2013 Collaborative Regional Education (CORE) i3 Study

Final Report

Implementation and Impact Study Results

December 2018

Submitted to:
Jacksonville State University

Submitted by:
ICF Evaluation Team

Table of Contents

I. Introduction 5
1. Program Background 5
2. Description of the Intervention 5
3. Evaluation Overview 6
4. Purpose of this Report 7

II. Impact Study 7
1. Impact Study Introduction 7
2. Research Questions 8
3. Impact Study Methodology 8
   3.1 Teacher Randomization 8
   3.2 Impact Study Data Sources 9
   3.3 Impact Study Test Administration 10
   3.4 Student Engagement and Self-Efficacy Scores for Exploratory Outcome Analysis 10
4. Impact Study Analysis 11
5. Impact Study Results 12
   5.1 Attrition and Baseline Equivalence 12
   5.2 Confirmatory Analysis of Program Impact on CWRA+ Outcomes 15
   5.3 Exploratory Analysis of Program Impact on Student Engagement and Self-Efficacy Scores 16
   5.4 Subgroup Impact Analysis 17
   5.5 Summary of Program Impact Analysis 18

III. Implementation Study 18
1. Implementation Study Introduction 18
2. Implementation Study Methodology 19
   2.1 Implementation Fidelity Measurement System 19
   2.2 Data Sources 20
   2.3 Data Collection 21
   2.4 Implementation Study Analysis 21
3. Implementation Study Results 22
   3.1 Implementation Fidelity by Key Component 22
   3.2 Implementation Fidelity by Indicator 26
   3.4 Correlation between Students' Average CWRA+ Scores and Teachers' Implementation Scores 34

IV. Discussion 34

V. Conclusions 36

VI. References 37

Appendix A: Logic Model 38

Appendix C: Implementation Study Tables 64

Appendix D: Correlation Analysis: Teachers' Program Implementation and Students' Average CWRA+ Scores 66

I. Introduction

1. Program Background

One in four rural students do not graduate from high school and only 17% of adults in rural areas have a college degree (Byun, Meece, & Irvin, 2012). Howley and Hambrick (2011) report that one of the most effective ways to address low graduation rates is by having teachers who are well-prepared for instruction and hold students to high standards.

Technology in classrooms provides opportunities for teacher preparation through collaboration, mentoring, and sharing tools and strategies for instruction. Because internet access in rural areas is sometimes limited or unreliable, rural schools often face obstacles to using technology in classrooms. However, simply making technology available in rural classrooms does not provide a solution, as teachers need scaffolded support to effectively use the technology resources made available to them (Blanchard, LePrevost, Tolin, & Gutierrez, 2016). To address this issue, Jacksonville State University (JSU) has partnered with a growing number of PK-12 school systems as a program developer for more than five years to implement the Collaborative Regional Education (CORE) project through a 2013 Investing in Innovation (i3) validation grant funded by US Department of Education, *Validating the Collaborative Regional Education (CORE) Comprehensive Model: Technology in Rural Classrooms*.

The goal of CORE is to have a positive impact on students' college and work readiness outcomes by improving teachers' use of classroom technology and project-based learning (PBL). Teachers are provided with support for integrating tools and strategies into instruction through intensive professional development (PD) opportunities and technology resources, ongoing mentoring, and networking opportunities with their peers. Strategic change-management support is also provided for building and school system administrators. The CORE model is comprehensive in that it includes technology integration in classrooms alongside PD that prepares teachers for using technology as a tool to support individualized student learning. Specifically, this project assesses the needs of school systems with regard to students' college readiness, assists them in making plans for technology access and hardware, and also provides resources (e.g., iPads and funding) to support the creation of learning environments that promote individualized learning and development of 21st century skills for students. Ultimately, students who have access to these learning opportunities are expected to be better prepared for college and careers.

2. Description of the Intervention

The CORE model is a comprehensive, systems-based approach that consists of seven components designed to build school capacity to better prepare students for college and career by enhancing their 21st century skills, such as critical thinking, problem-solving, technology skills, collaboration skills, and creativity. The seven CORE key components (KCs) are (1) collaboration among administrators and school system leaders, (2) professional learning communities (PLCs) for teachers, (3) provision of classroom technology equipment and resources and instructional support from Education Technology Assistants (ETAs), (4) CORE Active Learning Model (CALM)/PBL PD, (5) ongoing follow-up training and support, (6) support

and coaching in navigating the change-management process in participating schools, and (7) college-readiness advisement and support using the EdReady™ assessment tool.

CORE Academy is the medium through which the CALM instructional content is promoted, via ongoing PD workshops, support, mentoring, and collegial networking through content-focused PLCs. CALM was developed by JSU and promotes student engagement through learning-based teaching and differentiation of instruction. Change-management support is also provided to CORE schools to assist them with making the shift to new modes of instruction. Furthermore, providing support for college-readiness assessments and other resources is expected to directly impact students' college and career readiness—leading to positive long-term high school and college outcomes. The CORE program's effect on college and career readiness and non-cognitive skills outcomes is thought to be mediated by schools' use of CALM. These relationships are depicted in the study logic model (see Appendix A).

3. Evaluation Overview

Under this validation grant, the CORE project and its cornerstone PD offering, CORE Academy, has been evaluated with a rigorous impact and implementation study that included two phases: a two-year local phase and a two-year national phase. The two-phase study design allowed the project team an opportunity to identify a scalable CORE model in rural schools. Each study phase is a two-year randomized controlled trial (RCT) design wherein participant teachers remain in the study for both years, while students in their classes are likely to change over the course of the two study years.[1] RCT is considered the most rigorous design to assess program impact.

The local phase began in school year (SY) 2014–15, with JSU implementing CORE in Alabama by introducing elements of the program model (e.g., PD workshops, various instructional and technology support programs) to 48 schools within the state. Local and national partnerships were established and fostered during this phase to make the national study a viable option. The local-phase study also provided an opportunity to develop data collection and implementation tools that could be piloted locally and eventually scaled up for a study on a national scale. This phase concluded in SY 2015–16.

The national phase began in SY 2016–17, with the CORE program serving 63 rural middle and high schools in a total of seven states: Alabama, Arkansas, Georgia, Louisiana, Missouri, North Carolina, and Texas. These 63 schools were recruited and supported by eight regional university partners (RUPs): (1) JSU, (2) University of Arkansas at Little Rock (UALR), (3) Savannah State University (SSU), (4) Louisiana Tech University (LTU), (5) Southeastern Missouri State University (SEMO), (6) Fayetteville State University (FSU), (7) Tarleton State University (TSU), and (8) West Texas A&M University (WTAM). The RUPs were an integral component of the national-phase study, as these partners were leveraged to provide supports to the participating schools in implementing research-based practice and sustaining improvements through evaluation and change management. Randomization for the national phase validation study was completed in March 2016, and the end of SY 2017–18 marked the conclusion of

[1] Specifically, this means that teachers participated in the intervention and the study for two years while students only participated for one year.

JSU's 2013 CORE i3 grant study. As the study progressed, some school attrition occurred, and 11 schools dropped out of the study by the end of the national phase.

The national-phase study goes beyond the local-phase study by including teachers located at schools in multiple states, whereas the local phase was limited to selected schools in Alabama.[2] However, the basic methodology for the local-phase study remained the same during the national phase.

JSU contracted with ICF to conduct a federally mandated third-party implementation and impact evaluation of the 2013 i3 validation grant. Throughout the evaluation, the National Evaluation of i3 grant (NEi3) technical assistance and support team provided feedback to the evaluation team to ensure that the study meets the What Works Clearinghouse (WWC) (2013) standards and results in a high-quality fidelity of implementation (FOI) study. WWC standards include rules the WWC uses to evaluate the quality of studies for practitioners, researchers, and policymakers.

4. Purpose of this Report

The purpose of this report is to provide an overview of evaluation findings at the culmination of the CORE i3 2013 grant, including implementation and impact results from the local and national study phases. Previously, we provided separate annual reports to summarize findings from the first and second year of the local-phase study and the first year of the national phase study. Additionally, this report provides details on data collection procedures and results obtained during SY 2017–18, which represents the second year of the national-phase study and final year of the evaluation. Impact and implementation study findings from all grant years are presented separately in the following sections.

II. Impact Study

1. Impact Study Introduction

The 2013 validation study of CORE used a cluster RCT design to assess the impact of the CORE intervention on students' college and career readiness outcomes. The study addressed the confirmatory evaluation question of whether students' college and career readiness outcomes, as measured by the College and Work Readiness Assessment + (CWRA+), were higher for those teachers who were assigned to the treatment (i.e., CORE) condition than for those teachers who were assigned to the control (i.e., business-as-usual) condition. The study estimated the treatment effect of CORE at the cluster (i.e., teacher) level. Two contrasts were conducted as part of the local-phase study and again as part of the national-phase study. Contrast 1 tested the impact of CORE after one year of program implementation. Contrast 2 tested the impact of CORE after two full years of program implementation at the conclusion of the local-phase study in SY 2016–17 and again at the conclusion of the national-phase study in

[2] Note that the school samples for the local and national phases were mutually exclusive (no school that participated in the local phase also participated in the national phase).

SY 2017–18. This section of the report describes results of four years of program implementation in two study phases—local and national.

2. Research Questions

Implementation of the CORE model was expected to ultimately improve levels of college and work readiness among students by targeting teacher practices and instructional strategies. As part of the impact study, the CWRA+ was administered to students two times during each of the school years under study. The CWRA+ is designed to measure students' competency in critical thinking, analytic reasoning, and problem-solving and written communication skills—all 21st skills that the *Partnership for 21st Century Skills* has found necessary in college and work environments. The impact analysis focused on addressing the following four evaluation questions:

- Confirmatory Q1: Does one school year of the CORE program have an effect on the mean teacher-level college/career-readiness of grade 8–12 students compared to the mean college/career-readiness of grade 8–12 students in the business-as-usual condition?
- Exploratory Q1: Does one school year of the CORE program have an effect on the mean teacher-level student engagement levels and student self-efficacy levels of grade 8–12 students compared to the mean student engagement levels and student self-efficacy levels of grade 8–12 students in the business-as-usual condition?
- Exploratory Q2: After one school year of CORE program implementation, does the size of program impact vary significantly by student subgroups (based on grade-levels, gender, free/reduced lunch status as a proxy measure of poverty, achievement level, race, ethnicity, language proficiency, dual enrollment)?
- Exploratory Q3: After one school year of CORE program implementation, does the size of CORE program impact vary by the level of program implementation (based on teachers' use of classroom technology and PBL)?

3. Impact Study Methodology

3.1 Teacher Randomization

Participating teachers included in the local-phase study were recruited by JSU. Once the project was scaled-up, national-phase study teacher participants were recruited by the eight RUPs (JSU, UALR, SSU, LTU, SEMO, FSU, TSU, and WTAM). Following initial recruitment, each school provided a list of teachers who were interested in participating in the study and who met the following eligibility criteria: (1) teach primarily grade 8–12 students; (2) have little to no direct prior experience with technology-integrated instruction, PBL, or the CORE program; and (3) teach either science, social studies, humanities, English Language Arts (ELA)/English and/or math or humanities courses. For the local-phase study, 48 teachers from 16 districts in Alabama were identified and met the criteria. For the national phase, 120 teachers from 63 schools in multiple states were nominated and met the criteria.

Due to resource constraints, a single teacher in each school was randomly chosen to participate in the program, a model that was followed for both the local- and national-phase studies. The research team assigned each nominated teacher a random number, and sorted the list from

high to low, within each school. The teacher with the highest random number was selected to represent each school in the sampling frame.

Teachers for the local-phase study were recruited from a population of 72 CORE partnership schools situated in 18 Alabama school districts. CORE partnership schools/districts were defined as those that have historically had some degree of involvement in the CORE program prior to its being funded for study under the i3 grant. From this population, 48 schools representing 16 school districts were recruited based upon their willingness to participate in the CORE i3 validation study. Randomization was completed using the following school-level blocks to assign teachers to conditions: (a) middle school vs. high school and (b) prior exposure to elements of the CORE model. Using these blocks, we randomly selected one teacher per school, assigning 24 to participate in the treatment condition and the remaining 24 to the business-as-usual condition. The final sample included teachers representing 17 middle schools and 31 high schools. The associated school districts ranged in total K–12 enrollment from 1,459 to 8,963 students, had a median percentage non-White enrollment of ~20% for grades 8–12, a median free- and reduced-lunch eligibility rate of 66% for grades K–12, and 50% were classified as rural school districts.

Teachers for the national-phase study were recruited by eight RUPs (JSU, UALR, SSU, LTU, SEMO, FSU, TSU, and WTAM). Together, the eight institutions recruited 63 middle and high schools to participate in the two-year national-phase study (i.e., from SY 2016–17 to SY 2017–18). National-phase random assignment was then completed using the following blocks to assign 63 teachers to conditions: (a) the sponsoring RUP and (b) school level (i.e., middle school or high school). In March 2016, a total of 32 teachers were randomly assigned to participate in the treatment condition and an additional 31 to the "business-as-usual," or control condition, for the two-year duration of the study. The final sample for the national-phase validation study included 63 teachers representing 20 middle schools and 43 high schools, situated across seven southeastern states.

3.2 Impact Study Data Sources

The CWRA+ is a standardized assessment developed by the Council for Aid to Education (CAE), which is designed to measure student mastery of 21st century skills that are necessary for success in postsecondary education and workforce settings (e.g., critical thinking). The assessment includes both performance task (PT) and selected response questions (SRQs). Both of these were used for the first pretest administration in local phase Year 1; however, to minimize test administration time, only SRQs were administered for the remainder of the local phase and during the national phase. As a result, only students' CWRA+ SRQ scores were used as the pretest and posttest measures for all four years, for consistency. The SRQ score represents students' cumulative performance related to the following three 21st century skills: (1) scientific and quantitative reasoning, (2) critical reading and evaluation, and (3) critiquing an argument. These are all skills that were hypothesized to be positively influenced by teachers' exposure to and use of instructional strategies learned through participation in CORE. The SRQ has sufficient internal consistency (Cronbach's $\alpha = .73$). For more information about the CWRA+ and SRQ score measures, see http://cae.org/images/uploads/pdf/CWRA_Plus_Technical_FAQs.pdf.

3.3 Impact Study Test Administration

The CWRA+ SRQ items were administered to grade 8–12 students who were enrolled in the participating teachers' classrooms on or before the beginning of each school year at the participating schools. To select these students, the research team gathered course and student rosters from participating teachers. For each participating teacher, ICF selected a random sample of grade 8–12 classes to take part in data collection. These classes were selected from the universe of all grade 8–12 classes offered during the school year, and thus are representative of the full population of students who received i3-supported interventions.

For each teacher, ICF selected at random the number of grade 8–12 classes necessary to achieve a sample of at least 60 students per teacher. If a teacher had fewer than 60 students in her/his offered courses, all students were selected to participate in the assessment. All students in these randomly selected classes were asked to complete the study outcome measure during the fall (pretest) and spring (posttest) of 2014–15 and 2015–16 school years in the local phase and 2016–17 and 2017–18 school years in the national phase, except for those students for whom CWRA+ testing would violate their individualized education plans (IEPs).

Table 1 below provides a summary of the total number of teachers, clusters, and eligible students selected to participate in the study. These numbers are based on the initial list of participant teachers and students before CWRA+ data attrition was taken into consideration. The number of students, however, does not include students whose teachers had dropped out of the study. Descriptive statistics of the samples included in analysis are presented in a later section.

Table 1. Number of Teachers, Classes, and Students by Treatment Status Available for Study Participation

		Treatment Group		Control Group	
		Teachers	Students	Teachers	Students
Local Phase	Year 1	24	1,346	24	1,206
	Year 2		1,577		1,040
National Phase	Year 1	32	1,830	31	1,844
	Year 2		1,536		1,323

3.4 Student Engagement and Self-Efficacy Scores for Exploratory Outcome Analysis

When completing the CWRA+ at pretest and posttest, students were also asked to respond to four student engagement questions from the Consortium on Chicago School Research Academic Engagement Scale (CCSR/AES) (Consortium on Chicago School Research, 2007) and five self-efficacy questions from the Patterns of Adaptive Learning Scales (PALS, Midgley et al., 2000). It was hypothesized that students in classrooms where teachers were implementing CORE would report higher levels of engagement and self-efficacy than students in control classrooms. Table 2 presents the nine questions included. Students rated themselves on each item on a Likert-type scale of *strongly agree* to *strongly disagree*.

Table 2. Student Engagement and Self-Efficacy Questions Included with CWRA+

Survey Items	
Engagement	
Q1	I usually look forward to this class.
Q2	I work hard to do my best in this class.
Q3	Sometimes I get so interested in my work in this class that I don't want to stop.
Q4	The topics we are studying in this class are interesting and challenging.
Self-Efficacy	
Q5	I'm certain I can master the skills taught in this class this year.
Q6	I'm certain I can figure out how to do the most difficult work in this class this year.
Q7	I can do almost all the work in this class if I don't give up.
Q8	Even if the work is hard in this class, I can learn it.
Q9	I can do even the hardest work in this class if I try.

Average scores for these two scales were derived for each student and used as student engagement and self-efficacy outcome variables, respectively. The pretest and posttest student engagement reliability estimates as measured by Cronbach's Alpha were .71 and .70, respectively, and the pretest and posttest student self-efficacy reliability estimates were .84 and .81 for the pretest and posttest. These results demonstrate sufficient reliability.[3]

4. Impact Study Analysis

The research team conducted a multi-level statistical analysis of students' CWRA+ SRQ scores collected at the end of each school year (i.e., the posttest) separately for each of the four study years. Impact coefficients for treatment status were derived. Hierarchical linear modeling (HLM), a standard approach to data that are hierarchically structured (students nested within teachers), was used in analysis. A standard assumption of residuals being normally distributed was not attainable as the data were correlated by cluster. In the HLM model, students and teachers were specified, respectively, as level 1 and level 2 units. To address the clustering issue, the model estimated the intercepts (i.e., school effects) as random effects. To improve the precision of the impact coefficient, our analysis considered important covariates, including students' pretest CWRA+ SRQ scores, grade levels, gender, race and ethnicity, and parent education levels as proxy measures of socioeconomic status. Some of the local phase schools had some previous exposure to CORE program features prior to its being funded for study under the i3 program. The analysis team created a binary variable differentiating the local phase schools that were exposed to CORE-like elements prior to the intervention and those that were not. The middle school/high school distinction was incorporated into the model via the inclusion of student grade-level data. All covariates were included in the final models. The program effect was estimated as the coefficient of the treatment status (1 if treatment, 0 if control) and the standardized effect size was presented to facilitate interpretation. The standardized program effect was derived by rerunning the same statistical model using the z-score version of CWRA+ outcome scores. The following equation summarizes the model described above.

[3] It is generally accepted that a value greater than or equal to .70 for Cronbach's Alpha is considered sufficient reliability.

$$Posttest_{ij} = \beta_{00} + \beta_{10} * pretest_{ij} + \beta_{20} * treatment_j + ... + r_{ij} + u_j$$

where

- Posttest represents posttest outcome scores
- Pretest represents the baseline scores (of the posttest outcomes)
- Postscripts *i* and *j*, respectively represent student and teacher
- βs are parameters to be estimated
- The three ellipses (i.e., "…") indicate that the model will include multiple predictors and corresponding parameters; predictors are gender, grade levels (8th to 12th), race and ethnicity, parents' college education (if at least one parent earned BA degree 1; else 0), and school exposure to CORE-like elements prior to the intervention (applicable to local phase).
- Treatment represents the treatment status (1 if treatment group; 0 if control group)
- r and u are independently and identically distributed residuals with a mean of 0.

To compare the post-intervention difference between the treatment and control groups on these two exploratory outcomes, we used the same HLM framework used for the CWRA+ SRQ outcome analysis, but replaced the CWRA+ SRQ outcome with the engagement and self-efficacy outcomes, respectively, as previously described.

5. Impact Study Results

5.1 Attrition and Baseline Equivalence

High attrition at both teacher and student levels affects the integrity of the RCT design. ICF calculated both cluster- (teacher) and sub-cluster-level (student) attrition for all four years and examined the baseline equivalence based on CWRA+ SRQ pretest scores. Overall, the JSU and evaluation team worked collaboratively to maintain low attrition during both phases of the study. In the local-phase study, Year 1 samples were at low attrition. In the national-phase study, both Year 1 and Year 2 samples were at low attrition and well-executed. This outcome is a direct result of the strong collaborative effort by JSU, a third-party responsible for the CWRA+ assessment, CAE, and the ICF evaluation team to provide school-level support during the testing administration window. This is also a result of the institutional partnerships established by the RUPs with their school districts. Only during one year of the four-year study was there high attrition. The local-phase Year 2 sample suffered cluster-level attrition and baseline equivalence was therefore not established. ICF modified the local-phase Year 2 sample by using propensity score matching (PSM) to establish baseline equivalence and meet WWC standards with reservation. The local-phase Year 2 study is thus a quasi-experimental design (QED) study.

Attrition Calculation for Cluster- (Teacher/School) and Sub-Cluster-levels

Table 3 summarizes the result of overall cluster-level attrition rate and differential attrition rates from four study years.[4]

For local phase Year 1 and national phase Year 1 and 2, teacher-level attrition was kept within a threshold specified by WWC standards. Local phase Year 1 began with 24 teachers each in treatment (Tx) and control groups. These numbers reduced to 20 and 22, respectively, for treatment and control. The national phase began with 32 treatment teachers and 31 control teachers. The Year 1 analysis sample retained 29 treatment teachers and 28 control teachers. The Year 2 analysis sample retained 27 treatment teachers and 25 control teachers.

The local phase Year 2 sample did not pass the WWC test of cluster-level attrition; the number of treatment teachers decreased from 24 to 22, while that of control teachers reduced from 24 to 15. Thus, the local phase Year 2 sample was rated as a high-attrition sample by the WWC calculation. The later section will show how this sample also failed the baseline equivalence test.

Table 3. Summary of Sample Sizes and Cluster-level Attrition Information

Cluster level	Roster			Analysis Sample			Attrition (%s)					
	Tx	Control	Sub-total	Tx	Control	Sub-total	Overall	Tx	Control	Differential Attrition Rate (DAR)	WWC liberal boundary	Attrition level
Local Phase												
Year 1	24	24	48	22	20	42	0.13	0.08	0.17	0.08	0.11	Low
Year 2	24	24	48	22	15	37	0.23	0.08	0.38	0.29	0.10	High
National Phase												
Year 1	32	31	63	29	28	57	0.10	0.09	0.10	0.00	0.11	Low
Year 2	32	31	63	27	25	52	0.17	0.16	0.19	0.04	0.11	Low

Table 4 reports the results of student-level attrition rates. The results suggest that for the sub-cluster-level (student-level), all four study samples met the WWC attrition test as samples with low attrition.

Table 4. Sub-Cluster-level Attrition Calculations by School Year

Sub-cluster (student) Level	Roster			Analysis Sample			Attrition (%s)			DAR and Level		
	Tx	Control	Sub-total	Tx	Control	Sub-total	Overall	Tx	Control	DAR	WWC boundary	Attrition level
Local Phase												
Year 1	1,346	1,206	2,552	1,060	842	1,902	0.25	0.21	0.30	0.09	0.09	Low
Year 2	1,577	1,040	2,617	1,265	777	2,042	0.22	0.20	0.25	0.06	0.10	Low
National Phase												
Year 1	1,830	1,844	3,674	1,515	1,468	2,983	0.19	0.17	0.20	0.03	0.10	Low
Year 2	1,536	1,323	2,859	1,397	1,212	2,394	0.09	0.09	0.08	0.01	0.11	Low

Baseline Equivalence Tests

As mentioned above, the local-phase Year 2 sample suffered high attrition, while the other three samples were low-attrition RCT studies. The high-attrition study needed to show baseline equivalence to meet WWC standards with reservation. As shown in Table 5, however, this sample did not establish baseline equivalence based on the CWRA+ pretest. Since baseline equivalence was not achieved, the local-phase Year 2 sample was adjusted using a PSM approach. The treatment and control groups were matched based on grade levels as exact matching criterion and pretest CWRA+ scores.[5]

Table 5. Local Phase Year 2 Results of Baseline Equivalence Test of Pretest CWRA+ SRQ Scores

Baseline Equivalence Test Results (CWRA+ SRQ)			
	N	Mean	Std. Dev.
Control group pretest outcome	777	896	172
Treatment group pretest outcome	1,265	849	153
Standardized mean difference in pretest values	.29		
WWC threshold for baseline equivalence	.25		
Is baseline equivalence established?	No		

Table 6 shows the results of baseline equivalence tests on all four samples, including the modified local phase Year 2 sample whose baseline equivalence was established by the PSM method. All four samples showed that the CWRA+ SRQ pretest score differences between the treatment and control groups were within an acceptable range.

Table 6. Baseline Equivalence Test Results for CWRA+ SRQ

		Treatment Group			Control Group			WWC Baseline Test		
	Total N	N	Mean	SD	N	Mean	SD	Mean difference	Hedge g	Result
Local Phase										
Year 1	1,902	1,060	870.72	157.42	842	855.5	164.21	15.22	0.10	Satisfies BE with statistical adjustment
Year 2	1,948	1,265	866.76	156.71	683	875.8	163.3	-9.06	-0.06	Satisfies BE with statistical adjustment
National Phase										
Year 1	2,983	1,515	904.14	165.64	1,468	912.4	169.54	-8.21	-0.05	Satisfies BE
Year 2	2,609	1,397	899.01	166.89	1,212	914.0	164.54	-15.02	-0.09	Satisfies BE with statistical adjustment

The same conclusion was reached for student engagement outcome scores and student self-efficacy outcome scores. As shown in Tables 7 and 8, all samples established baseline

[5] Software R and MatchIT program were used for matching. To retain the largest number of matched control cases possible, ICF conducted one-to-many matching without replacement modeling (in logistic regression) for the control condition. A control student was matched with up to four individual treatment cases. All treated cases were retained (n=1265), while 88% of the control cases were matched (original n=777, final n=684). Because of the one-to-many matching, subjects have different probabilities of being included in the analysis sample. To account for this, all national phase Year 2 analysis used the probability-based weights generated by the MatchIT program.

equivalence by showing that the baseline pretest outcome differences were sufficiently small for both of the noncognitive scales.

Table 7. Baseline Equivalence Test Results for Student Engagement Scores

	Total N	Treatment Group			Control Group			WWC Baseline Test		
		N	Mean	SD	N	Mean	SD	Mean difference	Hedge g	Result
Local Phase										
Year 1	1,823	1,022	3.03	0.48	801	3.09	0.48	-0.06	-0.12	Satisfies BE with statistical adjustment
Year 2	1,770	1,144	3.04	0.52	626	3.04	0.52	0.00	0.01	Satisfies BE
National Phase										
Year 1	2,784	1,407	3.02	0.51	1,377	2.98	0.51	0.04	0.07	Satisfies BE with statistical adjustment
Year 2	2,472	1,344	3.03	0.51	1,128	3	0.55	0.03	0.07	Satisfies BE with statistical adjustment

Table 8. Baseline Equivalence Test Results for Student Self-Efficacy Scores

		Treatment Group			Control Group			WWC Baseline Test		
	Total N	N	Mean	SD	N	Mean	SD	Mean difference	Hedge g	Result
Local Phase										
Year 1	1,817	1,017	3.34	0.48	800	3.30	0.49	0.04	0.08	Satisfies BE with statistical adjustment
Year 2	1,770	1,144	3.30	0.49	626	3.29	0.53	0.01	0.02	Satisfies BE
National Phase										
Year 1	2,741	1,385	3.23	0.52	1,356	3.24	0.50	-0.01	-0.02	Satisfies BE
Year 2	2,113	1,131	3.22	0.53	982	3.22	0.55	0.00	0.01	Satisfies BE

5.2 Confirmatory Analysis of Program Impact on CWRA+ Outcomes

As mentioned, ICF used the HLM framework to conduct the multivariate regression modeling to estimate program impact. The derived program impact estimates were adjusted for pretest CWRA+ scores, grade levels, gender, race and ethnicity, and parents' college education. For the local phase, a school's prior exposure to CORE program-like experiences was used as a predictor. School differences were also adjusted in the model as random effects. For the calculation of effect sizes, the mean and the standard deviation estimates of the analysis sample were used.

Table 9 summarizes the program impact analysis. Program impact coefficients (shown in the effect column) and statistical test results show that none of the program effects were large (standardized effect sizes range from -0.05 to 0.03) and none were statistically significant.

Appendix tables B1–B8 contain the full results of HLM analysis and descriptive statistics of the analysis samples.

Table 9. Summary of Program Impact Analysis Results for CWRA+ Scores

		N. of teachers	N. of students	Effect	Std. error	p	Sig.	Standardized effect
Local Phase	Year 1 RCT study	42	1,902	-5.87	11.9	0.62	ns	-0.04
	Year 2 QED study	37	1,948	0.5	18.48	0.98	ns	0.00
National Phase	Year 1 RCT study	57	2,983	-8.91	12.37	0.47	ns	-0.05
	Year 2 RCT study	52	2,609	4.39	14.48	0.76	ns	0.03

Notes: Statistical significance (2-tail test): * = p<.05, ** = p<.01, *** = p<.001. "ns" means the results were not statistically significant. The standardized effect sizes were based on the analysis sample's mean and SD. See Appendix B, Table B25 for What Works Clearinghouse calculation of standardized effects (Hedge's g).

5.3 Exploratory Analysis of Program Impact on Student Engagement and Self-Efficacy Scores

Using the same HLM framework with the same set of covariates (pretest scores, grade level, gender, race and ethnicity, parents' college education; for local phase, school's prior exposure to CORE-like programs), the program impacts were estimated for student engagement and self-efficacy scores. The general expectation was that students in the treatment group had a higher level of posttest engagement and self-efficacy levels when controlling for various factors.

As shown in Table 10, the results were not consistent with this expectation. None of the program effects on the two outcomes were large (standardized effect sizes ranged from -0.08 to 0.09) and none of them were statistically significant. See Appendix tables B9–B24 for more information.

Table 10. Summary of Program Impact Analysis Results for Student Engagement and Self-Efficacy Scores

Student Engagement Score Analysis							
	N. of teachers	N. of students	Effect	Std. error	p	Sig.	Standardized effect
Year 1 RCT study	42	1,823	-0.04	0.03	0.14	Ns	-0.08
Year 2 QED study	37	1,770	0.01	0.04	0.87	Ns	0.01
Year 1 RCT study	57	2,784	0.04	0.03	0.23	Ns	0.08
Year 2 RCT study	52	2,472	0.05	0.04	0.18	Ns	0.09
Student Self-Efficacy Score Analysis							
	N. of teachers	N. of students	Effect	Std. error	p	Sig.	Standardized effect
Year 1 RCT study	42	1,817	-0.03	0.03	0.28	Ns	-0.06
Year 2 QED study	37	1,770	0.03	0.04	0.45	Ns	0.06
Year 1 RCT study	57	2,741	0.03	0.03	0.43	Ns	0.05
Year 2 RCT study	52	2,113	0.03	0.03	0.44	ns	0.04

Note: Statistical significance (2-tail test): * = p<.05, ** = p<.01, *** = p<.001. The standardized effect sizes were based on the analysis sample's mean and SD. See Appendix B, Table B25 for What Works Clearinghouse calculation of standardized effects (Hedge's g), which returned the same result.

5.4 Subgroup Impact Analysis

The program impact analysis from all study years did not return any results that were statistically significant. To explore the possibility that the program impact on students CWRA+ scores and student engagement and self-efficacy scores may exist within certain subgroups, we conducted a series of impact analyses on subgroups defined by the following variables:

- Grade level (8th, 9th, 10th, 11th, 12th)
- Gender (male, female)
- Race and ethnicity (Black, White, Hispanic)
- CWRA+ levels (equal to or above the average, below the average)
- Parents' education level (college graduates vs. others)
- Teachers' subject areas (humanities vs. STEM)

Taking advantage of the fact that the study collected data over a four-year time period, subgroup analysis was conducted using the local-phase sample (two study years combined), the national-phase sample (two study years combined), and the four-year sample (all four years combined).

The analysis generated 48 subgroups based on the factors listed above (e.g., grade levels, gender) and three types of data (local phase, national phase, and all four years combined). As shown in Table 11, only three subgroups returned program impacts that were statistically significant. The three subgroups were Black students from the national phase (negative program impact), 12th graders from the national phase, and 10th graders from all four years. Because these findings are only three out of 48 subsamples and the findings are not consistent (e.g., the finding regarding Black students is only present with the national phase sample, but not in the local sample or the four-year sample), we concluded that this analysis did not generate evidence to support the program impact on student subgroups.

Table 11: Selective Summary of Subgroup Impact Analysis Results

	Sample Name	N. of teachers	N. of students	Effect	Std. error	p	Sig	Standardized effect
CWRA+	National Phase: Black students	50	1,082	-43.43	11.95	0.00	**	-0.32
Student Engagement	National Phase: 12th grade students	41	931	0.12	0.05	0.02	*	0.22
Student Self-Efficacy	All 4 years: 10th grade students	58	1,825	0.08	0.04	0.03	*	0.15

Note: Statistical significance (2-tail test): * = $p<.05$, ** = $p<.01$, *** = $p<.001$. The standardized effect sizes were based on the analysis sample's mean and SD.

5.5 Summary of Program Impact Analysis

The following summarizes the program impact analysis section.

- Out of four years, local phase Year 2 suffered a significant attrition and did not establish baseline equivalence. We used PSM to force the two groups to be equivalent at baseline, turning the study into a QED study. Other samples (local Year 1, national Years 1 and 2) passed the WWC tests of attrition and baseline equivalence. They can be considered well-executed RCTs.
- Confirmatory analysis of four samples (addressing confirmatory question 1) did not find any statistically significant findings with respect to program impacts on CWRA+ SRQ scores.
- Exploratory analysis (addressing exploratory question 1) did not lead to any statistically significant findings with respect to student engagement and self-efficacy levels.
- The subgroup impact analysis (addressing exploratory question 2) found three subgroups of data in which program impacts were statistically significant; however, the results were not consistent (e.g., the program impact within the Black student sample for national phase data was statistically significant; however, it was not significant for the local phase sample).

III. Implementation Study

1. Implementation Study Introduction

An FOI study, required by the NEi3 project, was conducted to measure the extent to which the CORE model was implemented as intended in participating schools. The study was guided by seven evaluation questions aligned to the intervention KCs specified in the CORE program logic model (refer to Appendix A): (1) collaboration among administrators and school system leaders, (2) PLCs for teachers, (3) provision of classroom technology resources and support, (4) CALM/PBL PD, (5) ongoing follow-up training and support, (6) support and coaching in navigating the change-management process in participating schools, and (7) college-readiness advisement and support using the EdReady™ tool. This section of the report summarizes the implementation study findings for each two-year phase study. The timeframe and sample of teachers in the implementation study align with those previously described for the impact study.

2. Implementation Study Methodology

2.1 Implementation Fidelity Measurement System

FOI was tracked annually but measured across two school years of each phase of the study: 2014–16 for the local phase and 2016–18 for the national phase. FOI was assessed using a collaboratively developed measurement system that included 17 indicators aligned to the seven KCs of the CORE program logic model. Implementation study questions were directly linked to the KCs of the program (see Appendix C, Table C1 for details). In 2014, ICF and JSU identified each initial indicator and set implementation thresholds for the local-phase study of CORE (see Appendix C, Table C2 for details of the local-phase implementation fidelity system). In 2016, indicators were revised to reflect the goals of the national-phase study. Table 9 presents the KCs, revised fidelity indicators, and data sources for the national phase.

Table 9. KC, Indicator, and Data Sources for the FOI Study (National Phase)

Measuring Implementation Fidelity		
Key Component	**Indicator**	**Data Source**
KC1. Principals in new CORE schools connect with principals in existing partnership schools.	1.1 Principal collaboration with their peers and superintendents at the CORE Academy and three CORE Workshops	CORE event evaluation survey
KC2. CORE teachers participate in online learning communities.	2.1 Active participation in online PLCs	CANVAS activity (post) count
KC3. CORE teachers receive and adopt classroom technology resources and support.	3.1 Provision of laptops and iPads for CORE teachers	Technology equipment log
	3.2 Provision of classroom technology funds for CORE teachers	Financial disbursement log
	3.3 Provision of classroom technology assessments	Technology assessment log
	3.4 Provision of ongoing support from Education Technology Assistants (ETAs)	ETA assistance log
KC4. CORE teachers participate in CORE Academy CALM instructional PD.	4.1. Teacher attendance at the annual CORE Academy	CORE Academy attendance roster
	4.2. Quality of CORE Academy	CORE Academy event survey
	4.3. Relevance of CORE Academy	
	4.4. Usefulness of CORE Academy	
	4.5. Increased content knowledge	
	4.6. Use of CALM	CORE pre- and post-CALM assessment
	4.7. Use of technology	CORE pre- and post-technology assessment
KC5. CORE teachers participate in follow-up CALM PD workshops.	5.1. Participation in follow-up workshops	Workshop attendance roster
	5.2. Sharing of learning experiences	YouTube video log
KC6. CORE principals provide change-management support to CORE schools.	6.1. Participation in follow-up debriefings on change management	Change Diagnostic Index (CDI) report and Pivot Point debriefing log
KC7. CORE Math and English teachers use EdReady™ college assessment tool in schools.	7.1 EdReady™ utilization log-in count	EdReady™ log-in records

2.2 Data Sources

The implementation study drew data for the indicators associated with both phases of the study from the following sources: (1) event evaluation surveys, (2) CANVAS post counts, (3) equipment distribution logs, (4) ETA assistance logs, (5) fund disbursement records, (6) technology assessment logs, (7) attendance rosters, (8) YouTube video logs, (9) teacher pretest/posttest surveys, (10) the CDI report and debriefing log and (11) EdReady™ log-in records.

In addition to the above data sources, onsite and virtual site visits with a sample of CORE treatment schools provided additional sources of data. These visits took place in spring 2016 with a sample of schools participating in the local phase, and fall 2017 with a sample of schools participating in the national phase. The purpose of these visits was to better understand the details of CORE implementation and document best practices and challenges to implementation and sustainability. Qualitative data collected during site visits are described in further detail below.[6]

Principal and Teacher Interviews. A sample of eight teachers in March 2016 and twelve teachers in October and November 2017 participated in semi-structured onsite and virtual interviews. The interview protocols were consistent across sites and phases, and made connections between classroom observations (another data collection activity, discussed in more detail below) and overall program implementation. Specifically, teachers responded to pre- and post-observation interview questions. In addition, eight principal interviews were conducted. These interviews gathered valuable school-level information on implementation context—including challenges and supports—and provided an opportunity for principals to reflect on program benefits and challenges faced by teachers when implementing the CORE program.

Site Visit Observations. The ICF team collaborated with JSU to develop an observation rubric with ratings across the four CORE skills (technology use, collaborative environment, critical thinking, and problem-solving). For class observations, the team followed a protocol that required alternating descriptive note-taking with observation. That is, each observer took notes for five minutes, and then observed for ten minutes, repeating this method three times during each observation. The evaluation team collected one lesson plan that aligned with the class lesson being observed and any other non-confidential handouts distributed during the observed class.

ETA Focus Groups. Focus group interviews were conducted in June 2017 during the second national study CORE Academy PD event with all ETAs and all RUP representatives. Each group interview lasted 45 minutes to one hour. ICF staff conducted all interviews without CORE staff present, to foster an open environment where individuals were comfortable speaking candidly about their experiences. The goal of the focus groups was to learn more about CORE program implementation from the perspectives of ETAs and RUPs. Candid feedback from these groups was highly valuable to the program because it helped JSU staff improve supports for participating teachers.

[6] Note that the primary data sources for the KCs are those listed in Table 9. Site visit data sources are described in detail to provide context; however, this data collection was meant to supplement the data sources directly linked to study KC.

2.3 Data Collection

All data sources were developed and maintained by JSU, with consultation from ICF. JSU was responsible for data collection and implementation for each school year. Data were transmitted to ICF in December and July of each calendar year. After Year 1 of the local-phase study, the evaluation team consulted with the program developer and revised and consolidated data collection efforts to reflect different expectations for schools to better align with the program theory. Several changes were made to KCs and their associated indicators before launching the national phase of the study. These changes are summarized below by KC.

- KC1: Two indicators were associated with this KC in the first year of the local phase. For the second year of the local phase (and continuing into the national phase), this component was based on just one indicator. Measurement for this component shifted away from attendance (given the voluntary nature of the activity for principals) and focused instead on collaboration of those who attended.
- KC2: The threshold for the indicator was not changed; however the measurement tool was updated from Facebook to the CANVAS system, which allowed JSU to better measure participation.
- KC4 and KC5: The PBL PD and follow-up workshops provided were revised to CALM for the national phase.
- KC7: This KC was significantly revised because the dual enrollment aspect that defined this component during Year 1 of the local phase applied only to select CORE treatment schools. Revisions were made to allow the program to maintain its goal of providing college-readiness support to all CORE schools in a scalable way; specifically, it was revised to measure teacher use of a college-readiness assessment tool.

2.4 Implementation Study Analysis

Individual indicator implementation scores were calculated for each of the 27 treatment schools remaining in the study at the conclusion of SY 2017–18.[7] All 17 fidelity indicators were scored for each school. The resulting scores were then coded to represent the extent to which each school met the associated indicator's implementation threshold (typically measured as low, medium, or high). Once indicator implementation scores were derived, they were summed within each KC to arrive at a single KC implementation score for each treatment school (typically measured as low, medium, or high).

We then calculated the percentage of treatment schools meeting the criteria for "high" implementation for each KC and compared this to an established threshold for "high" fidelity at the sample level (e.g., greater than 80%). If the percentage of schools in the entire sample who met the criteria for "high" implementation met or exceeded this threshold, FOI was considered to be met for the KC at the sample level.

Special note. *The denominator for FOI calculations includes only those teachers/schools that were remaining in the treatment group at the end of each school year. For SY 2017–18 this includes 27 of the 32 teachers who were originally randomly assigned to the treatment group.*

[7] Three treatment schools were lost during Year 1 of the national study, and two treatment schools were lost during Year 2 of the national study, as shown in Table 3.

3. Implementation Study Results

As previously described, the fidelity study was supplemented by site visits conducted in both the local and national phases of the study. Data collected during these visits indicated mostly positive findings about the CORE experience, with many successes stemming from the CORE academies and trainings, regular Education Technology Assistant (ETA) meetings and support, and opportunities that come from implementing new technology in classroom instruction. Overall, teachers and principals participating in site visits reported CORE implementation within their respective schools to be successful as it related to their expectations and impact on classroom instruction and student engagement. Findings from the twelve CORE classroom observations showed teachers' progress in CORE model use supporting students' 21st century skills. The following sections discuss implementation findings related to each of the KCs. Section 3.1 provides an overview of fidelity of implementation by KC. Section 3.2 includes more detailed information on fidelity of implementation by indicator, so as to provide more specific information on where fidelity was achieved across the various program components.

3.1 Implementation Fidelity by Key Component

In this section, we provide a summary of KC level fidelity outcomes for both phases. Fidelity outcomes are based on the following numbers of teachers and/or their schools for each school year of the study:

- SY 2014–15 (local phase Year 1): 23 teachers and/or their schools;
- SY 2015–16 (local phase Year 2): 22 of the original teachers and/or their schools remaining at the end of the local-phase study;
- SY 2016–17 (national phase Year 1): 29 teachers and/or their schools; and
- SY 2017–18 (national phase Year 2): 27 of the original 29 teachers and/or their schools remaining at the end of the national-phase study.

The following findings are organized by KC and list the status as reported to the NEi3 in the VALID33 design summary template.[8] Table 10 summarizes these findings. The results are highlighted for both the local and national phases of the study. Overall, JSU CORE reached adequate fidelity on four out of seven components during the initial local-phase study and reached the threshold for adequate FOI during the national phase project level for all seven components.

KC1: High implementation fidelity to KC1, Principals in new CORE schools connect with principals in existing partnership schools, was not met for the local phase and was met with high fidelity for the national phase of the study.

To achieve high fidelity on KC1 during either phase, at least 81% of principals in CORE schools who attended a CORE-sponsored event during the school year needed to complete a post-event survey and agree or strongly agree that the event(s) they attended helped promote collaboration and networking among their peers.

[8] VALID33 refers to the i3 Study Design Plan Template completed for this study. This document contains study design details that are important to document for review by the i3 Analysis and Reporting team.

Calculated annually, but reported once over each two-year phase study, 5 of 6 principals (83%) met the threshold in SY 2014–15 and 3 of 4 principals (75%) met the threshold in SY 2015–16. When considering the local phase overall (reporting across the two years), 6 out of 8 principals (75%) met the threshold in the local phase,[9] and 8 out of 8 (100%) met the threshold in the national phase based on the number of principals that provided data for KC1. While fidelity was met for the national phase and not met for the local phase, it should be noted that the sample for either phase of the study is not representative of the number of principals remaining (local, N = 22 and national, N = 27).

KC2: High implementation fidelity to KC2, CORE teachers participate in online learning communities, was not met for the local-phase study and was met for the national-phase study.

KC2 refers to teacher PLC participation. Initially, the 2013 grant fidelity was based on Facebook posts, which would indicate a Professional Learning Network (PLN) rather than a PLC. It was met with high fidelity when at least 81% of treatment teachers enrolled in PLCs and showed online activity with at least three posts per month. Online activity was averaged over nine months (i.e., over the academic year). For the national phase, this data source was changed to professional learning activity in the learning management system (CANVAS/SmarterU) to better reflect active participation in online PLCs.

Calculated annually and reported once for each phase of the study, one of 23 teachers (4%) met the threshold in SY 2014–15 and three of 22 teachers (14%) met the threshold in SY 2015–16. Twenty-five out of 29 teachers (86%) met the threshold in Year 1 of the national phase. For Year 2 of the national phase, SY 2017–18 data show that 25 of 27 teachers (93%) met the threshold.

KC3: High implementation fidelity to KC3, CORE teachers receive and adopt classroom technology resources and support, was met with high fidelity for both phases of the study.

To achieve high fidelity on KC3 during either phase, at least 81% of CORE teachers had to participate and receive classroom support in four areas: (1) receive a MacBook Air Laptop and 21 iPad tablet computers for their classrooms; (2) receive $2,100 in classroom technology funds; (3) participate in an initial technology assessment; and (4) receive at least three site-based support visits/consultations annually from ETAs.

Calculated annually over each two-year phase study, 100% of 23 teachers met the threshold in SY 2014–15. A total of 20 of 22 (91%) met the threshold in Year 2 of the local phase SY 2015–16. In the national-phase study, a total of 28 of 29 teachers met the threshold (97%) in Year 1, followed by 100% of 27 of teachers who met the threshold in Year 2.

KC4: High implementation fidelity to KC4, CORE teachers participate in CORE Academy CALM/PBL PD, was met overall.

This KC consists of seven indicators, five of which focus on the participation and quality of PD workshops offered through the CORE program: content knowledge, relevancy and usefulness.

[9] For the local phase, five of six principals (83%) met the threshold in SY 2014-15. Three of four principals (75%) met the threshold in SY 2015-16.

The remaining two indicators focused specifically on the use of active learning methods and technology as reported in a pre-post assessment delivered to teachers annually. To achieve high fidelity for this component, teachers had to participate in three annual PD workshops held in August, November, and March as well as one CORE Academy held in the summer. Additionally, teachers surveyed had to report high levels of agreement with the quality, relevance, and usefulness of the trainings for their CORE implementation efforts, as well as agree that their content knowledge had increased. Lastly, teachers had to show gains in self-reported measures collected once annually on technology and CALM/PBL use.

Ninety-six percent of 23 teachers met "High" implementation in the first year of the local phase SY 2014–15. A total of 19 of 22 teachers (86%) met the threshold in Year 2 of the local phase, SY 2015–16. In Year 1 of the national phase, SY 2016–17, 24 of 28 teachers (86%) met the threshold based on the number of schools that provided data for this component. By the final year of the national phase, 24 of 27 teachers (89%) met the threshold. Overall, the majority of indicators relevant for KC4 were met with high fidelity for all four years and at the conclusion of both phases. The indicator that seemed to be particularly challenging was 4.6, use of CALM in the classroom.

KC5: High implementation fidelity to KC5, CORE teachers participate in follow-up CALM/PBL PD workshops, was met overall during both study phases.

KC5 consists of two indicators, and refers to teacher participation in three follow-up PBL/CALM PD workshops focused on project data collection, technology use, subject-specific support, and curriculum sharing (5.1) and sharing of learning experiences through presentations during CORE workshops or the CORE Academy (5.2). High fidelity was met when 81% of teachers attended 100% of the three follow-up workshops provided annually, and each CORE teacher actively participated in follow-up workshops as evidenced by providing at least one presentation at one of the three annual CORE workshops or at the annual CORE Academy by the conclusion of the two-year phase study.

The KC was measured each year during the grant cycle for both the local and national phase. During the local phase, 78% of 23 teachers met the threshold for SY 2014–15. Nineteen of 22 teachers (86%) met the threshold for SY 2015–16. During the national phase, SY 2016–17 data show that 97% of 29 teachers and 81% of 27 teachers met the workshop threshold in SY 2017–18.

KC6: High implementation fidelity to KC6, CORE principals provide change-management support to CORE schools, was met during both study phases.

KC6 fidelity refers to school administration participation in change management debriefing. High fidelity was met when 67% of principals participated in a debrief session for their results. The associated indicator 6.1 determined the participation of school administration in a debrief session about the change-management diagnostic tool administered in their school. The component was measured twice during the local phase and once during the grant cycle in Year 2 of the national phase.

During the local phase, 100% of seven schools participated in a debriefing session for change management in SY 2014–15. Three of four schools (75%) participated during SY 2015–16. Twenty schools participated in the Change Diagnostic Index assessment during SY 2016–17,

illustrating some level of school activity in the process, and 16 of 20 principals (80%) participated in a debriefing session held by PivotPoint and JSU in SY 2017–18.

KC7: Dual enrollment advisement and support to students in CORE schools (local phase Year 1) and CORE math and ELA (English language arts) teachers use EdReady™ college-readiness assessment tool in CORE schools (national phase), was not met during the local phase and was met during the national phase.

To achieve high fidelity during the first year of the local phase, dual enrollment students had to be provided with targeted support and guidance from JSU counselors. In 2015–16, this indicator was revised and operationalized as the number of math teachers that access and use the EdReady™ college-readiness tool. Two of five (40%) did so—fidelity was not met.

In the national phase, math and ELA teachers participating in the CORE treatment group needed to administer the EdReady™ college-readiness assessment tool. According to teacher log-in activity, 100% of 11 math and ELA teachers met the threshold in both Year 1 and Year 2 of the national study.

Table 10: FOI Status by KC and Cohort

KCs in the Logic Model	Definition of High Implementation	Definition of "implementation with Fidelity" at Program Level	Local Phase		National Phase	
			SY14–15	SY15–16	SY16–17	SY17–18
KC1. Principals in new CORE schools connect with principals in existing partnership schools.	Calculation based on 1 indicator at the end of each grant cycle	At least 81% of principals in CORE schools complete a post-event survey and agree or strongly agree with an item stating that the CORE event(s) they attended helped promote collaboration and networking among their peers.	83% Met	75% Not Met	100% Met	100% Met
KC2. CORE teachers participate in online learning communities.	Calculation based on 1 indicator at the end of each year of implementation	At least 81% of treatment teachers enroll in PLCs and show online activity with at least three posts per month. Online activity is averaged over nine months (academic year).	4% Not Met	14% Not Met	86% Met	93% Met
KC3. CORE teachers receive and adopt classroom technology resources and support.	Calculation based on 4 indicators (#3.1 through #3.4) 3.4 measured at the end of each year	At least 81% of CORE teachers receive technology hardware, classroom funds, and technical assistance (TA), as well as participate in a technology assessment.	100% Met	91% Met	97% Met	100% Met
KC4. CORE teachers participate in CORE Academy CALM/PBL PD.	Calculation based on 7 indicators (#4.1 – 4.7) at the end of each year of	At least 81% of teachers attend CORE Academy PD and complete the evaluation survey indicating the PD was of high quality, relevant,	96% Met	86% Met	86% Met	89% Met

	implementation	and useful. Teachers also report increased content knowledge, use of CALM in the classroom, and increased use of technology.				
KC5. CORE teachers participate in follow-up CALM/PBL PD workshops.	Calculation based on 2 indicators (#5.1 - #5.2) at the end of each year of implementation; national study Year 1 includes 5.1 only and Year 2 is both indicators	At least 81% of teachers attend CORE follow-up workshops and actively participate by presenting.	78% Not Met	86% Met	97% Met	81% Met
KC6. CORE principals provide change-management support to CORE schools.	Calculation based on one indicator measured at the end of each year (local phase) and once at the end of the grant cycle (national phase)	At least 67% of principals who elected to complete the CDI during each year will participate in a debrief session with PivotPoint regarding the CDI results.	100% Met	75% Met	N/A*	80% Met
KC7. CORE math and ELA teachers use EdReady™ college-readiness assessment tool in schools.	Calculation based on one indicator measured at the end of each year	Local phase: JSU counselors will provide targeted support to dual enrollment students. National phase: At least 81% of CORE math and English teachers will conduct an assessment using the EdReady™ tool.	0% Not Met	40% Not Met	100% Met	100% Met

*Measured only once in national two-year phase study

3.2 Implementation Fidelity by Indicator

Table 11 presents a breakdown of fidelity performance data comparing local and national study results by year. The table is organized by KC and lists the corresponding indicator(s) and scoring details and thresholds (e.g., low, medium, and high) as reported to the NEi3.

Fidelity to each indicator was assessed using the same scoring criteria established for each indicator's respective KC. For example, the threshold for high fidelity to KC3 is that 81% of the sample will achieve high implementation fidelity when data are aggregated across indicators 3.1–3.4. To make a fidelity determination separately for each individual indicator (i.e., 3.1, 3.2, 3.3, and 3.4), we first assessed what percentage of the sample met the criteria for "high" fidelity on each indicator. If at least 81% of the sample met the criteria for "high" fidelity at the indicator level, we determined fidelity was "met" for the indicator. When survey items are used for fidelity

indicators, denominators are determined by the actual number of respondents for that item rather than the number in the study.

Of the seven KCs and corresponding indicators highlighted in Table 11, indicator 3.4 and KC3 were implemented with high fidelity consistently across both phases of the study. Indicators 3.1, 3.2, and 3.3 measure the receipt by schools of technology (e.g., iPads and laptops) and financial resources to be used in CORE classrooms. Indicator 3.4 measured the one-on-one ongoing support ETAs provided to teachers. The goal of this support was to build teacher confidence and capacity to integrate technology in CORE classrooms. At least 10 support sessions and consultations were required to meet fidelity; but on average, teachers participated in 50 sessions. These sessions were held in person during the local phase and included virtual, live, and technical support offerings for the national phase. Teachers and ETAs shared in interviews that JSU CORE was successful in effectively providing technology resources and support via in-person and virtual means. This enabled the provision of customized, responsive classroom support to many CORE schools. While teachers varied in the amount of support requested and used from the ETAs, they commented on the effort by the JSU CORE team to provide this ongoing instructional peer support. This effort translated to strong outcomes for some.

Over the course of the four-year study, CORE made some changes to the program and subsequent data collection efforts. Specifically, changes to two components, KC2 (participation in PLC) and KC7 (use of college-readiness tools) led to increased performance on this fidelity threshold from the local phase to the national-phase study. Regarding KC2, teachers had to actively participate in an online PLC with colleagues, as evidenced by at least three monthly Facebook postings. Interviews revealed that teachers had meaningful communication and collaboration with colleagues during CORE Academy and PD workshops; however, performance with active participation in an online PLC in the local phase was low. To promote networking in the national phase, JSU CORE provided teachers with access to CANVAS, a new learning management system. KC7 was altered to provide the EdReady™ assessment tool for college readiness as a scalable tool for use in national rural schools, given policies for dual enrollment were inconsistently used. This change, although only applicable to math and ELA teachers, was more widely used in the national phase and as a result, performance increased.

Table 11. FOI for SY 2014–15 through SY 2017–18 by KC and Indicator

			FOI Local Phase		FOI National Phase	
KC	**Indicator**	**Scoring Details**	**SY 2014–15 (Year 1) Performance Data (N = 23)**	**SY 2015–16 (Year 2) Performance Data (N = 22)**	**SY 2016–17 (Year 1) Performance Data (N = 29)**	**SY 2017–18 (Year 2) Performance Data (N = 27)**
KC1. Principals in new CORE schools connect with principals in existing partnership schools.	1.1. Principal collaboration with their peers and superintendents at the CORE Academy and three CORE workshops	**1–2 = Low (0)** **3 = Med (1)** **4–5 = High (2)** Survey ratings	**83%** 5 of 6 "High" fidelity met	**75%** 3 of 4 "High" fidelity not met	**100%** 6 of 6 "High" fidelity met	**100%** 3 of 3 "High" fidelity met
KC2. CORE teachers participate in online learning communities.	2.1. Active participation in online PLCs	**0 = Low (0)** **1–2 = Med (1)** **>=3 = High (2)** # Monthly posts	**4%** 1 of 23 "High" fidelity not met	**14%** 3 of 22 "High" fidelity not met	**86%** 25 of 29 "High" fidelity met	**93%** 25 of 27 "High" fidelity met
KC3. CORE teachers receive and adopt classroom technology resources and support.	3.1. Provision of laptops and iPads for CORE teachers	**N = Low (0)** **Y = High (1)** Provided/distributed	**100%** 23 of 23 "High" fidelity met	**Measured once in Year 1 during 2-year grant cycle**	**100%** 29 of 29 "High" fidelity met	**Measured once in Year 1 during 2-year grant cycle**
	3.2. Provision of classroom technology funds for CORE teachers	**N = Low (0)** **Y = High (1)** Provided/distributed	**100%** 23 of 23 "High" fidelity met		**97%** 28 of 29 "High" fidelity met	
	3.3. Provision of classroom technology assessments	**N = Low (0)** **Y = High (1)** Provided/distributed	**100%** 23 of 23 "High" fidelity met		**97%** 28 of 29 "High" fidelity met	
	3.4. Provision of ongoing support from ETAs	**0–4 = Low (0)** **5–9 = Med (1)** **10 or more = High (2)** # ETSS Consultations	**100%** 23 of 23 "High" fidelity met	**91%** 20 of 22 "High" fidelity met	**97%** 28 of 29 "High" fidelity met	**100%** 27 of 27 "High" fidelity met
KC4. CORE teachers participate in CORE Academy CALM/PBL PD.	4.1. Teacher attendance at the annual CORE Academy	**0 = Low (0)** **1–2 = Med (1)** **3 = High (2)** # attendance days	**87%** 20 of 23 "High" fidelity met	**82%** 18 of 22 "High" fidelity met	**86%** 25 of 29 "High" fidelity met	**93%** 25 of 27 "High" fidelity met
	4.2. Quality of CORE Academy	**1–2 = Low (0)** **3 = Med (1)** **4–5 = High (2)** Survey ratings	**96%** 22 of 23 "High" fidelity met	**77%** 17 of 22 "High" fidelity not met	**79%** 22 of 28 "High" fidelity not met	**100%** 20 of 20 "High" fidelity Met
	4.3. Relevance of CORE Academy	**1–2 = Low (0)** **3 = Med (1)** **4–5 = High (2)** Survey ratings	**96%** 22 of 23 "High" fidelity met	**77%** 17 of 22 "High" fidelity not met	**81%** 22 of 27 "High" fidelity met	**95%** 19 of 20 "High" fidelity met

			FOI Local Phase		FOI National Phase	
KC	**Indicator**	**Scoring Details**	**SY 2014–15 (Year 1) Performance Data (N = 23)**	**SY 2015–16 (Year 2) Performance Data (N = 22)**	**SY 2016–17 (Year 1) Performance Data (N = 29)**	**SY 2017–18 (Year 2) Performance Data (N = 27)**
	4.4. Usefulness of CORE Academy	**1–2 = Low (0)** **3 = Med (1)** **4–5 = High (2)** Survey ratings	**96%** 22 of 23 "High" fidelity met	**77%** 17 of 22 "High" fidelity not met	**81%** 22 of 27 "High" fidelity met	**95%** 18 of 19 "High" fidelity met
	4.5. Increased content knowledge	**0–1.99 = Low (0)** **2–3.00 = Med (1)** **>3= High (2)** Average knowledge rating	**100%** **23 of 23** "High" fidelity met	**81%** **18 of 22** High" fidelity met	**81%** 22 of 27 "High" fidelity met	**100%** 19 of 19 "High" fidelity met
	4.6. Use of CALM	**N = Low (0)** **Y = High (1)** Change in average use	**63%** 10 of 16 "High" fidelity not met	**80%** 12 of 15 "High" fidelity not met	**65%** 11 of 17 "High" fidelity not met	**73%** 16 of 22 "High" fidelity not met
	4.7. Use of technology	**N = Low (0)** **Y = High (1)** Change in average use	**91%** 21 of 23 "High" fidelity not met	**100%** 15 of 15 "High" fidelity met	**85%** 23 of 27 "High" fidelity met	**90%** 19 of 21 "High" fidelity met
KC5. CORE teachers participate in follow-up CALM/PBL PD workshops.	5.1. Participation in follow-up workshops	**0 = Low (0)** **1–2 = Med (1)** **3 = High (2)** # Attendance days	**83%** 19 of 23 "High" fidelity met	**86%** 19 of 22 "High" fidelity met	**97%** 28 of 29 "High" fidelity met	**81%** 22 of 27 "High" fidelity met
	5.2. Sharing of learning experiences	**N = Low (0)** **Y = High (1)**	**100%** 23 of 23 "High" fidelity met	**86%** 19 of 22 "High" fidelity met	**Measured once in 2-year grant cycle**	**100%** 27 of 27 "High" fidelity met
KC6. CORE principals provide change-management support to CORE schools.	6.1. Participation in follow-up debriefings on change management	**N = Low (0)** **Y = High (1)** PivotPoint debriefing visits	**100%** 7 of 7 "High" fidelity met	**75%** 3 of 4 "High" fidelity met	**Measured once in 2-year grant cycle**	**80%** 16 of 20 "High" fidelity met
KC7. CORE Math and ELA teachers use EdReady™ college-readiness assessment tool in schools.	7.1. EdReady™ utilization log-in count	**N = Low (0)** **Y = High (1)** Log-in activity	**0%** 0 of 5 schools 0 of 7 counselors 43 of 179 scholars "High" fidelity not met	**40%** 2 of 5 "High" fidelity not met	**100%** 11 of 11 schools "High" fidelity met	**100%** 11 of 11 "High" fidelity met

3.3 Use of CALM/PBL and Technology by Intervention Group Status

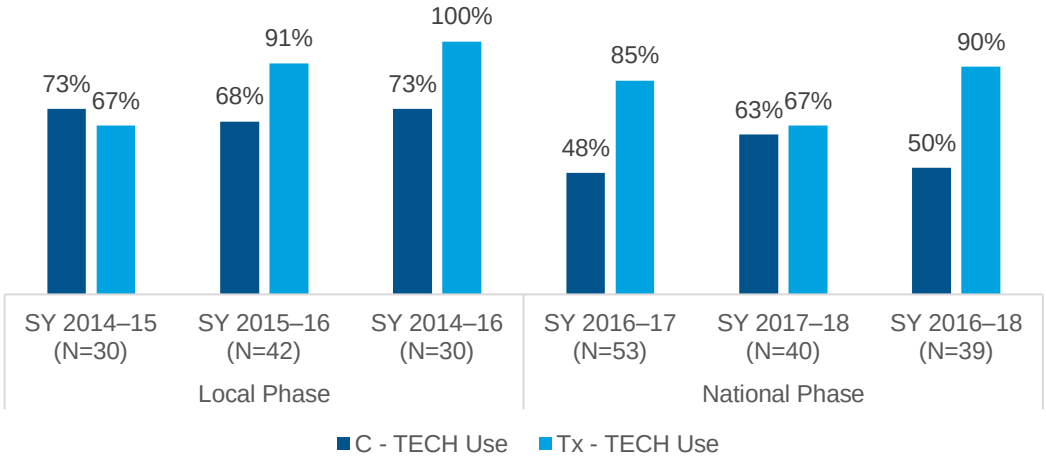
Fidelity indicators 4.6 and 4.7 tracked changes in CALM/PBL and technology usage at the teacher level. The data used to track these indicators were measured using a self-reported survey and collected from both treatment and control group teachers three times during each phase of the study: Baseline, Year 1, and Year 2. This provided an opportunity to compare longitudinal use of these CORE instructional practices for both groups of study participants. Given that the sample size was relatively small for this analysis, differences were analyzed using a descriptive methodology as opposed to statistical significance testing. For the FOI study, CALM/PBL and technology scores in the previous section of this report were evidenced by a positive change in self-reported gains, measured annually. In this section, we calculated the average score, by year, for only those teacher respondents who had scores available.

For both phases of the study, we determined increases in technology and CALM/PBL use by calculating the difference between teacher respondent scores at three time points. Both treatment and control group teachers were asked to complete final Year 2 versions of both CORE Methodology and Technology surveys to determine if there were any differences from their initial Baseline and Year 1 scores. Any reported positive differences between two time points were each given a value of "1" while all other values received no point value. These values were totaled for the numerator, and the denominator was the total number of matched teachers, excluding teachers from the originally randomized sample who did not complete both surveys.

In the local-phase study, as shown in Figure 1, 100% of 15 treatment group teachers who completed the SY 2014 and SY 2016 technology surveys showed an increase in their classroom technology use at the end of Year 2.[10] As expected, for control teachers this percentage was lower. Only 73% of 15 control group teachers reported increased technology use during the same period.

[10] Local Phase Technology Use: SY 2014 and SY 2015 (TxN=15), SY 2015 and SY 2016 (TxN=23), SY 2014 and SY 2016 (TxN=15). National Phase Technology Use: SY 2014 and SY 2015 (TxN=28), SY 2015 and SY 2016 (TxN=22), SY 2014 and SY 2016 (TxN=21).

Figure 1. Teacher-Reported Use of Technology

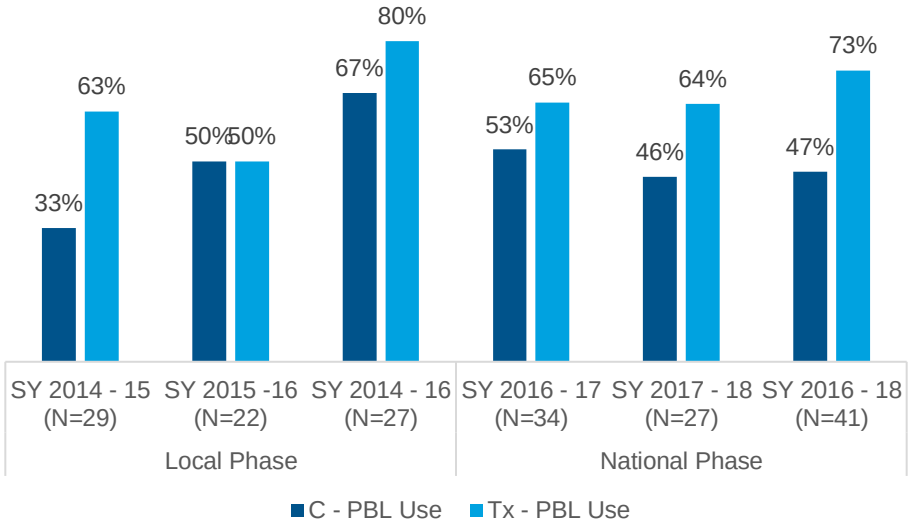


Notably, we found that 19 treatment group teachers of the 21 (90%) who completed both the baseline technology survey and the 2018 post-test technology survey showed an increase in their classroom technology use at the end of the two-year national study. As expected, for control teachers this percentage was lower. Only 50% (9 out of 18) of control group teachers who completed the survey at both time points reported increased technology use during the same period.

Regarding the CALM/PBL items, of the 12 control group teachers who completed both the SY 2014 and SY 2016 surveys, eight (67%) reported an increase in their CALM usage during Year 2 (see Figure 2). This is compared to 80% of the 15 treatment teachers who took both the SY 2014 and SY 2016 surveys.[11] Of the 19 control group teachers in the national-phase study who completed the SY 2016 survey, 9 (47%) reported an increase in their use of CALM instructional strategies during Year 2, SY 2018. This is compared to 73% of the 22 treatment teachers who took the SY 2016-18 surveys.

[11] Local Phase for PBL/CALM Use: SY 2014 and SY 2015 (TxN=15), SY 2015 and SY 2016 (TxN=10), SY 2014 and SY 2016 (TxN=17). National Phase Treatment Group for PBL/CALM Use: SY 2014 and SY 2015 (TxN=17), SY 2015 and SY 2016 (TxN=14), SY 2014 and SY 2016 (TxN=21).

Figure 2. Teacher-Reported Use of CALM/PBL



The section below compares treatment and control teachers in terms of their classroom technology and PBL approaches using inferential statistics. This is an exploratory analysis and the findings do not suggest causality; however, our general expectation is that the treatment teachers might have increased their reliance on classroom technology and PBL approaches. As described earlier in detail, participant teachers reported on the use of these two approaches through teacher surveys. Teachers in both phases of the study responded to the surveys three times during their respective study phase; surveys were always administered in June. The two groups were compared on three change scores: (a) Year 1 to Year 2 change, (b) Year 2 to Year 3 change, and (c) Year 1 to Year 3 change. While CORE teachers did not meet adequate fidelity for the indicator related to PBL/CALM scores, gains for both local and national phases showed treatment teachers, on average, were increasing in both knowledge and confidence in this area.

Figure 3 shows the self-reported gains in confidence for control and treatment group teachers. Specifically, during Year 1 to Year 2 of the local phase, the treatment group's change score (.73) was greater than the control group's (-0.09). This trend continued to Year 3 of the local-phase study. The Year 1 to Year 3 change score (.99) was greater than the control group's (0.10).

Figure 3. Matched Teacher Reported Gains: Use of CALM by Year and Study Phase

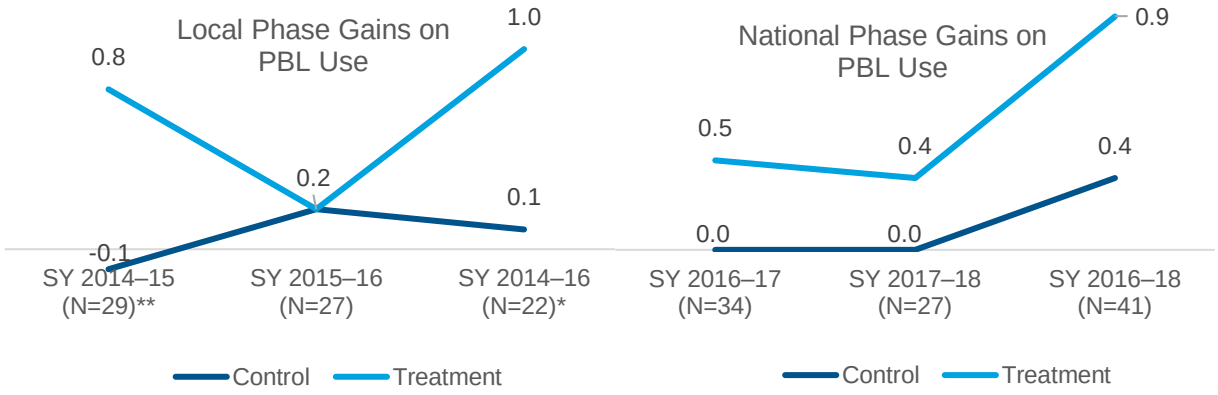
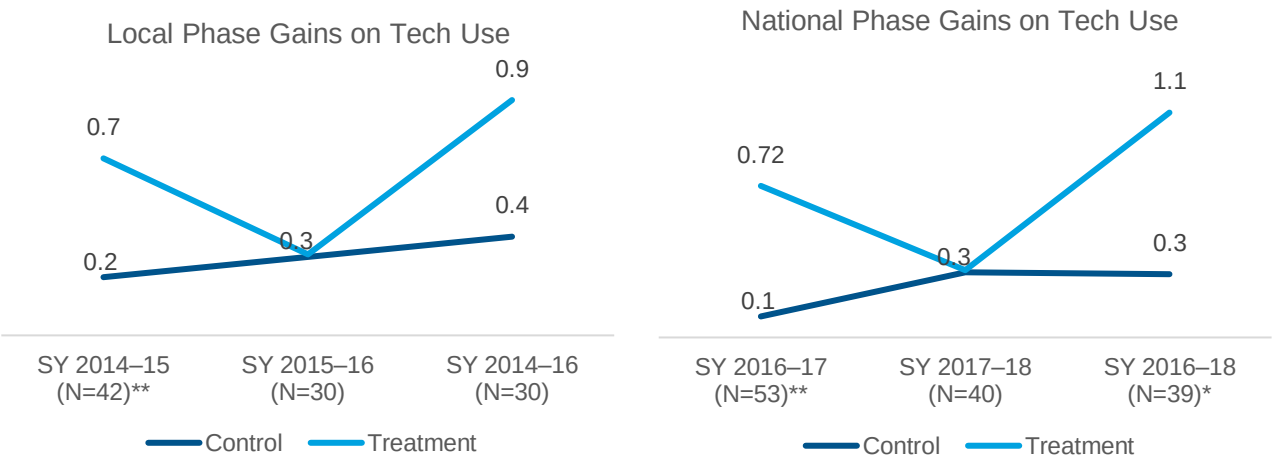


Figure 4 shows the self-reported gains in confidence for control and treatment group teachers. CORE teachers met adequate fidelity for the indicator related to technology use scores each year of the study except for the first year of the local phase. Gains for both local and national phases showed treatment teachers, on average, were increasing in both knowledge and confidence in this area. During the local-phase study, the treatment group's Year 1 to Year 3 change score (0.70) was greater than the control group's (.23) and the difference was statistically significant. This trend continued for the national phase, when the Year 1 to Year 2 classroom technology score was greater for the treatment group (.72) than for the control group (.1) and the difference (0.63) was statistically significant. In the final year of the study, the treatment group's Year 1 to Year 3 change score (1.07) was greater than the control group's (.30) and the difference (0.76) was statistically significant.[12]

Figure 4. Matched Teacher Reported Gains: Use of Technology by Year and Study Phase



[12] Statistical significance (2-tail test): * = p<.05, ** = p<.01, *** = p<.001. Further details are available in the appendix.

3.4 Correlation between Students' Average CWRA+ Scores and Teachers' Implementation Scores

This section addresses the exploratory question 4 and presents findings related to the correlation between teachers' program implementation and students' average CWRA+ scores. As presented earlier, the program impact analysis did not support the expectation that the CORE program improves students' CWRA+ scores or students' engagement and self-efficacy levels. However, the program effect may depend on how much teachers implement the programs in their everyday teaching through the use of classroom technology and PBL (CALM), as described in Table C3 and Table C4. To test this exploratory hypothesis, we examined the correlation between teachers' classroom technology use and PBL scores, and their students' average CWRA+ change scores. This exploratory analysis produced no statistically significant results.[13]

The unit of analysis was the teachers and thus teacher provided classroom technology and PBL scores by participating in surveys. For each teacher, we calculated the average change score based on students CRWA+ scores (by subtracting a posttest score from a pretest score at the student level and derived the average score per teacher). A limitation of this analysis is that the number of teachers was relatively small. We conducted the analysis using all teachers (treatment and control teachers combined) and treatment teachers alone. As our interest is primarily with the treatment teachers, we presented the results from only the sample of treatment teachers; however, the results were the same with either of the samples. None of the correlation statistics were statistically significant at alpha=5%.

Table D1 (In Appendix D) shows the results from local phase teachers. For both Year 1 and Year 2, the teacher average CWRA+ change scores and classroom technology scales expressed as pretest score, posttest score, and posttest – pretest score were not associated with statistical significance. The same was true for the PBL scale. The teacher average of students' change scores was not correlated with the PBL scales with statistical significance.

The results from the national phase was consistent with the results from the local phase (see Appendix D, Table D2). The correlation statistics obtained for national phase teachers from Year 1 and Year 2 were not statistically significant. Again, the sample consisted of only treatment teachers, but the results from the all teacher analyses were without any statistically significant findings (see Appendix D, Tables D3 and D4).

IV. Discussion

In this report, we provided formative and summative details from the evaluation of two studies—local and national—for the 2013 CORE project. After local-phase implementation, a strategic effort took place to revise the model to provide supports for rural schools that are scalable on a national level. Despite the achievement of adequate FOI of the national-phase study, JSU CORE did not produce statistical evidence of program impact for the CORE program. While the CORE program itself brought about various positive experiences among participating teachers, and many expressed growth in their PD as evidenced by the FOI findings, significant student-

[13] Statistical significance (2-tail test): * = p<.05, ** = p<.01, *** = p<.001

level impact was not evident. This section provides possible explanations for why this evidence was not found. There are a few important factors that may have contributed to the absence of student level impact: (1) principal participation, (2) exposure to the intervention, and (3) relevancy of PD. Each of these factors is discussed in detail in this section.

The CORE model was intended to support school-level change and promote principal collaboration around 21st century skills and college readiness. Since the CORE grant relies on one teacher and one principal per school to engage the entire school in the CORE model, it may be challenging for the school to fully adopt the model if there is a lack of consistency in leadership or staff turnover. Lack of continuity in the roles of those who serve as program champions could be detrimental for overall program success. Based on site visits and participant interviews, principals in both phases were enthusiastic about the participation of their teachers in the project, but many could not identify the CORE model or a way to expand its efforts beyond the one participating teacher. Thus, although KC1 (i.e., principals in new CORE schools connect with principals in existing partnership schools) and KC6 (i.e., change-management debriefing) were met with adequate fidelity, the number of principal respondents was low and not representative of the study sample. During the national-phase study, several principal debriefings were conducted by JSU in the second year of the study, however this may have occurred too late in the implementation phase to support schoolwide change.

Another factor that may have played a role in program impact was the targeted focus of the intervention. CORE components were provided to treatment teachers without coordinated efforts to reach and collaborate with other teachers at participating schools to demonstrate collective impact. The intervention could be more effective if it is shared with other teachers in the school and implemented collaboratively. For example, while students were taught by teachers participating in the intervention for part of the day, this limited exposure may not be sufficient to bring about change in school culture and detectable program impact. In fact, the CORE program was originally designed as a school-level intervention, requiring collaboration among teachers. Another grant, the 2016 Validating the Collaborative Regional Education (CORE) Comprehensive Model in Rural High Schools program evaluation is being conducted to allow the estimation of program impact at the school level.

Our interpretation of the null effects is that student-level impact was not apparent in the limited timeframe of the study. Statistically significant differences in CWRA+ outcomes may require more time to emerge. The confirmatory hypothesis focused on CWRA+ scores, which measured students' critical thinking and problem-solving abilities, important skills for students' college and work readiness. When preparatory analysis was conducted,[14] only half of students had a positive change from the time of pretest to the posttest. Students' skills and knowledge as measured on the tests may need more than one school year to manifest. Although each study phase was two years long, the students sampled were only exposed to the treatment teacher for one year. Future evaluations of the CORE program may need to consider the possibility of measuring longer-term change, or supplementing the study with other student outcomes, such as math and reading scores. The 2016 CORE intervention project that is currently underway is

[14] Analysis results available upon request.

following the same set of students for three years and thus their growth in CWRA+ scores may be more substantial and program impact may be easier to detect.

Finally, CORE program PD offerings focused on teacher CALM/PBL and technology skills through the CORE Academy and follow-up workshops. Indicators 4.2 and 4.3 measured teacher-reported agreement that the Academy was of high quality and relevant. Participating teachers gained knowledge in technical platforms that could be used as a part of general classroom instruction implementing the CORE model. Because the PD did not have narrow parameters on who could participate based on content area taught, teachers of different subject areas participated and were expected to translate the knowledge they gained into their subject-specific instruction (i.e., English, history, social studies, mathematics, and science). Because program content was meant to be adapted for multiple subject areas, there may not have been enough targeted focus on subject-specific content covered in workshops. Teachers participating in site visit interviews expressed that they did not find some sessions to be as relevant for their work. Future research may focus on one or two subject fields (e.g., English, math, science) and seek to tailor PD content to the targeted teachers.

V. Conclusions

Impact Study Conclusions: In collaboration with JSU, the ICF evaluation team conducted a cluster RCT evaluation to estimate the impact of CORE upon college and work-readiness outcomes for grade 8–12 students during a two-phase study (a two-year local study and a two-year national study). The local study began with teachers and students from 42 middle and high schools, and the national study began with teachers and students from 63 middle and high schools.[15] This evaluation was conducted as part of a comprehensive evaluation of CORE. The study exhibited low-cluster (i.e., teacher) and sub-cluster (i.e., student) attrition during three of the four study years. This resulted in the confirmatory analysis being eligible to receive a rating of *Meets WWC Evidence Standards, Without Reservations*, the highest rating available from the WWC.

After analyzing the data, the ICF evaluation team found no statistically significant or practically meaningful differences in average CWRA+ SRQ, student engagement, or student self-efficacy outcomes across the treatment and control group samples during all four study years. An exploration of impact on CWRA+ scores, student engagement, and student self-efficacy within sub-groups yielded significant positive findings for two groups: 12th graders of teachers participating in the national phase and 10th graders during all study years. Alternatively, a negative program impact was detected for Black students of teachers participating in the national phase. However, due to lack of consistency with these findings, the research team concluded that the data did not generate evidence to support program impact on student subgroups.

[15] Due to attrition, the Year 2 local study sample included 37 schools and the Year 2 national study sample included 52 schools.

Implementation Study Conclusions: A mixed-method FOI study was conducted to determine the extent to which seven KCs of the CORE program were implemented as intended. The study revealed that the majority of the seven KCs were delivered with high fidelity during the grant period. Challenges with fidelity were more apparent during the local-phase study, and in some cases, were remedied by revising indicators or measurement approaches to more accurately measure key activities. Specifically, KC1 and KC2 were not met with high fidelity during the local phase, but high fidelity was achieved during the national phase. Similarly, KC7 was met with high fidelity during the national phase only. KCs 3, 4, 5, and 6 were met with high fidelity during both study phases. KC3, KC4, and KC5 consisted of multiple indicators, a few of which were not met. However, high fidelity was achieved overall when considering the indicators for these three components as a whole. Taken together, these results provide evidence that the CORE intervention was implemented as intended. Although a few implementation challenges emerged, these were addressed by the end of the grant period.

VI. References

Blanchard, M. R., LePrevost, C. E., Tolin, A. D., & Gutierrez, K. S. (2016). Investigating technology-enhanced teacher professional development in rural, high-poverty middle schools. *Educational Researcher, 45*(3), 2016.

Byun, S., Meece, J. L., & Irvin, M. J. (2012). Rural–nonrural disparities in postsecondary educational attainment revisited. *American Educational Research Journal, 49*(3), 412–437. http://eric.ed.gov/?id=EJ968047

Consortium on Chicago School Research at the University of Chicago. (2007). *Sample high school detailed report.* Retrieved September 19, 2015 from http://ccsr.uchicago.edu

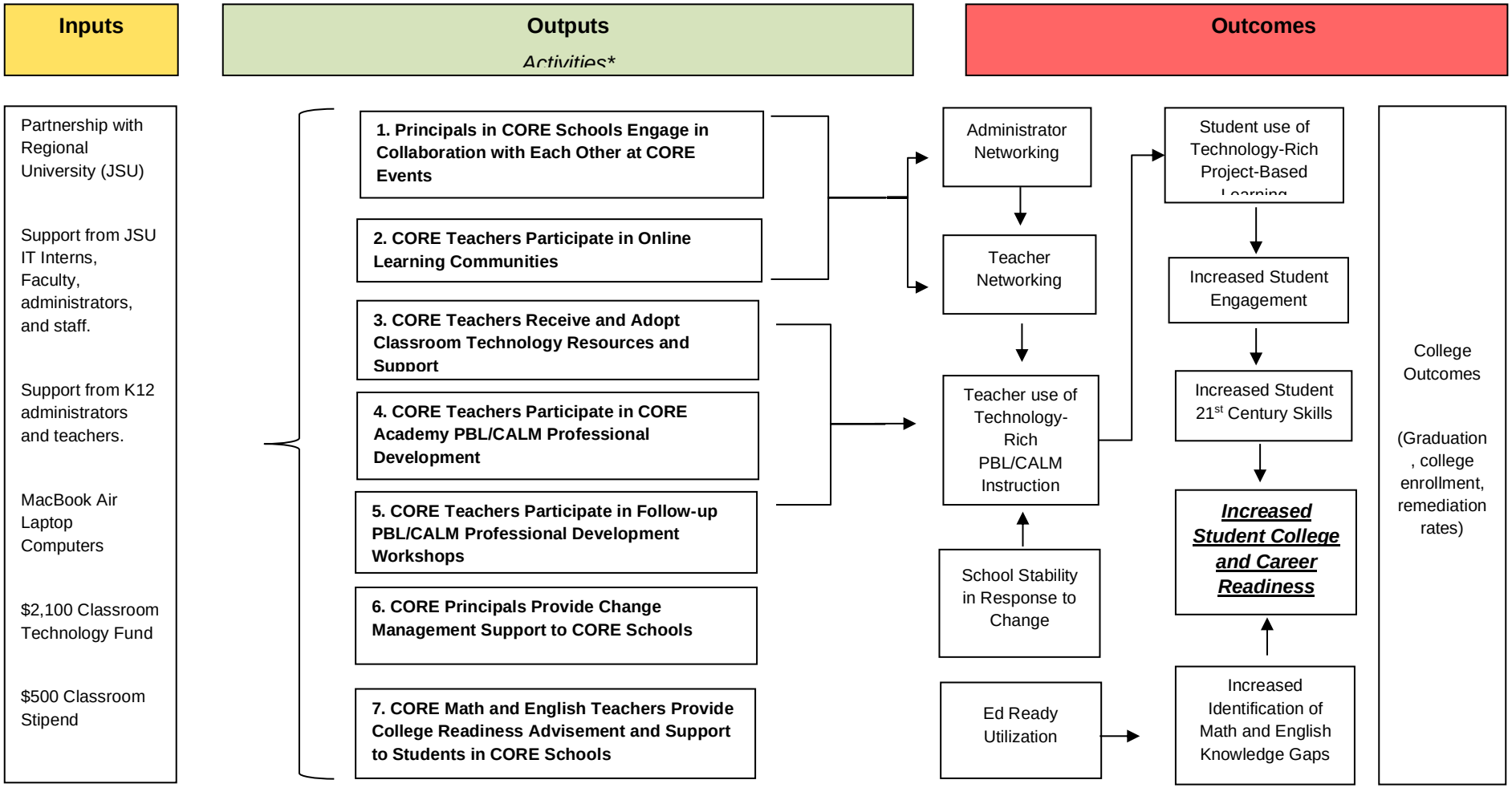
Howley, C. & Hambrick, K. (2011) *Attachment and aspiration: What influences rural youths' educational and residential plans?* White Paper. Retrieved October 10, 2018 from http://eric.ed.gov/ ERIC Number: ED532584

Midgley, C., Maehr, M. L., Hruda, L., Anderman, E. M., Freeman, K. E., Gheen, M., & Urdan, T. (2000). *Manual for the patterns of adaptive learning scales (PALS).* Ann Arbor, MI: University of Michigan.

What Works Clearinghouse. (2013). *What Works Clearinghouse: Procedures and standards handbook* (Version 3.0). Retrieved September 1, 2018 from http://whatworks.ed.gov

P21.org Partnership for 21st Century Skills. Web. 16 Sept 2018. http://www.p21.org/about-us/our-mission.

Appendix A: Logic Model



External Factors

Schools and districts support teachers' participation in professional development, learning communities, project-based learning, and technology use.

Schools and districts are supportive of changes to technology infrastructure and practice in participating schools.

Teacher and Student turnover are minimized.

Appendix B: Impact Study Tables

Table B1. Local Phase Year 1: HLM Results for the Assessment of CORE Program Impact on SY 2014–15 Student CWRA+ SRQ Scores

	Anova Model				Final Model				
	Coeff.	SE	p	Sig	Coeff.	SE	p	Sig	Standardized Effect
Intercept	853.94	8.43	0.00	***	852.11	14.26	0.00	***	-0.02
Treatment status					-5.87	11.9	0.62		-0.04
CWRA+ (pretest, centered)					0.36	0.02	0.00	***	0.00
Male					-5.91	6.33	0.35		-0.04
9th grader					24.31	16.53	0.15		0.16
10th grader					20.38	15.72	0.20		0.13
11th grader					16.8	15.69	0.29		0.11
12th grader					8.54	15.32	0.58		0.06
Other race groups					-28.44	10.08	0.00	**	-0.19
African American					-30.27	8.88	0.00	***	-0.2
Hispanic					-34.2	17.86	0.06		-0.22
Parent with college degree					12.53	6.79	0.07		0.08
Parent college info missing					0.00				0.00
School's exposure to CORE-like elements prior to the intervention					14.78	13.71	0.29		0.10
Model Summary Statistics									
Joint F-Test race groups					5.92		0.00	***	
Joint F-Test grade levels					0.77		0.55		
Level-1 variance	21,358			***	18,316			***	
Level-2 variance	2,433			***	937			**	
Intraclass correlation	0.10				0.05				
Level-1 variance explained					0.14				
Level-2 variance explained					0.61				

Note: Number of students: 1,902. Number of teachers: 42. Statistical significance (2-tail test): ns = not significant, * = $p<.05$, ** = $p<.01$, *** = $p<.001$. Joint F test: Race groups were statistically significant predictors. Grade levels were not statistically significant predictors. Omitted categories for grade levels and race categories were grade 8 and White students. Standardized effects were based on the analysis sample's outcome mean and SD.

Table B2. Local Phase Year 1: Descriptive Statistics for HLM Results for the Assessment of CORE Program Impact on SY 2014–15 Student CWRA+ SRQ Scores

	N. of cases	Minimum	Maximum	Mean	SD
CWRA+ test score	1,902	512.00	1,372.00	855.05	153.44
Treatment status	1,902	0.00	1.00	0.56	0.50
CWRA+ (pretest)	1,902	433.00	1,435.00	863.98	160.60
CWRA+ (pretest, centered)	1,902	-430.98	571.02	0.00	160.60
Male	1,902	0.00	1.00	0.50	0.50
8th grader	1,902	0.00	1.00	0.43	0.49
9th grader	1,902	0.00	1.00	0.14	0.34
10th grader	1,902	0.00	1.00	0.14	0.35
11th grader	1,902	0.00	1.00	0.14	0.35
12th grader	1,902	0.00	1.00	0.16	0.36
Other race groups	1,902	0.00	1.00	0.14	0.35
African American	1,902	0.00	1.00	0.20	0.40
Hispanic	1,902	0.00	1.00	0.03	0.18
White	1,902	0.00	1.00	0.62	0.48
Parent with college degree	1,902	0.00	1.00	0.35	0.48
Parent college info missing	1,902	0.00	0.00	0.00	0.00
School's exposure to CORE-like elements prior to the intervention	1,902	0.00	1.00	0.30	0.46

Table B3. Local Phase Year 2: HLM Results for the Assessment of CORE Program Impact on SY 2015–16 Student CWRA+ SRQ Scores

	Anova Model				Final Model				
	Coeff.	SE	p	Sig	Coeff.	SE	p	sig	Standardized Effect
Intercept	887.14	11.35	0.0	***	860.93	24.15	0.00	***	-0.13
Treatment status					0.50	18.48	0.98		0.00
CWRA+ (pretest, centered)					0.41	0.02	0.00	***	0.00
Male					2.38	6.51	0.71		0.01
9th grader					44.57	21.43	0.04	*	0.27
10th grader					25.17	20.85	0.23		0.15
11th grader					32.44	20.81	0.12		0.19
12th grader					31.04	21.38	0.15		0.19
Other race groups					-15.68	10.6	0.14		-0.09
African American					-27.91	9.14	0.00	**	-0.17
Hispanic					-17.7	17.23	0.30		-0.11
Parent with college degree					23.71	7.02	0.00	***	0.14
Parent college info missing					-9.81	16.71	0.56		-0.06
School's prior exposure to CORE-like elements					7.87	21.1	0.71		0.05
Model Summary Statistics									
Joint F-Test race groups					3.36		0.02	*	
Joint F-Test grade levels					1.21		0.31		
Level-1 variance	23,762			***	19,594			***	
Level-2 variance	4,185			***	2,374			***	
Intraclass correlation	0.15				0.11				
Level-1 variance explained					0.18				
Level-2 variance explained					0.43				

Note: Number of students: 1,948. Number of teachers 37. Statistical significance (2-tail test): ns = not significant, * = p<.05, ** = p<.01, *** = p<.001. Joint F test: Race groups were statistically significant predictors. Grade levels were not statistically significant predictors. Omitted categories for grade levels and race categories are grade 8 and White students. Standardized effects were based on the analysis sample's outcome mean and SD.

Table B4. Local Phase Year 2: Descriptive Statistics for HLM Results for the Assessment of CORE Program Impact on SY 2015–16 Student CWRA+ SRQ Scores

	N. of cases	Minimum	Maximum	Mean	SD
CWRA+ test score	1,948	506.00	1,508.00	882.52	166.54
Treatment status	1,948	0.00	1.00	0.65	0.48
CWRA+ (pretest)	1,948	467.00	1,493.00	869.94	159.07
CWRA+ (pretest, centered)	1,948	-402.94	623.06	0.00	159.07
Male	1,948	0.00	1.00	0.50	0.50
8th grader	1,948	0.00	1.00	0.26	0.44
9th grader	1,948	0.00	1.00	0.18	0.39
10th grader	1,948	0.00	1.00	0.15	0.35
11th grader	1,948	0.00	1.00	0.20	0.40
12th grader	1,948	0.00	1.00	0.21	0.41
Other race groups	1,948	0.00	1.00	0.12	0.32
African American	1,948	0.00	1.00	0.20	0.40
Hispanic	1,948	0.00	1.00	0.04	0.19
White	1,948	0.00	1.00	0.65	0.48
Parent with college degree	1,948	0.00	1.00	0.34	0.47
Parent college info missing	1,948	0.00	1.00	0.04	0.20
School's exposure to CORE-like elements prior to the intervention	1,948	0.00	1.00	0.30	0.46

Table B5. National Phase Year 1: HLM Results for the Assessment of CORE Program Impact on SY 2016–17 Student CWRA+ SRQ Scores

	Anova Model				Final Model				
	Coeff.	SE	p	Sig	Coeff.	SE	p	Sig	Standardized Effect
Intercept	906.56	9.63	0.0	***	919.53	12.94	0.00	***	0.06
Treatment status					-8.91	12.37	0.47		-0.05
CWRA+ (pretest, centered)					0.46	0.02	0.00	***	0.00
Male					-11.83	5.31	0.03	*	-0.07
9th grader					-11.32	12.65	0.37		-0.06
10th grader					15.39	12.72	0.23		0.09
11th grader					2.68	12.77	0.83		0.02
12th grader					0.39	13.33	0.98		0.00
Other race groups					-44.9	19.28	0.02	*	-0.26
African American					-43.33	9.24	0.00	***	-0.25
Hispanic					-22.54	10.57	0.03	*	-0.13
Parent with college degree					22.32	5.88	0.00	***	0.13
Parent college info missing					14.2	13.09	0.28		0.08
Model Summary Statistics									
Joint F-Test race groups					8.94		0.00	***	
Joint F-Test grade levels					1.65		0.16		
Level-1 variance	25,929.2			***	20,473.4			***	
Level-2 variance	4,686.74			***	1,704.43			***	
Intraclass correlation	0.15				0.08				
Level-1 variance explained					0.21				
Level-2 variance explained					0.64				

Note: Number of students: 2,983. Number of teachers 57. Statistical significance (2-tail test): ns = not significant, * = $p<.05$, ** = $p<.01$, *** = $p<.001$. Joint F test: Race groups were statistically significant predictors. Grade levels were not statistically significant predictors. Omitted categories for grade levels and race categories are grade 8 and White students. Standardized effects were based on the analysis sample's outcome mean and SD.

Table B6. National Phase Year 1: Descriptive Statistics for HLM Results for the Assessment of CORE Program Impact on SY 2016–17 Student CWRA+ SRQ Scores

	N. of cases	Minimum	Maximum	Mean	SD
CWRA+ test score	2,983	498.00	1,529.00	908.80	175.01
Treatment status	2,983	0.00	1.00	0.51	0.50
CWRA+ (pretest)	2,983	498.00	1,455.00	908.18	167.59
CWRA+ (pretest, centered)	2,983	-410.18	546.82	0.00	167.59
Male	2,983	0.00	1.00	0.50	0.50
8th grader	2,983	0.00	1.00	0.23	0.42
9th grader	2,983	0.00	1.00	0.17	0.37
10th grader	2,983	0.00	1.00	0.23	0.42
11th grader	2,983	0.00	1.00	0.21	0.41
12th grader	2,983	0.00	1.00	0.17	0.38
Other race groups	2,983	0.00	1.00	0.02	0.14
African American	2,983	0.00	1.00	0.20	0.40
Hispanic	2,983	0.00	1.00	0.11	0.31
White	2,983	0.00	1.00	0.68	0.47
Parent with college degree	2,983	0.00	1.00	0.35	0.48
Parent college info missing	2,983	0.00	1.00	0.05	0.21

Table B7. National Phase Year 2: HLM Results for the Assessment of CORE Program Impact on SY 2017–18 Student CWRA+ SRQ Scores

	Anova Model				Final Model				
	Coeff.	SE	p	Sig	Coeff.	SE	p	Sig	Standardized Effect
Intercept	914.79	10.54	0.0	***	928.02	15.57	0.00	***	0.06
Treatment status					4.39	14.48	0.76		0.03
CWRA+ (pretest, centered)					0.5	0.02	0.00	***	0.00
Male					-15.7	5.58	0.00	**	-0.09
9th grader					7.39	14.94	0.62		0.04
10th grader					5.55	14.79	0.71		0.03
11th grader					14.17	15.22	0.35		0.08
12th grader					12.73	15.62	0.42		0.07
Other race groups					-22.91	8.71	0.01	**	-0.13
African American					-47.86	9.69	0.00	***	-0.27
Hispanic					-22.92	10.74	0.03	*	-0.13
Parent with college degree					0.01	6.18	1.00		0.00
Parent college info missing					-5.95	13.43	0.66		-0.03
Model Summary Statistics									
Joint F-Test race groups					9.27		0.00	***	
Joint F-Test grade levels					0.37		0.83		
Level-1 variance	25,279			***	19,171			***	
Level-2 variance	5,619			***	2,248			***	
Intraclass correlation	0.18				0.10				
Level-1 variance explained					0.24				
Level-2 variance explained					0.57				

Note: Number of students: 2,609. Number of teachers: 52. Statistical significance (2-tail test): ns = not significant, * = $p<.05$, ** = $p<.01$, *** = $p<.001$. Joint F test: Race groups were statistically significant predictors. Grade levels were not statistically significant predictors. Omitted categories for grade levels and race categories are grade 8 and White students. Standardized effects were based on the analysis sample's outcome mean and SD.

Table B8. National Phase Year 2: Descriptive Statistics for HLM Results for the Assessment of CORE Program Impact on SY 2017–18 Student CWRA+ SRQ Scores

	N. of cases	Minimum	Maximum	Mean	SD
CWRA+ test score	2,609	498.00	1,513.00	918.19	174.43
Treatment status	2,609	0.00	1.00	0.54	0.50
CWRA+ (pretest)	2,609	533.00	1,559.00	905.98	165.94
CWRA+ (pretest, centered)	2,609	-372.98	653.02	0.00	165.94
Male	2,609	0.00	1.00	0.49	0.50
8th grader	2,609	0.00	1.00	0.18	0.38
9th grader	2,609	0.00	1.00	0.16	0.36
10th grader	2,609	0.00	1.00	0.26	0.44
11th grader	2,609	0.00	1.00	0.22	0.41
12th grader	2,609	0.00	1.00	0.19	0.39
Other race groups	2,609	0.00	1.00	0.13	0.34
African American	2,609	0.00	1.00	0.19	0.39
Hispanic	2,609	0.00	1.00	0.10	0.30
White	2,609	0.00	1.00	0.57	0.49
Parent with college degree	2,609	0.00	1.00	0.33	0.47
Parent college info missing	2,609	0.00	1.00	0.05	0.22

Table B9. Local Phase Year 1: HLM Results for the Assessment of CORE Program Impact on SY 2014–15 Student Engagement Scores

	Anova Model				Final Model				
	Coeff.	SE	p	Sig	Coeff.	SE	p	Sig	Standardized Effect
Intercept	3.02	0.02	0.0	***	3.07	0.04	0.00	***	0.10
Treatment status					-0.04	0.03	0.14		-0.08
Engagement level (pretest centered)					0.51	0.02	0.00	***	0.95
Male					-0.09	0.02	0.00	***	-0.16
9th grader					-0.03	0.04	0.41		-0.06
10th grader					0.00	0.04	0.99		0.00
11th grader					0.00	0.04	0.93		-0.01
12th grader					-0.04	0.04	0.37		-0.07
Other race groups					0.04	0.03	0.28		0.07
African American					0.01	0.03	0.73		0.02
Hispanic					0.03	0.06	0.63		0.06
Parent with college degree					0.01	0.02	0.57		0.03
School's exposure to CORE-like elements prior to the intervention					0.04	0.03	0.19		0.08
Model Summary Statistics									
Joint F-Test race groups					0.44		0.73		
Joint F-Test grade levels					0.38		0.82		
Level-1 variance	0.27			***	0.22			***	
Level-2 variance	0.01			**	0.00				
Intraclass correlation	0.04				0.01				
Level-1 variance explained					0.20				
Level-2 variance explained					0.80				

Note: Number of students: 1,823. Number of teachers: 42. Statistical significance (2-tail test): ns = not significant, * = p<.05, ** = p<.01, *** = p<.001. Joint F test: Race groups were not statistically significant predictors. Grade levels were not statistically significant predictors. Omitted categories for grade levels and race categories are grade 8 and White students. Standardized effects were based on the analysis sample's outcome mean and SD.

Table B10. Local Phase Year 1: Descriptive Statistics for HLM Results for the Assessment of CORE Program Impact on SY 2014–15 Student Engagement Scores

	N. of cases	Minimum	Maximum	Mean	SD
Engagement level	1,823	1.00	4.00	3.02	0.53
Treatment status	1,823	0.00	1.00	0.56	0.50
Engagement level (pretest,centered)	1,823	-2.06	0.94	0.00	0.48
Male	1,823	0.00	1.00	0.50	0.50
8th grader	1,823	0.00	1.00	0.43	0.50
9th grader	1,823	0.00	1.00	0.14	0.34
10th grader	1,823	0.00	1.00	0.14	0.35
11th grader	1,823	0.00	1.00	0.14	0.35
12th grader	1,823	0.00	1.00	0.15	0.36
Other race groups	1,823	0.00	1.00	0.14	0.35
African American	1,823	0.00	1.00	0.20	0.40
Hispanic	1,823	0.00	1.00	0.04	0.18
White	1,823	0.00	1.00	0.62	0.48
Parent with college degree	1,823	0.00	1.00	0.35	0.48
Parent college info missing	1,823	0.00	0.00	0.00	0.00
School's exposure to CORE-like elements prior to the intervention	1,823	0.00	1.00	0.30	0.46

Table B11. Local Phase Year 2: HLM Results for the Assessment of CORE Program Impact on SY 2015–16 Student Engagement Scores

	Anova Model				Final Model				
	Coeff.	SE	p	Sig	Coeff.	SE	p	Sig	Standardized Effect
Intercept	3.00	0.03	0.00	***	3.00	0.06	0.00	***	0.00
Treatment status					0.01	0.04	0.87		0.01
Engagement level (pretest,centered)					0.48	0.02	0.00	***	0.87
Male					-0.04	0.02	0.13		-0.06
9th grader					-0.02	0.05	0.76		-0.03
10th grader					-0.07	0.05	0.19		-0.13
11th grader					0.02	0.05	0.70		0.04
12th grader					0.03	0.05	0.52		0.06
Other race groups					0.05	0.04	0.23		0.08
African American					0.09	0.03	0.01	**	0.16
Hispanic					0.00	0.06	0.99		0.00
Parent with college degree					-0.01	0.03	0.75		-0.01
Parent college info missing					-0.06	0.07	0.37		-0.11
School's exposure to CORE-like elements prior to the intervention					0.00	0.05	0.98		0.00
Model Summary Statistics									
Joint F-Test race groups					2.77		0.04	*	
Joint F-Test grade levels					1.47		0.22		
Level-1 variance	0.29			***	0.23			***	
Level-2 variance	0.02			**	0.01			*	
Intraclass correlation	0.06				0.03				
Level-1 variance explained					0.19				
Level-2 variance explained					0.60				

Note: Number of students: 1,770. Number of teachers: 37. Statistical significance (2-tail test): ns = not significant, * = $p<.05$, ** = $p<.01$, *** = $p<.001$. Joint F test: Race groups were statistically significant predictors. Grade levels were not statistically significant predictors. Omitted categories for grade levels and race categories are grade 8 and White students. Standardized effects were based on the analysis sample's outcome mean and SD.

Table B12. Local Phase Year 2: Descriptive Statistics for HLM Results for the Assessment of CORE Program Impact on SY 2015–16 Student Engagement Scores

	N. of cases	Minimum	Maximum	Mean	SD
Engagement level	1,770	1.00	4.00	3.00	0.56
Treatment status	1,770	0.00	1.00	0.65	0.48
Engagement level (pretest-centered)	1,770	-2.04	0.96	0.00	0.52
Male	1,770	0.00	1.00	0.49	0.50
8th grader	1,770	0.00	1.00	0.26	0.44
9th grader	1,770	0.00	1.00	0.18	0.38
10th grader	1,770	0.00	1.00	0.14	0.35
11th grader	1,770	0.00	1.00	0.20	0.40
12th grader	1,770	0.00	1.00	0.21	0.41
Other race groups	1,770	0.00	1.00	0.11	0.31
African American	1,770	0.00	1.00	0.19	0.39
Hispanic	1,770	0.00	1.00	0.04	0.20
White	1,770	0.00	1.00	0.66	0.47
Parent with college degree	1,770	0.00	1.00	0.34	0.48
Parent college info missing	1,770	0.00	1.00	0.03	0.18
School's exposure to CORE-like elements prior to the intervention	1,770	0.00	1.00	0.31	0.46

Table B13. National Phase Year 1: HLM Results for the Assessment of CORE Program Impact on SY 2016–17 Student Engagement Scores

	Anova Model				Final Model				
	Coeff.	SE	p		Coeff.	SE	p		Standardized Effect
Intercept	3.03	0.02	0.0	***	2.99	0.04	0.00	***	-0.05
Treatment status					0.04	0.03	0.23		0.08
Engagement level (pretest,centered)					0.44	0.02	0.00	***	0.86
Male					-0.05	0.02	0.00	**	-0.1
9th grader					0.00	0.04	0.90		-0.01
10th grader					0.00	0.04	0.95		0.00
11th grader					-0.02	0.04	0.67		-0.03
12th grader					0.04	0.04	0.25		0.09
Other race groups					-0.05	0.06	0.37		-0.11
African American					0.11	0.03	0.00	***	0.21
Hispanic					0.07	0.03	0.03	*	0.14
Parent with college degree					0.01	0.02	0.76		0.01
Parent college info missing					0.03	0.04	0.47		0.06
Model Summary Statistics									
Joint F-Test race groups					6.2		0.00	***	
Joint F-Test grade levels					1.12		0.35		
Level-1 variance	0.24			***	0.19			***	
Level-2 variance	0.03			***	0.01			***	
Intraclass correlation	0.11				0.05				
Level-1 variance explained					0.19				
Level-2 variance explained					0.63				

Note: Number of students: 2,784. Number of teachers: 57 teachers. Statistical significance (2-tail test): ns = not significant, * = p<.05, ** = p<.01, *** = p<.001. Joint F test: Race groups were statistically significant predictors. Grade levels were not statistically significant predictors. Omitted categories for grade levels and race categories are grade 8 and White students. Standardized effects were based on the analysis sample's outcome mean and SD.

Table B14. National Phase Year 1: Descriptive Statistics for HLM Results for the Assessment of CORE Program Impact on SY 2016–17 Student Engagement Scores

	N. of cases	Minimum	Maximum	Mean	SD
Engagement level	2,784	1.00	4.00	3.02	0.51
Treatment status	2,784	0.00	1.00	0.51	0.50
Engagement level (pretest,centered)	2,784	-2.00	1.00	0.00	0.51
Male	2,784	0.00	1.00	0.49	0.50
8th grader	2,784	0.00	1.00	0.22	0.42
9th grader	2,784	0.00	1.00	0.17	0.37
10th grader	2,784	0.00	1.00	0.23	0.42
11th grader	2,784	0.00	1.00	0.21	0.41
12th grader	2,784	0.00	1.00	0.17	0.38
Other race groups	2,784	0.00	1.00	0.02	0.14
African American	2,784	0.00	1.00	0.19	0.39
Hispanic	2,784	0.00	1.00	0.11	0.31
White	2,784	0.00	1.00	0.69	0.46
Parent with college degree	2,784	0.00	1.00	0.35	0.48
Parent college info missing	2,784	0.00	1.00	0.04	0.20

Table B15. National Phase Year 2: HLM Results for the Assessment of CORE Program Impact on SY 2017–18 Student Engagement Scores

	Anova Model				Final Model				
	Coeff.	SE	p	Sig	Coeff.	SE	p	Sig	Standardized Effect
Intercept	3.01	0.03	0.0	***	2.92	0.05	0.00	***	-0.17
Treatment status					0.05	0.04	0.18		0.09
Engagement level (pretest,centered)					0.43	0.02	0.00	***	0.75
Male					-0.08	0.02	0.00	***	-0.14
9th grader					0.03	0.05	0.56		0.05
10th grader					0.14	0.05	0.00	**	0.24
11th grader					0.10	0.05	0.03	*	0.18
12th grader					0.13	0.05	0.01	*	0.22
Other race groups					-0.01	0.03	0.76		-0.02
African American					0.09	0.03	0.01	*	0.15
Hispanic					0.01	0.04	0.89		0.01
Parent with college degree					-0.01	0.02	0.75		-0.01
Parent college info missing					0.06	0.05	0.23		0.11
Model Summary Statistics									
Joint F-Test race groups					2.45		0.06		
Joint F-Test grade levels					3.20		0.01	*	
Level-1 variance	0.30			***	0.25			***	
Level-2 variance	0.03			***	0.01			***	
Intraclass correlation	0.08				0.06				
Level-1 variance explained					0.17				
Level-2 variance explained					0.47				

Note: Number of students: 2,472. Number of teachers: 52 teachers. Statistical significance (2-tail test): ns = not significant, * = $p<.05$, ** = $p<.01$, *** = $p<.001$. Joint F test: Race groups were not statistically significant predictors. Grade levels were statistically significant predictors. Omitted categories for grade levels and race categories are grade 8 and White students. Standardized effects were based on the analysis sample's outcome mean and SD.

Table B16. National Phase Year 2: Descriptive Statistics for HLM Results for the Assessment of CORE Program Impact on SY 2017–18 Student Engagement Scores

	N. of cases	Minimum	Maximum	Mean	SD
Engagement level	2,472	1.00	4.00	3.02	0.57
Treatment status	2,472	0.00	1.00	0.54	0.50
Engagement level (pretest)	2,472	1.00	4.00	3.02	0.53
Engagement level (pretest,centered)	2,472	-2.01	0.99	0.00	0.53
Male	2,472	0.00	1.00	0.49	0.50
8th grader	2,472	0.00	1.00	0.18	0.38
9th grader	2,472	0.00	1.00	0.16	0.36
10th grader	2,472	0.00	1.00	0.27	0.44
11th grader	2,472	0.00	1.00	0.21	0.41
12th grader	2,472	0.00	1.00	0.19	0.39
Other race groups	2,472	0.00	1.00	0.13	0.34
African American	2,472	0.00	1.00	0.19	0.39
Hispanic	2,472	0.00	1.00	0.11	0.31
White	2,472	0.00	1.00	0.58	0.49
Parent with college degree	2,472	0.00	1.00	0.33	0.47
Parent college info missing	2,472	0.00	1.00	0.05	0.21

Table B17. Local Phase Year 1: HLM Results for the Assessment of CORE Program Impact on SY 2014–15 Student Self-Efficacy Scores

	Anova Model				Final Model				
	Coeff.	SE	p	Sig	Coeff.	SE	p	Sig	Standardized Effect
Intercept	3.18	0.02	0.00	***	3.21	0.03	0.00	***	0.07
Treatment status					-0.03	0.03	0.28		-0.06
Self-efficacy level (pretest,centered)					0.44	0.02	0.00	***	0.87
Male					-0.06	0.02	0.00	**	-0.12
9th grader					-0.05	0.04	0.19		-0.10
10th grader					-0.02	0.04	0.67		-0.03
11th grader					-0.02	0.04	0.58		-0.04
12th grader					-0.06	0.04	0.09		-0.12
Other race groups					0.00	0.03	0.97		0.00
African American					0.01	0.03	0.79		0.02
Hispanic					-0.04	0.06	0.54		-0.07
Parent with college degree					0.05	0.02	0.05	*	0.09
Parent college info missing					0.00				0.00
School's exposure to CORE-like elements prior to the intervention					0.05	0.03	0.08		0.10
Model Summary Statistics									
Joint F-Test race groups					0.17		0.92		
Joint F-Test grade levels					0.94		0.45		
Level-1 variance	0.25			***	0.21			***	
Level-2 variance	0.01			*	0.00				
Intraclass correlation	0.03				0.01				
Level-1 variance explained					0.17				
Level-2 variance explained					0.83				

Note: Number of students: 1,817. Number of teachers: 42 teachers. Statistical significance (2-tail test): ns = not significant, * = $p<.05$, ** = $p<.01$, *** = $p<.001$. Joint F test: Race groups were not statistically significant predictors. Grade levels were not statistically significant predictors. Omitted categories for grade levels and race categories are grade 8 and White students. Standardized effects were based on the analysis sample's outcome mean and SD.

Table B18. Local Phase Year 1: Descriptive Statistics for HLM Results for the Assessment of CORE Program Impact on SY 2014–15 Student Self-Efficacy Scores

	N. of cases	Minimum	Maximum	Mean	SD
Self-efficacy level	1,817	1.00	4.00	3.17	0.51
Treatment status	1,817	0.00	1.00	0.56	0.50
Self-efficacy level (pretest,centered)	1,817	-2.32	0.68	0.00	0.48
Male	1,817	0.00	1.00	0.50	0.50
8th grader	1,817	0.00	1.00	0.43	0.50
9th grader	1,817	0.00	1.00	0.13	0.34
10th grader	1,817	0.00	1.00	0.14	0.35
11th grader	1,817	0.00	1.00	0.14	0.35
12th grader	1,817	0.00	1.00	0.15	0.36
Other race groups	1,817	0.00	1.00	0.14	0.35
African American	1,817	0.00	1.00	0.20	0.40
Hispanic	1,817	0.00	1.00	0.04	0.18
White	1,817	0.00	1.00	0.62	0.48
Parent with college degree	1,817	0.00	1.00	0.35	0.48
Parent college info missing	1,817	0.00	0.00	0.00	0.00
School's exposure to CORE-like elements prior to the intervention	1,817	0.00	1.00	0.30	0.46

Table B19. Local Phase Year 2: HLM Results for the Assessment of CORE Program Impact on SY 2015–16 Student Self-Efficacy Scores

	Anova Model				Final Model				
	Coeff.	SE	p	Sig	Coeff.	SE	P	Sig	Standardized Effect
Intercept	3.15	0.03	0.0	***	3.11	0.06	0.00	***	-0.07
Treatment status					0.03	0.04	0.45		0.06
Self-efficacy level (pretest,centered)					0.46	0.02	0.00	***	0.85
Male					-0.02	0.02	0.50		-0.03
9th grader					-0.06	0.06	0.31		-0.11
10th grader					-0.08	0.06	0.17		-0.14
11th grader					0.04	0.05	0.47		0.07
12th grader					0.05	0.06	0.38		0.09
Other race groups					0.01	0.04	0.83		0.01
African American					0.00	0.03	0.91		-0.01
Hispanic					0.01	0.06	0.88		0.02
Parent with college degree					0.02	0.02	0.35		0.04
Parent college info missing					-0.07	0.07	0.28		-0.13
School's exposure to CORE-like elements prior to the intervention					0.10	0.05	0.06		0.18
Model Summary Statistics									
Joint F-Test race groups					0.03		0.99		
Joint F-Test grade levels					2.51		0.04	*	
Level-1 variance	0.28			***	0.23			***	
Level-2 variance	0.02			***	0.01			**	
Intraclass correlation	0.08				0.04				
Level-1 variance explained					0.18				
Level-2 variance explained					0.63				

Note: Number of students: 1,770. Number of teachers: 37 teachers. Statistical significance (2-tail test): ns = not significant, * = $p<.05$, ** = $p<.01$, *** = $p<.001$. Joint F test: Race groups were not statistically significant predictors. Grade levels were statistically significant predictors. Omitted categories for grade levels and race categories are grade 8 and White students. Standardized effects were based on the analysis sample's outcome mean and SD.

Table B20. Local Phase Year 2: Descriptive Statistics for HLM Results for the Assessment of CORE Program Impact on SY 2015–16 Student Self-Efficacy Scores

	N. of cases	Minimum	Maximum	Mean	SD
Self-efficacy level	1,770	1.00	4.00	3.15	0.55
Treatment status	1,770	0.00	1.00	0.65	0.48
Self-efficacy level (pretest,centered)	1,770	-2.30	0.70	0.00	0.50
Male	1,770	0.00	1.00	0.49	0.50
8th grader	1,770	0.00	1.00	0.26	0.44
9th grader	1,770	0.00	1.00	0.18	0.38
10th grader	1,770	0.00	1.00	0.14	0.35
11th grader	1,770	0.00	1.00	0.20	0.40
12th grader	1,770	0.00	1.00	0.21	0.41
Other race groups	1,770	0.00	1.00	0.11	0.31
African American	1,770	0.00	1.00	0.19	0.39
Hispanic	1,770	0.00	1.00	0.04	0.20
White	1,770	0.00	1.00	0.66	0.47
Parent with college degree	1,770	0.00	1.00	0.34	0.48
Parent college info missing	1,770	0.00	1.00	0.03	0.18
School's exposure to CORE-like elements prior to the intervention	1,770	0.00	1.00	0.31	0.46

Table B21. National Phase Year 1: HLM Results for the Assessment of CORE Program Impact on SY 2016–17 Student Self-Efficacy Scores

	Anova Model				Final Model				
	Coeff.	SE	p	Sig	Coeff.	SE	P	Sig	Standardized Effect
Intercept	3.13	0.02	0.0	***	3.12	0.04	0.00	***	-0.04
Treatment status					0.03	0.03	0.43		0.05
Self-efficacy level (pretest,centered)					0.40	0.02	0.00	***	0.81
Male					-0.03	0.02	0.09		-0.06
9th grader					0.01	0.04	0.77		0.02
10th grader					0.04	0.04	0.23		0.09
11th grader					0.02	0.04	0.64		0.03
12th grader					0.06	0.04	0.10		0.13
Other race groups					-0.07	0.06	0.26		-0.14
African American					0.00	0.03	0.91		-0.01
Hispanic					-0.02	0.03	0.48		-0.05
Parent with college degree					0.00	0.02	0.86		0.01
Parent college info missing					-0.05	0.04	0.28		-0.09
Model Summary Statistics									
Joint F-Test race groups					0.55		0.65		
Joint F-Test grade levels					1.10		0.36		
Level-1 variance	0.23			***	0.19			***	
Level-2 variance	0.02			***	0.01			***	
Intraclass correlation	0.07				0.05				
Level-1 variance explained					0.17				
Level-2 variance explained					0.40				

Note: Number of students: 2,741. Number of teachers: 57 teachers. Statistical significance (2-tail test): ns = not significant, * = $p<.05$, ** = $p<.01$, *** = $p<.001$. Joint F test: Race groups were not statistically significant predictors. Grade levels were not statistically significant predictors. Omitted categories for grade levels and race categories are grade 8 and White students. Standardized effects were based on the analysis sample's outcome mean and SD.

Table B22. National Phase Year 1: Descriptive Statistics for HLM Results for the Assessment of CORE Program Impact on SY 2016–17 Student Self-Efficacy Scores

	N. of cases	Minimum	Maximum	Mean	SD
Self-efficacy level	2,741	1.00	4.00	3.14	0.50
Treatment status	2,741	0.00	1.00	0.51	0.50
Self-efficacy level (pretest,centered)	2,741	-2.24	0.76	0.00	0.51
Male	2,741	0.00	1.00	0.49	0.50
8th grader	2,741	0.00	1.00	0.22	0.42
9th grader	2,741	0.00	1.00	0.16	0.37
10th grader	2,741	0.00	1.00	0.23	0.42
11th grader	2,741	0.00	1.00	0.21	0.41
12th grader	2,741	0.00	1.00	0.17	0.38
Other race groups	2,741	0.00	1.00	0.02	0.14
African American	2,741	0.00	1.00	0.19	0.39
Hispanic	2,741	0.00	1.00	0.11	0.31
White	2,741	0.00	1.00	0.69	0.46
Parent with college degree	2,741	0.00	1.00	0.35	0.48
Parent college info missing	2,741	0.00	1.00	0.04	0.20

Table B23. National Phase Year 2: HLM Results for the Assessment of CORE Program Impact on SY 2017–18 Student Self-Efficacy Scores

	Anova Model				Final Model				
	Coeff.	SE	p	Sig	Coeff.	SE	p	Sig	Standardized Effect
Intercept	3.14	0.02	0.00	***	3.09	0.04	0.00	***	-0.09
Treatment status					0.02	0.03	0.44		0.04
Self-efficacy level (pretest,centered)					0.4	0.02	0.00	***	0.72
Male					-0.08	0.02	0.00	***	-0.15
9th grader					0.05	0.05	0.31		0.09
10th grader					0.08	0.04	0.05		0.15
11th grader					0.05	0.04	0.28		0.09
12th grader					0.09	0.04	0.06		0.16
Other race groups					-0.05	0.03	0.17		-0.09
African American					0.01	0.03	0.76		0.02
Hispanic					-0.03	0.04	0.49		-0.05
Parent with college degree					0.06	0.02	0.01	**	0.11
Parent college info missing					0.06	0.05	0.23		0.11
Model Summary Statistics									
Joint F-Test race groups					0.83		0.48		
Joint F-Test grade levels					1.27		0.28		
Level-1 variance	0.3			***	0.25			***	
Level-2 variance	0.01			**	0.01			*	
Intraclass correlation	0.03				0.02				
Level-1 variance explained					0.16				
Level-2 variance explained					0.45				

Note: Number of students: 2,113. Number of teachers: 52 teachers. Statistical significance (2-tail test): ns = not significant, * = $p<.05$, ** = $p<.01$, *** = $p<.001$. Joint F test: Race groups were not statistically significant predictors. Grade levels were not statistically significant predictors. Omitted categories for grade levels and race categories are grade 8 and White students. Standardized effects were based on the analysis sample's outcome mean and SD.

Table B24. National Phase Year 2: Descriptive Statistics for HLM Results for the Assessment of CORE Program Impact on SY 2017–18 Student Self-Efficacy Scores

	N. of cases	Minimum	Maximum	Mean	SD
Self-efficacy level	2,113	1.00	4.00	3.14	0.55
Treatment status	2,113	0.00	1.00	0.54	0.5
Self-efficacy level (pretest)	2,113	1.00	4.00	3.22	0.54
Self-efficacy level (pretest,centered)	2,113	-2.21	0.79	0.01	0.54
Male	2,113	0.00	1.00	0.49	0.5
8th grader	2,113	0.00	1.00	0.18	0.39
9th grader	2,113	0.00	1.00	0.14	0.35
10th grader	2,113	0.00	1.00	0.28	0.45
11th grader	2,113	0.00	1.00	0.22	0.41
12th grader	2,113	0.00	1.00	0.19	0.39
Other race groups	2,113	0.00	1.00	0.13	0.34
African American	2,113	0.00	1.00	0.18	0.39
Hispanic	2,113	0.00	1.00	0.11	0.31
White	2,113	0.00	1.00	0.58	0.49
Parent with college degree	2,113	0.00	1.00	0.33	0.47
Parent college info missing	2,113	0.00	1.00	0.05	0.22

Table B25. Descriptive Statistics of Outcome Variables and Hedge's g by Group Status

	N of cases		Unadjusted Means		Adjusted Means		Unadjusted SD		
	Treatment (T)	Control (C)	T	C	T	C	T	C	Hedge's g
CWRA+ SRQ (posttest)									
Local phase Year 1	1060	842	854.26	856.05	846.24	852.11	152.47	154.73	-0.04
Local phase Year 2	1265	683	879.18	888.70	861.43	860.93	164.53	170.16	0.00
National phase Year 1	1515	1468	903.85	913.91	910.62	919.53	176.05	173.85	-0.05
National phase Year 2	1397	1212	917.91	918.51	932.41	928.02	178.84	169.28	0.03
Student engagement (posttest)									
Local phase Year 1	1022	801	2.99	3.06	3.03	3.07	0.53	0.54	-0.08
Local phase Year 2	1144	626	3.00	2.99	3.00	3.00	0.55	0.57	0.01
National phase Year 1	1407	1377	3.04	3.00	3.03	2.99	0.52	0.51	0.08
National phase Year 2	1344	1128	3.05	2.98	2.98	2.92	0.57	0.58	0.09
Student efficacy (posttest)									
Local phase Year 1	1017	800	3.17	3.17	3.18	3.21	0.52	0.5	-0.06
Local phase Year 2	1144	626	3.15	3.13	3.14	3.11	0.54	0.56	0.06
National phase Year 1	1385	1356	3.15	3.12	3.14	3.12	0.49	0.5	0.05
National phase Year 2	1131	982	3.15	3.12	3.11	3.09	0.54	0.56	0.04

Note: Hedge's g was based on adjusted means and unadjusted SDs.

Appendix C: Implementation Study Tables

Table C1. Implementation Study Questions Aligned with the CORE i3 Program Logic Model

Implementation Study Question	Logic Model Component
ISQ1. To what extent do principals in the 32 participating CORE schools engage in meaningful collaborative partnerships with other administrators at CORE events as intended?	1
ISQ2. To what extent do the 32 participating CORE teachers establish and participate actively in content-based online learning communities as intended?	2
ISQ3. To what extent were technology resources and support provided to the participating 32 CORE teachers as intended?	3
ISQ4. To what extent do the participating 32 CORE teachers engage in in-depth CORE Active Learning Methods (CALM) PD as intended?	4
ISQ5. To what extent do the 32 participating CORE teachers engage in follow-up CORE workshops as intended?	5
ISQ6. To what extent do the 32 participating CORE schools receive active support to navigate the change-management process as intended?	6
ISQ7. To what extent do CORE Math and English teachers in the 32 participating CORE schools provide students with college-readiness advisement and support through use of the Ed Ready tool in participating CORE schools as intended?	7

Table C2. KC, Indicator, and Data Sources for the FOI Study (Local Phase Year 1)

Measuring Implementation Fidelity SY 2014–15		
KC	**Indicator**	**Data Source**
KC1. Principals in new CORE schools connect with principals in existing partnership schools.	1.1 Principal attendance at the CORE Academy and three CORE workshops that connect superintendents and principals	CORE Academy attendance roster
	1.2 Principal collaboration with their peers and superintendents at the CORE Academy and three CORE workshops	CORE event evaluation survey
KC2. CORE teachers participate in online learning communities.	2.1 Teacher enrollment in content-based PLCs	PLC post-query tracking using Facebook and CANVAS
	2.2 Active participation in online PLCs	
KC3. CORE teachers receive and adopt classroom technology resources and support.	3.1 Provision of laptops and iPads for CORE teachers	Technology equipment log
	3.2 Provision of classroom technology funds for CORE teachers	Financial disbursement log
	3.3 Provision of classroom technology assessments	Technology assessment log
	3.4 Provision of ongoing support from Education Technology Assistants (ETAs)	ETA assistance log
KC4. CORE teachers participate in CORE Academy PBL PD.	4.1. Teacher attendance at the annual CORE Academy	CORE Academy attendance roster
	4.2. Quality of CORE Academy	

Measuring Implementation Fidelity SY 2014–15		
KC	**Indicator**	**Data Source**
	4.3. Relevance of CORE Academy	CORE Academy event survey
	4.4. Usefulness of CORE Academy	
	4.5. Increased content knowledge	
	4.6. Use of CALM	CORE pre- and post-CALM assessment
	4.7. Use of technology	CORE pre- and post-technology assessment
KC5. CORE teachers participate in follow-up PBL PD workshops.	5.1. Participation in follow-up workshops	Workshop attendance roster
	5.2. Sharing of learning experiences	YouTube video log
KC6. CORE principals provide change-management support to CORE schools.	6.1. Participation in follow-up debriefings on change management	CDI report
KC7. JSU counselors provide dual enrollment advisement and support to students in CORE schools.	7.1 Provision of targeted support and guidance to dual enrollment students	Memo report and Counselor Tracking Sheet

Table C3. Technology Use Score by Status and Phase

	Time Period	N	T Mean	SD	C Mean	SD	Diff	p	Sig
Local Phase	SY 2014–15	C(N=19); T(N=23)	0.70	0.56	0.23	0.45	0.47	0	**
	SY 2015–16	C(N=15); T(N=15)	0.31	0.63	0.30	0.49	0.01	0.96	-
	SY 2014–16	C(N=15); T(N=15)	0.93	0.84	0.39	0.70	0.55	0.05	-
National Phase	SY 2014–15	C(N=25); T(N=28)	0.72	0.82	0.10	0.73	0.63	0.00	**
	SY 2015–16	C(N=18); T(N=22)	0.32	0.59	0.31	0.58	0.01	0.95	-
	SY 2014–16	C(N=18); T(N=21)	1.07	1.03	0.30	0.92	0.76	0.01	*

Table C4. PBL Score by Status and Phase

	Time Period	N	T Mean	SD	C Mean	SD	Diff	p	Sig
Local Phase	SY 2014–15	C(N=13); T(N=16)	0.73	1.11	-0.09	0.53	0.82	0	**
	SY 2015–16	C(N=10); T(N=12)	0.19	0.82	0.17	0.99	0.02	0.96	-
	SY 2014–16	C(N=12); T(N=15)	0.99	1.16	0.10	0.99	0.89	0.05	*
National Phase	SY 2014–15	C(N=17); T(N=17)	0.52	1.20	-0.05	1.03	0.58	0.13	
	SY 2015–16	C(N=13); T(N=14)	0.44	0.96	-0.02	0.77	0.46	0.16	
	SY 2014–16	C(N=19); T(N=22)	0.90	1.43	0.40	1.27	0.50	0.24	

Appendix D: Correlation Analysis: Teachers' Program Implementation and Students' Average CWRA+ Scores

Table D1. Local Phase analysis—Correlation Statistics between Pretest–Posttest CWRA+ Score Change and Implementation Scales (Treatment Teachers Only)

Local Phase Year 1						
	Classroom Technology Scale			PBL Scale		
	Time 1 (pre)	Time 2 (post)	Time 2 – 1 change	Time 1 (pre)	Time 2 (post)	Time 2 – 1 change
Correlation w/ CWRA+ score change	.13 ns	0.08 ns	-0.10 ns	-0.03 ns	0.33 ns	0.26 ns
N. of teachers	22	22	22	21	15	14
Local Phase Year 2						
	Classroom Technology Scale			PBL Scale		
	Time 2 (pre)	Time 3 (post)	Time 3 – 2 change	Time 2 (pre)	Time 3 (post)	Time 3 – 2 change
Correlation w/ CWRA+ score change	0.16 ns	-0.37 ns	-0.43 ns	-0.20 ns	-0.40 ns	-0.26 ns
N. of teachers	22	15	15	15	15	15

Note: Statistical significance (2-tail test): * = p<.05, ns=not significant. See Appendix D3 for descriptive statistics used for the correlational analysis. Time 1, 2, and 3 correspond to June 2014, June 2015, and June 2016, respectively.

Table D2: National Phase analysis—Correlation Statistics between Pretest–Posttest CWRA+ Score change and Implementation Scales (Treatment Teachers Only)

National Phase Year 1						
	Classroom Technology Scale			PBL Scale		
	Time 1 (pre)	Time 2 (post)	Time 2 – 1 change	Time 1 (pre)	Time 2 (post)	Time 2 – 1 change
Correlation w/ CWRA+ score change	-0.23 ns	-0.10 ns	0.15 ns	-0.18 ns	-0.45 ns	-0.17 ns
N. of teachers	27	28	27	26	18	16
National Phase Year 2						
	Classroom Technology Scale			PBL Scale		
	Time 2 (pre)	Time 3 (post)	Time 3 – 2 change	Time 2 (pre)	Time 3 (post)	Time 3 – 2 change
Correlation w/ CWRA+ score change	-0.01 ns	-0.04 ns	-0.10 ns	0.9 ns	0.24 ns	-0.08 ns
N. of teachers	27	22	22	17	23	14

Note: Statistical significance (2-tail test): * = p<.05, ns=not statistically significant. See Appendix D4 for descriptive statistics used for the correlational analysis. Time 1, 2, and 3 correspond to June 2016, June 2017, and June 2018, respectively.

Table D3. Descriptive Statistics for the Correlation Analyses between Students' Average CWRA+ Scores and Teachers' Implementation Scores (Treatment Teachers Only)

Local Phase Year 1					
	N	Mean	SD	Min	Max
Average CWRA+ score change (pre to post)	22	-18.20	35.12	-112.00	31.68
Classroom technology scale Time 1 (pre)	23	2.38	0.87	1.33	4.69
Classroom technology scale Time 2 (post)	24	3.03	0.75	1.78	4.35
Classroom technology scale Time 2 - 1 change	23	0.70	0.56	-0.34	1.75
PBL scale Time 1 (pre)	23	2.90	1.18	1.00	4.50
PBL scale Time 2 (post)	17	3.48	0.69	2.00	5.00
PBL scale Time 2 - 1 change	16	0.73	1.11	-0.80	3.00
Local Phase Year 2					
	N	Mean	SD	Min	Max
Average CWRA+ score change (pre to post)	22	18.93	49.70	-96.38	113.60
Classroom technology scale Time 2 (pre)	24	3.03	0.75	1.78	4.35
Classroom technology scale Time 3 (post)	15	3.50	0.86	2.05	5.00
Classroom technology scale Time 3 - 2 change	15	0.31	0.63	-1.00	1.75
PBL scale Time 2 (pre)	17	3.48	0.69	2.00	5.00
PBL scale Time 3 (post)	15	3.89	0.93	2.00	5.00
PBL scale Time 3 - 2 change	12	0.19	0.82	-1.00	1.20

Table D4. Descriptive Statistics for the Correlation Analyses between Students' Average CWRA+ Scores and Teachers' Implementation Scores (Treatment Teachers Only)

National Phase Year 1					
	N	Mean	SD	Min	Max
Average CWRA+ score change (pre to post)	28	-1.76	54.17	-84.11	229.45
Classroom technology scale Time 1 (pre)	28	2.36	0.89	1.05	4.73
Classroom technology scale Time 2 (post)	29	3.08	0.75	1.68	4.90
Classroom technology scale Time 2 - 1 change	28	0.72	0.82	-0.73	3.43
PBL scale Time 1 (pre)	28	2.94	1.25	1.00	5.00
PBL scale Time 2 (post)	19	3.57	0.97	1.80	5.00
PBL scale Time 2 - 1 change	17	0.52	1.20	-1.80	2.50
National Phase Year 2					
	N	Mean	SD	Min	Max
Average CWRA+ score change (pre to post)	27	15.96	66.80	-56.44	251.30
Classroom technology scale Time 2 (pre)	29	3.08	0.75	1.68	4.90
Classroom technology scale Time 3 (post)	22	3.35	0.78	1.95	5.00
Classroom technology scale Time 3 - 2 change	22	0.32	0.59	-0.90	1.48
PBL scale Time 2 (pre)	19	3.57	0.97	1.80	5.00
PBL scale Time 3 (post)	23	3.67	0.71	2.00	5.00
PBL scale Time 3 - 2 change	14	0.44	0.96	-0.70	3.00

Note: The number of cases depends on a specific correlational analysis involving two variables at a time.