making a meaningful difference

January 2014

SUBMITTED TO

Mary Schearer, PhD

Director of Research

Classroom, Inc.

SUBMITTED BY

Michael Scuello, Senior Associate

Artis Bergman, Research Associate

Transition to High School

Outcome Study

metis associates

120 Wall Street
21st Floor
New York, New York 10005
212-425-8833
www.metisassoc.com

Table of Contents

Introduction

Classroom, Inc., (CI) a NYC-based non-profit educational organization, is a pioneer in the creation of virtual workplaces, where students learn the direct link between learning in school and a successful future. Currently providing programs that use ten virtual workplaces, CI has been implemented in over 20 states and over 1,000 individual sites, serving over 700,000 students.[1] CI's mission is to close the achievement gap by using technology and the world of work to engage, teach and inspire adolescents.

In 2010, CI developed a program focused on helping low-income students transition from middle school to high school. Supported by the lead New York Life Foundation and other generous supporters, this program—the Transition to High School (TTHS) initiative—aims to "strengthen the academic skills of low-achieving students before and during their transition to high school, so that they are prepared for high school level work." [1]

Overall, Classroom, Inc.'s TTHS program is designed to help rising eighth-grade students:

- Increase high school readiness,
- Increase literacy skills,
- Increase workplace readiness skills, and
- Support Common Core-aligned content area instruction in reading, math, science, and social studies.

Participation in the TTHS initiative primarily consisted of the use of teacher-guided interdisciplinary computer-based "simulations" of workplace environments. Each simulation includes a series of episodes that use a particular simulated workplace environment to build competency in core academic areas and develop workplace readiness.[1] For example, a student might be presented with a simulation where they play the role of a lawyer in a "small, general practice law firm" through which, they would learn "about civil and criminal situations and explore American democracy and legal principles"[2]. Students must use academic skills as an adult would in the workplace – on the computer, on individually completed worksheets and through team activities – within the context of tackling scenarios that typically occur in the law firm.

During the summer of 2011 and the 2011–2012 school year, the TTHS program was used in six New York City public high schools. Students at these schools participated either during the summer 2011 only, during both the summer and the school year, or during the 2011–2012 school year only. Some brief notes about the differing characteristics of summer and school-year implementation follow to help interpretation of results.

Summer implementation for CI is generally very different from school-year implementations. The summer program is intense, short-term, continuous, and focused—and a much greater proportion

[1] http://classroominc.org
[2] http://classroominc.org/curriculum/our-simulations

of the available instructional hours are spent on CI. Often, CI is the only instructional program to which summer students are exposed. For these reasons, CI has historically found more regular, reliable test score gains in the summer than during the school year.

During the school year, it is often a challenge for schools to fit CI into the high school schedule. As a result, there is often more varied scheduling and much more time between uses of the program (e.g., when only used for one period a week or every other week), making it more challenging for students to experience continuity and focus. In addition, during the school year, CI competes with other ELA programs and teachers and students are pressured to use school-specific curriculum while preparing for state-wide standardized tests. Furthermore, for TTHS school-year students in this study, two of six schools elected to use CI's very brief *The Finance Center* simulation, which represents only about five instructional hours—a very different experience than the full summer programs.

In the summer of 2011, Classroom, Inc. hired Metis Associates (Metis), a national research and evaluation firm, to conduct a rigorous outcome study of the TTHS program to determine if the intended outcomes were being achieved, and if so, whether there were differences in outcomes between students who participated in the summer only, both the summer and school year, or the school year only. The outcomes Metis examined were students' academic performance and high school attendance. There is a considerable body of evidence from educational researchers that both academic performance and attendance in the freshman year of high school are predictors of students' persistence in high school versus dropping out.

To assess the possible effects of TTHS programming on student achievement, two distinct studies were conducted, each with a different purpose and using different outcomes measures. The first, a rigorous quasi-experimental comparison group design (QED), was conducted on high school outcomes available for both participating students and prospective comparisons (i.e., students not participating in the TTHS program). The intention of this analysis was to determine whether TTHS participation resulted in better achievement results via comparison to similarly situated students not receiving TTHS intervention.

The second study measured student performance based on assessments purchased by CI and administered to participating students prior to and after program attendance. Essentially a within-group pretest/posttest design, the additional data on participants permitted further disaggregation of results to include comparisons among subgroups of participants, such as by program site. The purpose of this study was to determine if academic growth occurred among TTHS participants after they experienced the program.

The following sections describe each methodology in greater detail and present each study's findings, as well as conclusions and recommendations.

Quasi-Experimental Design

Participant rosters for summer 2011 and school year 2011–2012 were provided to Metis by Classroom, Inc. Metis combined these rosters to form three (3) distinct subgroups denoting the amount of TTHS "treatment" received: *summer only* (students present only on the summer roster), *summer and school year* (students present on both rosters), and *school year only* (students present only on the school year roster). Participant information was matched to the New York City Department of Education (NYC DOE) database to obtain student demographics, school attendance records, and achievement data for the TTHS students. Of the 1,245 unduplicated records in the TTHS rosters, Metis was able to match 1,236 students to the NYC DOE data (a 99% match rate).

To determine the possible impacts of TTHS programming on student achievement and attendance, Metis employed a rigorous propensity score matching (PSM) approach to generate an equivalent or closely matched comparison group to the TTHS students from the entire public school pool of NYC DOE 8th grade students in school year 2010–2011[3]. The citywide PSM approach was capable of achieving a high level of internal validity. Thus, treatment and comparison students were balanced on observed covariates so that the net differences observed in outcomes could be attributed to TTHS programming. In other words, PSM ensured that the treatment and comparison groups were statistically equivalent on each of the following covariates obtained from DOE 8th grade data[4]:

- Gender
- Ethnicity
- Eligibility for free/reduced price lunch
- Special education status
- English Language Learner (ELL) status
- Average daily attendance
- NYS English Language Arts (ELA) achievement
- NYS Mathematics achievement
- School borough
- Summer school enrollment between 8th and 9th grade
- Years overage for grade in 8th grade

The first five covariates above are No Child Left Behind (NCLB) subgroups. Metis also used ELA and Math 8th grade achievement, attendance data, borough and student age to ensure matched students were similar to TTHS students prior to summer school. While details about the types of programs the matched comparison students attended during the summer were unknown, Metis

[3] Under the PSM framework (Rosenbaum & Rubin, 1983, 1985; Rosenbaum, 2002), a propensity score for a participant is the conditional probability of assignment to a particular treatment versus non-treatment given the observed covariates.
[4] Note that data from the 8th grade were used to establish baseline equivalence of the two groups of students.

ensured that the students – like the CI students – were attending summer school for academic enrichment[5]. All covariates were used to control for the possible confounding effects they might have on the measured outcomes.

PSMs conducted for reading and attendance/credit accumulation outcomes are shown below in Table 1. Of the 1,236 TTHS students for whom records were found in the NYC DOE database, 1,100 (89%) were successfully matched to a comparison student based on the availability of outcomes. Although there were some slight differences in the key matching variables between the groups, none of these differences were statistically significant.

Table 1 – Baseline Characteristics, TTHS Participants and Matched Comparison Students

Characteristics		TTHS (N = 1,100)		Comparison (N = 1,100)	
		N	%	N	%
Gender	Male	607	55.2%	642	58.4%
	Female	493	44.8%	458	41.6%
Ethnicity	Hispanic	478	43.5%	482	43.8%
	Black, not of Hispanic Origin	342	31.1%	359	32.6%
	Other Ethnicities	280	25.5%	259	23.5%
English Language Learner (ELL) SY2010-2011	No	1,024	93.1%	1,024	93.1%
	Yes	76	6.9%	76	6.9%
Special Education Student SY2010-2011	No	1,031	93.7%	1,037	94.3%
	Yes	69	6.3%	63	5.7%
Eligibility for Free/Reduced Price Lunch SY2010-2011	Not eligible	429	39.0%	439	39.9%
	Eligible	671	61.0%	661	60.1%
Overage for grade in SY2010-2011	Not overage	848	77.1%	859	78.1%
	One year	209	19.0%	207	18.8%
	Two years	38	3.5%	31	2.8%
	Three years	5	0.5%	3	0.3%
School Borough SY2010-2011	Brooklyn	187	17.0%	177	16.1%
	Manhattan	9	0.9%	12	1.1%
	Queens	886	80.5%	899	81.7%
	Staten Island	1	0.1%	0	0.0%
	Bronx	17	1.5%	12	1.1%
"Treatment" Subgroup	Summer Only	207	18.8%	207	18.8%
	Summer and School Year	151	13.7%	151	13.7%
	School Year Only	742	67.5%	742	67.5%
SY2010-2011 (8th Grade) Average Daily Attendance (ADA)		1,100	90.9%	1,100	90.6%

[5] Possible comparison students for the summer sessions were selected based on having attempted mathematics and/or ELA credits. Possible comparisons of students attending summer school but attempting *non-academic* credits only (e.g., physical education) were eliminated from the comparison pool.

Characteristics	TTHS (N = 1,100)		Comparison (N = 1,100)	
	N	Scale Score (Level)	N	Scale Score (Level)
Mean 8th Grade NYS ELA Score (Mean ELA Performance Level)	1,100	648.1 (2)	1,100	647.8 (2)
Mean 8th Grade NYS Math Score (Mean Math Performance Level)	1,100	671.4 (2)	1,100	671 (2)

Analyses of the covariates were also done separately for the three treatment subgroups. Those analyses, which appear in Appendix A, revealed a number of differences between the TTHS participants and the matched comparison students. It is important to note that for both the Summer and Summer/School Year subgroups, the TTHS samples had a higher proportion of Special Education students and students eligible for free/reduced lunch:

- Summer only: There was a higher proportion of: Special Education students in the TTHS group (8.7%) than the comparison group (2.9%); more TTHS students eligible for free/reduced lunch (54.1%) than in comparison (46.4%); and greater Average Daily Attendance (ADA) in 8th grade for TTHS (94.3%) than comparison (90.9%). Differences between groups were only significant for the proportion of special education students, 8th grade ADA and 8th grade mathematics scores.

- Summer and School Year: There was a lower proportion of males in the TTHS sample (47.7%) than the comparison sample (55%); more Special Education students in the TTHS sample (12.6%) than the comparison (8.6%); a much greater proportion of TTHS students eligible for free/reduced price lunch (62.9%) than the comparison group (49.7%); and a higher ADA in 8th grade for TTHS (93.1%) than for comparison students (90.8%). Differences between groups were only significant for the proportion of students eligible for free/reduced price lunch and 8th grade ADA.

- Differences between the School Year Only TTHS and comparison students on the covariates were negligible.

After the comparison groups were successfully generated, outcomes were analyzed using independent *t*-tests and multiple regression analysis. The former provided a global overview of the differences between group outcomes, while the latter was utilized to investigate the relationship between outcomes and their potential predictors. Multiple regression analyses were selected because of their ability to explain the variation in a given outcome based on a set of predictors. As it can be assumed that intervention group membership would not be the only predictor of outcomes, the multiple regressions provided a more in-depth look at the interaction between other factors, such as subgroup characteristics, prior performance, and participation in the program, to successfully predict an outcome.

Outcome analysis

In collaboration with CI, Metis selected several outcome measures to determine the possible effects of TTHS programming on high school readiness and English Language Arts skills. High school

credit accumulation was selected as an important measure of academic performance, as students were required to pass courses to obtain the credits; students who accumulate more credits are assumed to be faring better academically than students who accumulate fewer credits. Further, there is a substantial body of research evidence that the amount of credits accumulated in 9th grade is a strong predictor of academic persistence and successful graduation from high school.

For attendance metrics, summer term and full school year attendance were selected to determine whether TTHS participants attended school more days than their comparison counterparts. An additional attendance metric representing the first 40 days of the school year (i.e., September and October) was included to determine whether possible attendance effects were temporary or sustained throughout the school year. Attendance for the first 30[6] days of high school has also been viewed as an important indicator of persistence toward high school graduation.

Because there are no 9th grade ELA standardized exams in NYC, and since TTHS was literacy-focused, the QED focused on other performance variables, while the second study presented below addresses performance in reading.

Three distinct treatment and comparison subgroups (*summer only*, *summer and school year*, *school year only*) were created and the metrics were selected based on the time period for which they were measured. For example, summer credit accumulation was only compared for TTHS students who attended during the summer and their matched peers. In total, five outcomes were examined, listed in Table 2 by treatment subgroup:

Table 2 – Summer 2011 and School Year 2011–2012 outcomes examined by treatment subgroup

Outcome	Summer Only	Summer and School Year	School Year Only
Average Daily Attendance Summer 2011	✓	✓	
Average Daily Attendance 1st 40 days 2011-2012	✓	✓	✓
Average Daily Attendance 2011-2012	✓	✓	✓
Credits Earned Summer 2011	✓	✓	
Credits Earned 2011-2012	✓	✓	✓

[6] NYC DOE attendance data was available only for the first 40 days, not the 30 days used in national research studies.

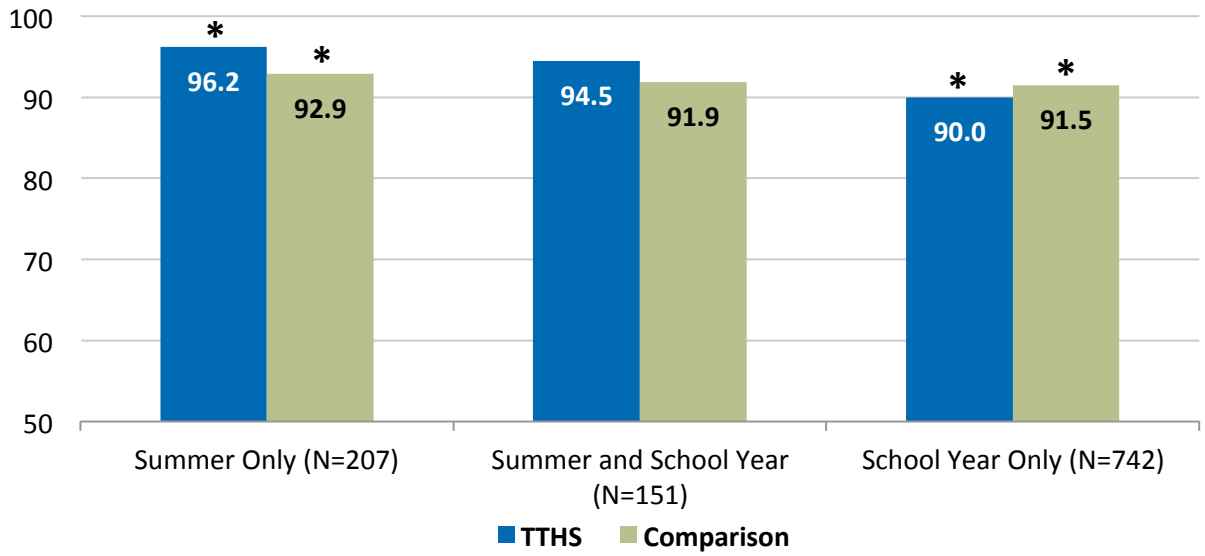
Findings from the Quasi-Experimental Study

This section presents the evaluation findings for the QED study by outcome and treatment subgroup. In the figures, an asterisk above the bars denotes a statistically significant finding. The overall number of matched students for each group presented in these figures is: *Summer Only* (N=207), *Summer and School Year* (N=151), and *School Year Only* (N=742).

Analyses of Attendance Outcomes

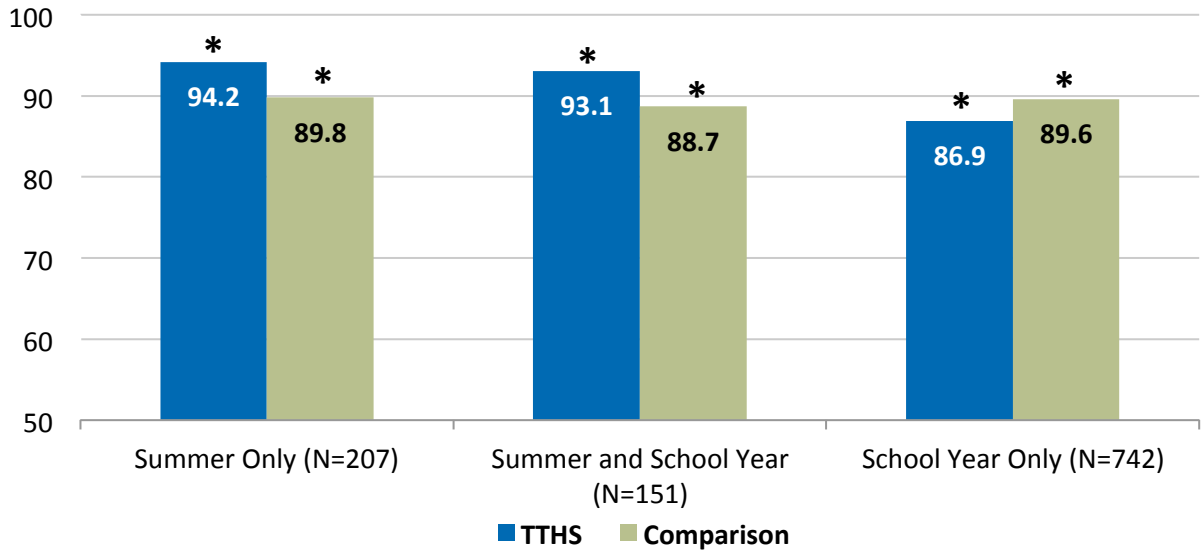
As seen in Figure 1, looking at the first 40 days of attendance during the school year, students participating in TTHS during the *summer only* and *summer and school year* had higher attendance rates than their comparison counterparts. *School year only* TTHS students attended school slightly less often than comparison students during the first 40 days of the school year. Differences for the *summer only* and *school year only* groups were statistically significant.

Figure 1 – Average Daily Attendance (ADA) for 1st 40 days of School Year 2011–2012, by Participant Group and Treatment Subgroup



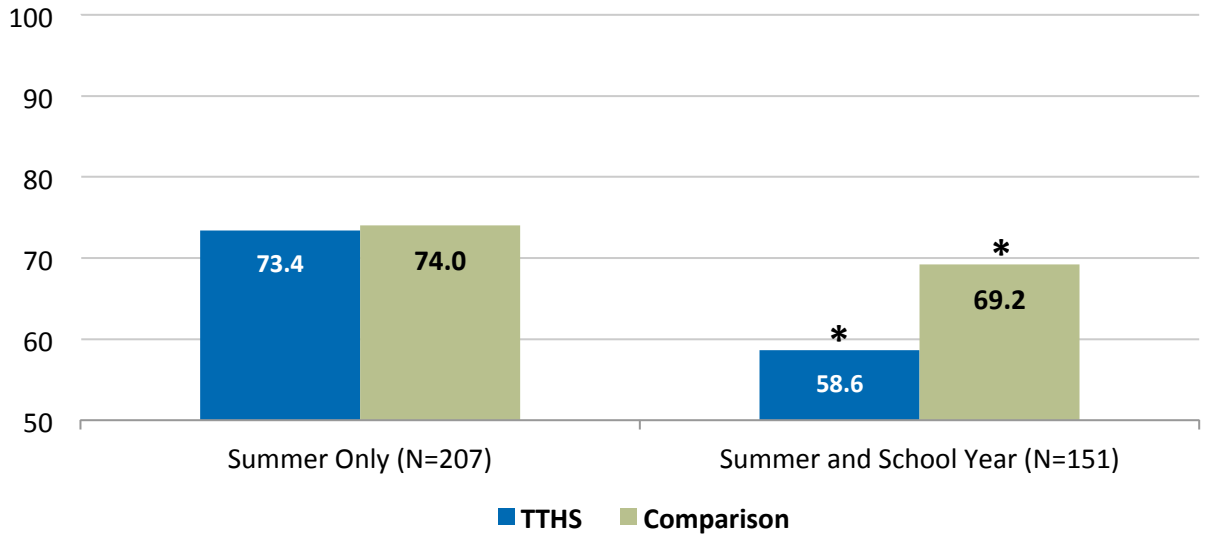
A similar pattern was observed when looking at results for school year attendance. TTHS participants during the *summer only* and *summer and school year* maintained higher school year attendance rates than their comparison peers, but students participating in TTHS activities during the *school year only* attended school less than their comparison counterparts. Differences for all three groups were statistically significant.

Figure 2 –Average Daily Attendance (ADA) for School Year 2011–2012, by Participant Group and Treatment Subgroup



As seen in Figure 3, summer school attendance rates showed that students enrolled in the TTHS program during the *summer and school year* attended significantly less than comparisons, although no significant difference in summer attendance was observed between TTHS students and their counterparts who attended during the *summer only*. To better understand why this discrepancy in attendance rates occurred, Metis looked at additional comparison group characteristics. Over two-thirds (69%) of the comparison group's students either attended summer school in their own middle school or another school that was not a high school, strongly suggesting these students were *mandated* to attend summer school as a precondition for graduation from 8th grade – a likely contributor to increased attendance. In contrast, all students in TTHS attended on a voluntary basis.

Figure 3 – Summer 2011 Average Daily Attendance, by Participant Group and Treatment Subgroup*

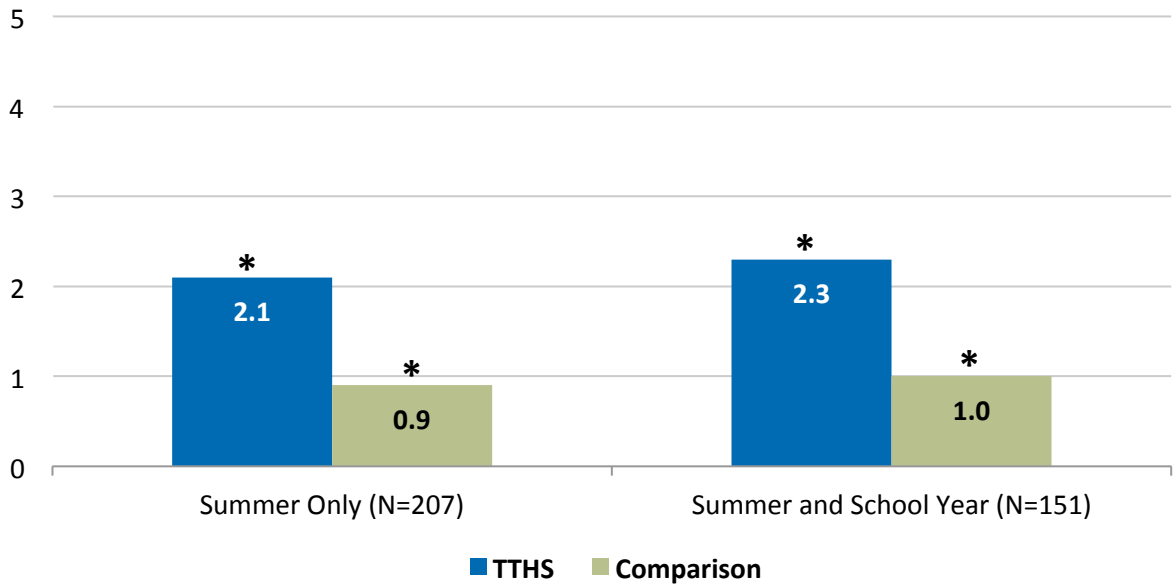


* Only available for *Summer Only and Summer and School Year* participants.

Analyses of Achievement Outcomes

When comparing credits earned in summer 2011, both *summer only* and *summer and school year* TTHS groups significantly outperformed their comparison counterparts, earning at least one credit more on average (Figure 4). This is consistent with the TTHS program where students could earn *high school credits* in the summer (e.g., elective or ELA credit). In contrast, those 8th grade comparison students mandated to attend summer school would not be earning high school credits but instead, were making up required 8th grade work. The rest of the students in the comparison sample for summer school were from academic programs across a wide array of high schools and there were no available data about the credit or attendance policies.

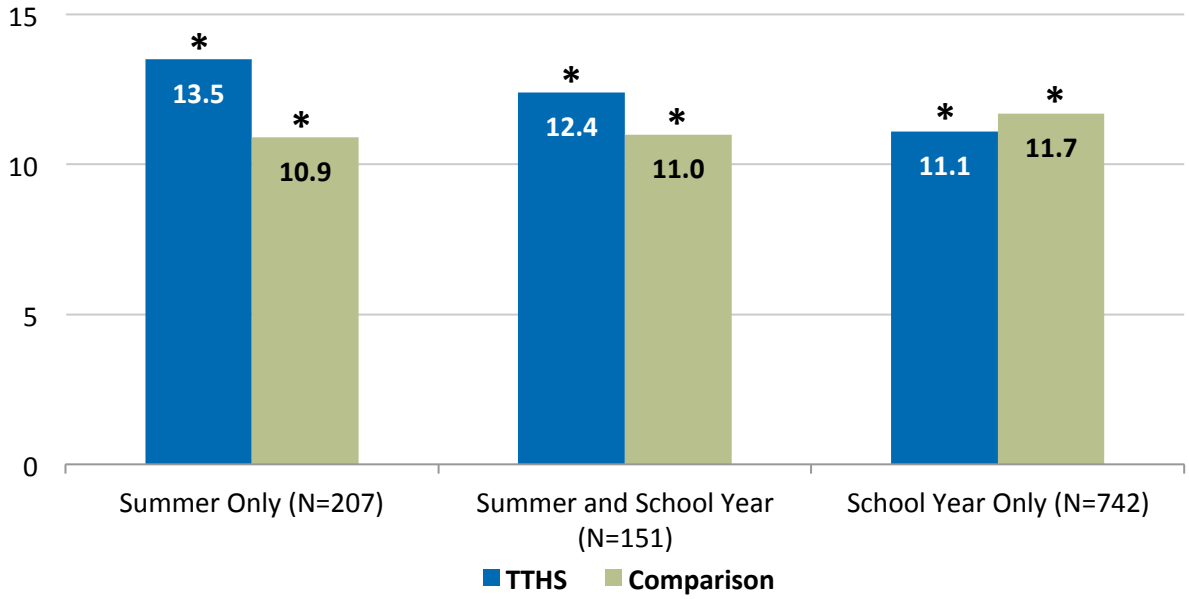
Figure 4 – Credits Earned in Summer 2011, by Participant Group and Treatment Subgroup*



* Only available for *Summer and School Year* participants and *Summer Only* participants.

For the academic year achievement metrics, mixed results were again observed between TTHS students and comparisons. Figure 5 shows that during school year 2011–2012, TTHS students participating during the *summer only* and *summer and school year* earned significantly more credits on average than their comparison counterparts, but students participating in TTHS during the *school year only* earned 0.6 credits less on average than comparison students, also a statistically significant difference.

Figure 5 – Credits Earned in School Year 2011–2012, by Participant Group and Treatment Subgroup



Summary of t-Test Results

Table 3 summarizes the results of all of the *t*-tests for each outcome by treatment subgroup. Significant results have been highlighted in green to denote positive results and pink to denote negative results. Of particular note is the lack of relationship between summer attendance and credit accumulation for the *summer and school year* participants—the *summer and school year* TTHS group attended summer sessions less frequently than their comparison group counterparts, yet they earned more credits. This finding is not unexpected because most TTHS participants, based on prior agreement with the participating high school, were able to earn elective or ELA high school credits based on their performance. Attendance at TTHS was voluntary. As mentioned earlier, we can infer that over two-thirds of the comparison students were mandated to attend summer school. That typically means they are making up 8th grade course work needed to graduate from middle school so they would have high attendance but would not earn high school credits. Irrespective of attendance, note that the *summer only* TTHS participants also earned more credits than their comparisons.

Table 3 – *t*-Test Results by Treatment Subgroup

Treatment Subgroup	Outcome	TTHS		Comparison		*p-value* [a]
		N	Mean	N	Mean	
Summer Only	Average Daily Attendance Summer 2011	170	73.4	207	74.0	0.789
	Average Daily Attendance 1st 40 days SY 2012	207	96.2	207	92.9	0.002*
	Average Daily Attendance SY 2012	207	94.2	207	89.8	0.001*
	Credits Earned Summer 2011	140	2.1	155	0.9	0.000*
	Credits Earned 2012	207	13.5	207	10.9	0.000*
Summer and School Year	Average Daily Attendance Summer 2011	108	58.6	149	69.2	0.001*
	Average Daily Attendance 1st 40 days SY 2012	149	94.5	149	91.9	0.088
	Average Daily Attendance SY 2012	149	93.1	149	88.7	0.011*
	Credits Earned Summer 2011	111	2.3	105	1.0	0.000*
	Credits Earned 2012	150	12.4	150	11.0	0.004*
School Year Only	Average Daily Attendance 1st 40 days SY 2012	735	90.0	733	91.5	0.036*
	Average Daily Attendance SY 2012	736	86.9	736	89.6	0.001*
	Credits Earned 2012	740	11.1	740	11.7	0.006*

[a] An asterisk (*) in this column denotes statistical significance of *t*-test statistic, indicating that the difference between the TTHS group and comparison group is statistically significant.

While the *t*-tests provided an overview of the outcome differences between groups, the regression analyses should provide a more detailed account of the other factors that may have an effect on attendance and achievement, and whether accounting for these factors would provide the same conclusions.

Results of Multiple Regression Analyses

Separate stepwise multiple regression[7] analyses were carried out for each subgroup for the pertinent attendance and achievement outcomes. Table 4 presents the predictors included in each regression model, along with variable values for the categorical predictors. Note that throughout these analyses, group (TTHS or comparison) is the primary predictor of interest. When coefficient values for group are positive, this denotes that TTHS is a significant predictor of positive performance. Conversely, when coefficient values for group are negative, TTHS can be seen as a significant predictor of negative or poorer performance.

Table 4 – Regression Predictor Variables, by Subgroup

Variable	Treatment Subgroup			Variable Values
	Summer Only	Summer and School Year	School Year Only	
Group	✓	✓	✓	0=Comparison, 1=TTHS
ELA Scale Score 2011	✓	✓	✓	
Math Scale Score 2011	✓	✓	✓	
Gender	✓	✓	✓	1=Male, 2=Female

[7] Stepwise regression methods only include statistically significant predictors in the final data models.

Variable	Treatment Subgroup			Variable Values
	Summer Only	Summer and School Year	School Year Only	
Black vs Other Ethnicities	✓	✓	✓	0=Other, 1=Black
Hispanic vs Other Ethnicities	✓	✓	✓	0=Other, 1=Hispanic
Free and Reduced Lunch Flag 2011	✓	✓	✓	0=Not eligible, 1=Eligible
ELL Flag 2011	✓	✓	✓	0=Not ELL, 1=ELL
Special Ed Flag 2011	✓	✓	✓	0=Not Special Ed, 1=Special Ed
Average Daily Attendance 2011	✓	✓	✓	

Tables 5 through 9 present the results based on the final regression models retained following stepwise procedures[8]. For each table, the R Square (R^2) value shows the proportion of variation in the outcome explained by the model. The R Square change values show the percent of variation in the outcome explained by adding a given predictor to the model that had already included a subset of predictors in the stepwise procedure. Finally, the standardized coefficient betas represent the change in outcome standard deviation units as a result of a one standard deviation change in the predictor while all other predictors are held constant. Note that all standardized beta coefficients presented in these tables are statistically significant with *p*-values less than 0.05.

From Table 5, it can be seen that while several variables combined to significantly predict approximately half of the variation in school year attendance for each of the treatment subgroups, TTHS participation (group) appeared as a significant negative predictor of attendance just for the *school year only* group. As expected, prior school year attendance was the strongest predictor of subsequent school year attendance, followed by a variety of other factors including prior achievement, free and reduced lunch eligibility, and ethnicity.

Table 5 – Regression Results for School Year 2011–2012 ADA

Significant Predictors	Standardized Coefficient Beta [a]	R Square Change	R Square
Summer Only (N=414)			
Average Daily Attendance 2011	0.730	0.541	0.547
Black vs Other Ethnicities	-0.078	0.006	
Summer and School Year (N=301)			
Average Daily Attendance 2011	0.688	0.528	0.537
Math Scale Score 2011	0.105	0.009	
School Year Only (N=1,479)			
Average Daily Attendance 2011	0.614	0.421	0.469
ELA Scale Score 2011	0.068	0.009	

[8] Note that while regression analyses were conducted for summer 2011 attendance, the R Square values of the models were small (less than 0.1), suggesting that the variable set used as predictors were not a good fit for predicting the outcomes. Thus, results for these analyses are not presented.

Significant Predictors	Standardized Coefficient Beta [a]	R Square Change	R Square
Group	-0.053	0.003	
Black vs Other Ethnicities	-0.044	0.002	
Math Scale Score 2011	0.050	0.002	

[a] The beta coefficients are the standardized regression coefficients and they represent the amount of standard deviation units the dependent variable changes when the independent variable changes one standard deviation and all other independent variables are held constant.

For the outcome of first 40 days attended during the school year, none of the models included group as a significant predictor (Table 6). Again, prior attendance proved to be the strongest predictor, followed by prior achievement, ethnicity, free and reduced price lunch eligibility, and special education status.

Table 6 – Regression Results for School Year 2011–2012 ADA for 1st 40 days

Significant Predictors	Standardized Coefficient Beta [a]	R Square Change	R Square
Summer Only (N=414)			
Average Daily Attendance 2011	0.705	0.497	0.497
Summer and School Year (N=301)			
Average Daily Attendance 2011	0.684	0.478	0.486
ELA Scale Score 2011	0.094	0.008	
School Year Only (N=1,480)			
Average Daily Attendance 2011	0.564	0.337	0.356
ELA Scale Score 2011	0.114	0.012	
Hispanic vs Other Ethnicities	-0.095	0.002	
Black vs Other Ethnicities	-0.071	0.003	
Special Ed Flag 2011	0.044	0.002	

[a] The beta coefficients are the standardized regression coefficients and they represent the amount of standard deviation units the dependent variable changes when the independent variable changes one standard deviation and all other independent variables are held constant.

Table 7 shows that TTHS participation was a positive predictor for the *summer only* subgroup and a negative predictor for the *school year only* subgroup. Other significant predictors included prior attendance and achievement, ethnicity and gender.

Table 7 – Regression Results for Credits Earned in School Year 2011–2012

Significant Predictors	Standardized Coefficient Beta [a]	R Square Change	R Square
Summer Only (N=414)			
Average Daily Attendance 2011	0.375	0.243	0.370
Math Scale Score 2011	0.223	0.062	
Group	0.205	0.040	
Black vs Other Ethnicities	-0.125	0.012	

Significant Predictors	Standardized Coefficient Beta [a]	R Square Change	R Square
Gender	0.112	0.013	
Summer and School Year (N=300)			
Average Daily Attendance 2011	0.387	0.233	0.298
Math Scale Score 2011	0.239	0.048	
Gender	0.133	0.017	
School Year Only (N=1,480)			
Average Daily Attendance 2011	0.389	0.207	0.270
Math Scale Score 2011	0.110	0.034	
ELA Scale Score 2011	0.113	0.011	
Black vs Other Ethnicities	-0.142	0.008	
Gender	0.073	0.005	
Hispanic vs Other Ethnicities	-0.073	0.003	
Group	-0.051	0.002	

[a] The beta coefficients are the standardized regression coefficients and they represent the amount of standard deviation units the dependent variable changes when the independent variable changes one standard deviation and all other independent variables are held constant.

Group did appear as a significant positive predictor of credits earned during the summer for both *summer only* and *summer and school year* groups (Table 8). Also predictive of summer credit accumulation were prior achievement and attendance, special education status, and ethnicity.

Table 8 – Regression Results for Credits Earned in Summer 2011

Significant Predictors	Standardized Coefficient Beta [a]	R Square Change	R Square
Summer Only (N=295)			
Group	0.483	0.221	0.281
Math Scale Score 2011	0.197	0.049	
Black vs Other Ethnicities	-0.112	0.011	
Summer and School Year (N=217)			
Group	0.455	0.252	0.439
Math Scale Score 2011	0.388	0.143	
Special Ed Flag 2011	0.153	0.025	
Average Daily Attendance 2011	0.148	0.019	

[a] The beta coefficients are the standardized regression coefficients and they represent the amount of standard deviation units the dependent variable changes when the independent variable changes one standard deviation and all other independent variables are held constant.

Table 9 presents a summary of the presence of group as a significant predictor in all of the conducted regressions. Positive findings were found for the *summer and school year* and *summer only* subgroups, while negative findings were limited to the *school year only* subgroup.

Table 9 – Presence of Group as a Significant Predictor in Regression Models

Outcome	Summer Only	Summer and School Year	School Year Only
Average Daily Attendance 2012	N	N	Y-
Average Daily Attendance 1st 40 days 2012	N	N	N
Average Daily Attendance Summer 20118	--	--	
Credits Earned 2012	Y+	N	Y-
Credits Earned Summer 2011	Y+	Y+	

N: Not a significant predictor; Y+: Significant positive predictor; Y-: Significant negative predictor

Taken together, the findings from the quasi-experimental study suggest that students participating in TTHS during the summer—irrespective of whether they continue to participate during the school year or not—fare better than those who participate only during the school year. Implementation data may provide some explanation for these differences. A closer look at the participant group via the pretest/posttest design provides insight into academic gains students may have made as a result of TTHS programming.

Within-Group Pretest/Posttest Design

Three student subgroups were created for the purposes of this comparison based on student exposure to CI programming. As with the quasi-experimental design, the groups were defined according to when a student participated in the TTHS program, independent of how many simulation episodes they completed[9]. Table 10 lists the three groups and presents the number of students who submitted matched pre-and post-test data for reading (*Reading-Level Indicator, RLI*).

Table 10 – Definition and Size of Groups – *RLI*

Group		Participation		Population (Sample Size for Analyses)
		Summer 2011	SY 2011-2012	Reading
(A)	Summer only	✓		188
(B)	Summer and school year	✓	✓	94
(C)	School year only		✓	386
			Total TTHS Participants	**668**

[9] Classroom Inc. simulations vary in the number of episodes that a student can complete. For example, *The Sports Network* includes 12 episodes while *The Finance Center* includes 5 episodes. Depending on the site and period of participation, students experienced different simulations.

It should be noted that the three groups varied in the CI simulations they used as well as in what students experienced outside of the program. Students across and within groups also experienced different levels of program dosage—as measured by the number of simulation episodes completed by each student—as well as varied use of the array of TTHS program components, depending upon teachers' assessment of student needs. In addition, a sizeable body of educational research indicates that a measurable level of learning *loss* should have taken place over the summer (relevant only to Group A and Group B). In contrast, students who participated during the school year would be expected to show gains since the vast majority of their instruction in English occurred during the traditional school day with TTHS being only a supplemental part of their learning experiences.

Outcome analysis

To assess student performance in reading, teachers implementing the program administered a reading assessment, the *Reading-Level Indicator (RLI),* before and after the program was delivered following a pretest/posttest design. This assessment was developed by the American Guidance Service, Inc. (now Pearson) and has been documented to be valid and reliable measures of students' reading performance.

To prepare for the pre/post analyses, raw test scores were converted into scale scores following the W-ability scale score conversion provided in the test manual. Paired-samples *t*-tests were conducted on the scale scores to determine whether there were statistically significant differences in students' reading scores from pretest to posttest. Students who responded to less than half of the items on either the pretest or the posttest were excluded from the analyses. The results for the reading assessments are disaggregated first by group and then by school and gender. All students were either rising eighth-graders or in the ninth grade. Effect sizes were also calculated using Cohen's *d* to determine the magnitude of the differences. According to Cohen (1988), effect sizes of 0.2 are generally considered small, 0.5 are medium and 0.8 and higher are large.

Findings from the Within-Group Pretest/Posttest Study

Reading-Level Indicator (RLI) analysis

Composition of groups. From summer 2011 through spring 2012, a total of 668 ninth-grade students participated in one of five TTHS programs and completed both the pretest and posttest reading assessment. Most students included in the *RLI* analyses attended two schools: High Schools C and D. Across all three groups, there were relatively equal numbers of male and female students. Table 11 shows the distribution of students included in the reading analyses by school, gender and group.

Table 11 – Distribution of Students by School, Gender and Group – *RLI*

Characteristic	Group		
	(A) Summer only	(B) Summer and school year	(C) School year only
School (Percent of each group enrolled at each school)			
High School A	-	2.1%	1.8%
High School B	17.6%	3.2%	7.3%
High School C	46.8%	57.4%	40.7%
High School D	12.8%	30.9%	50.3%
High School E	22.9%	6.4%	-
Gender (Percent of each group that is male or female)			
Male	50.5%	47.9%	51.0%
Female	49.5%	52.1%	49.0%

Data in Table 12 confirm that the students who had the TTHS program in both summer and school year also used the largest number of episodes (Mean=11.4). For students attending the TTHS program only during the school year, 71 percent used between six to nine episodes (Mean=8.1) while summer only students averaged 4.7 episodes. However, note that this table does not address the related issue of intensity of the TTHS experience. With summer only students, TTHS would typically be the *only instructional* program used, and, for each episode, students in the summer often use additional program components made available within CI programs. In contrast, during the school year TTHS is usually done in addition to regular ELA curriculum and so while more episodes may be used, the depth to which they are used may well be less than in the summer.

Table 12 – Distribution of Students by Dosage and Group – RLI

Group	Total Simulation Episodes				Dosage Group		
	N	Mean	Min	Max	0 to 5 episodes	6 to 9 episodes	10 or more episodes
(A) Summer only	188	4.7	3	6	77.1%	22.9%	0.0%
(B) Summer and school year	94	11.4[a]	6	20	5.4%	30.1%	64.5%
C) School year only	386	8.1	5	10	3.6%	71.0%	25.4%

[a] Total simulation episodes based on the combination of the two time periods, thus the higher mean number of episodes completed.

To make the *RLI* results more understandable, the average scale scores at pretest and posttest were converted into grade equivalents (GE) of students' instructional reading level.[10] For example, according to the test publisher, a reading scale score at pretest of 118.63 would correspond to an instructional reading level of 7.8, or seventh grade, eighth month, while a reading scale score at posttest of 119.44 would correspond to an instructional reading level of 9.5, or ninth grade, fifth month, an increase in instructional reading level of one year, seven months over the summer. The following tables present *RLI* scale score and effect size data for the TTHS students (Table 13) and corresponding changes in instructional level (Table 14). The same data disaggregated by school and gender are presented in Tables B1 through B4 in the report appendix.

Table 13 – Changes in *RLI* Scale Scores-by Group

Group	Matched N	Pretest Scale Score	Posttest Scale Score	Change in Scores (Posttest-Pretest)	*t* (Sig.)[a]	Effect Size (*Cohen's d*)[b]
A) Summer only	188	118.63	119.44	0.81	*1.626* (0.106)	-
B) Summer and school year[c]	94	119.09	124.06	4.97	*6.955* (0.000)*	0.72
C) School year only	386	119.52	122.61	3.08	*9.236* (0.000)*	0.47

[a] An asterisk in this column denotes a statistically significant difference at the p≤.05 level based on a paired-samples *t*-test.
[b] Effect size is a measure of the magnitude of the gains or losses, expressed in gain score standard deviation units. According to Cohen (1988), effect sizes of 0.2 are considered small, 0.5 are considered medium, and 0.8 are considered large. Effect sizes are only presented for statistically significant differences.
[c]The pre-test *RLI* for this group was administered at the beginning of summer school and the post-test, at the end of the school year. This was the longest time period between pre- and post-testing.

Table 14 – *RLI* Grade Level Equivalencies (Instructional Reading Levels) and Change

Group	Matched N	Pretest Scale Score	Posttest Scale Score	Pretest Instructional Reading Level	Posttest Instructional Reading Level	Change in Instructional Reading Level (months)
A) Summer only	188	118.63	119.44	7.8	9.5	+ 1 year, 7 months
B) Summer and school year	94	119.09	124.06	9.5	12.2	+ 2 years, 7 months
C) School year only	386	119.52	122.61	9.5	10.4	+ 9 months

[10] As described in the publisher's manual, "grade equivalents are referred to as *developmental norms* because they place an individual along a span or continuum of development. Grade-equivalent values are presented in tenths of a grade." To calculate GEs, average scale scores were first converted to raw scores and then into a GE following the publisher's conversion tables provided in the manual. The average scale score at pretest (119.21) corresponds to a raw score of 28; the average scale score at posttest (121.92) corresponds to a raw score of 29. Note that GEs have many limitations. Since they are not equal-interval scales of measurement, they cannot be manipulated arithmetically (e.g., averaged) or used for direct longitudinal comparisons.

From tables 13 and 14, the *RLI* data revealed that all groups showed substantial increases in average instructional reading levels (the measure of grade level equivalency):

- For students who participated during the ***summer only*** (Group A) there was ***no summer learning loss***, which would have been the norm for low-income students such as those participating in TTHS. Further, RLI scale scores increased from pretest to posttest. Although the average change in scale score was not statistically significant, students improved, on average, from an instructional reading level of seventh grade, fifth month, to an instructional reading level of ninth grade, fifth month. The one year, seven month improvement over the summer meant students began the ninth grade in a much better position to read at their grade level, a critical skill to smooth their transition to high school.
- Students who participated during the ***summer and school year*** (Group B) or during the ***school year only*** (Group C) experienced ***significant scale score gains with moderate to large effect sizes,*** respectively. Group B students showed, on average, a large gain of two years, seven months, moving from an average instructional reading level of grade nine, five months at pretest to an instructional reading level of twelfth grade, four months, at posttest. Students in Group C went, on average, from an instructional reading level of ninth grade, fifth month, to an instructional reading level of tenth grade, four months, an increase, on average, of nine months. This is comparable to the level of growth expected during a typical nine-month school year.

Analyses of covariance were conducted to determine whether there were statistically significant differences between these three groups after controlling for pretest scores. As seen in Table 15, differences in posttest scores by group were statistically significant beyond a 0.01 level of probability with an effect size of 1.77, suggesting that the duration and type of exposure to the TTHS program (e.g., school year or summer or both) has had an impact on the reading gains made by students.

Table 15 – Comparison of *RLI* Test Results, by Group

Number of Episodes	Total N	Posttest Adjusted Mean Score [a]	F (Sig.)[b]	Effect size [c]	Post Hoc Comparisons
A) Summer only	188	119.85	17.274 (0.000)*	1.77	Group B > Group A Group B > Group C Group C > Group A
B) Summer and school year	94	124.15			
C) School year only	386	122.39			

[a] Posttest mean scores were adjusted to take into account pretest differences in W-ability scores.
[b] An asterisk denotes a statistically significant difference at the .05 level based on an analysis of covariance.
[c] Effect size is a measure of the magnitude of the gains or losses. Effect sizes of 0.2 are considered small, 0.5 are considered medium, and 0.8 are considered large.

As the post hoc comparisons indicate, students in Group B, who participated in the *summer and school year*, had higher posttest scores than students in the other two groups after taking into account baseline differences. Results also show that students who participated during the *school year only* showed greater gains than students who participated during the *summer only*. This would be expected because school year students had an entire ELA curriculum over a nine-month period in addition to TTHS to boost their RLI performance whereas summer students only had TTHS over a four or five week period as academic enrichment. In sum, students participating in summer 2011 *in addition* to the following school year (2011–2012) were found to make the greatest gains in reading achievement, above and beyond the gains made during the school year alone.

Findings by student characteristics and program site. While an analysis of changes in scale scores by school is limited by small sample sizes for three of the schools, data (see Table B2 in Appendix) indicate that gains and corresponding effect sizes differ by school. Specifically, in relation to the two schools with the highest concentration of students in the *summer and school year* and *school year only* groups (High Schools C and D), High School C shows the highest levels of gains as measured by the RLI.

Results also indicate that overall, gains made by male and female students are largely similar within each group (see Table B3) but that in Groups B and C, where gains were found to be statistically significant, male students do have slightly higher gains than female students.

Conclusions and Recommendations

The quasi-experimental study found a positive impact of TTHS programming on participant academic and attendance outcomes for students who participated during the summer—whether or not they continued to participate during the school year. Both the independent samples *t*-tests and multiple regression analyses revealed positive impacts. These findings suggest that students participating in TTHS during the summer—irrespective of whether or not they continue to

participate during the school year—fare better than those who participate only during the school year. For students who participated *only* during the school year, in comparison to a group of similarly situated students, the results were negative. This may imply that summer programming is required for TTHS to be successful, or it may be due to a difference between populations served, service delivery (e.g., fidelity to service delivery models), or student reaction to services delivered during the summer and school year. For instance, teachers may have been more motivated to use TTHS during summer school since it was typically the ONLY instructional program used, and they were trained to use it with weekly support by CI staff or consultants. Further, the simulated workplaces and learning in context using computers may have been perceived by students as different and more fun than typical instructional programs. In contrast, during the school year, teachers often used TTHS as a supplement in addition to their regular required curriculum and test preparation, and students in the school year across all 9th grade classes have common ELA curriculum tests and standards.

Results from the within-group pretest/posttest study using the *Reading-Level Indicator* (*RLI)* analyses suggest that, overall, TTHS participants experienced significant moderate gains in their reading performance. The average grade level equivalency of students who participated in summer alone increased substantially after participation, a sharp contrast from the typical summer learning loss that occurs for students from low-income neighborhoods. Students who participated in the summer in addition to the school year showed the largest gains in *RLI* scale scores. This indicates that the strongest TTHS intervention begins with a summer program and continues use of CI during the school year as an ongoing supplemental learning experience for freshman in high school. Participating in TTHS during the school year only also had a positive impact on student achievement but not any greater than would be expected over the nine-month period from any ELA curriculum.

Observational data provided by Classroom Inc. indicates that service delivery may have been the key to the differences observed in the quasi-experimental study. While program implementation varied across program sites, the TTHS summer implementation was structured to deliver no fewer than 36 hours and as much as 50 hours of instruction. It was typically the only curriculum students had so teachers could focus on using diverse components and instructional strategies united within the simulated workplace. Summer TTHS provides a concentrated learning experience designed to remind students of the practical value of academic learning. In contrast, more difficulties were reported with implementation during the school year—specifically in the larger program sites—which may have hampered the program's effectiveness, particularly for teachers and students not already familiar with how CI's program works. For example, implementation at one of the high schools started late on a truncated timetable and at another was limited to a single simulation offered during the last period of the day, also on a truncated timetable. Two of the schools chose to use CI's brief five-episode *The Finance Center* simulation because it matched their perceived needs even though it shortened students' exposure to TTHS. Unfortunately, there were also a couple of instances during the school year where teachers had poor classroom management skills that seriously

hampered sound delivery of TTHS. Factors such as these may have contributed to the relative poorer showing by the students participating only during the school year.

Furthermore, although all participants were carefully matched by Metis for comparisons on baseline demographic characteristics, achievement, and attendance, the TTHS sample had a greater proportion of special education and low-income students than the comparison sample in the *summer only* and *summer and school year* subgroups, respectively. In addition, non-cognitive factors such as motivation, problem-solving skills, and confidence in academic ability could not be accounted for due to a lack of data. Students may have therefore been deemed equivalent on most observed characteristics, but other pertinent factors that may have affected performance were not controlled for. While Classroom, Inc. surveys TTHS participants and can obtain some of this information, it would be prohibitively costly to find the same information for comparison students. However, the findings do suggest the benefit of conducting more in-depth analyses of the CI survey data or collecting more systematic implementation data at the time of program delivery to systematically determine whether experiential and/or implementation differences may exist. Survey instruments may also need to be revised to collect more detailed information about students' non-cognitive factors.

In sum, the two methodologies measured impact in different ways. The quasi-experimental study provided a view of how well participants fared on achievement metrics when compared to a similarly situated group of students while the pretest/posttest study provided insight into changes in participant performance over time. Overall, the message from the findings is that TTHS programming can successfully bolster student achievement in reading. Not only was summer learning loss prevented, but students in TTHS summer programs demonstrated important gains in their RLI instructional reading levels and earned credit accepted by their high schools. In fact, all participating students achieved higher instructional reading levels from pretest to posttest as measured by the RLI irrespective of when they participated and implementation difficulties. As mentioned above, perhaps more detail could help to interpret the results from the quasi-experimental study, but they also seem to point to the success of TTHS when the program is implemented in a consistent manner as evidenced by the positive results observed for those students participating during the summer and during the summer and the school year. Finally, the results for the students who participated during the school year only seem to suggest that the summer school component is central to a successful TTHS program.

References

Alexander, K. L., Entwisle D. R., & Olson L. S. (2007a). Lasting consequences of the summer learning gap. *American Sociological Review, 72*, 167-180.

Child Trends (2013). *High school dropout rates.* http://www.childtrends.org/?indicators=high-school-dropout-rates

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, 66, 227-268.

Rosenbaum, P. R. (2002). *Observational studies* (2nd ed.). New York: Springer.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41-55.

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

Rumberger, R. W. (2004). Why students drop out of school. In Gary Orfield (Ed.). *Dropouts in America: Confronting the graduation rate crisis*. Cambridge, MA: Harvard Education Press.

Appendix

Table A1 – Baseline Characteristics, TTHS Participants and Matched Comparison Students Summer Only Subgroup

Characteristics		TTHS (N = 207)		Comparison (N = 207)	
		N	%	N	%
Gender	Male	110	53.1%	115	55.6%
	Female	97	46.9%	92	44.4%
Ethnicity	Hispanic	66	31.9%	71	34.3%
	Black, not of Hispanic Origin	77	37.2%	77	37.2%
	Other Ethnicities	64	30.9%	59	28.5%
English Language Learner (ELL) SY2010-2011	No	198	95.7%	198	95.7%
	Yes	9	4.3%	9	4.3%
Special Education Student SY2010-2011[a]	No	189	91.3%	201	97.1%
	Yes	18	8.7%	6	2.9%
Eligibility for Free/Reduced Price Lunch SY2010-2011	Not eligible	95	45.9%	111	53.6%
	Eligible	112	54.1%	96	46.4%
Overage for grade in SY2010-2011	Not overage	164	79.2%	166	80.2%
	One year	37	17.9%	37	17.9%
	Two years	4	1.9%	3	1.4%
	Three years	2	1.0%	1	0.5%
School Borough SY2010-2011	Brooklyn	25	12.1%	36	17.4%
	Manhattan	3	1.4%	2	1.0%
	Queens	177	85.5%	164	79.2%
	Staten Island	--	0.0%	--	0.0%
	Bronx	2	1.0%	5	2.4%
SY2010-2011 (8th Grade) Average Daily Attendance (ADA)[b]		207	94.3%	207	90.9%
		N	Scale Score (Level)	N	Scale Score (Level)
Mean 8th Grade NYS ELA Score (Mean ELA Performance Level)		207	649.2 (2)	207	647.9 (2)
Mean 8th Grade NYS Math Score (Mean Math Performance Level)[c]		207	674.4 (2)	207	668.2 (2)

[a] Yates $\chi^2(1, N=414) = 5.352$, $p = 0.021$

[b] $t(307.9) = 4.253$, $p < 0.001$

[c] $t(363.3) = 2.373$, $p = 0.018$

Table A2 – Baseline Characteristics, TTHS Participants and Matched Comparison Students Summer and School Year Subgroup

Characteristics		TTHS (N = 151)		Comparison (N = 151)	
		N	%	N	%
Gender	Male	72	47.7%	83	55.0%
	Female	79	52.3%	68	45.0%
Ethnicity	Hispanic	54	35.8%	54	35.8%
	Black, not of Hispanic Origin	56	37.1%	54	35.8%
	Other Ethnicities	41	27.1%	43	28.5%
English Language Learner (ELL) SY2010-2011	No	133	88.1%	133	88.1%
	Yes	18	11.9%	18	11.9%
Special Education Student SY2010-2011	No	132	87.4%	138	91.4%
	Yes	19	12.6%	13	8.6%
Eligibility for Free/Reduced Price Lunch SY2010-2011[a]	Not eligible	56	37.1%	76	50.3%
	Eligible	95	62.9%	75	49.7%
Overage for grade in SY2010-2011	Not overage	119	78.8%	123	81.5%
	One year	26	17.2%	24	15.9%
	Two years	5	3.3%	4	2.6%
	Three years	1	0.7%	--	0.0%
School Borough SY2010-2011	Brooklyn	38	25.2%	45	29.8%
	Manhattan	2	1.3%	1	0.7%
	Queens	111	73.5%	104	68.9%
	Staten Island	--	0.0%	--	0.0%
	Bronx	--	0.0%	1	0.7%
SY2010-2011 (8th Grade) Average Daily Attendance (ADA)[b]		151	93.1%	151	90.8%
		N	Scale Score (Level)	N	Scale Score (Level)
Mean 8th Grade NYS ELA Score (Mean ELA Performance Level)		151	651.0 (2)	151	650.8 (2)
Mean 8th Grade NYS Math Score (Mean Math Performance Level)		151	677.3 (2)	151	672.7 (2)

[a] Yates $\chi^2(1, N=302) = 4.858$, $p = 0.028$

[b] $t(249.7) = 2.016$, $p = 0.045$

Table A3 – Baseline Characteristics, TTHS Participants and Matched Comparison Students School Year Only Subgroup

Characteristics		TTHS (N = 742)		Comparison (N = 742)	
		N	%	N	%
Gender	Male	425	57.3%	444	59.8%
	Female	317	42.7%	298	40.2%
Ethnicity	Hispanic	358	48.2%	357	48.1%
	Black, not of Hispanic Origin	209	28.2%	228	30.7%
	Other Ethnicities	175	23.6%	157	21.2%
English Language Learner (ELL) SY2010-2011	No	693	93.4%	693	93.4%
	Yes	49	6.6%	49	6.6%
Special Education Student SY2010-2011	No	710	95.7%	698	94.1%
	Yes	32	4.3%	44	5.9%
Eligibility for Free/Reduced Price Lunch SY2010-2011	Not eligible	278	37.5%	252	34.0%
	Eligible	464	62.5%	490	66.0%
Overage for grade in SY2010-2011	Not overage	565	76.1%	570	76.8%
	One year	146	19.7%	146	19.7%
	Two years	29	3.9%	24	3.2%
	Three years	2	0.3%	2	0.3%
School Borough SY2010-2011	Brooklyn	114	15.4%	106	14.3%
	Manhattan	7	0.9%	6	0.8%
	Queens	611	82.3%	618	83.3%
	Staten Island	--	0.0%	1	0.1%
	Bronx	10	1.3%	11	1.5%
SY2010-2011 (8th Grade) Average Daily Attendance (ADA)		742	89.4%	742	90.5%
		N	Scale Score (Level)	N	Scale Score (Level)
Mean 8th Grade NYS ELA Score (Mean ELA Performance Level)		742	647.0 (2)	742	647.0 (2)
Mean 8th Grade NYS Math Score (Mean Math Performance Level)		742	669.0 (2)	742	671.0 (2)

Table B1 – *RLI* Grade Level Equivalencies (Instructional Reading Levels) and Change, by School

Group	School	Matched N	Pretest Scale Score	Posttest Scale Score	Pretest Instructional Reading Level	Posttest Instructional Reading Level	Change in Instructional Reading Level (months)
A) Summer only	School B	33	120.33	121.00	9.5	10.4	+ 9 months
	School C	88	118.73	119.65	7.8	9.5	+1 year, 7 months
	School D	24	120.40	119.97	9.5	9.5	no change
	School E	43	116.14	117.52	6.9	7.8	+ 9 months
	Total	**188**	**118.63**	**119.44**	7.8	9.5	+ 1 year, 7 months
B) Summer and school year	School A	2	- [a]	- [a]	- [a]	- [a]	- [a]
	School B	3	- [a]	- [a]	- [a]	- [a]	- [a]
	School C	54	118.80	124.83	7.8	>12.2	+ 4 years, 4 months or more
	School D	29	120.90	123.93	10.4	12.2	+ 1 year, 8 months
	School E	6	- [a]	- [a]	- [a]	- [a]	- [a]
	Total	**94**	**119.09**	**124.06**	9.5	12.2	+ 2 years, 7 months
C) School year only	School A	7	- [a]	- [a]	- [a]	- [a]	- [a]
	School B	28	114.50	117.34	6.3	6.8	+ 5 months
	School C	157	119.84	124.66	9.5	>12.2	+ 2 years, 7 months or more
	School D	194	119.98	121.57	9.5	10.4	+ 9 months
	Total	**386**	**119.52**	**122.61**	9.5	10.4	+ 9 months

[a] Analyses are not reported for groups with fewer than 10 students.

Table B2– Changes in *RLI* Scale Scores, Total and by Program Site and Group

Comparison Group	School	Matched	Pretest Scale Score	Posttest Scale Score	Change in Scores (Posttest-Pretest)	*t* (Sig.)[a]	Effect Size (*Cohen's d*)[b]
A) Summer only	School B	33	120.33	121.00	0.67	*0.65* (0.520)	-
	School C	88	118.73	119.65	0.93	*1.133* (0.260)	-
	School D	24	120.40	119.97	-0.44	*-0.334* (0.742)	-
	School E	43	116.14	117.52	1.37	*1.502* (0.141)	-
	Total	**188**	**118.63**	**119.44**	**0.81**	***1.626* (0.106)**	**-**
B) Summer and school year	School A	2	-[c]	-[c]	-[c]	-[c]	-[c]
	School B	3	-[c]	-[c]	-[c]	-[c]	-[c]
	School C	54	118.80	124.83	6.03	*6.058* (0.000)*	0.82
	School D	29	120.90	123.93	3.03	*2.73* (0.011)*	0.51
	School E	6	-[c]	-[c]	-[c]	-[c]	-[c]
	Total	**94**	**119.09**	**124.06**	**4.97**	***6.955* (0.000)***	**0.72**
C) School year only	School A	7	-[c]	-[c]	-[c]	-[c]	-[c]
	School B	28	114.50	117.34	2.83	*2.825* (0.009)*	0.53
	School C	157	119.84	124.66	4.83	*8.263* (0.000)*	0.66
	School D	194	119.98	121.57	1.59	*3.926* (0.000)*	0.28
	Total	**386**	**119.52**	**122.61**	**3.08**	***9.236* (0.000)***	**0.47**

[a] An asterisk in this column denotes a statistically significant difference at the p≤.05 level based on a paired-samples *t*-test.
[b] Effect size is a measure of the magnitude of the gains or losses, expressed in gain score standard deviation units. According to Cohen (1988), effect sizes of 0.2 are considered small, 0.5 are considered medium, and 0.8 are considered large. Effect sizes are only presented for statistically significant differences.
[c] Analyses were not conducted for groups with fewer than 10 students.

Table B3 – Changes in RLI Scale Scores, Total and by Gender and Group

Comparison Group	Gender	Matched	Pretest Scale Score	Posttest Scale Score	Change in Scores (Posttest-Pretest)	*t* (Sig.)[a]	Effect Size (*Cohen's d*)[b]
A) Summer only	Male	95	118.78	119.39	0.61	*0.780* (0.437)	-
	Female	93	118.48	119.49	1.02	*1.641* (0.104)	-
	Total	**188**	**118.63**	**119.44**	**0.81**	***1.626* (0.106)**	**-**
B) Summer and school year	Male	45	118.52	123.90	5.38	*5.241* (0.000)*	0.78
	Female	49	119.61	124.21	4.60	*4.583* (0.000)*	0.65
	Total	**94**	**119.09**	**124.06**	**4.97**	***6.955* (0.000)***	**0.72**
C) School year only	Male	197	119.29	122.98	3.70	*7.282* (0.000)*	0.52
	Female	189	119.77	122.21	2.44	*5.731* (0.000)*	0.42
	Total	**386**	**119.52**	**122.61**	**3.08**	***9.236* (0.000)***	**0.47**

[a] An asterisk in this column denotes a statistically significant difference at the p≤.05 level based on a paired-samples *t*-test.
[b] Effect size is a measure of the magnitude of the gains or losses, expressed in gain score standard deviation units. According to Cohen (1988), effect sizes of 0.2 are considered small, 0.5 are considered medium, and 0.8 are considered large. Effect sizes are only presented for statistically significant differences.

Table B4 – Direction of Change in RLI Scores, by Participant Characteristics

Participant Characteristics		N	Percentage of Students, Pretest to Posttest		
			% Declined	% No Change	% Improved
Total TTHS Participants (Reading)		**668**	**30.4%**	**4.6%**	**65.0%**
Group	A) Summer only	188	42.0%	4.3%	53.7%
	B) Summer and school year	94	18.1%	3.2%	78.7%
	C) School year only	386	27.7%	5.2%	67.1%
Gender	Male	337	29.7%	3.0%	67.4%
	Female	331	31.1%	6.3%	62.5%
School	CLOTH	9	-[a]	-[a]	-[a]
	Cobble Hill	64	35.9%	7.8%	56.3%
	Flushing HS	299	26.1%	4.7%	69.2%
	John Adams HS	247	33.2%	4.0%	62.8%
	Queens Vocational HS	49	38.8%	4.1%	57.1%
Dosage Group	0 to 5 episodes	164	41.5%	4.3%	54.3%
	6 to 9 episodes	345	26.4%	4.9%	68.7%
	10 or more episodes	158	27.8%	4.4%	67.7%

[a] Analyses were not conducted for groups with fewer than 10 students.