*, *emerging*, and *planning* implementation of each of the framework steps and factors. To make the framework more functional for the community schools enrolled in the initiative, each of the steps and factors outlined in Figure 1 were organized into four domains:

1. **Domain 1.** Establish and maintain essential structures and resources needed for an effective community school (Integrates: Implementation Steps 1, 2, 3, 4, 5, and 6 and Sustainability Factor A).
2. **Domain 2.** Establish and maintain community school programs and services (Integrates: Implementation Steps 7, and 8 and Sustainability Factors B and F).
3. **Domain 3.** Establish and maintain continuous improvement structures (Integrates: Implementation Step 9 and Sustainability Factors C and E).

4. **Domain 4.** Develop strategies and commit resources to financially and organizationally sustain the community school (Integrates: Implementation Step 2 and Sustainability Factors A and D).

Within each domain, elements were defined to articulate the practices and approaches that needed to be in place in order for there to be effective implementation of a given framework implementation step or sustainability factor. Each of these elements is presented in Table 1 across each of the four primary domains.

**Table 1. Summary of Implementation Framework Elements by Domain**

<b>Domain 1: Establish and maintain essential structures and resources needed for an effective community school.</b>	
<b>ISPS Framework</b>	<b>Rubric Elements</b>
<p><b>Implementation Step 1.</b> Principal commits to community school model.</p> <p><b>Implementation Step 2.</b> Designate school–LPA collaboration.</p> <p><b>Implementation Step 3.</b> Develop a shared vision statement.</p> <p><b>Implementation Step 4.</b> Identify and acclimate resource coordinator.</p> <p><b>Implementation Step 5.</b> Form an advisory committee.</p> <p><b>Implementation Step 6.</b> Conduct needs and resources assessment.</p> <p><b>Sustainability Factor A.</b> Develop an infrastructure to support community schooling.</p>	<p><b>Element 1A.</b> Develop and implement a clear, written, and mutually agreed upon collaborative agreement (formal written agreements, collaborative goal setting, and shared decision-making process) between the school, the LPA, and other partners providing services at the school.</p> <p><b>Element 1B.</b> Obtain, communicate, and use information from a comprehensive needs and resources assessment to guide community school programs and services.</p> <p><b>Element 1C.</b> Hire and orient a qualified person to serve as the community school resource coordinator.</p> <p><b>Element 1D.</b> Establish and maintain an advisory committee with representatives from the school, LPA, community partners, and the community at large.</p> <p><b>Element 1E.</b> Develop, communicate, and execute policies and procedures to support implementation of community school programming.</p> <p><b>Element 1F.</b> Develop and communicate a shared vision among stakeholders for the community schools.</p>
<b>Domain 2: Establish and maintain community school programs and services.</b>	
<b>ISPS Framework</b>	<b>Rubric Elements</b>
<p><b>Implementation Step 7.</b> Develop and organize community school activities.</p> <p><b>Implementation Step 8.</b> Conduct community school activities.</p> <p><b>Sustainability Factor B.</b> Integrate out-of-school time and school time.</p> <p><b>Sustainability Factor F.</b> Recruit participants and increase access to out-of-school time programming.</p>	<p><b>Element 2A.</b> Establish academic programs that are aligned to community school needs and that reflect best practices.</p> <p><b>Element 2B.</b> Establish nonacademic enrichment activities that are aligned to student interests and reflect best practices.</p> <p><b>Element 2C.</b> Establish family and community programs and services that are aligned to community and family needs and resources.</p> <p><b>Element 2D.</b> Develop a strategic schedule that optimizes time and is adequate for the planned activities.</p> <p><b>Element 2E.</b> Establish a structure and procedures for the recruitment, retention, and management of quality out-of-school time staff.</p> <p><b>Element 2F.</b> Identify and use targeted recruitment and enrollment</p>

	strategies for students and families and any special populations to be served.
<b>Domain 3: Establish and maintain continuous improvement structures.</b>	
<b>ISPS Framework</b>	<b>Rubric Elements</b>
<p><b>Implementation Step 9.</b> Revisit and adjust community school activities.</p> <p><b>Sustainability Factor C.</b> Evaluate programs and utilize feedback.</p> <p><b>Sustainability Factor E.</b> Develop a PLC.</p>	<p><b>Element 3A.</b> Identify and collect various sources of information to inform continuous improvement processes associated with implementation of community school programs and services.</p> <p><b>Element 3B.</b> Use various sources of information and involve multiple stakeholders to inform continuous improvement processes associated with implementation of community school programs and services (e.g., mental health services, family programs).</p> <p><b>Element 3C.</b> Provide professional learning opportunities for community school staff and stakeholders.</p> <p><b>Element 3D.</b> Establish a PLC that includes community school program staff and providers and school-day staff.</p>
<b>Domain 4: Develop strategies and commit resources to financially and organizationally sustain the community school.</b>	
<b>ISPS Framework</b>	<b>Rubric Elements</b>
<p><b>Implementation Step 2.</b> Designate school–LPA collaboration (<i>Note:</i> Also included in Domain 1).</p> <p><b>Sustainability Factor A.</b> Develop an infrastructure to support community schooling (<i>Note:</i> Also included in Domain 1).</p> <p><b>Sustainability Factor D.</b> Build relationships and engage stakeholders.</p>	<p><b>Element 4A.</b> Establish and maintain relationships with the LPA and other stakeholders to pursue and allocate funding sources and other resources to establish and sustain community school services.</p> <p><b>Element 4B.</b> Identify and obtain diverse resources to support the community school infrastructure.</p>

*Note:* This table reflects the configuration of elements as of the development of this report. Subsequent reports will likely reflect changes in the elements.

In the self-assessment tool developed by the evaluation team, now known as the CSI Self-Assessment Quality Improvement Rubric, criteria have been developed for each element. The resulting rubrics aim to help a community school determine whether its implementation of a given element is at a certain level: *planning*, *emerging*, *proficient*, or *exemplary*. Guiding questions were developed for each element to help school stakeholders complete the self-assessment and make more meaningful and accurate judgments about how well they are doing in implementing a given element.

The development of the CSI Implementation Framework and aligned CSI Self-Assessment Quality Improvement Rubric represented a watershed moment in the initiative’s history by providing schools with very detailed criteria on what constituted effective implementation of the strategy. With these in place, steps could be taken to explore the relationship between implementation of the strategy and youth outcomes.

## Measuring CSI Implementation

During the course of 2011 and into the early part of 2012, the evaluation team set about the task of constructing a series of surveys that were oriented at measuring how well a sample of community schools were implementing the criteria represented in the CSI Implementation Framework. The goal of this approach was to obtain perceptions from those key school staff and stakeholders who are critical to implementation of the community school strategy at a given school. A total of five different key stakeholder surveys were constructed:

- *Resource coordinator survey.* As the individuals who coordinate the design and delivery of community school activities within a building, resource coordinators play a pivotal role in the implementation of the community school model. To determine how well the framework was adopted by a school, it was imperative to learn how resource coordinators cultivated the partnerships and served as coordinators for activities and services provided to students, their families, and the community at large. The resource coordinator survey also contained scales designed to assess the degree of support the resource coordinator received from school-day staff and administrators as well as their parent organization.
- *Principal survey.* Sustained principal support is a prominent feature of many of the elements identified in the CSI Implementation Framework. The principal survey was designed to assess the degree to which a principal was committed and involved in the implementation of CSI at their school and the degree to which the principal exercised leadership in facilitating the components that are critical to the successful implementation of the community school model. In addition, questions on the principal survey were designed to assess the principal's perceptions regarding how well other key elements of the framework were functioning; these elements include the school's relationship with the LPA and the degree to which the resource coordinator had been integrated into school operations.
- *LPA survey.* Representatives from the LPAs were asked about communication and relationships with school administrators and staff and the LPA role in activities addressed in the framework (i.e., needs assessment, advisory group formulation and convening, and evaluation).
- *School-day teacher survey.* The teacher survey contained scales that measured the extent to which teachers engaged in practices that facilitate implementation of the framework, as well as the extent to which normative factors (e.g., collaboration, collective efficacy and responsibility, and trust) have been established in the building. Like the principal survey, the survey included questions that asked teachers to assess how well framework elements were present in the building.
- *Activity leader survey.* Given the prominence of afterschool and summer programming for the delivery of CSI services and activities, an activity leader survey was developed to collect information on practices that the research defines as effective and that are aligned with the framework. Survey questions also asked activity leaders about their perceptions on how well key implementation factors were present in the school.

The surveys were administered to a sample of 52 community schools that received funding from the 21st Century Community Learning Centers (21st CCLC) program in the spring of 2012. The

schools included in the sample represented all community schools funded by 21st CCLC (the funding source for the evaluation) that had not been involved in efforts to pilot and refine the surveys in early 2012. Given the critical role that the resource coordinator plays in implementing the community school strategy at a given school, it was deemed critical that a complete resource coordinator survey be received for a school in order to be included in analyses oriented at exploring implementation of framework elements. Of the 52 schools included in the original sample, completed resource coordinator surveys were received from 41 schools (79 percent). These 41 schools became the basis for all subsequent analyses related to assessing the implementation of framework elements and for the domain of analyses related to assessing the impact of CSI on youth outcomes detailed in this report. Response rates across each of the surveys associated with these 41 schools by survey type are presented in Table 2.

**Table 2. Summary of Survey Response Rates by Respondent Type**

Survey Type	Number of Surveys Completed	Number of Surveys Administered	Response Rate
Resource coordinator	41	41	100.0%
Principal	36	41	87.8%
Teacher*	558	888	62.8%
Activity leader*	375	502	74.7%
Lead partner manager	32	41	78.0%

\*Some teachers in a given school also worked as activity leaders in delivering out-of-school time programming and were administered questions containing both teacher and activity leader questions. The individuals have been counted in both the teacher and activity leader response categories outlined in the table.

When designing the surveys, an intentional effort was made to ask the same set of questions across different surveys, thereby getting the perspective of different stakeholders on how well the school was doing in implementing a given element represented in the framework. This approach allowed the evaluation team to classify schools into one of three implementation categories for most of the elements represented in the CSI Implementation Framework:

- All respondent types suggested a *higher than average* implementation in relation to a given framework element.
- All respondent types suggested a *lower than average* level of implementation in relation to a given framework element.
- Responses among respondent types indicated a *divergence* in the perceptions of CSI implementation, with some respondents indicating higher than average implementation, while others indicated lower than average implementation.

Both hierarchical cluster analysis and other descriptive techniques were utilized to assign schools to one of these three groups.<sup>1</sup> Given the goal of exploring the impact of higher implementing schools on youth outcomes, the evaluation team was particularly interested in schools categorized in the group where there was agreement among all respondent types of *higher than average* implementation on a particular framework element. The percentage of schools falling in this category by framework element examined by means of the surveys can be found in Table 3.

As shown in Table 3, the percentage of schools in the sample classified as following in the *higher than average* group varied substantially from one framework element to another. For example, while 61 percent of the schools represented in the sample fell in the *higher than average* group in terms of *resources available to support community school efforts* (Element 2A–B), only 14 percent fell in this group in terms of the *presence of a professional learning community* (Element 3C–D).

No elements pertaining to *Domain 4: Develop strategies and commit resources to financially and organizationally sustain the community school* were examined by means of the surveys. In addition, as shown in Table 3, the number of schools with responses that were included in classification analyses varied considerably as well, ranging from 11 to 41. This variation was related to (a) the use of cluster analysis techniques to classify programs, which requires complete data sets across all variables used in the analysis (i.e., schools missing survey data from a stakeholder group on the element being examined were dropped from the analysis) or (b) survey questions that were asked only when a previous question was answered in the affirmative. For example, questions related to the functioning of the advisory committee (Element 1D) would be asked only if the school actually had an advisory committee in place (Element 1B).

**Table 3. Percentage of Sampled Schools in Higher Than Average Group by Framework Element**

CSI Implementation Framework Domain and Element	Number in Higher Than Average Group/Total Schools With Responses	Percentage in Higher Than Average Group
<b>Domain 1: Establish and maintain essential structures and resources needed for an effective community school.</b>		
Element 1A: Shared goals and expectations	13/28	46.4%
Element 1A: Coordination between the school and LPA	12/28	42.9%
Element 1A: LPA engagement	11/35	31.4%
Element 1A: Principal engagement	9/27	33.3%

<sup>1</sup> For a fuller description of how survey data were analyzed to assign schools to different implementation categories, please see Norbury, H., Naftzger, N., Sparr, M., Crecelius, S., & Diehl, D. (2013). *Chicago Public School Community Schools Initiative Evaluation: 2011–12 Implementation Report*. Chicago, IL: American Institutes for Research.

<b>CSI Implementation Framework Domain and Element</b>	<b>Number in Higher Than Average Group/Total Schools With Responses</b>	<b>Percentage in Higher Than Average Group</b>
Element 1B: Needs and resources assessment	3/22	13.6%
Element 1C: RC received enough job training	15/41	36.6%
Element 1C: RC attends school events	8/41	19.5%
Element 1C: RC receives support in performing his or her job	15/41	36.6%
Element 1C: RC Perceived as effective	8/36	22.2%
Element 1D: Advisory committee role in decision making	5/11	45.5%
Element 1D: Advisory committee functions well	4/12	33.3%
<b>Domain 2: Establish and maintain community school programs and services.</b>		
Element 2A–B: Linkages to the school day and adoption of practices supportive of academic skill building	7/37	18.9%
Element 2A–B: Adoption of practices supporting youth development	10/38	26.3%
Element 2A–B: Resources available to support community schooling efforts	19/31	61.3%
Element 2A–B: Perception of student success in programming	3/33	9.1%
Element 2C: Perception of adult family success in programming	9/30	30.0%
Element 2E: Absence of staffing challenges	10/40	25.0%
<b>Domain 3: Establish and maintain continuous improvement structures.</b>		
Element 3A–B: Evaluation results used to inform current programming	12/40	30.0%
Element 3A–B: Youth outcomes examined through evaluation efforts	11/21	52.4%
Element 3C–D: RC indicates having high-quality professional development opportunities	6/32	18.8%
Element 3C–D: Number of professional development events focused on improving program delivery	11/39	28.2%
Element 3C–D: Number of professional development events focused on adoption of the community school	10/39	25.6%

CSI Implementation Framework Domain and Element	Number in Higher Than Average Group/Total Schools With Responses	Percentage in Higher Than Average Group
framework		
Element 3C–D: Activities leaders report receiving training on best practices in the design and delivery of programming	8/34	23.5%
Element 3C–D: Presence of PLC	4/28	14.3%

## Refining the List of Implementation Elements

A total of 24 different indicators of school implementation of elements associated with the CSI Implementation Framework are outlined in Table 3. Although the *relevance* of each of these elements to the implementation of community schooling in Chicago was established through the work conducted by the CPS internal evaluation team in constructing the framework, there was still interest on the part of CPS to better understand the relative *importance* of these indicators to ensuring a successful implementation of the strategy.

In order to explore the relative *importance* of the framework elements, a decision was made by the evaluation team to focus on how implementation of a given framework element related to the extent to which the community school had been able to engage in *leveraged partnerships*. Leveraged partnerships represent new providers that the resource coordinator recruits to contribute to the community schooling effort based on results from the needs and resources assessment to provide additional, in-kind services to the school. The goal is that leveraged partnerships result in substantive new resources flowing into the school to meet the needs of enrolled students and their families, which go beyond the funds provided to the lead partner agency to run 21st CCLC–funded afterschool and summer learning programs. The concept of leveraged partnerships is at the heart of what the CSI in Chicago is trying to accomplish, and it is why CPS has opted to utilize 21st CCLC funds to support the placement of resource coordinators from community-based organizations and providers in community schools.

The only effort to estimate the value of programming and services provided through leveraged partnerships was conducted by AIR at the conclusion of the 2009–10 school year (Vinson, 2010). As part of this effort, the value of leveraged partnerships was estimated at the community school level two ways:

1. Programming provided to enrolled youth and adult family members
2. Other services/contributions not specifically linked to direct service provision, including, but not limited to, professional development, staffing, evaluation services, etc.

Even though there was an approximate two-year difference between when data on leveraged partnerships and the implementation surveys were collected, the evaluation team opted to explore whether any substantive differences could be found between schools in the *higher than*

*average* group on a given element and schools in the other implementation groups in terms of the value derived from leveraged partnerships. In this sense, the question we were interested in answering was: Could a subset of framework elements be identified where assignment to the *higher than average* group was associated with a greater total value for leveraged partnerships?

A total of 36 of the 41 schools represented in the sample were found to have data associated with the 2010 study. Among these 36 schools, the average value of programming provided on an in-kind basis was \$52,511, while the average value of other services and contributions was \$70,047.

Although few differences across framework elements were found when considering the value of other services/contributions not specifically linked to direct service provision (Item 2 above), there were a total of nine elements from Table 3 where substantive differences were noted in terms of the value of programming provided through leveraged partnerships (Item 1 above). These elements can be found in Table 4, which outlines the mean value of programming provided through leveraged partnerships both for schools classified in the *higher than average* group and for those schools that were not.

As shown in Table 4, more than half of the elements where meaningful differences were found to exist in the average value of programming are associated with Domain 1 of the framework—*establish and maintain essential structures and resources needed for an effective community school*. Having a set of shared goals and expectations between the school and the LPA, a needs and resources assessment in place, a resource coordinator who has been oriented and trained for the role and who feels supported in carrying out the work, and an advisory committee that has a meaningful role in decision making all are positively related to the value derived from programming provided by leveraged partnerships. Across these five elements, schools in the *higher than average* group had on average a leveraged partnership programming value that was 150 percent higher than schools that were not classified in the *higher than average* group on these elements.

Similar differences between groups were found in relation to two elements related to Domain 2—*establish and maintain community school programs and services*: (1) adoption of instructional practices by activity leaders that support positive youth development and (2) an absence of staffing challenges related to providing afterschool and summer learning programs. In this case, schools in the *higher than average* group had an average leveraged partnership programming value that was 114 percent higher than schools that were not classified in the *higher than average* group on these elements. Less substantial, but still notable, differences were found in relation to two elements associated with Domain 3: (1) the number of professional development events focused on the adoption of the community school framework and (2) the presence of a professional learning community that includes community school program staff and providers and school-day staff. Here, schools in the *higher than average* group had on average a leveraged partnership programming value that was 51 percent higher than schools that were not classified in the *higher than average* group on these elements.

**Table 4. Average Value of Programming Provided Through Leveraged Partnerships by Framework Elements and School Status in a Higher Than Average Implementation Group**

<b>Framework Element</b>	<b>Mean Value for Schools <i>Not</i> in the Higher Than Average Group</b>	<b>Mean Value for Schools in the Higher Than Average Group</b>
Element 1A: Shared Goals and Expectations	\$39,313 (n=15)	\$96,832 (n=9)
Element 1B: Needs and Resources Assessment	\$27,852 (n=15)	\$80,500 (n=2)
Element 1C: RC Received Enough Job Training	\$39,498 (n=25)	\$82,085 (n=11)
Element 1C: RC Receives Support in Performing His or Her Job	\$40,366 (n=23)	\$73,999 (n=13)
Element 1D: Advisory Committee Role in Decision Making	\$47,333 (n=6)	\$152,588 (n=5)
Element 2A–B: Adoption of Practices Supporting Youth Development	\$44,752 (n=24)	\$90,256 (n=9)
Element 2E: Absence of Staffing Challenges	\$40,781 (n=28)	\$92,606 (n=7)
Element 3C–D: Number of professional development events focused on adoption of the community school framework	\$44,844 (n=25)	\$69,471 (n=9)
Element 3C–D: Presence of professional learning community	\$41,787 (n=22)	\$61,430 (n=3)

### **Using Attendance Metrics to Explore the Quality of Implementation**

In addition to data obtained from the 2012 surveys related to implementation of the framework and the 2010 leveraged partnership survey, steps were taken to explore the quality of implementation through an examination of program attendance data collected in YouthServices.net. All schools enrolled in CSI are required to track youth participation in programming, including what activities the youth were enrolled in and whether or not they attended that activity on a given day that it was provided. Data on program attendance is key to both exploring how afterschool and summer learning activities impact youth and the relative quality of programming.

In light of the goals and objectives of CSI, a series of measures were crafted based on data housed in YouthServices.net that were seen as indicators of effective implementation of framework elements, particularly Domain 2, which focuses on the design and delivery of youth programming.

1. *Percentage of youth participating in two consecutive years of programming (2010–11 and 2011–12).* There is ample evidence in the afterschool literature that greater program effects are likely to be witnessed when youth participate regularly in programming across multiple years (Fredricks & Eccles, 2006; Huang et al., 2007; Pierce, Auger, & Vandell, 2013). Among the 41 schools represented in the sample, the average percentage of youth served in 2011–12 that had also participated in CSI programming in 2010–11 was 42 percent, although this percentage ranged from 7 percent to 78 percent.
2. *Mean number of minutes of programming attended during 2011–12 school year.* Youth cannot benefit from programming if they are not attending programming regularly. Among the 41 schools represented in the sample, the average number of minutes of programming attended by youth during the 2011–12 school year was 4,932 minutes of programming or about 82 hours of programming. The average number of minutes of programming attended ranged from 1,255 to 12,525 among the schools represented in the sample.
3. *Duration of participation (in days) during the 2011–12 school year.* Not only is retention in programming across school years important, but it is also important that youth be engaged in programming during the span of the school year as well. Among the 41 schools represented in the sample, the average duration of time spent in the program was 150 days during the 2011–12 school year or about five months. The average duration ranged from 47 days to 241 days.
4. *Percentage of school population enrolled in programming during the 2011–12 school year.* Given that CSI is designed to be a whole school effort, it seemed appropriate to assess the degree to which the student body of the community school was enrolled in programming provided through the initiative as a way to measure the depth and penetration of implementation. Among the 41 schools represented in the sample, the average percentage of the student body enrolled in programming was 55 percent and ranged from 18 percent to 100 percent.

Like all of the data used to assess the quality of implementation relative to the framework, the attendance data submitted through YouthService.net is self-report data provided by program staff, typically the resource coordinator. As a result, there is likely to be some unknown level of error associated with these data that has not been taken into account in the analyses undertaken in this report. The reader is encouraged to keep this mind, both in relation to the data obtained from YouthServices.net and the surveys related to framework implementation and leveraged partnerships.

## Assigning Schools to Higher and Lower Implementing Groups

Having identified a series of measures that allow for some assessment to be made of how well the 41 schools represented in the sample were implementing the CSI Implementation Framework, the next step taken by the evaluation team was to combine these measures in a way that would allow us to assign schools in the sample to higher and lower implementing groups. The following school-level measures were used to support this process:

- Measures obtained from the 2010 leveraged partnership survey:
  - Value of programming provided through leveraged partnerships
  - Value of other services and supports provided through leveraged partnerships
- Measures obtained from YouthServices.net for program attendance during the 2011–12 school year:
  - Percentage of youth participating in two consecutive years of programming (2010–11 and 2011–12)
  - Mean number of minutes of programming attended during 2011–12 school year
  - Mean duration of participation (in days) during the 2011–12 school year
  - Percentage of school population enrolled in programming during the 2011–12 school year

In addition, measures of elements of the CSI Implementation Framework derived from the 2012 key stakeholder surveys were used in the classification process. However, in order to simplify the classification task, the nine elements found to be related to the value of programming derived from leveraged partnerships (see Table 4) were combined into a single implementation index on a 0 to 100 scale, with higher values indicative of greater implementation of these practices. Index scale values had an average of 32.0 and ranged from 9.6 to 90.0. In order to assess the correlative value of this index, the bivariate correlation with the total programming value derived from the leverage partnership survey was calculated. A moderately strong correlation was found to exist between the index created to represent high implementation of these elements and the value of programming provided through leveraged partnerships ( $r = .456, p < .01$ ).

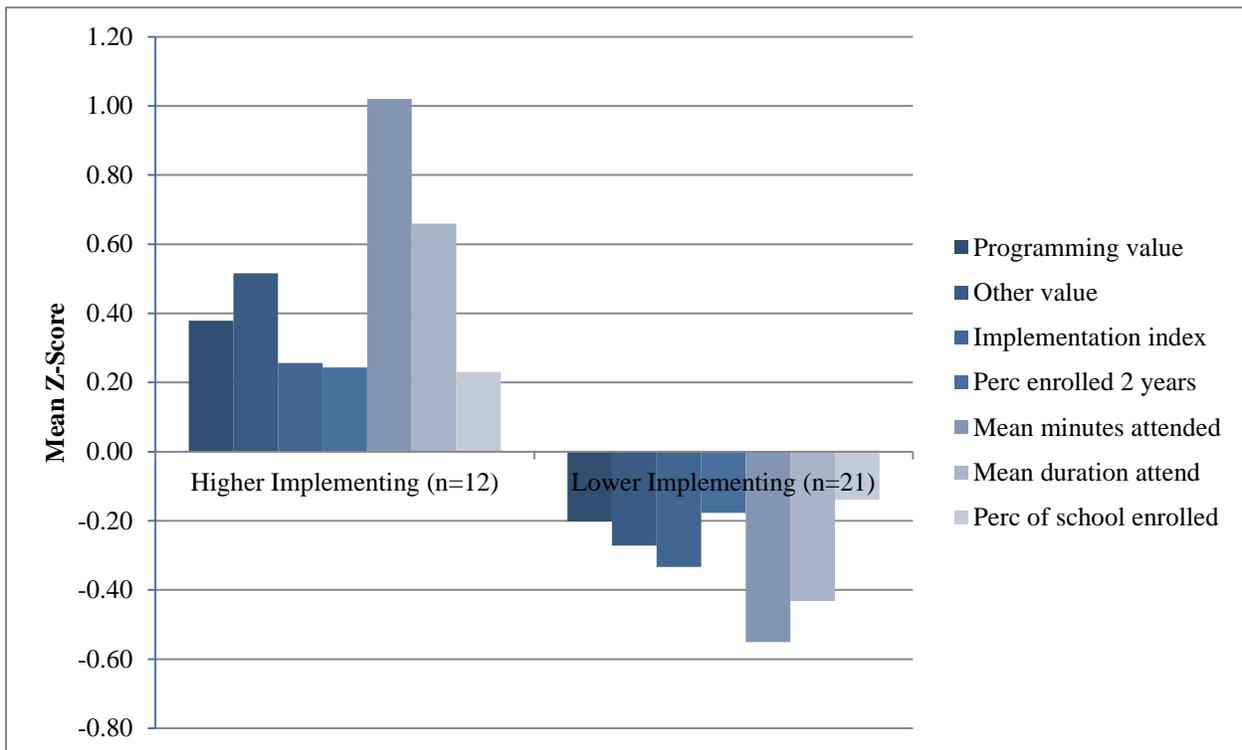
With the measures selected, the next step in the process was to classify the 41 centers represented in the sample into different groups using hierarchical cluster analysis. Typically, cluster analysis is employed to combine cases (or, in this case, schools) into groups using a series of variables as criteria to determine the degree of similarity between individual cases, and it is particularly well suited when there is a desire to classify a large number of cases into a smaller domain of discrete groupings. Our goal was to use cluster analysis to form initial groupings based on the aforementioned measures related to implementation of the CSI Implementation Framework.

Steps were taken to run a series of cluster analyses where the number of groups specified ranged from 2 to 4. The implementation survey index and variables obtained from the leverage partner

survey and YouthService.net were all standardized prior to inclusion in the analysis using  $z$  scores. Generally, the three-cluster solution was deemed to yield the most appropriate grouping of schools, although the group with consistently high levels of performance across all measures was found to comprise only four schools. Because of this, the highest performance cluster was merged with another cluster where schools demonstrated especially high performance on the minutes and duration of youth participation in programming and a moderately high level of performance on the implementation index constructed from key elements measured from the stakeholder surveys. This resulted in a two-group scheme outlined in Figure 2, with the higher implementing group comprising 12 schools and the lower implementing group consisting of 21 schools. A total of eight schools were dropped from the analysis because of missing data on one or more of the measures used to perform the classification process.

As shown in Figure 2, for the higher implementing group, the average on each of the measures used in the classification process was above average (signified by the fact that all of the bars in the figure are above the 0 line, which represents the mean standardized score), while the mean for each of the measures for the lower implementing group was found to be below average (thereby, all of the columns are below the 0 line of the figure). The difference between higher and lower implementing schools was especially stark on the average minutes of programming youth attended and the duration of participation.

**Figure 2. Initial School Classification Into Higher and Lower Implementing Groups**



In order to further maximize the contrast between higher and lower implementing groups, a series of schools were dropped from the lower implementing group, which had scores on some measures that overlapped with the range of scores on those measures among schools in the

higher implementing group. In this sense, steps were taken to minimize any score overlap between the higher and lower implementing groups. In addition, one school initially classified in the lower implementing group was moved to the higher implementing group given the high level of performance on a subset of attendance-related metrics. Based on these adjustments, the final sample of higher and lower implementing schools each comprised 13 schools.

The average difference between schools in the final higher and lower implementing groups is summarized in Table 5. As anticipated, schools in the higher implementation group demonstrated higher levels of performance across each of the measures used to classify programs into higher and lower implementation groups.

**Table 5. Mean Level of Performance on Metrics Used to Classify Schools Into Lower and Higher Implementing Groups**

<b>Metrics</b>	<b>Lower Implementing (n=13)</b>	<b>Higher Implementing (n=13)</b>
<i>From the 2010 Leveraged Partner Survey</i>		
Value of programming provided through leveraged partnerships	\$32,622	\$138,475
Value of other services and supports provided through leveraged partnerships	\$10,097	\$88,900
<i>Participation Data From YouthService.net</i>		
Percentage of youth participating in two consecutive years of programming (2010–11 and 2011–12)	38.69%	43.55%
Mean number of minutes of programming attended during 2011–12 school year	3,288	6,844
Duration of participation (in days) during the 2011–12 school year	133	155
Percentage of school population enrolled in programming during the 2011–12 school year	51.13%	63.18%
<i>From the 2012 Evaluation Surveys</i>		
Index created to represent performance on ISPS elements found to be correlated with the value of programming provided through leveraged partnerships*	26.68	30.83
<i>Data Available Only for Schools Enrolled in the Pilot Coaching Initiative**</i>		
Short-form YPQA	2.68	3.93
CSI Self-Assessment Quality Improvement Rubric	2.12	3.16

\*These elements related to shared goals and expectations; conducting a needs and resources assessment; the RC receiving sufficient job training and support; advisory committee role in decision making; the adoption of instructional practices related to youth development; the absence of staffing challenges; professional development on the community school model; and the presence of a professional learning community.

\*\*A total of 10 of the 26 schools represented in the table had these data available—five from the lower and five from the higher implementing groups.

For a subset of schools in each group (five in each group), both self-assessment data from school completion of the Self-Assessment Quality Improvement Rubric and data from the short-form version of the Youth Program Quality Assessment (YPQA) were available in relation to the 2013–14 school year. The short-form YPQA is an observational measure developed by the David P. Weikart Center for Youth Program Quality, which contains criteria relating to what constitutes instructional quality in youth-serving programs at the point of service. These 10 schools were part of another study being conducted by the evaluation team at AIR and Diehl Consulting of a pilot coaching process anchored to the Self-Assessment Quality Improvement Rubric in 15 schools. As part of this pilot, the Self-Assessment Quality Improvement Rubric was completed as a self-assessment by the resource coordinator and other key school staff, and the short-form YPQA was scored based on observations of programming conducted by the coach assigned to a given school. Although these data are available only for a subset of higher and lower implementing schools and were collected later than the 2011–12 school year under consideration in this report, a decision was made to include them in this report to explore whether they further reinforced the difference we expected to see between the higher and lower implementing groups.

As shown in Table 5, schools in the higher implementing group scored a full point higher on the short-form YPQA on average than schools in the lower implementing group (based on a 1–5 scale). The mean score for the higher implementing group is especially notable since the mean value of 3.93 exceeds the target performance threshold of 3.90 recommended by the Weikart Center, the point at which substantive change has been shown in the level at which youth demonstrate engagement in the programming being provided based on past research by the Weikart Center (Akiva, Pearson, Sugar, Peck, Smith, & Denault, 2010).

In addition, schools in the higher implementing group had an average Self-Assessment Quality Improvement Rubric score that was also a full point higher than the average for schools in the lower implementing group, putting the higher implementing group in the *proficient* range of the scale in terms of implementation of elements represented in the rubric, while the average for lower implementing schools was in the *emerging* portion on the scale.

The results in Table 5 appear to confirm that there were substantive and meaningful differences between the higher and lower implementing schools fashioned through the methods described previously. In the sections that follow, we explore how youth outcomes varied among youth attending higher and lower implementing schools as well as with similar youth enrolled in CPS schools not part of CSI.

## Assessing the Impact of Participation in a Higher Implementing Community School

As articulated in the introduction, the primary purpose of this report is to describe what was found when steps were taken to assess whether youth enrolled in programming in higher implementing schools demonstrated better functioning on a variety of academic, behavioral, and learning experience outcomes relative to (a) youth enrolled in programming in lower implementing schools and (b) similar youth attending non-CSI schools.

The comparison with lower implementing schools is directly aligned with the efforts of CSI to develop a series of tools and processes like the Self-Assessment Quality Improvement Rubric that are oriented toward helping lower quality community schools progress to higher levels of implementation quality. By answering this question, CPS will have additional information about the possible impact of moving lower implementing schools to a higher level of implementation quality.

The second question is meant to demonstrate the promise of high-implementing community schools in a broader sense by demonstrating the impact the initiative can have relative to similar students enrolled in nonparticipating schools. This information will be helpful to CPS in demonstrating what higher implementing community schools can achieve in the way of youth outcomes and how these outcomes may justify continued efforts to support the development of effective implementation of the strategy in the full domain of schools enrolled in the initiative.

### Youth Outcomes Examined

A variety of academic, behavioral, and learning experience outcomes were examined in order to assess how enrollment in programming provided at higher implementing schools was impacting youth outcomes. All outcomes examined pertained to the 2011–12 school year. Given that the domain of outcomes varied somewhat by grade level, separate analyses were done for youth in Grades K–3, 4–8, and 9–12. Behavioral outcomes like school-day absences, school-day misconducts, and school-day suspensions were examined for all grade levels. In addition, for students in Grades 4–8, Illinois Standards Achievement Test (ISAT) reading and mathematics scores were examined, and for students in Grades 6–8, various scales obtained from the *My Voice, My School* survey were also examined pertaining to both youth experiences in school and their general well-being. For youth in Grades 9–12, academic outcomes like cumulative weighted grade point average (GPA) and the number of credits earned during the year were taken into consideration, as well as the same domain of scales from the *My Voice, My School* survey utilized in conducting impact analyses for youth in Grades 6–8. Some analyses at this grade level also involved ACT and Prairie State Achievement Examination (PSAE) assessment scores.

The following behavioral outcomes were examined for *all* grade levels:

- School-day absences
- School-day misconducts
- School-day suspensions

The following additional outcomes were examined for youth in *Grades 4–8*:

- ISAT mathematics scale scores
- ISAT reading scale scores
- *My Voice, My School* survey scale scores (Grades 6–8 only):
  - Emotional Health
  - Academic Engagement
  - Academic Personalism
  - Human and Social Resources in the Community
  - Parent Support
  - Psychological Sense of School Membership
  - Safety
  - Inquiry-Based Science Instruction
  - Student-Teacher Trust

The following additional outcomes were examined for youth in *Grades 9–12*:

- Cumulative weighted grade point average (GPA)
- Credits earned
- PSAE mathematics (viable only for analyses involving youth attending non-CSI schools)
- PSAE reading (viable only for analyses involving youth attending non-CSI schools)
- ACT mathematics (viable only for analyses involving youth attending non-CSI schools)
- ACT science (viable only for analyses involving youth attending non-CSI schools)
- ACT reading (viable only for analyses involving youth attending non-CSI schools)
- ACT English (viable only for analyses involving youth attending non-CSI schools)
- *My Voice, My School* survey scale scores:
  - Emotional Health
  - Academic Engagement
  - Academic Personalism
  - Human and Social Resources in the Community
  - Parent Support
  - Psychological Sense of School Membership
  - Safety
  - Inquiry-Based Science Instruction
  - Student-Teacher Trust

## Analytic Approach

In any evaluation of a program where participants are not randomly assigned to treatment or control groups, the problem of selection is paramount. Youth from higher implementing schools may differ from those from lower implementing schools or schools not enrolled in CSI in a variety of ways. These differences can bias estimates of impact. If we were simply to compare youth from higher implementing schools to those from lower implementing schools or schools not enrolled in the initiative, we would not be able to disentangle the effect of the program implementation level from the preexisting difference between these two groups of youth.

To address these potential confounders of estimating the effect of attending a higher implementing school, the research team employed a propensity score stratification approach. Propensity score stratification is a statistical method that allows researchers to more closely estimate the causal effect of interest by creating a comparison group that looks like the treatment group on all observable characteristics. This approach has two main components. First, the research team used a propensity score stratification approach to construct a comparison group of youth from lower implementing schools or schools not enrolled in CSI (depending upon the analysis) but was similar to students from higher implementing schools on observable characteristics. Then the research team examined whether youth from higher implementing schools outperformed youth in the matched comparison group on the outcomes of interest.

To create the comparison group, the propensity that each of these youth from higher implementing schools is calculated based on the available observable characteristics. The outcome of interest in modeling propensity scores is treatment status (1 for attending a higher implementing school, 0 for attending a lower implementing school or school not enrolled in CSI). To account for this binary outcome, logistic regression was used to model the logit (or log-odds) of student group assignment status. The propensity score was formulated as follows:

$$\text{logit}(Z_{ij}) = \alpha + \mathbf{X}'_{ij}\boldsymbol{\beta} + \mathbf{W}'_j\boldsymbol{\beta} ,$$

where  $Z_{ij}$  indicates the treatment status for student  $i$  who attended school  $j$  ( $Z_{ij} = 1$  for students from a higher implementing school,  $Z_{ij} = 0$  for students from a lower implementing school or school not enrolled in CSI), and a student's logit is a linear function of a vector of individual student characteristics,  $\mathbf{X}'_{ij}$ , and a vector of characteristics of the school the student attends,  $\mathbf{W}'_j$ . These student-level and school-level characteristics correlate to the outcomes of interest and the group assignment such as student prior year reading and mathematics achievement, English proficiency level, special education status, school prior year reading and mathematics achievement, and the portion of a school's students who qualify for free or reduced price lunch, etc.

Once a propensity score was calculated for each student, we stratified the sample on the propensity score and calculated strata-specific treatment on the treated weights. This approach essentially forms a matched comparison group of youth from lower implementing schools but are similar to youth from higher implementing schools on all observable characteristics.

Given the nested structure of the data (students within schools), we then used a hierarchical modeling approach to examine the effect of attending a higher implementing school on the outcomes of interest. The model was formulated as follows:

**Level 1—Students:**

$$y_{ij} = \pi_{0j} + \pi_{1j}Z_{ij} + \sum \pi_{pj}x_{ij} + e_{ij}$$

where  $Y_{ij}$  is the outcome for student  $i$  in school  $j$ ,  $Z_{ij}$  is an indicator of whether student  $i$  attended a higher implementing school  $j$ , and  $x_{ij}$  is a vector of student-level characteristics.<sup>2</sup>

**Level 2—Schools:**

$$\pi_{0j} = \beta_{00} + \sum \beta_{01}W_j + r_{0j}$$

$$\pi_{1j} = \beta_{10}$$

$$\pi_{pj} = \beta_{p0}$$

where the intercept is modeled as random and treatment effect and all other covariates are modeled as fixed effects and  $W_j$  is a vector of school characteristics.

The following section describes the results for the domain of analyses undertaken using this approach.

## **Results—Higher Implementing Versus Lower Implementing Schools**

It is important to note before reviewing the results of the impact analyses performed comparing higher and lower implementing schools that steps were taken to include all youth attending CSI activities at higher and lower implementing schools during the 2011–12 school year in the impact analyses. This is different than what we typically do as part of our 21st CCLC evaluation work where we set thresholds for adequate dosage (usually at 30 and 60 days of participation). The evaluation team felt this was the more appropriate approach to take here given the theory that greater implementation of the CSI Implementation Framework was likely to result in activities more aligned to youth needs and a broader array of activities resulting in greater overall participation in programming. In this sense, higher implementing CSI should be better at getting youth involved in activities over a longer period than lower implementing schools. To set a threshold for participation akin to what we typically do when conducting other 21st CCLC evaluation work would seem to work against detecting this effect thought to be associated with higher implementation. This choice on the part of the evaluation should be kept in mind when reviewing the following results.

---

<sup>2</sup> Although the propensity score matching was the primary method for control differences between treatment and comparison students, covariates that either were not sufficiently balanced between treatment and control and covariates that are seen as particularly important for theoretical reasons can be included as additional controls in the outcomes model, making the modeling approach “doubly robust.”

## Grades K–3

There are 1,782 students in Grades K–3 in the initial sample that participated in CSI programming at higher and lower implementing schools during the 2011–12 school year. As noted previously, the outcomes of interest for this group included only the number of unexcused absence days, the number of misconducts, and the number of days suspended. Table 6 provides a descriptive picture of students from both higher and lower implementing schools.

When using a propensity score stratification approach, one thing the analyst is especially concerned about is the baseline level of equivalence of the groups—that is, how similar they are on key characteristics that may be related to treatment status (in this case, treatment is enrollment in a higher implementing school). As shown in Table 6, the baseline equivalence of the two groups (students from higher implementing schools and students from lower implementing schools) is largely demonstrated on most student- and school-level characteristics and outcomes (What Works Clearinghouse, 2008). One notable area of difference relates to the percentage of youth identified as *limited English proficient (LEP)*, where students enrolled in higher implementing schools were more likely to have LEP status. This was the only characteristic where there was a significant difference between students in the higher and lower implementing groups.

The research team tried several propensity score matching methods such as nearest available neighbor with replacement and using overlap of the propensity score distributions and failed to find the matched comparison group that has better baseline equivalence than the initial comparison group. Therefore, the research team decided to use all 1,782 students as the analytic sample and adjust the covariates on which the baseline equivalence was not found in the impact analysis model.

**Table 6. Higher and Lower Implementing Group Characteristics—Grades K–3**

	Higher Implementing	Lower Implementing	Standardized Mean Difference
<i>Student-Level Characteristics</i>			
Number of absences 2010–11	7.19	7.88	-0.07
Number of misconducts 2010–11	0.04	0.06	-0.04
Number of times suspended 2010–11	0.03	0.05	-0.07
Special education	8.90%	9.45%	0.4
Limited English proficient	16.87%	11.57%	3.2 <sup>3</sup>
<i>School-Level Characteristics</i>			
Total enrollment	477	462	0.09
Percent limited English proficient	0.15	0.10	0.22
Percent Black	0.68	0.83	-0.38
Percent Hispanic	0.30	0.09	0.63
Percent special education	0.15	0.15	-0.05

<sup>3</sup> Z-value of the difference between the two proportions is 3.2 and p-value is 0.002. There is a statistically significant difference between the two proportions.

Results assessing the impact of higher implementing schools on youth outcomes for students in grades K–3 are outlined in Table 7.<sup>4</sup> As shown in Table 7, students from higher implementing schools had a lower number of school-day absences and a fewer number of suspensions but a higher number of misconducts, although none of these differences was statistically significant. In this sense, no significant differences were found between higher and lower implementing schools for the outcomes examined for students in Grades K–3.

**Table 7. Estimated Effect of Attending a Higher Implementing School—Grades K–3**

Outcome	Estimated Effect	Standard Error	Event Ratio (Treatment/Comparison)
Number of school-day absences	0.06	0.09	0.92
Number of misconducts	0.07	0.14	0.91
Number of times suspended	0.05	0.06	0.84

### Grades 4–8

There are 3,088 students in Grades 4–8 in the initial sample that participated in CSI programming at higher and lower implementing schools during the 2011–12 school year. For this set of analyses, we included only those youth who had taken the ISAT in the 2010–11 school year (prior year) and who had outcome data in the 2011–12 school year (outcome year). Although this reduced our analytic sample from 3,088 to 2,904 students (about a 6 percent reduction), this ensured that we were able to control for past academic achievement, an important predictor of student outcomes (Bifulco, 2012).

Table 8 provides a descriptive picture of students from both higher and lower implementing schools. The standardized mean difference column demonstrates the difference in values on a given characteristic after propensity score matching approaches have been used to create a similar comparison group. The baseline equivalence of the two groups is demonstrated on most student- and school-level characteristics and outcomes, especially in relation to the academic outcomes measured by mathematics ISAT scores and reading ISAT scores for 2010–11. One notable area of difference relates to the percentage of youth identified as LEP, where students enrolled in higher implementing schools were significantly more likely to have LEP status.

---

<sup>4</sup> All estimated effect coefficients represented in the tables appearing in this section of the report have been standardized and can be interpreted as effect sizes. Event ratios have also been calculated in relation to effects associated with school-day absences, misconducts, and suspensions. For example, when interpreting event ratios, values of 1.0 indicate no mean difference on the outcomes. A value above 1.0, say 1.35, would indicate that the mean of the higher implementing CSI student group was larger than the control group by that amount (i.e., was 1.35 times the size of the control group’s mean). Values below 1.0 indicate that higher implementing CSI student group’s mean was smaller (e.g., a value of .50 would indicate a treatment group mean being half the size of the control group mean).

**Table 8. Higher and Lower Implementing Group Characteristics—Grades 4–8**

	Higher Implementing	Lower Implementing	Standardized Mean Difference
<i>Student-Level Characteristics</i>			
Mathematics ISAT 2010–11	227.99	229.14	-0.04
Reading ISAT 2010–11	216.04	215.12	0.03
Number of absences 2010–11	3.86	4.58	-0.13
Special education	14.97%	16.2%	-0.89
Limited English proficient	10.73%	6.26%	4.23 <sup>5</sup>
<i>School-Level Characteristics</i>			
Total enrollment	473	470	0.02
Percent limited English proficient	0.15	0.09	0.31
Percent Black	0.67	0.84	-0.47
Percent Hispanic	0.31	0.09	0.70
Percent special education	0.15	0.15	0.07

Table 9 displays the estimated difference in the outcomes of interest between students from higher implementing schools and their counterparts from lower implementing schools. Contrary to what was hypothesized, students from lower implementing schools performed better than students from higher implementing schools on both ISAT mathematics and reading tests in 2011–12. For both content areas, these effects were small. However, students from lower implementing schools had a higher number of school-day absences and suspensions but fewer misconducts, although none of these differences was statistically significant.

As noted previously, a series of scales from the *My Voice, My School* survey were also considered for youth in Grades 6–8. Among these outcomes, only the scale score related to the emotional health scales was found to be significant but not in the hypothesized direction. In this sense, youth enrolled in lower implementing schools had significantly higher scale scores on this scale of the *My Voice, My School* survey.

**Table 9. Estimated Effect of Attending a Higher Implementing School—Grades 4–8**

Outcome	Estimated Effect	Standard Error
Mathematics ISAT 2011–12 achievement	-0.23*	0.10
Reading ISAT 2011–12 achievement	-0.18*	0.08
Number of school-day absences	-0.22 (0.78) <sup>1</sup>	0.12
Number of misconducts	0.07 (1.01) <sup>1</sup>	0.19

<sup>5</sup> Z-value of the difference between the two proportions is 4.23 and p-value is 0.00. There is a statistically significant difference between the two proportions.

Outcome	Estimated Effect	Standard Error
Number of times suspended	-0.08 (0.93) <sup>1</sup>	0.07
Emotional Health	-0.46*	0.22
Academic Engagement	0.02	0.20
Academic Personalism	0.06	0.28
Human and Social Resources in the Community	0.22	0.21
Parent Support	-0.05	0.14
Psychological Sense of School Membership	-0.20	0.18
Safety	0.06	0.25
Inquiry-Based Science Instruction	0.21	0.23
Student-Teacher Trust	0.28	0.18

Note: \* indicates statistically significant at 0.05 level.

<sup>1</sup> Represents an event ratio—treatment/comparison.

## Grades 9–12

The high school sample consisted of 705 students across Grades 9–12 who participated in CSI programming at higher and lower implementing schools during the 2011–12 school year in four different high schools (two higher implementation schools and two lower implementation schools). Despite attempts to do so, the small sample size prevented us from using a propensity score matching algorithm to facilitate causal inferences. As shown in Table 10, there was substantial imbalance on a number of characteristics associated with youth participating in higher and lower implementing schools, including academic achievement and demographic data. Statistically significant differences were found to exist between students enrolled in higher and lower implementing schools in terms of *special education* status (a higher percentage of students with this classification were enrolled in lower implementing schools) and LEP (a higher percentage in higher implementing schools).

**Table 10. Higher and Lower Implementing Group Characteristics—Grades 9–12**

	Higher Implementing	Lower Implementing	Standardized Mean Difference
<i>Student-Level Characteristics</i>			
Cumulative weighted GPA 2009–10	2.45	2.25	0.18
Cumulative weighted GPA 2010–11	2.72	2.18	0.54
ACT reading score 2010–11	17.03	14.16	0.85
ACT math score 2010–11	16.45	15.03	0.64
ACT science score 2010–11	17.64	15.36	0.71
ACT English score 2010–11	16.90	12.29	1.13
Number of absences 2010–11	9.23	20.21	-0.80

	Higher Implementing	Lower Implementing	Standardized Mean Difference
Number of misconducts 2010–11	0.88	1.11	-0.13
Number of times suspended 2010–11	0.31	0.96	-0.51
Special education	9.09%	27.18%	-6.36 <sup>6</sup>
Limited English proficiency	4.91%	1.68%	2.30 <sup>7</sup>
<b>School-Level Characteristics</b>			
Total enrollment	463.05	612.41	-2.28
Percent limited English proficiency	0.07	0.006	4.04
Percent Black	0.74	0.98	-8.78
Percent Hispanic	0.19	0.01	7.30
Percent special education	0.13	0.20	-2.47

In light of our inability to use propensity score approaches, our analytical approach was to control for student prior academic achievement in different subject areas. The available data required that we strategically select student- and school-level characteristics to control for in our comparisons of higher and lower implementing schools. Because our sample of schools is so small, two schools in each group, there was limited variability in many of the school-level demographic characteristics, which were subsequently highly collinear with implementation classification (i.e., higher versus lower). To retain a parsimonious model, we focused on controlling for prior student achievement in GPA and in ACT scores. Prior achievement scores tend to be strong model covariates because they covary with many student and school characteristics. We specified, for each of our outcomes, the following mixed model:

$$Y_{ij} = \gamma_{00} + \gamma_{01}(HighImpl_j) + \gamma_{10}(CumulativeGPA2011_{ij}) + \gamma_{20}(CumulativeGPA2010_{ij}) + \gamma_{30}(ACTReading2011_{ij}) + \gamma_{40}(ACTMath2011_{ij}) + \gamma_{50}(ACTScience2011_{ij}) + \gamma_{60}(ACTEnglish2011_{ij}) + u_{0j} + r_{ij},$$

where  $\gamma_{01}$  is the difference between higher and lower implementing schools,  $\gamma_{10}$  through  $\gamma_{60}$  are prior achievement fixed effects,  $u_{0j}$  is the school-level error term with a mean of 0 and a variance equal to  $\tau_{00}^2$ , and  $r_{ij}$  is the school-level residual error term with a mean of 0 and variance equal to  $\sigma^2$ . Positive values of  $\gamma_{01}$  indicate that higher implementing schools have a larger mean outcome,  $Y_{ij}$ , than lower implementing schools, controlling for prior academic performance. All prior achievement covariates were grand mean centered.

Table 11 presents the differences between higher and lower implementing schools using this approach. As shown in Table 11, the only outcome where higher and lower implementing schools significantly differed was in number of absences, with youth participating in CSI

<sup>6</sup> Z-value of the difference between the two proportions is 6.36 and p-value is 0.00. There is a statistically significant difference between the two proportions.

<sup>7</sup> Z-value of the difference between the two proportions is 2.30 and p-value is 0.02. There is a statistically significant difference between the two proportions.

programming at higher implementing schools having fewer absences than lower implementing schools. More specifically, youth attending programming at higher implementing schools demonstrated 59 percent fewer absences than their peers in lower implementing schools. Although these results provide a preliminary examination of differences between higher and lower implementing CSI programs, they should not be viewed as causal given the less robust approach used to conduct these analyses.

**Table 11. Estimated Effect of Attending a Higher Implementing School<sup>8</sup>—Grades 9–12**

Outcome	Estimated Effect	Standard Error
Number of school-day absences	-0.79* (0.41) <sup>1</sup>	0.12
Number of school-day misconducts	0.04 (1.31) <sup>1</sup>	0.17
Number of school-day suspensions	-0.20 (0.66) <sup>1</sup>	0.17
Cumulative weighted GPA	-0.02	0.02
Percentage of credits earned	0.20%	0.07
Emotional Health	-0.11	0.21
Academic Personalism	-0.27	0.22
Human and Social Resources in the Community	0.11	0.22
Parent Support	-0.17	0.22
Psychological Sense of School Membership	-0.13	0.22
Safety	0.001	0.56
Inquiry-Based Science Instruction	-0.30	0.86
Student-Teacher Trust	-0.08	0.22

<sup>1</sup> Represents an event ratio—treatment/comparison.

## Summary of Findings—Higher Implementing Versus Lower Implementing Schools

The primary goal in undertaking these analyses was to explore the answer to the following question: What was the impact on youth outcomes when youth participate in programming provided at a higher versus lower implementing CSI school? It was hypothesized that higher implementing schools would have a variety of positive effects on the academic achievement, behavioral, and learning experience outcomes under consideration. Unfortunately, the analyses employed to answer this question provided very little in the way of evidence to support this hypothesis. The only finding where this was found to be true was in relation to youth in Grades 9–12 attending higher implementing programs who demonstrated significantly fewer unexcused absences than their peers in lower implementing programs.

<sup>8</sup> All outcomes other than those indicated by (Standardized) were kept on the original metric. Standardized outcomes facilitate a conditional standardized mean difference comparison between higher and lower implementing schools (i.e., differences in the standard deviation metric).

More problematic was the finding that participation in lower implementing programs among youth in Grades 4–8 resulted in small, but positive, effects on ISAT reading and mathematics scores as compared to higher implementing schools, and a moderately strong, positive effect on youth emotional health. These results were certainly unexpected. Two possible explanations warrant exploration here. The first pertains to the types of activities delivered at higher and lower implementing schools. It is possible that lower implementing schools spent more time delivering activities meant to cultivate youth reading and mathematics skills or focused on social and emotional learning. Steps were not taken by the evaluation team to explore these differences, and these should be considered in the future.

Second, as previously mentioned, unlike most evaluations exploring the impact of 21st CCLC–funded participation on youth outcomes, all youth participants in programming were included in the analysis, as opposed to limiting impact analyses to only those youth participating for some minimum threshold. To explore whether selecting a subset of youth that participated in programming more intensively made a difference in terms of impact analysis results, the models were rerun including only youth who participated for 60 hours or more of CSI programming during the 2011–12 school year. The 60-hour threshold is similar to the 30-day requirement defined by the U.S. Department of Education for what constitutes a regular program attendee for the 21st CCLC program. When this was done, the significant impacts for ISAT reading and mathematics and emotional health associated with lower implementing schools for youth in Grades 4–8 went away. The positive finding related to fewer unexcused absences for youth in Grades 9–12 enrolled in programming in higher implementing schools, however, was retained when only youth attending 60 hours or more were included in the analysis.

In light of these results, it can generally be concluded that enrollment in programming at higher implementing schools had no appreciable impact on youth outcomes in the manner hypothesized relative to lower implementing schools, except on reducing unexcused absences among youth in Grades 9–12 where youth attending higher implementing programs missed fewer days of school.

It is important to note that the analyses conducted in this study were limited to a relatively small number of students who attended either a higher or lower implementing schools in the 2011–12 academic year. This small sample limits the statistical power of the study to detect effects. As in any study where random assignment to treatment and control conditions is not feasible, it is possible that there were unobservable differences between students who attended higher implementing schools and lower implementing schools that we were unable to control for. This is especially true where propensity score matching was not possible. To the extent that unobserved differences are related to student outcomes, these differences would bias the estimated differences between higher and lower implementing schools. To that end, these analyses provide initial evidence about the impact of attending higher implementing afterschool programs but should not necessarily be considered equivalent to experimental studies that have stronger internal validity.

## **Results—Higher Implementing Versus Non-CSI Schools**

Before reviewing results oriented at assessing the impact of participation in CSI programming at higher implementing schools relative to youth attending schools not enrolled in the initiative, additional information needs to be provided, both about the youth from higher implementing

schools selected for inclusion in these analyses and the schools from which comparison students were drawn.

Unlike the previous set of analyses, here we are interested in assessing the impact of a higher dose of participation in programming provided in higher implementing schools as compared to youth not participating in CSI-supported programming. In this case, a higher dose was defined as either attending 60 hours or more of programming both during the 2010–11 and 2011–12 school years or attending 120 hours or more of programming during the 2011–12 school year.

Youth for the non-CSI group were drawn from 70 schools that demonstrated similar student demographic and achievement characteristics as CSI schools and that were not:

1. Schools providing programming funded by After-School All-Stars
2. Schools without standard school IDs
3. Charter schools
4. Magnet schools
5. Child-parent centers
6. Former community schools
7. Special focus schools (selective enrollment, cognitive disability, gifted magnet, military academy)
8. Schools that opened after 2006 (because none of the CSI schools in our sample opened after 2006)
9. Schools associated with other 21st CCLC grants
10. Schools with missing demographic data between years 2007 and 2012

### Grades K–3

We first looked at students in Grades K–3, with a total of 11,310 students attending both higher implementing and non-CSI comparison schools. Here again, the outcomes of interest for this subset were number of school-day absences, number of misconducts, and number of suspensions. As shown in Table 12, higher implementing CSI schools and non-CSI schools were balanced on most of the propensity score model covariates, although we additionally controlled for grade in the impact models. A statistically significant difference was also noted between groups in terms of the percentage of students identified as LEP, with a higher percentage of students in non-CSI schools receiving this classification.

**Table 12. Higher Implementing and Non-CSI Group Characteristics—Grades K–3**

	Higher Implementing	Non-CSI	Standardized Mean Difference
<i>Student-Level Characteristics</i>			
Number of absences 2010–11	6.10	6.06	0.01
Number of misconducts 2010–11	0.06	0.12	-0.09

	Higher Implementing	Non-CSI	Standardized Mean Difference
Number of times suspended 2010–11	0.02	0.10	-0.14
Special education	11.90%	9.72%	1.28
Limited English proficient	9.32%	29.85%	-7.85 <sup>9</sup>
<i>School-Level Characteristics</i>			
Total enrollment	451	466	-0.07
Percent limited English proficient	0.09	0.10	-0.04
Percent Black	0.83	0.82	0.04
Percent Hispanic	0.15	0.17	-0.04
Percent special education	0.15	0.14	0.14

Table 13 presents the results of these models. Higher implementing schools were lower on each of the outcomes compared to non-CSI schools, significantly in number of misconducts and number of suspensions. More specifically, youth attending higher implementing schooling had 55 percent fewer misconducts and 53 percent fewer suspensions than similar youth attending non-CSI schools.

**Table 13. Estimated Effect of Attending a Higher Implementing School—Grades K–3**

Outcome	Estimated Effect	Standard Error	Event Ratio (Treatment/Comparison)
Number of school-day absences	-0.17	0.20	(0.94)
Number of misconducts	-0.31 <sup>+</sup>	0.16	(0.45)
Number of suspensions	-0.29 <sup>+</sup>	0.16	(0.47)

Note: <sup>+</sup> indicates statistically significant at 0.10 level.

### Grades 4–8

There were 14,396 student records in Grades 4–8 in the initial sample. After matching, 12,664 student records were kept in the analytic sample. The treatment group and the matched comparison group have a good balance on the important covariates such as student prior reading and mathematics achievement, special education status, and English proficiency level, etc. (see Table 14). The covariates that were not sufficiently balanced and important for theoretical reasons were included as additional controls in the hierarchical linear modeling model. This was particularly the case with LEP classification, where a statistically significant difference was also noted between groups with a higher percentage of students in the non-CSI schools receiving this classification.

<sup>9</sup> Z-value of the difference between the two proportions is -7.85 and p-value is 0.00. There is a statistically significant difference between the two proportions.

**Table 14. Higher Implementing and Non-CSI Group Characteristics—Grades 4–8**

	Higher Implementing	Non-CSI	Standardized Mean Difference
<i>Student-Level Characteristics</i>			
Mathematics ISAT 2010–11	226.16	223.29	0.06
Reading ISAT 2010–11	214.67	212.86	0.04
Number of absences 2010–11	3.11	3.31	-0.05
Number of misconducts 2010–11	0.26	0.30	-0.05
Number of times suspended 2010–11	0.19	0.26	-0.08
Special education	14.46%	16.16%	-1.10
Limited English proficient	9.69%	17.22%	-4.76 <sup>10</sup>
<i>School-Level Characteristics</i>			
Average mathematics ISAT 2010–11	234.50	232.25	0.24
Average reading ISAT 2010–11	219.52	218.27	0.19
Total enrollment	441	446	-0.03
Percent limited English proficient	0.16	0.14	0.10
Percent Black	0.68	0.72	-0.10

Table 15 displays the estimated difference in the outcomes of interest between students from higher implementing schools and served intensively in the program and their counterparts from non-CSI schools. Students from higher implementing schools performed better in the 2011–12 ISAT mathematics test and worse in the 2011–12 ISAT reading test than students from non-CSI schools. However, the differences were trending toward zero, indicative of no significant difference. Students from non-CSI schools had a higher number of school-day absences, a higher number of misconducts, and a higher number of times suspended, although none of the differences was statistically significant. For the *My Voice, My School* survey scales, only one difference between the two groups was statistically significant. In this case, students from higher implementing schools achieved higher scores on the Emotional Health scale of the survey. This could be deemed a moderate effect (effect size of .43).

**Table 15. Estimated Effect of Attending a Higher Implementing School—Grades 4–8**

Outcome	Estimated Effect	Standard Error
Mathematics ISAT 2011–12 achievement	0.02	0.06
Reading ISAT 2011–12 achievement	-0.01	0.05
Number of school-day absences	-0.04 (.88) <sup>1</sup>	0.06
Number of misconducts	-0.04 (.65) <sup>1</sup>	0.10

<sup>10</sup> Z-value of the difference between the two proportions is -4.76 and p-value is 0.00. There is a statistically significant difference between the two proportions.

Outcome	Estimated Effect	Standard Error
Number of times suspended	-0.06 (.56) <sup>1</sup>	0.05
Emotional Health	0.43*	0.22
Academic Engagement	0.26	0.18
Academic Personalism	0.10	0.22
Human and Social Resources in the Community	0.10	0.13
Parent Support	0.17	0.14
Psychological Sense of School Membership	0.26	0.17
Safety	0.25	0.16
Inquiry-Based Science Instruction	0.06	0.23
Student-Teacher Trust	0.28	0.16

Note: \* indicates statistically significant at 0.05 level.

<sup>1</sup> Represents an event ratio—treatment/comparison.

## Grades 9–12

After matching, 10,687 student records were kept in the analytic sample for students in Grades 9–12. The treatment group and the matched comparison group were found to have a good balance on the important covariates such as student prior reading and mathematics achievement, special education status, and English proficiency level, etc. (see Table 16). No significant difference was found between the higher implementing and non-CSI comparison group after propensity score matching approaches were employed.

**Table 16. Higher Implementing and Non-CSI Group Characteristics—Grades 9–12**

	Higher Implementing	Non-CSI	Standardized Mean Difference
<i>Student-Level Characteristics</i>			
Cumulative weighted GPA 2010–11	2.30	2.26	0.03
ACT reading score 2010–11	4.40	4.18	0.03
ACT mathematics score 2010–11	4.26	4.14	0.02
ACT science score 2010–11	4.55	4.36	0.02
ACT English score 2010–11	4.37	4.13	0.03
PSAE reading score 2010–11	38.23	36.98	0.02
PSAE mathematics score 2010–11	37.47	36.43	0.02
Number of absences 2010–11	6.81	7.41	-0.09
Number of misconducts 2010–11	0.69	0.85	-0.06
Number of times suspended 2010–11	0.23	0.52	-0.24
Special education	10.33%	14.16%	-1.48
Limited English proficiency	5.98%	8.16%	-1.07

	Higher Implementing	Non-CSI	Standardized Mean Difference
<i>School-Level Characteristics</i>			
Total enrollment	469	1141	-1.23
Percent limited English proficiency	0.06	0.05	0.06
Percent Black	0.79	0.69	0.28
Percent Hispanic	0.14	0.27	-0.42
Percent special education	0.13	0.14	-0.25

\* Because of the data manipulation (i.e., creating missing indicators and replacing missing values with 0) during the matching process, the mean may not reflect the actual number in the original data set.

Table 17 presents the results of the impact analysis for Grades 9–12. Students in higher implementing CSI schools had significantly fewer absences. More specifically, youth attending higher implementing schooling had 61 percent fewer unexcused absences than similar youth attending non-CSI schools. In addition, youth attending higher implementing CSI schools scored significantly higher on the following survey constructs: Academic Personalism, Psychological Sense of Membership, Inquiry-Based Science Instruction, and Student-Teacher Trust compared to non-CSI school students. These survey-related effects were in the small to moderate range, with effect sizes ranging from .24 to .42.

**Table 17. Estimated Effect of Attending a Higher Implementing School—Grades 9–12**

	Estimated Effect	Standard Error
Number of school-day absences	-0.68* (0.39) <sup>1</sup>	0.28
Number of school-day misconducts	0.13 (0.93) <sup>1</sup>	0.36
Number of school-day suspensions	-0.09 (0.51) <sup>1</sup>	0.39
Cumulative weighted GPA	0.14	0.29
Number of credits earned	0.07	0.14
PSAE mathematics	0.16	0.29
PSAE reading	0.25	0.22
ACT mathematics (Standardized)	0.17	0.29
ACT science (Standardized)	0.22	0.25
ACT reading (Standardized)	0.30	0.19
ACT English (Standardized)	0.18	0.26
Emotional Health (Standardized)	0.10	0.13
Academic Engagement	0.13	0.10
Academic Personalism (Standardized)	0.24*	0.11

	Estimated Effect	Standard Error
Human and Social Resources in the Community (Standardized)	0.17	0.17
Parent Support (Standardized)	-0.16	0.11
Psychological Sense of School Membership (Standardized)	0.42*	0.10
Safety (Standardized)	0.12	0.13
Inquiry-Based Science Instruction (Standardized)	0.41*	0.18
Student-Teacher Trust (Standardized)	0.25*	0.09

Note: \* indicates statistically significant at 0.05 level.

<sup>1</sup> Represents an event ratio—treatment/comparison.

## Summary of Findings—Higher Implementing Versus Non-CSI Schools

The analyses undertaken in this report represent one of the first studies to carefully assign community schools to higher implementing and lower implementing categories and assess how youth participation in higher implementing schools impacts education-related outcomes. Part of the innovation here is using a quasi-experimental design that better facilitates causal inferences about the impact of CSI on youth outcomes. Findings summarized in this section of the report include the following:

- The effects of being in a higher implementing school in Grades K–3 appear to have a moderately significant impact on the frequencies of suspensions and misconducts, with fewer occurrences in the higher implementing CSI schools compared to matched non-CSI schools (53 percent and 55 percent fewer, respectively).
- Attending higher implementing schools and participating intensively in afterschool programming and services did not have much of an impact on Grades 4–8 student academic achievement in mathematics and reading and other nonacademic aspects including number of school-day absences, number of misconducts, and number of times suspended. The only significant positive effect found here was that youth in the higher implementing schools scored significantly higher on the Emotional Health scale of the *My Voice, My School* survey. This was a moderate effect (effect size of .43).
- In the upper grades, higher implementing CSI schools had significantly fewer absences (61 percent fewer absences) than similar youth in non-CSI schools. In addition, student survey responses indicated that youth in higher implementing CSI schools provided responses to the *My Voice, My School* survey that were indicative of a more positive school climate: Academic Personalism, Psychological Sense of Membership, Inquiry-Based Science Instruction, and Student-Teacher Trust. These survey-related effects were in the small to moderate range (ranging from .24 to .42).

Again, it is important to note that the analyses conducted in this study were limited to a relatively small number of treatment students who attended higher implementing schools and participated intensively in the afterschool program in the 2011–12 academic year (1086 total, 311 in Grades K–3, 588 in Grades 4–8, and 187 in Grades 9–12). This small treatment sample may limit the statistical power of the study to detect effects. As in any study where random assignment to

treatment and control conditions is not feasible, it is possible that there were unobservable differences between students who attended higher implementing schools and served intensively in the program and those from non-CSI schools that we were unable to control for. To the extent that unobserved differences were related to student outcomes, these differences would bias the estimated differences between higher implementing and non-CSI schools. To that end, these analyses provide initial evidence about the impact of attending higher implementing afterschool programs but should not necessarily be considered equivalent to experimental studies that have stronger internal validity.

## Conclusion and Recommendations

The goal of the analyses undertaken in this report were to (a) develop a method for classifying a sample of community schools active during the 2011–12 school year into higher and lower implementing groups based on criteria outlined in the CSI Implementation Framework and (b) explore how youth participation in CSI programming provided in higher implementing schools impacted a wide variety of youth outcomes.

The key findings resulting from these efforts fall within three general categories:

1. A subset of measures related to elements represented in the CSI Implementation Framework was found to be especially related to the value of programming provided through leveraged partnerships. The concept of leveraged partnerships is critical to successful CSI implementation, and having some information to connect certain measures of implementation with more successful outcomes in this regard is an important development. Practices referenced by how they appear in the CSI Self-Assessment Quality Improvement Rubric especially correlated with the value of leveraged partnerships are outlined as follows:

- Element 1A: Shared Goals and Expectations
- Element 1B: Needs and Resources Assessment
- Element 1C: RC Received Enough Job Training
- Element 1C: RC Receives Support in Performing His or Her Job
- Element 1D: Advisory Committee Role in Decision Making
- Element 2A–B: Adoption of Practices Supporting Youth Development
- Element 2E: Absence of Staffing Challenges
- Element 3C–D: Number of professional development events focused on adoption of the community school framework
- Element 3C–D: Presence of professional learning community

Steps should be taken as part of future evaluation efforts to more closely measure these elements and explore how they may serve to drive the creation of partnerships that bring additional programming and services into community schools that are aligned with the needs of the student population and their families.

2. Analyses oriented at assessing whether youth enrolled in higher implementing schools demonstrated better outcomes than youth enrolled in lower implementing schools largely showed no significant difference, except in terms of school-day absences for youth in Grades 9–12 where youth enrolled in higher implementing schools demonstrated significantly fewer absences. This finding parallels findings from the 2007–08 evaluation where the most substantive outcome associated with participation in CSI programming related to reductions in school-day absences.
3. Analyses that explored the effect of participating intensively in higher implementing schools relative to youth enrolled in similar non-CSI schools were more promising. Not

only was the effect related to fewer school-day absences replicated for youth in Grades 9–12 from higher implementing schools (where youth enrolled in higher implementing schools having 61 percent fewer absences), but participation in higher implementing schools also was associated with higher scores on a variety of scales related to the *My Voice, My School* survey, including Academic Personalism, Psychological Sense of Membership, Inquiry-Based Science Instruction, and Student-Teacher Trust (with effects being in the small to moderate range). In this sense, youth in Grades 9–12 enrolled in higher implementing community schools were found to be having a more positive educational experience than similar youth in non-CSI schools on these scales. In addition, youth in Grades K–3 enrolled in higher implementing schools had significantly fewer suspensions and misconducts than similar youth in non-CSI schools (53 and 55 percent fewer, respectively), while youth in Grades 4–8 demonstrated higher scores on the Emotional Health scale of the survey than their counterparts from non-CSI schools.

Many of the findings related to youth outcomes suggest the benefits associated with participating in a higher implementing community school (based on how implementation has been measured to date) are more likely to be witnessed on youth behaviors (like absences, misconducts, and suspensions) and experiences in school as opposed to substantial effects on academic achievement, at least in the one- to two-year period of participation in CSI programming examined in this report. Increasingly, however, attention is being paid to the role youth behavior, school bonding, and social and emotional functioning play in ensuring youth success, both in school and in life more generally, so the positive effects noted here are important and relevant.

In addition, there are at least three limitations associated with the current study that should try to be addressed in future evaluation efforts:

1. *More careful measurement of implementation at the point-of-service.* No steps were taken during the study conducted here to observe programming in order to assess the level of point-of-service quality. This must be done in the future to validate the functioning of higher implementing schools in providing supportive, interactive, and engaging programming for youth, which must be present if programming is going to have an optimal impact on participating youth.
2. *Steps should be taken to increase the sample size.* An argument can be made that some of the analyses performed in this report were underpowered to detect significant effects. This is especially the case for the analyses done with students in Grades 9–12. Future analyses should focus on a larger sample of higher implementing schools if possible.
3. *Examine youth outcome trajectories over a longer period of time.* Because most of the positive effects detected in the study were associated with changes in youth behaviors and experiences in school, it seems appropriate to try to explore the cumulative effect of supporting the development of these skills and experiences on academic achievement several years out from the youth's participation in CSI programming, particularly among those youth that participate in programming throughout a number of years. It may be the case that achievement-related outcomes may be more likely to be present under such a scenario as opposed to looking for such effects only after a year or two of participation in the programming.

Finally, when key stakeholders associated with the initiative articulate what they have witnessed in terms of how community schools have a positive impact on the lives of enrolled youth and their families, what is often referenced are a series of youth development outcomes (e.g., self-efficacy, self-management and regulation, social skills, positive mind-sets, etc.) that are not represented in the school-related youth outcome data or the school climate data explored in this report. There is a substantial need to better measure and document how CSI is impacting these important components of youth development. Only when this is done well and carefully will the full impact of the Community Schools Initiative in Chicago truly be understood.

## References

- Akiva, T., Pearson, L., Sugar, S. A., Peck, S. C., Smith, C., & Denault, A. (2010). *Linking after-school instructional practices to youth engagement: A pattern-centered approach*. Paper presented at the Society for Research on Adolescence Biennial Meeting.
- Bifulco, R. (2012). Can nonexperimental 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.
- Fredricks, J. A., & Eccles, J. S. (2006). Extracurricular involvement and adolescent adjustment: Impact of duration, number of activities, and breadth of participation. *Applied Developmental Science*, 10(3), 132–146.
- Huang, D., Coordt, A., La Torre, D., Leon, S., Miyoshi, J., Pérez, P., et al. (2007). *The afterschool hours: Examining the relationship between afterschool staff-based social capital and student engagement in LA's BEST* (CSE Technical Report No. 712). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Center for the Study of Evaluation, Graduate School of Education and Information Studies, University of California.
- McCrary, J., & Hoffman, A. (2010). *Chicago Public Schools Community Schools Initiative: 2007–08 student outcomes evaluation interim technical report*. Atlanta, GA: WestEd.
- Pierce, K. M., Auger, A., & Vandell, D. L. (2013, April). *Narrowing the achievement gap: Consistency and intensity of structured activities during elementary school*. Paper presented at the Society for Research in Child Development Biennial Meeting, Seattle, WA.
- Vinson, M. (2010). *Chicago Public Schools Community Schools Initiative: Analysis of resources leveraged by lead partner agencies—2009–2010*. Naperville, IL: Learning Point Associates.
- Whalen, S. P. (2007). *Three years in to Chicago's Community Schools Initiative: Progress, challenges, and emerging lessons*. Chicago, IL: University of Illinois at Chicago.
- Whalen, S. P., Fujimoto, K., & Xiong, Y. (2008). *Hours of OST participation and developmental benefits for student participants in the Chicago Public Schools' Community Schools Initiative (CSI): A report of the UIC Community Schools Evaluation Project*. Chicago, IL: College of Education, University of Illinois at Chicago.
- What Works Clearinghouse. (2008). *Procedures and standards handbook (version 2)*. Washington, DC: Author. Retrieved from [http://ies.ed.gov/ncee/wwc/pdf/reference\\_resources/wwc\\_procedures\\_v2\\_standards\\_handbook.pdf](http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v2_standards_handbook.pdf)

Zander, K., Burnside, E., & Poff, M. (2010). *The development of an implementation and sustainability process strategy (ISPS) for the Chicago Public Schools Community Schools Initiative: Findings and recommendations*. Chicago, IL: Chicago Public Schools.

## ABOUT AMERICAN INSTITUTES FOR RESEARCH

Established in 1946, with headquarters in Washington, D.C., American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally. As one of the largest behavioral and social science research organizations in the world, AIR is committed to empowering communities and institutions with innovative solutions to the most critical challenges in education, health, workforce, and international development.

## LOCATIONS

### Domestic

Washington, D.C.  
Atlanta, GA  
Baltimore, MD  
Chapel Hill, NC  
Chicago, IL  
Columbus, OH  
Frederick, MD  
Honolulu, HI  
Indianapolis, IN  
Naperville, IL  
New York, NY  
Rockville, MD  
Sacramento, CA  
San Mateo, CA  
Waltham, MA

### International

Egypt  
Honduras  
Ivory Coast  
Kyrgyzstan  
Liberia  
Tajikistan  
Zambia



AMERICAN INSTITUTES FOR RESEARCH®

1120 East Diehl Road, Suite 200  
Naperville, IL 60563-1486  
630.649.6500 | 800.356.2735

**[www.air.org](http://www.air.org)**

*Making Research Relevant*