The Evaluation Group
Dynamic Program Evaluation

Impact Evaluation of *INSPIRE: Infusing Innovative STEM Practices Into Rigorous Education*

U.S Department of Education
Investing in Innovation Fund (i3)
August 2018

Presented by:

Karyl Askew, Ph.D.
Olivia Stevenson, M.S.
The Evaluation Group
www.evaluationgroup.com

Bridget Jones
Program Choice Coordinator Cabarrus County Schools
www.cabarrus.k12.nc.us

INSPIRE
Final Impact Evaluation Report

Lead Agency:	Cabarrus County Schools 4401 Old Airport Road Concord, NC 28025
Project Director:	Bridget Jones
Third-Party Evaluator:	Karyl Askew, Ph.D. Olivia Stevenson, M.S. The Evaluation Group
Contact Information:	Bridget Jones: 704-783-5939, Bridget.Jones@cabarrus.k12.nc.us Karyl Askew: 803-542-2343, karyl@evaluationgroup.com Olivia Stevenson: 803-331-1679, olivia@evaluationgroup.com

Grant Period:	01/01/2014 – 12/31/2017
Funder:	US Department of Education, Office of Innovation and Improvement (OII), Investing in Innovation (i3) Fund
PR Award #:	U411C130073
i3 Cohort Year:	2014
i3 ID:	DEV66
Keywords:	Problem-Based Learning, Project-Based Learning, Curriculum Redesign, Elementary School Mathematics, Middle School Mathematics, High School Mathematics, Middle School Science, High School Science, Randomized Control Trial

Funding for this report came from the U.S. Department of Education under its Investing in Innovation (i3) initiative within the Office of Innovation and Improvement through Grant U411C130073 to Cabarrus County Schools. The i3 grant award required an independent evaluation of the implementation and impacts of the program. The opinions and findings and conclusions in this report do not necessarily represent the official positions or policies of the funders. Correspondence regarding this document should be addressed to Karyl Askew, The Evaluation Group, 169 Laurelhurst Ave., Columbia SC 29210. Phone: (803) 542-2343; email karyl@evaluationgroup.com. For information about The Evaluation Group, see our website: www.evaluationgroup.com.

Contents

1. ABSTRACT

INSPIRE is an Investing in Innovation (i3) development grant funded by the Office of Innovation and Improvement, U.S. Department of Education. *INSPIRE* provides an innovative integrated K-12 STEM pipeline approach focused on STEM course content and instructional redesign. The INSPIRE model was implemented in Cabarrus County Schools (CCS), which is among the largest school systems in North Carolina, serving nearly 30,000 students in 39 schools. The impact evaluation included two studies that examined the effect of *INSPIRE* on mathematics and science achievement as measured by North Carolina standardized End-of-Grades assessments. The elementary study (Study 1) used a three-year, longitudinal, single-cohort quasi-experimental design (QED) to assess the impact of INSPIRE on math achievement at the end of 5th grade after two years of program exposure. The secondary study (Study 2) used an individual-level, longitudinal, randomized controlled trial (RCT) with blocking by school level and cohort to assess the effects of INSPIRE on math and science achievement at the end of 7th and 10th grades after two years of program exposure. For both studies, we compared the outcomes of INSPIRE students with similar students from schools that did not offer a STEM program. For the elementary study, propensity score matching (PSM) was used to match INSPIRE elementary students and comparison student samples at baseline (on pre-test math achievement scores, gender, minority status, and economically disadvantaged status) and baseline equivalence was established all pre-test assessment measures; this study met What Works Clearinghouse (WWC) Group Design Standards with Reservations. For the secondary study, the overall and differential attrition rates were low based on the WWC attrition standards (WWC, 2017); this study met WWC Group Design Standards without Reservations. The results of the elementary study indicated a statistically significant difference between the *INSPIRE* treatment group and the business-as-usual comparison group on the math achievement outcome. Comparison students reported a statistically significant higher increase in math achievement than *INSPIRE* students. Results of the secondary study indicated no statistically significant difference between the INSPIRE treatment group and the business-as-usual comparison group on the math and science achievement outcome. The duration of students' exposure to INSPIRE, fidelity of implementation, alignment between PBL instruction and NC standardized assessments, and contextual factors that might have weakened the intervention strength relative to business-as-usual conditions are discussed as possible factors that account for these findings. The report concludes with suggestions for future research and implications for education policy.

2. INTRODUCTION

INSPIRE, Infusing Innovative STEM Practices into Rigorous Education, was awarded to Cabarrus County Schools (CCS) as a development grant under Absolute Priority 3: Improving STEM Education: Subpart (a). Cabarrus County Schools (CCS) is among the largest school systems in North Carolina serving nearly 30,000 students in 39 schools. CCS has a history of implementing large initiatives with measurable gains in student outcomes. INSPIRE was developed to provide the district's highest need students the opportunity to participate in a K-12 STEM magnet school pipeline, through intentional outreach designed to engage and retain minority and low-income students. CCS aimed to advance its STEM education model from a promising strong theory to an evidence-based practice to address a national need to develop and validate an integrated K-12 STEM pipeline that serves as a model for STEM course content and instructional redesign (National Research Council, 2011). Using problem-based learning (PBL) as the core of our curriculum redesign, *INSPIRE* was novel because the approach addressed limitations in STEM education by: 1) providing early, continuous engagement via a STEM pipeline starting in Kindergarten and channeling students into integrated STEM magnet programs at both middle and high school levels; 2) reducing selection bias and student interest factors by intentionally and automatically placing low-income, minority students in elementary STEM magnet schools located in their neighborhood; 3) linking PBL with interdisciplinary STEM course content connected to the Common Core; 4) integrating PBL units across all courses to enable students to explore the same issue in every subject with student real-world tethers (connections beyond the classroom); 5) connecting PBL and digital course content to personalized and tech-enabled instructional practice to impact achievement and engagement; and 6) revolutionizing the teacher role from transmitter to a facilitator of knowledge via substantive transformations in the way curricula, pedagogy, and assessments are conceptualized and implemented (Asghar et al., 2012).

2.1 Program Description

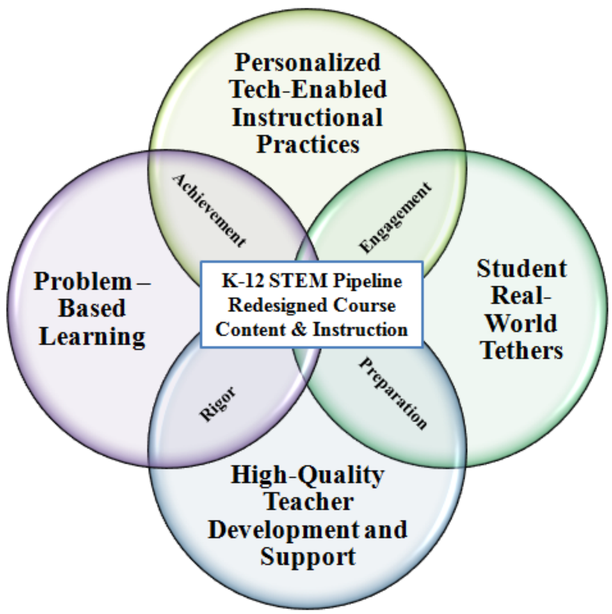
Rather than piecemeal STEM education or a sequence of STEM courses as has been done previously, the INSPIRE offered an integrated K-12 pipeline approach (Figure 1) consisting of four key strategies:

1) Developing rigorous PBL curriculum units to support STEM course content connected across all subjects (Asghar et al., 2012; Lou, Shih, Diez, & Tseng, 2011; National Research Council, 2011);
2) Designing STEM instructional practices that connect PBL course content to tech-enabled personalized learning strategies through digital content integration and a 1:1 technological device to student ratio (Brown, et al., 2013).;

3) Creating a high-quality teacher development and support process to sustain innovative STEM course content and instructional practices through STEM Coaches, STEMersion, Innovation Showcase, and Curriculum Review Week (Geier, et al., 2008);

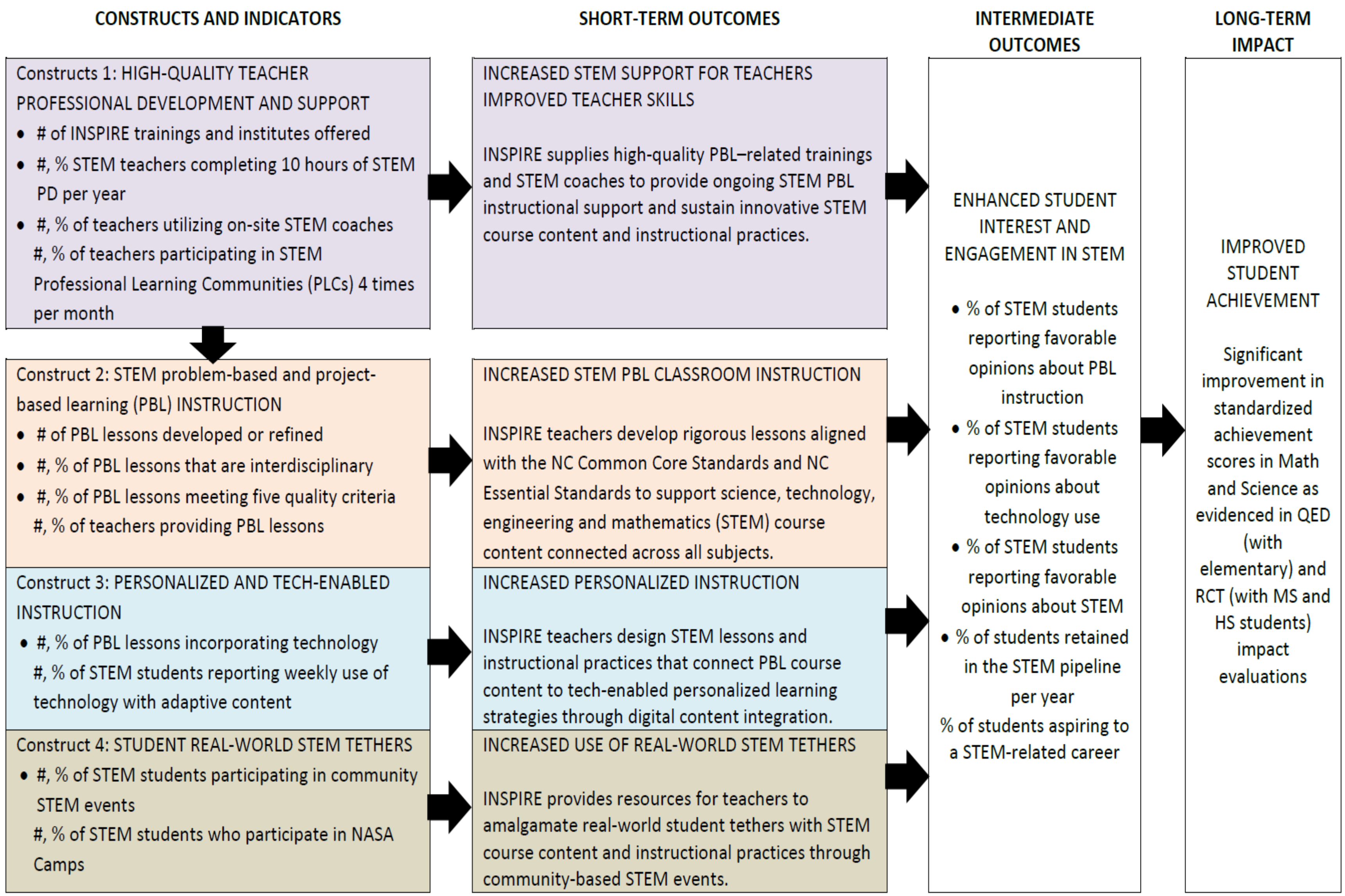
4) Amalgamating real-world student tethers with STEM course content and instructional practices through a NASA summer camp, student competitions, alignment of STEM content to annual field trips, and quarterly STEM events (Kwan, 2009; Strobel & van Barneveld, 2009; Van Berkel & Dolmans, 2006).

Figure 1: The INSPIRE Program Model



As illustrated by the logic model shown in Figure 2, the theory underlying INSPIRE is that training teachers to implement STEM-based problem-based learning (PBL) supported by personalized, tech-enabled instruction and connections to real-world tethers, will significantly increase student interest and engagement in STEM, leading to higher achievement in math and science (Asghar et al., 2012; Brown, et al., 2013; Geier, et al., 2008; National Research Council, 2011; Strobel & van Barneveld, 2009; Van Berkel & Dolmans, 2006; Walker & Leary, 2009).

Figure 1. The *INSPIRE* Logic Model



The development of CCS' initial K-9 continuum began with the design of a complete STEM magnet elementary school so that every student in the school's attendance zone was provided automatic acceptance into the middle and high school STEM magnet schools, as long as they scored proficient on NC end-of-grade tests in math and reading. However, on the whole, STEM education within and across CCS STEM classrooms and schools was incongruent. Table 1 below presents pre-award implementation and post-award implementation within Cabarrus County Schools. A glossary of key post-award activities appears in Appendix A.

Table 1. INSPIRE Pre-award and Post-award Implementation

Key Strategies	Pre-award Implementation	Post-award Implementation
PBL Lessons	PBL units implemented in grades K – 8. • Isolated and inconsistent implementation of PBL units and lessons within grade-levels. • Limited interdisciplinary focus. • Few opportunities for vertical alignment across grade levels. • No systematic development and review process of PBL lessons.	PBL units implemented in grades K – 12. • PBL focused on high-quality, interdisciplinary, horizontal and vertical alignment. • Development and documentation of a systematic PBL unit design and review process using the INSPIRE PBL Framework and Assessment of Quality Rubric (Cabarrus County School, 2014a). • Creation of a database of lessons aligned to the common core that can be shared nationally and internationally.
Personalized Tech-Enabled Instructional Practices	Technology used to varying degrees by teachers at all grade levels to supplement instruction. • Technology not systematically integrated into PBL units. • Multiple devices, platforms, and software programs. • Use of tablet in a 1:1 ratio only in grades 9 - 12. • Limited integration of blended-learning, use of data to personalize instruction, and	Systematic connection of technology to PBL units in K -12 • Use of MacBook Airs (for high school) and Chromebooks and tablets in a 1:1 ratio to deliver personalized instruction with students in grades K-12 • Instruction designed to be student driven, competency-based, hands-on. • Face-to-face teaching time; real-time student feedback.

Table 1. INSPIRE Pre-award and Post-award Implementation		
Key Strategies	**Pre-award Implementation**	**Post-award Implementation**
	use of technology to support learning beyond classroom instruction.	• Teacher and student access to meaningful data. • Technology to support the learning process beyond the classroom.
Teacher Development and Support	STEM-focused professional development offered during the academic year. • Specialized, yet limited, PD opportunities such as 5-day STEMersion. • Limited capacity of PLCs to offer in-class coaching and on-going support for instruction, development of PBL unit lessons and assessments, progress towards vertical and horizontal alignment of units.	Sustained year-round development and support. • Collaboration through weekly PLC meetings. • Ongoing support structures (STEM coaches provide unit guidance, coordinate K-12 pipeline, in-class coaching, aligned formative assessments). • Intentional coaching guided by the INSPIRE STEM PBL Instructional Coaching Tool (Cabarrus County School, 2014b) • Additional capacity for professional development (school in-service, PBL Framework and Assessment of Quality Rubric for Lesson Plan development, Innovation Showcases, an annual 5-day STEMersion, partnerships with STEM businesses).
Student Real-world Tethers	Limited and inconsistent use of tethers to connect STEM content to instructional practices. • All schools have and offer STEM-focused extracurricular clubs. • Schools sponsor students to attend STEM student competitions and annual field trips.	Using a community extension approach, real-world tethers are intentionally and systematically connected to STEM content and instructional practices. • Tethers incorporated into PBL units as a result of STEMersion community partnerships, STEM coach support, annual Innovation

Table 1. INSPIRE Pre-award and Post-award Implementation		
Key Strategies	Pre-award Implementation	Post-award Implementation
	• Authentic assessment through Senior Capstone Projects.	Showcases, and intentional curriculum design and planning using PLCs. • Schools offer new quarterly STEM events (guest speakers) and NASA summer camp. • Additional funding to increase school-sponsored student involvement in STEM competitions. • Increased the number of STEM-related high school clubs. • Created STEM elementary student mentoring by STEM high school students through science fair and STEM clubs.

2.2 Fidelity of Implementation

INSPIRE was evaluated on fidelity of implementation, the extent to which actual project implementation aligned with proposed project implementation. The INSPIRE Annual Fidelity Index, presented in Appendix B, was designed in the first year of implementation. The index is comprised of four components (aligned to the four key strategies), and 12 indicators. The numbers of indicators within each component ranged from two to four and leveraged multiple data sources (surveys, STEM implementation rubrics, lesson plans, administrative records, interviews, focus groups).

A fidelity score was calculated for each indicator and for each construct. Annual thresholds or targets were established a priori using baseline data and the project director's recommendations. The INSPIRE lead team utilized these annual thresholds as benchmarks to assess progress toward their long-term program goals.

Fidelity data were collected for three school years, SY 2014-2015 through SY 2016-2017. In SY 2014-2015, the first year of implementation, three of four *INSPIRE* components were implemented with fidelity. In Years 2 and 3, all four *INSPIRE* components were implemented with fidelity.

In the first year of implementation, the expected level of fidelity fell short on two of four indicators within Component 2: Problem-Based Learning, related to PBL lesson quality and implementation (Indicators 2.3 and 2.4). In the first year, a foundation was laid by INSPIRE hiring and training four school-based STEM coaches and developing standards for PBL lessons. Through the direction and support of the project director and independent evaluator, STEM coaches and curriculum specialist developed the INSPIRE Problem-Based/Project-Based Learning (PBL) Lesson Framework and Assessment of Quality Rubric (Cabarrus County Schools, 2014a). To advance an integrated and interdisciplinary K-12 STEM pipeline approach, INSPIRE used the Assessment of Quality Rubric to score each INSIRE PBL lesson on eight elements using a three-point scale and to select the strongest lessons for inclusion in a searchable INSPIRE PBL database. As of SY2017/18, the INSPIRE PBL database included over 370 interdisciplinary PBL lessons. Considerable time was spent on professional development and coaching to support a common understanding of and proficiency to use the framework, particularly in the first year. Teacher survey data suggested that the framework and rubric was transformational the development process. The fidelity scores represent upfront learning curves and gradual increase in performance on Indicators 2.3 and 2.4.

3. STUDY 1: IMPACT STUDY DESIGN

The elementary level impact study, Study 1, used a three-year, longitudinal, single-cohort *quasi-experimental design* (QED) to assess the impact of INSPIRE on math achievement at the end of 5th grade after three years of program exposure. Math achievement was measured using recognized assessments with proven reliability and validity, including NC End-of-Grade (EOG) and Discovery Education's (DE) Common Core Math Assessments. The sample included two STEM elementary schools that served the district's most ethnically diverse students. The outcomes of INSPIRE students were contrasted with similar students from four comparison elementary schools that did not offer a STEM program.

3.1 Samples

Treatment Sample

The INSPIRE model was tested at two full STEM elementary magnet schools. A snapshot of the high-need population for these schools is provided in Table 2.

Table 2. INSPIRE Schools Enrollment Characteristics							
Target School	N	Black	Asian	White	Hispanic	Other	Low-Income
Elementary School 1 (Title 1)	484	24%	5%	42%	23%	6%	37%
Elementary School 2	893	18%	3%	65%	9%	5%	24%

In CCS's initial K-9 initiative, elementary school 1 was selected as the pilot school. This school holds Title I status, and serves the district's lowest income, most ethnically diverse students. These students are traditionally underrepresented in STEM, making it an ideal environment in which to reach underserved students. In 2011, elementary school 1 was transformed into a complete STEM magnet school, so that every student in the school's attendance zone was reached and could automatically continue in the STEM magnet middle schools as long as they scored at least proficient on NC EOG tests in math and reading. Elementary school 2 serves students from the southern part of Cabarrus County. In school year (SY) 2013-2014, elementary school 2 transitioned from a traditional school to a STEM magnet school. Beginning in 2014-2015, the first year of the grant, elementary school 2 became a full STEM school. All Grade 3 students enrolled in treatment schools in SY 2014-2015 comprised the sample of treatment students. Students were followed through Grade 5.

Comparison Sample

The outcomes of INSPIRE students were compared with similar students from four comparison elementary schools that did not offer a STEM program. The four comparison schools were chosen because they were similar on demographic characteristics with respect to the representation of race/ethnicity and economic disadvantage. See Table 3 for details on the demographic make-up of the comparison elementary schools.

Table 3: Comparison Schools Enrollment Characteristics							
Comparison School	Demographic Characteristics of Student Population (2014-2015)						
	Total Enrollment	Black	Asian	White	Hispanic	Other	Low-income
	N	Percent Representation (%)					
School 1	925	13%	2%	59%	21%	5%	49%
School 2	727	17%	3%	46%	31%	3%	42%
School 3	809	15%	4%	70%	8%	3%	23%
School 4	803	12%	2%	72%	9%	5%	26%

The sample of similar students in comparison schools was identified using propensity score matching (PSM). The PSM included the following variables: Grade 3 pre-test math achievement scores (DE Common Core Math Assessment), gender, minority status, and free and reduced-price lunch. The sample of 200 students from comparison schools were also followed through Grade 5.

3.2 Study 1 Question

The Study 1 confirmatory question was: What is the effect of INSPIRE on 5th grade math achievement for INSPIRE 3rd grade students after 3 program years of exposure compared to the math achievement of 5th grade students in the business as usual condition?

3.3 Analysis and Results

3.3.1 Baseline Analytic Model

The following two-level model was used to estimate the baseline mean difference between the intervention and comparison groups.

$$y_{ij} = \alpha + \beta_1 T_j + e_j + \varepsilon_{ij}$$

Where,

y_{ij} = the math pre-test score for student *i* in school *j*

α = intercept

β_1 = covariate-adjusted impact of INSPIRE (i.e., the difference between the mean outcome for treatment schools and the mean outcome for comparison schools)

T_j = 1 for treatment school and 0 for comparison school

e_j = a random error term for school *j*

ε_{ij} = a random error term for student *i* in school *j*

3.3.1.1 Baseline Analytic Model Specifics

All baseline equivalence testing was done on the analysis sample, i.e., student with both pre- and post-test data. No outcome data and no pre-test data were imputed. Baseline equivalence was assessed for INSPIRE students and comparison students on a pre-test measure of math ability in third grade (DE Common Core Math Assessment) administered in the first nine weeks of Fall 2014, as well as all other propensity score matching variables. The treatment and comparison groups were equivalent at baseline (Appendix C).

We calculated the standardized baseline mean difference between the INSPIRE intervention group and the comparison group by dividing the baseline treatment-comparison difference (β_1 in the level-2 equation) by the student-level pooled standard deviation of pre-test mathematics test scores. Given that we included students' third grade baseline mathematics score in our impact analysis model, we considered baseline equivalence to be established if the standardized mean difference between treatment and comparison on pre-test math scores was less than 0.25. Baseline equivalence was met for the model after statistical adjustment (g = .12).

3.3.2 Confirmatory Analytic Model

$$y_{ij} = \alpha + \beta_1 T_{ij} + \beta_2 Pretest_{ij} + \beta_3 Gender_{ij} + \beta_4 MinorityStatus_{ij} + \beta_5 FRLunch_{ij} + e_j + \varepsilon_{ij}$$

Where,

y_{ij} = the outcome for student i in school j

α = intercept

β_1 = covariate-adjusted impact of INSPIRE (i.e., the difference between the mean outcome for treatment schools and the mean outcome for comparison schools)

T_{ij} = 1 for treatment school and 0 for comparison school

β_2 = parameter estimate for the effect of the student pre-test

$Pretest_{ij}$ = pre-test measure for each student i in school j

β_3 = parameter estimate for the effect of student gender

$Gender_{ij}$ = student gender values, 1 = female student, 0 = male student

β_4 = parameter estimate for the effect of student minority status

$MinorityStatus_{ij}$ = student minority status values, 1 = minority student, 0 = non-minority student

β_5 = parameter estimate for the effect of free or reduced-price lunch eligibility

$FRLunch_{ij}$ = 1= eligible for free or reduced-price lunch, 0 = not eligible

e_j = a random error term for school j

ε_{ij} = a random error term for student i in school j

3.3.3 Analytic Model and Sample Specifics

Math achievement of fifth grade students in treatment schools was compared to math achievement of fifth grade students in comparison schools, after three years of exposure to either the INSPIRE intervention or the "business-as-usual" STEM instruction, controlling for baseline math achievement, measured in the fall of third grade and child demographics (gender, FPRL status, ethnicity).

Students were designated to the treatment and comparison condition at the school level (two treatment schools, four comparison schools). Student characteristics were measured at the student level, as were the outcomes. The regression equation represents the nested structure of the data, in which students were nested within the six schools.

Impacts were estimated using a two-level model with the treatment impact estimated at the school-level. The effect of the INSPIRE intervention is represented by the level-two parameter estimate, β1. The parameter estimate quantifies the difference in the student outcome for INSPIRE schools compared to the outcome for "business-as-usual" comparison schools. If the p-value for the parameter estimate is less than 0.05, we would conclude that there is a statistically significant effect of the INSPIRE program on the given student outcome.

The analysis sample included students with non-missing data on all variables, including pre-test, demographic, and outcome data. Case deletion was used and no data were imputed.

3.3.4 Results for Study 1

Results indicated a statistically significant difference between the *INSPIRE* treatment group and the business-as-usual comparison group on the math achievement outcome. Comparison students demonstrated a statistically significant higher increase in math achievement than *INSPIRE* students. Two other aspects of the model were statistically significant: higher baseline scores in math achievement predicted higher outcome scores in math achievement and individuals who qualified for free and reduced-price lunch reported less growth in math achievement than individuals who did not qualify. Table 4 includes our regression model output.

Table 4. INSPIRE Study 1 Results

Variable	Estimate	Standard Error	*t*-value	*p*-value
Intercept	0.22	0.07	3.29	0.001
Baseline Math Score	0.64	0.03	19.06	< 0.001

Table 4. INSPIRE Study 1 Results				
Variable	Estimate	Standard Error	*t*-value	*p*-value
Condition	-0.22	0.07	-3.37	0.028
Gender	0.03	0.07	0.41	0.681
Minority Status	-0.01	0.07	-0.15	0.878
FRL	-0.27	0.08	-3.54	< 0.001

4. STUDY 2: IMPACT STUDY DESIGN

The secondary level impact study, Study 2, used an individual-level, longitudinal, randomized controlled trial (RCT) with blocking by school level (math achievement) and cohort (science achievement only) to assess the effects of INSPIRE on secondary students' math and science achievement after two years of program exposure. Use of a RCT design effectively minimized selection bias and ensured that the treatment and control groups were equitable at baseline in terms of background, demographic, and pre-program factors such as motivation (Appendix C). Math achievement was measured using recognized assessments with proven reliability and validity, including NC End-of-Grade (EOG) and End-of-Course (EOC) tests.

4.1 Samples

The INSPIRE secondary model was tested at one magnet middle school with a STEM track and one high School (magnet with STEM track). The impact study focused on the effects of an integrated K-12 STEM pipeline after two years of program exposure. A snapshot of the high-need population for these schools is provided in in Table 5.

Table 5. INSPIRE Schools Enrollment Characteristics							
Target School	**N**	**Black**	**Asian**	**White**	**Hispanic**	**Other**	**Low-Income**
Magnet Middle School	484	24%	5%	42%	23%	6%	37%
Magnet High School	893	18%	3%	65%	9%	5%	24%

The specific schools that students in the control group attended were identified after the random assignment of students was completed. Students who were randomized to the control condition attended their zoned middle school or high school. Consequently, students in the control group attended an array of secondary schools in the district, all of which implemented STEM instruction as usual. See Table 6 for a complete list of middle schools or high schools that students in the control condition were zoned for with details on their demographic make-up.

Table 6: Comparison Schools Enrollment Characteristics

Comparison School	Demographic Characteristics of Student Population (2014-2015)						
	N	Black	Asian	White	Hispanic	Other	Low-income
	Percent Representation (%)						
CCS Middle Schools							
Middle School 1	1091	23%	1%	51%	22%	3%	45%
Middle School 2	1185	20%	5%	63%	8%	4%	13%
Middle School 3	1164	20%	3%	67%	5%	5%	15%
Middle School 4	644	4%	1%	88%	5%	2%	33%
Middle School 5	906	22%	1%	54%	19%	4%	40%
Middle School 6	978	18%	2%	50%	28%	2%	42%
CCS High Schools							
High School 1	1252	28%	2%	51%	17%	2%	41%
High School 2	1477	22%	5%	63%	6%	4%	14%
High School 3	232	22%	4%	50%	18%	6%	37%
High School 4	1573	18%	4%	69%	5%	4%	15%
High School 5	1339	23%	2%	52%	20%	3%	33%
High School 6	1200	22%	1%	57%	16%	4%	35%

A carefully orchestrated lottery process was used to randomly assign eligible 6th grade and 9th grade lottery applicants to open INSPIRE middle school and high school slots, or to a non-INSPIRE control group in the middle school or high school for which the student was zoned. Eligible 6th grade applicants were randomly assigned to a treatment group at JN Fries Middle or to the control group. The same process was repeated to randomly assign each cohort of eligible 9th grade applicants to a treatment group at Central Cabarrus High or to the control group.

To be eligible to apply for an opening in the INSPIRE program, students had to be residents of the Cabarrus County School District, meet the criteria for admission, express a desire to attend and expend the effort to succeed in the magnet program, and submit the application to the Education Center (CCS central office) by 12:00 p.m on the last day of January in 2014 or 2015. The following were criteria for admission: a student's scores had to demonstrate a two-year pattern of proficiency on EOGs (e.g.g, score at Level 3 or higher two years prior applying, at or above the 50th percentile on EOGs in the prior academic year, and perform at the 75th percentile or higher on a nationally normed standardized achievement test scores in reading and math in the prior academic year).

The application process opened on the first school day of winter semester (e.g., January 2) and applications were due to the CCS Education Center on the last business day in January, with notification of program acceptance by the first week of March. CCS adhered to a strict application submission deadline. CCS exempted students from random assignment if they did not meet the aforementioned criteria, submitted their application after the deadline, or qualified for priority placement. Priority placement was granted if the applicant had a sibling attending the school or was the child of a CCS employee; priority placement cases were excluded from the lottery and automatically assigned a slot.

The CCS district office managed the application and selection process. When the application period closed, applications meeting the submission deadline were entered into a database. School representatives were consulted to identify priority placements (sibling and CCS employee priorities). After identification of priority placements, the remaining students were randomly assigned to open INSPIRE slots using an Excel random sort function.

This sample identification, selection, and assignment process was used in the 2014/15 school year to identify a cohort of incoming students. The cohort was part of the impact evaluation and followed for the duration of the grant. To increase analytical power, data were pooled across levels (middle school and high school).

In SY 2014-2015, 165 students applied for 66 open INSPIRE slots (37 slots at the middle school level and 29 slots at the high school level). After 18 priority placements, the total study sample of 147 students was entered into the lottery; 47 were assigned to the INSPIRE treatment group and 100 were assigned to the control group. Analysis assessed the effect of INSPIRE on math and science achievement for INSPIRE school students in Grades 8 and 10 after two program years of exposure.

4.2 Study 2 Questions

Study 2 addressed two confirmatory research questions:

1. What was the pooled effect of INSPIRE on middle and high school math achievement (Gr7&10) for INSPIRE students after two program years of program exposure compared to the math achievement of middle school and high school students (Gr7&10) in the “business as usual” condition?

2. What was the pooled effect of INSPIRE on middle and high school science achievement (Gr7&10) for INSPIRE students after two program years of program exposure compared to the science achievement of middle school and high school students (Gr7&10) in the "business as usual" condition?

For contrast 1 (math), analyses were run using pooled samples of students in the treatment condition after two years of exposure to INSPIRE (6th and 9th grade students: fall 2014 cohort tracked for two years with outcomes collected in spring 2016) and students in the comparison condition after two years of exposure to business-as-usual (6th and 9th grade students: fall 2014 cohort tracked for two years with outcomes collected in spring 2016).

For contrast 2 (science), analyses were run using pooled samples of two cohorts of students in the treatment condition after two years of exposure to INSPIRE (6th and 9th grade students: fall 2014 and fall 2015 cohorts tracked for two years with outcomes collected in spring 2016 and spring 2017) and students in the comparison condition after two years of exposure to business-as-usual (6th and 9th grade students: fall 2014 and fall 2015 cohorts tracked for two years with outcomes collected in spring 2016 and spring 2017).

4.3 Analysis and Results

4.3.1 Analysis of Attrition

The study followed an intent-to-treat model on a stable sample of students. The baseline sample of students included all eligible students who applied to attend an INSPIRE school for 6th or 9th grade and were randomized to either the treatment or control condition. The analysis sample was defined as all eligible students who were randomized and had no missing pre-test or post-test data.

The overall attrition rate was calculated as the number of students randomized and not included in the impact analysis divided by the number of students randomized. The attrition rate was calculated separately for the treatment group and comparison group. The differential attrition rate was calculated as the difference between the treatment group attrition rate and the control group attrition rate. The overall and differential attrition rates for each evaluation sample (i.e., for each outcome examined at each time point) were acceptable compared to the NEi3 attrition thresholds (based on the WWC liberal attrition standards). Attrition calculations appear in Appendix C.

4.3.2 Analytic Model and Sample Specifics

Students were randomly assigned to the treatment and comparison conditions at the student level, within the participating institutions. Impacts of the INSPIRE program were estimated using an ordinary least squares (OLS) regression model with school-level indicators to represent the blocked design. Student characteristics and outcomes were measured at the student level. Note that the coefficient β_1 represented the difference between the posttest mean scores of treatment and comparison students. The test of the null hypothesis that β_1 is equal to zero is a test of no intervention effect. If we were able to reject the null hypothesis (i.e., if the p-value is <0.05), then we would conclude that the INSPIRE intervention has an impact on the given outcome measure.

We created a standardized measure of the intervention impact where the estimated difference β_1 was divided by the student-level standard deviation of the outcomes and pooled across the intervention and comparison group standard deviations.

The analysis sample included students with non-missing outcome data and non-missing pre-test data. Case-deletion was used, and no outcome data and no pre-test data were imputed.

4.3.3 Confirmatory Analytic Model

$$y_{ij} = \alpha + \beta_1 T_{ij} + \beta_2 Pretest_{ij} + \beta_3 Gender_{ij} + \beta_4 MinorityStatus_{ij} + \beta_5 FRLunch_{ij} + e_j + \varepsilon_{ij}$$

Where,

y_{ij} = the outcome for student i in school j

α = intercept

β_1 = covariate-adjusted impact of INSPIRE (i.e., the difference between the mean outcome for treatment schools and the mean outcome for comparison schools)

T_{ij} = 1 for treatment school and 0 for comparison school

β_2 = parameter estimate for the effect of the student pre-test

$Pretest_{ij}$ = pre-test measure for each student i in school j

β_3 = parameter estimate for the effect of student gender

$Gender_{ij}$ = student gender values, 1 = female student, 0 = male student

β_4 = parameter estimate for the effect of student minority status

$MinorityStatus_{ij}$ = student minority status values, 1 = minority student, 0 = non-minority student

β_5 = parameter estimate for the effect of free or reduced-price lunch eligibility

$FRLunch_{ij}$ = 1= eligible for free or reduced-price lunch, 0 = not eligible

e_j = a random error term for school j

ε_{ij} = a random error term for student i in school j

4.3.4 Results for Study 2

4.3.4.1 Results for Math Outcome (Study 2)

Results indicated no statistically significant difference between the *INSPIRE* treatment group and the business-as-usual comparison group on the math achievement outcome. Also, baseline math achievement significantly predicted higher outcome math achievement. Table 7 includes our regression model output.

Table 7. INSPIRE Study 2 Math Achievement Model

Variable	Estimate	Standard Error	*t*-value	*p*-value
Intercept	0.09	0.24	0.38	0.704
Baseline Math Score	0.70	0.07	10.82	< 0.001
Condition	-0.14	0.14	-1.03	0.307
Gender	0.06	0.13	0.47	0.641
Ethnicity	0.18	0.15	1.23	0.222
FRL	-0.32	0.18	-1.80	0.075
School Level	0.11	0.30	0.38	0.713

4.3.4.2 Results for Science Outcome (Study 2)

Results indicated no statistically significant difference between the *INSPIRE* treatment group and the business-as-usual comparison group on the science achievement outcome. Two statistically significant findings arose from our model: higher baseline science achievement predicted higher outcome science achievement and individuals who qualified for free and reduced-price lunch reported less growth in science achievement than individuals who did not qualify. Table 8 includes our regression model output.

Table 8. INSPIRE Study 2 Science Achievement Model				
Variable	Estimate	Standard Error	*t*-value	*p*-value
Intercept	-0.11	0.48	-0.22	0.824
Baseline Science Score	0.53	0.05	10.53	< 0.001
Condition	0.03	0.10	0.25	0.800
Cohort	-0.08	0.41	-0.20	0.844
Gender	-0.04	0.10	-0.42	0.675
Ethnicity	0.16	0.11	1.44	0.151
FRL	-0.29	0.14	-2.12	0.035
School Level	0.04	0.06	0.59	0.554

5. DISCUSSION

INSPIRE was designed to address a national need to develop and validate an integrated K-12 STEM pipeline that serves as a model for STEM instructional redesign and a strategy for intentional outreach to engage and retain minority and low-income students in STEM. The *INSPIRE* impact study found statistically significant negative effects for math achievement at the elementary level and no statistically significant effects in mathematics and science achievement at the secondary level. For this study, mathematics and science achievement were assessed using the Discovery Education's (DE) Common Core Math Formative Assessment (an elementary baseline measure) and NC state's End-of-Grade and End-of-Course standardized exams. To contextualize these findings, the duration of students' exposure to INSPIRE, alignment between PBL instruction and NC standardized assessments, and contextual factors that might have narrowed the divide between INSPIRE and non-INSPIRE conditions are considered below.

Instructional redesign of the breadth and scope undertaken by INSPIRE required a substantial investment of time in designing infrastructure, teacher professional development, and coaching. In the first implementation year, foundational work was done in hiring and training four school-based STEM coaches and developing standards for PBL lessons. Based on the fidelity of implementation findings, INSPIRE teachers continued to strengthen their ability to develop and deliver high-quality, interdisciplinary PBLs with real-world relevance throughout the three years of implementation. Further, as INSPIRE's integrated approach interfaced with a K-12 continuum, it was hypothesized to provide an additive impact to help accelerate student growth. It is possible that the length of exposure was insufficient to detect impact on achievement as measured by standardized state exams. Secondly, high fidelity of implementation study findings provide evidence that INSPIRE developed and implemented high-quality, interdisciplinary PBLs that mapped to NC State Standards, incorporated technology, and centered on real-world problems. In fact, INSPIRE partnerships with STEM organizations and professionals played a key role in the development of PBLs and teachers' perceived efficacy in implementing a PBL approach. NC standardized state exams appeared to be well-aligned measures for the INSPIRE intervention because the tests in mathematics and science have an increased focus on processing information and higher-order thinking. However, in large part, the emphasis of the standardized exams was on computations and basic understandings (NC Department of Instruction, 2018). PBL classrooms and lessons looked very different from traditional business-as-usual classrooms because of an emphasis on implementing hands-on, inquiry-based, and industry-relevant lessons on a monthly basis. It is possible that students in business-as-usual conditions experienced greater levels of classroom instruction, exercises, and activities that were more closely aligned with the skills necessary to achieve on multiple-choice state assessments, especially in the first year of implementation (Strobel & van Barneveld, 2009). It is possible that INSPIRE elementary students developed their aptitude for standardized test-taking at a slower rate in relation to students in the comparison group.

Another important caveat to consider in interpreting the study findings relate to the degree to which the non-INSPIRE schools in the study narrowed the divide in STEM instruction. INSPIRE STEM schools and teachers were being recognized by North Carolina state officials and community members as leaders in STEM PBL instruction that was both rigorous and engaging. By SY 2016-17, all four INSPIRE schools were recognized by NC Department of Instruction as NC STEM Schools of Distinction on the Model Level. During the study period, the district was experiencing increased demand for our STEM magnet program. The CCS School Board responded to the community demand by proposing an expansion of the STEM program in School Year 2016-2017 to two elementary schools within the district. In fact, one of the two schools was originally selected as comparison schools for our elementary level study and had to be replaced. While the INSPIRE team was careful to control dissemination of key strategies within the district, it is likely that teachers within non-INSPIRE schools were narrowing the divide between INSPIRE and traditional STEM education. As suggested by a growing body of literature, this study would have been strengthened by measuring achieved relative strength, that is the difference in what is implemented and experienced in the treatment and untreated groups (Cordray, 2010; Hulleman, Rimm-Kaufman, Abry, 2013; Nelson, Cordray, Hulleman, Darrow, & Sommer, 2012).

In conclusion, this study makes significant contributions to the literature that have important programmatic and policy implications. Based on the fidelity of implementation findings, we conclude that it takes at least two academic years and ongoing job-embedded professional development to establish the infrastructure, procedures, and tools that support comprehensive curriculum redesign. This study adds to prior research that has documented that three to five years may be needed to see the effects of major system changes (Borman, Hughes, Overman, and Brown, 2003). The biggest challenge noted by program implementers was having patience to realize that making large impact changes required a substantial time investment and many conversations with all levels of stakeholders including administrators, teachers, school board members, parents, and community leaders.

Additionally, curriculum redesign that is accompanied by a review of district policies may be most fruitful in supporting pipeline persistence. Based on this study, it may take more than two or three years for comprehensive curricular reform to take hold and translate into significant student gains on standardized assessments. Furthermore, students that are immersed in PBL-rich environments may develop test-taking proficiency at slower rates. The over reliance on standardized assessment scores to make placement decisions at key transition points (particularly at the elementary to middle school transition point) may work at cross-purposes to bolstering engagement and persistence in learning, especially for students from low-income, minority backgrounds. Future research should explore the contribution of policy reform changes aimed at reducing barriers to pipeline persistence over and above curriculum reform. The district will pursue this line of research towards the goal of expanding magnet options and widening educational reach for its underserved communities (US Department of Education, Office of Innovation and Improvement (OII), Investing in Innovation (i3) Fund, PR Award # U411C160019).

REFERENCES

Asghar, A., Ellington, R., Rice, E., Johnson, F., & Prime, G. (2012). Supporting STEM education in secondary science contexts. *Interdisciplinary Journal of Problem-Based Learning, 6*(2), 85-125. Retrieved from http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1349&context=ijpbl

Bayer Corporation. (2010). *Planting the seeds for a diverse U.S. STEM pipeline: A Compendium of best practice K-12 STEM education programs*. Elkhart, Indiana: Bayer. Retrieved from www.bayerus.com/MSMS/web_docs/Companion_Guide.pdf

Borman, G., Hughes, G., Overman, L., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73*, 125-230. Retrieved from https://doi.org/10.3102%2F00346543073002125

Brown, S., Lawless, K., & Boyer, M. (2013). Promoting positive academic dispositions using a web-based PBL environment: The GlobalEd 2 project. *Interdisciplinary Journal of Problem-based Learning, 7*(1), 67-90. Retrieved from http://dx.doi.org/10.7771/1541-5015.1389

Cabarrus County Schools (2014a). INSPIRE Problem-Based/Project-Based Learning (PBL) Lesson Framework and Assessment of Quality Rubric. Retrieved from https://www.cabarrus.k12.nc.us/Page/68765

Cabarrus County Schools (2014b). INSPIRE STEM PBL Instructional Coaching Tool. Retrieved from https://www.cabarrus.k12.nc.us/Page/68765

Cordray, D.S. (2010). Assessing implementation fidelity and achieved relative strength in RCTs: Concepts and methods. *Presentation at The Nebraska Center for Research on Children, Youth, Families and Schools.*

Durlak, J., & Dupre, E. (2008). Implementation matters: A review of the research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350. Retrieved from http://link.springer.com/article/10.1007/s10464-008-9165-0

Geier, R., Blumenfeld, P. C., Marx, R. W., Krajcik, J. S., Fishman, B., Soloway, E., & Clay-Chambers, J. (2008). Standardized test outcomes for students engaged in inquiry-based science curricula in the context of urban reform. *Journal of Research in Science Teaching, 45*(8), 922-939. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/tea.20248/abstract;jsessionid=ADDE22323F1AB86FE24C9A3E05DC0AF7.d01t01?deniedAccessCustomisedMessage=&userIsAuthenticated=false

Hulleman, C.S., Rimm-Kaufman, S.E., & Abry, T. (2013). Chapter 3: Innovative methodologies to explore implementation. Whole-part-whole: Construct validity measurement and analytical issues for intervention fidelity assessment in education research. In *Applying Implementation Science in Childhood Settings*. Eds. Halle, T. G., Metz, A. J., & Martinez-Beck, I. Baltimore, MD: Brookes.

Kwan, A. (2009). Problem-based learning. In M. Tight, K. Mok, J. Huisman, & C. Morphew, *The Routledge international handbook of higher education* (pp. 91-107). Routledge. Retrieved from http://books.google.com/books?hl=en&lr=&id=WjenckTFlUIC&oi=fnd&pg=PA91&dq=problem+based+learning&ots=Oq79R-clOT&sig=mg8zgd2qX1U4wlWcJknDl6QEjds#v=onepage&q=problem%20based%20learning&f=false

Lou, S., Shih, R., Diez, C., & Tseng, K. (2011). The impact of problem-based learning strategies on STEM knowledge integration and attitudes: An exploratory study among female Taiwanese senior high school students. *International Journal of Technology and Design Education, 21*(2), 195-215. Retrieved from http://dx.doi.org/10.1007/s10798-010-9114-8

National Research Council. (2011). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics.* Retrieved from National Academies Press website: http://www.nap.edu/catalog.php?record_id=13158

Nelson, M.C., Cordray, D.S., Hulleman, C.S., Darrow, C.L., & Sommer, E.C. (2012). A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. *Journal of Behavioral Health Services and Research, 394*, 374-396.

NC Department of Insruction (2018, August). *North Carolina Testing Program*. Retreived from http://www.ncpublicschools.org/accountability/testing/

Strobel, J., & van Barneveld, A. (2009). When is PBL more effective? A meta-synthesis of meta-analyses comparing PBL to conventional classrooms. *Interdisciplinary Journal of Problem-based Learning, 3*(1), 44-58. Retrieved from http://dx.doi.org/10.7771/1541-5015.1046

Van Berkel, H., & Dolmans, D. (2006). The influence of tutoring competencies on problems, group functioning and student achievement in problem-based learning. *Medical Education, 40*(8), 730-736. Retrieved from http://dx.doi.org/10.1111/j.1365-2929.2006.02530.x

Walker, A., & Leary, H. (2009). A problem-based learning meta analysis: Differences across problem types, implementation types, disciplines, and assessment levels. *Interdisciplinary Journal of Problem-based Learning, 3*(1), 12-43. Retrieved from http://dx.doi.org/10.7771/1541-5015.1061

What Works Clearinghouse (2017). *Procedures and standards handbook version 4.0*. Retrieved from https://ies.ed.gov/ncee/wwc/handbooks

APPENDIX A
Glossary of Key INSPIRE Strategies

Curriculum Writing Days	Curriculum writing days were paid professional development days attended by teachers, typically three to five days during the summer to support curriculum development and teacher collaboration. Curriculum writing days was reported by teacher survey respondents as the second most effective support for enhancing instructional practice.
INSPIRE PBL Lesson Framework	CCS STEM Coaches and curriculum specialists developed the INSPIRE Problem-Based/Project-Based Learning (PBL) Lesson Framework to advance the mission of creating an integrated K-12 STEM pipeline approach (Cabarrus County Schools, 2014a). With the goal of preparing students for STEM careers, the PBL lesson framework places emphasis on real-world connections and interdisciplinary study. The PBL Lesson Framework served as a guide for developing rigorous PBL lessons that map to NC State Standards, incorporate technology, center on real-world problem, and encourage reflection from teachers and students.
INSPIRE PBL Assessment of Quality Rubric	Developed through a collaboration with STEM coaches, curriculum specialists, and the external evaluator, the INSPIRE PBL Assessment of Quality Rubric (Cabarrus County Schools, 2014a) is a guide for assessing PBL lessons. The INSPIRE Assessment of Quality Rubric includes six Essential Core Criteria and two Criteria of Excellence: Vertical/Horizontal Alignment, Common Core Alignment, Multifaceted Assessment, Real-world Connections, Curricular Integration, Collaborative Inquiry, Student Engagement, and Design Cycle/21st Century Skills, respectively. The tool can be used by teachers, instructional coaches, and educational administrators to support an ongoing process for developing rigorous problem-based or project-based learning instructional support in the nine areas.
INSPIRE STEM PBL Instructional Coaching Tool	Developed through a collaboration with STEM coaches, curriculum specialists, the external evaluator, INSPIRE STEM PBL Instructional Coaching Tool (Cabarrus County Schools, 2014b) is a professional development planning tool. This tool was designed to meet the following INSPIRE goals: development of an individualized plan to support growth in identified priority areas; identification of best practices that can be disseminated to others; and demonstration of proficiency in PBL-related instructional practices. The tool can be used by teachers, instructional coaches, and educational administrators to support an ongoing process for problem-based or project-based learning instructional support in nine areas.
INSPIRE Innovation Showcase	INSPIRE Innovation showcase allowed STEM teachers to share innovative practices. It served as a key dissemination strategy for INSPIRE throughout the grant, within and between schools. The Innovation Showcase was restricted to INSPIRE teachers in project Years 1 – 3 (spanning Spring 2014 – spring 2016) and open to all teachers in the district in the final year of the grant (fall 2017), after data collection for the impact study had been completed.
NASA Camp	NASA Camp was an annual summer camp for INSPIRE K – 12 students. It was a partnership with NASA experts who provided training to INSPIRE teachers.
Real-world Student Tethers	INSPIRE offered a variety of experiences during and outside of school hours to support student STEM engagement and achievement. The events included

	the third annual INSPIRE NASA Summer Camp focused on astronautics, field trips, clubs, and school-sponsored STEM events.
STEM Coach	STEM Coaches were school-based veteran educators. STEM coaches focused on introducing and training staff on the INSPIRE PBL lesson framework, instructional technologies, and design cycles based on a plan tailored for each school. STEM coaches worked to identify and address gaps in classroom instruction. For areas identified as gaps, STEM coaches used a combination of individual consultation, grade-level PLC team meeting support, and STEM-focused professional development to increase the proficiency of their teachers. STEM PD delivered by STEM coaches was reported by teacher survey respondents as the most effective support for enhancing instructional practice. As the program matured, the role of STEM coaches evolved from providing direct student/classroom support to providing leadership and building capacity toward the goals of rigor, relevance, and cross-curricular integration.
STEMersion	STEMersion was a five-day summer immersion in local STEM-related Cabarrus County businesses and industries. STEMersion provided STEM K-12 classroom educators the opportunity to visit local Cabarrus County businesses and industries to take a behind the scenes look at how they operate. Mini-STEMersion occurred through the school year. Teachers used STEMersion to enhance their classroom lessons with real world examples.
Vertical alignment days	Vertical alignment days were professional development days devoted to within-grade-band alignment, as well as, cross-school resource sharing and alignment. Providing student synergy and engagement, the aim was to integrate PBL lessons across courses, so students explore the same problem in every subject area.

APPENDIX B
INSPIRE Fidelity Index and Study Findings

Fidelity Indicator & Measure	Year 1 (SY 2014/15) Threshold/Actual/ Level (Score)	Year 2 (SY 2015/16) Threshold/Actual/ Level (Score)	Year 3 (SY 2016/17) Threshold/Actual/ Level (Score)
Component 1. High Quality Teacher Development & Support			
1. ***Dosage:*** # of INSPIRE-sponsored STEM trainings delivered by program per year. *Measure:* INSPIRE PD Tracking Tool	Target: 3 Actual: 12 ***Level: High (2)***	Target: 3 Actual: 23 ***Level: High (2)***	Target: 3 Actual: 31 ***Level: High (2)***
2. ***Reach:*** % of STEM teachers who complete at least 10 hours of STEM-related PD per school year. *Measure:* INSPIRE PD Tracking Tool	Target: 80% Actual: 84% **Level: High (2)**	Target: 80% Actual: 100% **Level: High (2)**	Target: 80% Actual: 99% **Level: High (2)**
3. ***Reach:*** % of STEM teachers who receive 20 or more sessions with on-site STEM coaches during the school year. *Measure*: STEM coaching logs (*INSPIRE* Teacher Survey in YR1 and YR2)	Target: 80% Actual: 71% ***Level: Moderate (1)***	Target: 80% Actual: 73% ***Level: Moderate (1)***	Target: 80% Actual: 94% ***Level: High (2)***
4. ***Reach:*** % of STEM teachers who meet with their PLCs per month per school year. *Measure*: PLC attendance logs	Target: 75% Actual: 96% ***Level: High (2)***	Target: 75% Actual: 91% ***Level: High (2)***	Target: 75% Actual: 97% ***Level: High (2)***
Component Level Score	**Target: 4 of 8** **Actual: 7 of 8** ✓	**Target: 5 of 8** **Actual: 7 of 8** ✓	**Target: 5 of 8** **Actual: 8 of 8** ✓
Component 2. Problem-Based Learning			
1. ***Dosage:*** Total number of lessons developed by the program per school year. *Measure*: Completed PBL Lesson Frameworks	Target: ES=48; MS=56; HS=60 Actual: CWES=63; PES=52; MS=59; HS=89 ***Level: High (1)***	Target: ES=48; MS=56; HS=60 Actual: CWES=62; PES=53; MS=64; HS=143 ***Level: High (1)***	Target: ES=48; MS=56; HS=60 Actual: CWES=63; PES=52; MS=68; HS=78 ***Level: High (1)***

Fidelity Indicator & Measure	Year 1 (SY 2014/15) Threshold/Actual/ Level (Score)	Year 2 (SY 2015/16) Threshold/Actual/ Level (Score)	Year 3 (SY 2016/17) Threshold/Actual/ Level (Score)
2. ***Quality:*** % of PBL Lessons developed by the program per school year that meet at least 5 of 6 quality criteria. *Measure:* PBL Framework Quality Assessments	Target: 75% Actual: 73% ***Level: Moderate/Low (0)***	Target: 75% Actual: 96% ***Level: High (1)***	Target: 75% Actual: 96% ***Level: High (1)***
3. ***Dosage:*** % of STEM teachers who implement at least one PBL lesson per month per school year. *Measure*: Coaches Implementation Spreadsheet (YR1 based on YR2 *INSPIRE* Teacher Survey; YR3 based on coaching logs).	Target: 80% Actual: 78% ***Level: Moderate/Low (0)***	Target: 80% Actual: 82% ***Level: High (1)***	Target: 80% Actual: 97% ***Level: High (1)***
Component Level Score	***Target: 4 of 4*** ***Actual: 2 of 4***	***Target: 4 of 4*** ***Actual: 4 of 4*** ✓	***Target: 4 of 4*** ***Actual: 4 of 4*** ✓
Component 3. Personalized Tech-enabled Instruction			
1. ***Dosage***: Percentage of PBL Lessons that incorporate technology developed per school year. *Measure:* Completed PBL Lesson Framework; YR3 Coaching Logs	Target: 80% Actual: 99% ***Level: High (1)***	Target: 80% Actual: 97% ***Level: High (1)***	Target: 80% Actual: 89% ***Level: High (1)***
2. ***Reach:*** % of eligible STEM students reporting use of technology with software that adjusts to students' level of mastery on a weekly basis during the school year. *Measure*: INSPIRE Year-End INSPIRE Student Survey (MS & HS) and Dreambox Usage Data (ES)	Target: 70% Actual: 92% ***Level: High (1)***	Target: 70% Actual: 86% ***Level: High (1)***	Target: 70% Actual: 88% ***Level: High (1)***
Component Level Score	***Target: 2 of 2*** ***Actual: 2 of 2*** ✓	***Target: 2 of 2*** ***Actual: 2 of 2*** ✓	***Target: 2 of 2*** ***Actual: 2 of 2*** ✓
Component 4. Student Real-World STEM Tethers			
1. ***Dosage***: % of STEM students who participate in at least 2 school-sponsored STEM events per school during the school year. *Measure*: School administrative records and permission slips	Target: ES=50%; MS=70%; HS=90% Actual: All Schools = 100% ***Level: High (1)***	Target: ES=50%; MS=70%; HS=90% Actual: CWES=80%; PES=80%; MS=100%; HS=94%	Target: ES=50%; MS=70%; HS=90% Actual: ES: 98% MS and HS Schools = 100%

		Level: High (1)	***Level: High (1)***
Fidelity Indicator & Measure	**Year 1 (SY 2014/15) Threshold/Actual/ Level (Score)**	**Year 2 (SY 2015/16) Threshold/Actual/ Level (Score)**	**Year 3 (SY 2016/17) Threshold/Actual/ Level (Score)**
2. ***Reactions:*** % of STEM students participating in NASA Camp annually during the summer months. *Measure*: Camp enrollment records and Camp attendance records	Target: 10% Actual: 15% ***Level: High (1)***	Target: 10% Actual: 13% ***Level: High (1)***	Target: 10% Actual: 10% ***Level: High (1)***
Component Level Score	***Target: 2 of 2*** ***Actual: 2 of 2*** ✓	***Target: 2 of 2*** ***Actual: 2 of 2*** ✓	***Target: 2 of 2*** ***Actual: 2 of 2*** ✓

APPENDIX C
Baseline Characteristics of the INSPIRE Study Sample

Table 1. Baseline Characteristics of the INSPIRE Study Sample			
Characteristics	INSPIRE Sample	Comparison/Control Group	Standardized Difference[b]
Study 1 (Elementary Math):	*n* = 200	*n* = 200	
% Gender	42	44	-0.04
% Minority	37	37	-0.01
% Economically Disadvantaged	35	37	-0.05
% Proficient[a]	29	29	0.00
Standardized Sample Mean (SD)[c]	0.31 (1.21)	0.17 (0.90)	0.12
Study 2 (Secondary Math):	*n* = 42	*n* = 89	
% Gender	38	44	-0.14
% Minority	79	73	0.18
% Economically Disadvantaged	17	17	-0.01
% High School Students	48	25	0.61
% Proficient[a]	95	89	-0.16
Standardized Sample Mean (SD)[c]	0.25 (1.03)	-0.12 (0.96)	0.37
Study 2 (Secondary Science):	*n* = 102	*n* = 176	

Table 1. Baseline Characteristics of the INSPIRE Study Sample			
Characteristics	INSPIRE Sample	Comparison/Control Group	Standardized Difference[b]
% Gender	38	43	-0.14
% Minority	74	69	0.12
% Economically Disadvantaged	15	18	-0.11
% High School Students	33	22	0.34
% Proficient[a]	96	93	0.00
Standardized Sample Mean (SD)[c]	0.12 (1.05)	-0.01 (0.96)	0.13

NOTES:

[a]Hedges' *g* was used to calculate the standardized differences. Baseline equivalence was established if the standardized difference between treatment and comparison groups was less than 0.25. Statistical adjustment was included for all key variables.

[b]Students scoring at proficiency levels 3, 4, or 5 are considered to be performing at grade level for state math and science exams. For technical information on North Carolina assessments, visit http://www.ncpublicschools.org/accountability/testing/technicalnotes.

[c]The standardized baseline mean difference between the INSPIRE students and comparison/control group students were calculated by dividing the baseline treatment-comparison difference by the student-level pooled standard deviation of pre-test scores.

Table 2. (Study 2) Student Attrition Rates for Treatment and Control Groups by Contrast					
Contrast	N of Students Randomized	N of Students in Analysis[a]	Group Attrition Rate	Overall Student Attrition Rate[b]	Differential Student Attrition Rate[c]
Secondary Math					
Treatment	47	42	10.6%	10.9 %	0.4%
Control	100	89	11.0%		
Secondary Science					
Treatment	121	102	15.7%	18.5%	4.3%
Control	220	176	20.0%		

NOTES:

[a] Missing data was the reason for loss. The analysis sample included students with non-missing outcome data and non-missing pre-test data.

[b] The overall attrition rate was calculated as the number of students randomized and not included in the impact analysis divided by the number of students randomized.

[c] The differential attrition rate was calculated as the difference between the treatment group attrition rate and the control group attrition rate.

[d] Middle School Math, High School Math, and Science topic area standards for attrition were applied and met. Topic area standards were retrieved from https://ies.ed.gov/ncee/wwc/Handbooks#protocol