Impact Evaluation of G²ROW STEM: Girls and Guys Realizing Opportunities with STEM

SUBMITTED TO:
Deanna Privette
G²ROW STEM
Metro Nashville Public Schools

SUBMITTED BY:
Catherine Snyder, Ph.D.
Olivia Stevenson, M.S.
The Evaluation Group
www.evaluationgroup.com

June 2021

G^2ROW STEM

Final Impact Evaluation Report

Lead Agency: Metro Nashville Public Schools
2601 Bransford Ave.
Nashville, TN 37204

Project Director: Deanna Privette

Third-Party Evaluator: Catherine Snyder, Ph.D.
Olivia Stevenson, M.S.
The Evaluation Group

Contact Information: Deanna Privette: 615-259-4636, Deanna.Privette@mnps.org>
Catherine Snyder: 803-605-2690, cathi@evaluationgroup.com
Olivia Stevenson: 803-331-1679, olivia@evaluationgroup.com

Grant Period: 01/01/2016 – 12/31/2020, no-cost extension until 6/30/2021

Funder: US Department of Education, Office of Innovation and Improvement (OII), Investing in Innovation (i3) Fund

PR Award #: U411C150082

I3 Cohort Year: 2015

i3 ID: DEV104

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 U411C150082 to Metro Nashville Public 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 Catherine Snyder, The Evaluation Group, 2607 Woodruff Rd., Suite E#112, Simpsonville, SC 29681. Phone: (803) 605-2690; email cathi@evaluationgroup.com. For information about The Evaluation Group, see our website: www.evaluationgroup.com.

Contents

EXECUTIVE SUMMARY 4
1. INTRODUCTION 5
1.1 Program Description 5
STEM Extended Learning. 6
STEM Professional Development. 7
STEM Industry Experiences. 8
Community STEM Awareness. 9
1.2 Fidelity of Implementation 9
Logic Model. 9
Fidelity Index. 10
2. IMPACT STUDY DESIGN 11
2.1 Research Questions 12
2.2 Samples 13
2.3 Baseline Equivalence 17
Baseline Analytic Model. 17
Baseline Analytic Model Specifics. 17
2.4 Data Elements 20
3. ANALYSIS AND RESULTS 22
3.1 Confirmatory Analytic Model 22
Confirmatory Analytic Model Specifics 22
3.2 Exploratory Analytic Model 23
Exploratory Analytic Model Specifics. 23
3.3 Results for Confirmatory One-Year Math Analysis 23
3.4 Results for Confirmatory Two-Year Math Analysis 25
3.5 Results for Confirmatory One-Year Science Analysis 26
3.6 Results for Exploratory Analyses 28
4. DISCUSSION 30

EXECUTIVE SUMMARY

G^2ROW STEM is an Investing in Innovation (i3) development grant funded by the Office of Innovation and Improvement, U.S. Department of Education. *G^2ROW STEM* targets high-need middle school students, particularly females and minorities, that tend to be underrepresented in STEM careers. *G^2ROW STEM* focuses on providing students with engaging, hands-on, project-based, extended-learning experiences to inspire interest in STEM and improve achievement. The impact evaluation used a quasi-experimental design (QED) to examine the effect of *G^2ROW STEM* on academic achievement in science and math. *G^2ROW STEM* student enrollment began during the 2016-17 school year and students were followed for three years, through the 2018-19 school year. Academic achievement outcomes for *G^2ROW STEM* students were compared to a matched sample of students within the same grade in the same schools who participated in business-as-usual, traditional academic instruction with no extended learning experiences. Academic achievement for comparison students was also tracked over a three-year period. Results showed statistically significant differences between treatment and comparison groups in achievement test scores across both math and science. Overall, treatment students gained seven NCE score points more than the comparison group students did after up to three years in the *G^2ROW STEM* program. This is roughly equivalent to one-third of a standard deviation. Significant differences were also noted for student subgroups with non-minority students realizing slightly higher gains in NCE scores than minority students by the end of the program (over all analyses a difference of 2.25 NCE points).

1. INTRODUCTION

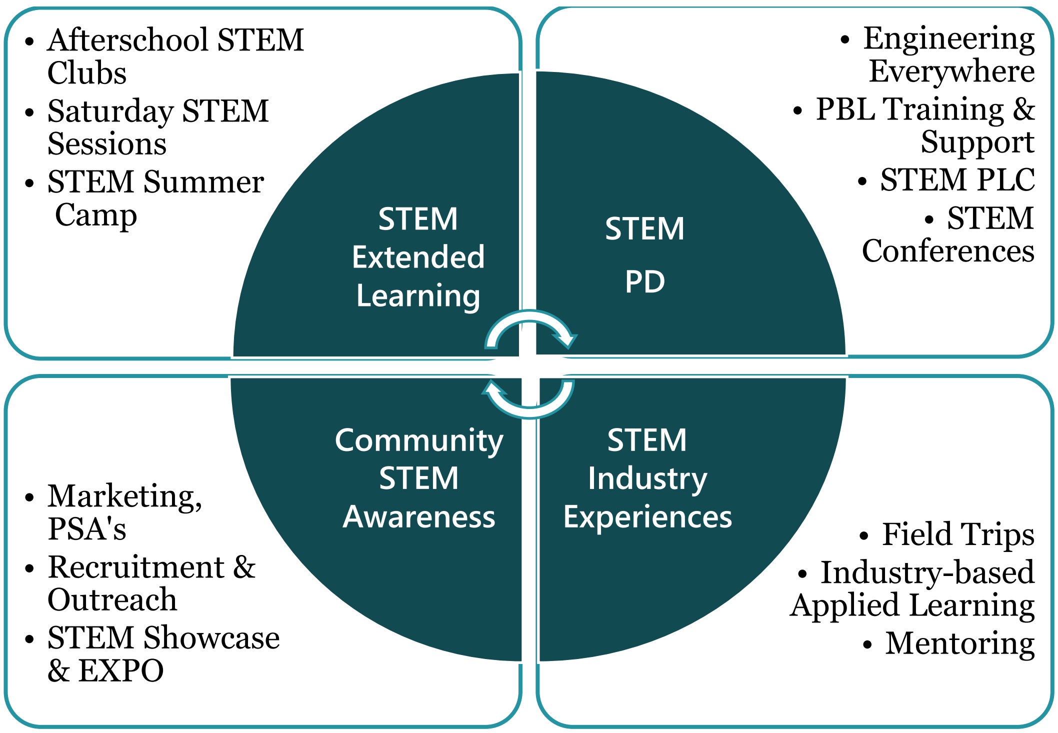
The *G²ROW STEM* program was developed to respond to a growing STEM workforce crisis that threatens the nation's economic well-being. According to the US Bureau of Labor Statistics, STEM career opportunities are projected to reach more than 9 million jobs by 2022 (Charette, 2016.) Recognizing the need to better prepare students for college and future careers, every Metro Nashville Public Schools (MNPS) high school offers career academies aligned with industry job projections, ensuring students are provided with the advanced skills they need to be successful in college, career, and life. Several academies have a STEM focus, such as engineering, information technology, digital design, national safety/security, and aviation. Students learn in a hands-on environment through project-based learning (PBL) and real-world application (including industry field trips, job shadowing, and a senior capstone project). The Academies of Nashville is nationally accredited through the National Career Academy Coalition and recognized as a Ford Next Generation Learning Community. Through *G²ROW STEM*, MNPS extended the success of the career academy concept to the middle schools and developed STEM extended learning and career-focused applied experiences for middle school students. With the identified shortages in STEM jobs, it is imperative that strategic interventions be made during the middle school years to engage high-need, underrepresented students (including females and minorities) in meaningful, real-world STEM learning experiences that will help boost their achievement, build confidence, and inspire them to pursue STEM academic pathways. *G²ROW STEM* was designed to address these needs through PBL, recognized as a promising practice by What Works Clearinghouse (Institute of Education Sciences, 2020), and a variety of experiences in STEM to increase students' interest, academic achievement, and future pursuits of STEM careers. Research indicates that the most effective STEM education models infuse classroom instruction that is based on a rigorous curriculum with frequent exposure to applied learning experiences through lab work, workplace activities, and supportive technology (Hanover, 2011). Further, continuous exposure to real-world STEM activities increases STEM engagement and learning (Bayer, 2010). Students who participate in career-focused programs that relate schooling to careers achieve higher levels of educational attainment and better labor market outcomes (Bridgeland, Balfanz, Moore, & Friant, 2010).

1.1 Program Description

G²ROW STEM was implemented and evaluated at seven Title 1 middle schools that serve some of the highest need students in MNPS. Target schools enroll 4,332 students with 43.6% economically disadvantaged, 38.7% African American, 30.8% Hispanic, 3.9% Asian, and 16.3% ELL. These high-need students still face challenges: on average, only 18.9% of target students scored at on target or mastery levels in math, 18.9% scored at on target or mastery levels in ELA (Tennessee Department of Education, 2020).

G²ROW STEM integrated four core strategies to support the impact of high-quality STEM extended learning and career-focused, applied experiences on academic achievement, STEM engagement, and STEM career aspirations for low-income, underserved middle school students. The *G²ROW STEM* program model is summarized in Figure 1.

Figure 1. The G²ROW STEM Program Model



STEM Extended Learning. MNPS identified a rigorous and relevant STEM curriculum, *Engineering Everywhere (EE)*, to use as the foundation for extended learning. Developed by the Museum of Science in Boston (Engineering is Elementary, 2020; Museum of Science, 2015), *EE* is an out-of-school time (OST) curriculum for middle school students that is grounded in PBL, which research shows can increase interest in STEM related fields and improve academic motivation and achievement (Cunningham, 2009; Strobel & van Barneveld, 2009). The curriculum was designed to improve academic achievement and youths' attitudes about their abilities to engineer, problem solve, and think creatively. Each unit focused on a real-world problem and underwent a rigorous pilot test with middle school students during development. Research shows that providing high quality, relevant, and more frequent STEM exposure in extended learning programs is critical to increase interest in STEM, improve academic achievement, and increase enrollment in more rigorous science and math classes (Krishnamurthi, Bevan, Coulon, & Rinehart, 2013). Compared to traditional instruction, PBL raises long-term retention of content, helps students perform as well or better than traditional learners in high-stakes tests, and improves problem-solving and attitudes towards learning (Strobel & van Barneveld, 2009; Walker & Leary, 2009). Further, PBL instruction has been shown to improve student engagement, motivation, and academic achievement in STEM, particularly among female students (Liu, Lou, & Shih, 2014). To ensure students received broad exposure to the range of STEM opportunities, MNPS supplemented *EE* with additional STEM units developed through collaborations with community partners to develop authentic, real-world activities that are research-based and grounded in PBL. These STEM units were offered during monthly Saturday sessions and were selected based on students' interests and authentic, real-world connections to the afterschool curriculum. Table 1 outlines extended learning strategies, which provided students with **222 hours of STEM extended learning** annually.

Table 1: *G²ROW STEM* Extended Learning Strategies

Afterschool STEM Clubs	Each school offered an Afterschool STEM Club for up to 30 students (2 days/week, 2 hours/day, 28 weeks; **112 total contact hours during the school year**). Led by two lead STEM teachers, the afterschool clubs provided broad exposure to hands-on, PBL STEM instruction using *Engineering Everywhere*. Students explored a broad range of STEM topics (i.e., aeronautical, agricultural, biomechanical, biomedical, chemical, environmental, green, materials, and mechanical engineering).
Saturday STEM Sessions	*G²ROW STEM* offered monthly Saturday STEM sessions for middle school students at three feeder high schools (15 students from each middle school, 5 hours/day, 7 sessions/year; **35 contact hours annually**). Sessions built on afterschool activities and provided more in-depth exposure and authentic connections to STEM with applied instruction provided by community partnerships, STEM field trips, STEM speakers, and STEM mentoring (peers in STEM academies, STEM universities and industry professionals).
STEM Summer Camps	*G²ROW STEM* offered theme-based STEM Summer Camps at two feeder high school sites (15 students from each middle school, 5 hours/day, 2 weeks in Year 2, 3 weeks in Years 3-4, and cancelled in Year 5 due to COVID-19, and 2 weeks during a no-cost extension period; **75 contact hours annually**). *G²ROW STEM* built on the afterschool program and Saturday sessions by exposing students to a specific camp theme (i.e., engineering, solar eclipses, healthy living using energy and sustainability, robotics, aerospace engineering, and forensics/crime scene investigations). Each camp included an introduction to the theme with PBL activities using community partnerships; STEM field trips; STEM speakers; and STEM mentoring. Themes were adjusted each year based on student interest.

STEM Professional Development. Each school identified two lead STEM teachers to lead afterschool, Saturday, and summer STEM instruction. Lead teachers played a critical role in aligning extended learning with school curriculum, fostering STEM collaborations with other teachers, and building a STEM culture in the school to sustain program impact. Research shows STEM programs are most effective when teachers are properly trained and have a strong STEM grounding (Howard-Brown & Martinez, 2012). A study funded by the U.S. Department of Education (U.S. ED) revealed that the use of PBL can have a positive impact on the way teachers view students' academic abilities and can reveal previously unseen academic potential in reluctant learners (Gallagher & Gallagher, 2013). Researchers have also recognized that science teacher effectiveness is correlated with future science achievement and pursuit of science-related careers (Bolshakova, Johnson, & Czerniak, 2011). To build STEM content knowledge and instructional skills, *G²ROW STEM* incorporated several professional development (PD) strategies, as Table 2 outlines.

Table 2: *G²ROW STEM* Professional Development Strategies

Engineering Everywhere PD and STEM PLC	STEM lead teachers participated in *EE* PD workshops, led by the director of the Center for STEM Education for Girls at Harpeth Hall School. Beyond initial training, lead teachers were given ongoing guidance from the district's grant coordinator and lead instructional coach implementation and support for sustainability. STEM lead teachers also participated in STEM PLC's, where they collaborated virtually each week and met in person monthly for continuous networking and support.
STEM Conferences	To increase exposure to best practices in STEM education and to disseminate i3 results to a broader audience, one STEM lead teacher from each school and program staff attended at least one STEM conference annually, such as the National Science Teachers Association Annual STEM Forum, Center for STEM Education for Girls conferences, and Tennessee STEM Innovation Hub conferences.
PBL Training and Support	MNPS provided PBL training via the Buck Institute for Education in all MNPS high schools to support rigorous and relevant curriculum and student engagement. Beginning in the 2015-16 school year, MNPS extended PBL training from the Buck Institute into its middle schools. STEM lead teachers participated in this training, which further developed their PBL instructional skills and help build a culture for STEM and PBL instruction across the district.

STEM Industry Experiences. *G²ROW STEM* supplemented extended learning experiences with career-focused applied learning opportunities provided by college/university faculty and students as well as industry representatives from an extensive network of business partners. Through Saturday STEM sessions and STEM Summer Camps, mentoring experiences were also provided that played a critical role in helping students envision themselves in STEM careers and encourage pursuit of rigorous STEM-related academic courses.

Community STEM Awareness. In our target schools, families have limited awareness of STEM opportunities, which further impedes students' interest in STEM. To help shift this culture, *G²ROW STEM* staff held family STEM events for students' parents to inform them of program activities and events. Staff also collaborated with the high school academy coaches to engage parents and students through informational sessions about the various STEM academies offered at the feeder high schools. MNPS hosted recruitment sessions at the seven middle schools and their feeder elementary schools to raise STEM awareness and encourage students to apply for the program. The *G²ROW STEM* project director also provided updates about events and program successes at quarterly advisory meetings that a number of community partners regularly attended. The *G²ROW STEM* program hosted an annual STEM Showcase allowing students to present on STEM topics and build excitement for STEM throughout their school and community. In addition, the project director, lead instructional coach, and lead teachers participated in a National Science Foundation STEM Showcase. During this PD, the project directors, lead instructional coach, and lead teachers designed and created an informational video detailing the *G²ROW STEM* program components to share with people across the country.

1.2 Fidelity of Implementation

Logic Model. The *G²ROW STEM* logic model (Table 3) provided a sound theoretical foundation from which to conduct both process and outcome evaluations, as well as the final impact research study. The logic model shows the linkages between the core strategies, mediators, and outcomes that guide program implementation. It also served to focus continuous improvement efforts resulting from data monitoring and tracking progress towards short- and long-term outcomes. The logic model was used to assess fidelity between *G²ROW STEM* theory and implementation plans, as well as help ensure that program activities were planned with a sufficient level of frequency, intensity, and duration to produce the desired outcomes.

Table 3. Metropolitan Nashville Public Schools *G²ROW STEM* Logic Model

Core Strategies & Activities	Mediators	Short-term Outcomes	Impacts
Stem Student Curriculum ▪ Afterschool STEM Club using *EE* curriculum ▪ Saturday Sessions ▪ Summer Camp ▪ Mentoring experiences ▪ STEM Showcase	**Students** ▪ Increase middle school students' interest in and aspiration to STEM careers ▪ Increase student participation up to three years in program activities ▪ Increase freshmen enrollment in a STEM Career Academy ▪ Increase freshmen enrollment in STEM advanced courses	▪ Increase students' math and science content knowledge and applications ▪ Increase student academic motivation and achievement ▪ Decrease achievement gap of underrepresented students	**Improved Student Achievement** Significant improvement in standardized achievement in Math and Science End of Grade State Exams
Teacher Professional Development ▪ STEM PBL pedagogy ▪ *Engineering Everywhere* curriculum ▪ Monthly STEM PLC ▪ STEM conference presentations and attendance	**Teachers** ▪ Increase application of PBL pedagogy ▪ Increase application of STEM tools & research ▪ Increase collaboration among STEM teachers ▪ Increase value added teacher evaluation scores	▪ Increase students' math and science content knowledge and applications ▪ Increase student academic motivation and achievement ▪ Decrease achievement gap of underrepresented students	**Improved Student Achievement** Significant improvement in standardized achievement in Math and Science End of Grade State Exams

Fidelity Index. The *G²ROW STEM* fidelity indices were developed to regularly monitor the activities of the *G²ROW STEM* program, specifically the extent to which actual project implementation aligned with what was proposed. The degree to which the program implemented the intervention in accordance with the program model and theory directly affected the intended outcome. Specific details of which program outcomes were not met informed appropriate continuous improvement efforts.

One fidelity index was created for monitoring **student-centered activities** and consisted of five components aligned to the extended-learning strategies and industry experiences. Supporting data came from multiple sources (attendance records, mentoring experiences, and presentations at STEM showcase events). A fidelity score was calculated for each component, first based on student-level implementation, then based on school-level implementation. Thresholds were set for adequate implementation at both the school and program levels. These thresholds were established a priori using the project director and other STEM expert recommendations. Fidelity data were collected for three school years, SY 2016-17 through SY 2018-2019. In SY 2016-17, the first year of implementation, the fidelity score was 57.1% with only four of the seven *G²ROW STEM* schools meeting the threshold for adequate implementation. The full program failed to meet the overall fidelity threshold. In Year 2, five of the seven *G²ROW STEM* schools met the threshold for adequate implementation, resulting in a fidelity score of 71.4%. The program met the overall fidelity threshold. Finally, in Year 3, all seven schools met the threshold for adequate implementation resulting in a fidelity score of 100%.

The second fidelity index was created for monitoring **teacher-centered activities** and consisted of four components aligned to the teacher professional development and training. Supporting data came from multiple sources (attendance and training records, conference participation and presentations). A fidelity score was calculated for each component first based on teacher-level implementation, then based on program-level implementation. Thresholds were set for adequate implementation at both the school and program levels. These thresholds were established a priori using the project director and other STEM expert recommendations. Fidelity data were collected for three school years, SY 2016-17 through SY 2018-2019. In SY 2016-17, the first year of implementation, the fidelity score was 62.5%. The threshold was set at 70% so the program did not meet the criteria for adequate implementation. In Year 2, the fidelity score was 87.5, exceeding the threshold. In Year 3, the fidelity scores was 75% and did not meet the threshold of 80%. Table 4 summarizes the student- and teacher-centered fidelity scores for the three-year period.

Table 4. Metropolitan Nashville Public Schools *G²ROW STEM* Fidelity Index Scores

Fidelity Index	SY 2016-17	SY 2017-18	SY 2018-19
Student-Centered	**57.1%** **Did not meet** threshold of 71.4% for the program	**71.4%** **Met** threshold of 71.4% for the program	**100.0%** **Met** threshold of 71.4% for the program
Teacher-Centered	**62.5%** **Did not meet** threshold of 70.0% for the program	**87.5%** **Met** threshold of 80.0% for the program	**75.0%** **Did not meet** threshold of 80.0% for the program

2. IMPACT STUDY DESIGN

This study assessed the impact of the *G²ROW STEM* program (student extended learning experiences and teacher professional development) on student outcomes using a rigorous, longitudinal, quasi-experimental design (QED) with student-level, multi-cohort comparisons. Students in Cohort 1 received up to three years of exposure to the *G²ROW STEM* program. Those in Cohort 2 received up to two years of exposure. A fixed-effect linear regression model with schools as a blocking variable was used for the analyses. Within schools, students were assigned to the treatment or comparison groups. The model assessed the treatment and control groups for differences in achievement outcomes on standardized math and science academic achievement assessments. Tennessee Comprehensive Assessment Program TNReady math and science normal curve equivalents (NCEs) served as the continuous dependent variables to objectively measure program outcomes. This design supported reliable and valid assessment of the impact of *G²ROW STEM* interventions on middle school students. *G²ROW STEM* evaluation results also contribute to existing research on *Engineering Everywhere (EE)* and supplemental STEM extended learning strategies by evaluating their impact on high-need, middle school students.

2.1 Research Questions

Table 5 presents the confirmatory and exploratory research questions and related outcome variables for the *G²ROW STEM* impact study.

Table 5: *G²ROW STEM* Research Questions

Research Question	Type of Question	Outcome Variable
What is the impact of up to three years of exposure to *G²ROW STEM* on middle school students' mathematics skills compared to a business-as-usual condition?	Confirmatory	TNReady end-of-grade math NCE
What is the impact of up to three years of exposure to *G²ROW STEM* on middle school students' science skills compared to a business-as-usual condition?	Confirmatory	TNReady end-of-grade science NCE
What is the impact of *G²ROW STEM* on female (male) students' mathematics skills compared to female (male) students in the business-as-usual condition?	Exploratory	TNReady end-of-grade math NCE
What is the impact of *G²ROW STEM* on female (male) students' mathematics skills compared to female (male) students in the business-as-usual condition?	Exploratory	TNReady end-of-grade science NCE

2.2 Samples

The research design for the *G²ROW STEM* program used matched comparison groups to contrast outcomes between *G²ROW STEM* students and non-participants from within the same schools. *G²ROW STEM* **(treatment)** students participated in the STEM afterschool program, Saturday events, summer camps, and mentoring experiences while the **comparison** students were exposed to business-as-usual traditional academic programs during the school day in the same middle school environment. Comparison group students did not have access to the afterschool STEM extended learning enrichments, Saturday events, summer camp, or the mentoring experiences provided in the *G²ROW STEM* program. In addition, comparison group students had only limited access to teachers that received PD in the teaching of STEM activities.

All students attending the seven targeted middle schools were eligible to enroll in the *G²ROW STEM* program. Students completed an application to determine their commitment to the program. Each school had 30 slots available for the afterschool extended learning and 15 slots for both the Saturday events and summer camps. Table 6 details the demographic characteristics of the *G²ROW STEM* students and that of the full population across the seven schools.

There was a higher percentage of girls in the *G²ROW STEM* program than there was in the total population of students across the seven schools. This partially resulted from efforts taken by school staff and *G²ROW STEM* teachers to recruit females. There were also fewer Hispanic students and English language learners in the *G²ROW STEM* program than in the full target school population. Table 6 also shows that *G²ROW STEM* students were high-need in terms of academic content mastery in ELA and math.

Table 6: *G²ROW STEM* Student Characteristics

Demographic Category	Percent of Grow Stem Students	Percent of Target School Population
Gender--Female	57.4%	46.7%
Race/Ethnicity--		
African American	40.8%	38.7%
Hispanic	19.2%	30.8%
White	32.5%	26.3%
Asian	5.8%	3.9%

Demographic Category	Percent of Grow Stem Students	Percent of Target School Population
Economically Disadvantaged	43.7%	43.6%
English Language Learners	7.5%	16.3%
On Track or Mastery Level ELA	18.9%	26.1%
On Track or Mastery Level Math	18.9%	26.1%

The original implementation design planned for analyses over three cohorts and four years of consistent exposure to the *G²ROW STEM* program (see Table 7). Due to interruptions in program implementation resulting from school closures and remote learning requirements of COVID-19, as well as the lack of state TNReady science assessments due to the development of new standards and file testing, the actual implementation design had to be modified (see Table 8).

Table 7: *G²ROW STEM* Implementation Plan with Cohorts and Years of STEM Exposure

Study Year	Cohort 1	Cohort 1	Cohort 1	Cohort 1	Cohort 2	Cohort 2	Cohort 2	Cohort 2	Cohort 3	Cohort 3	Cohort 3	Cohort 3
Year 1 (2016-17) Initial Implementation	baseline 4th grade	baseline 5th grade	baseline 6th grade	baseline 7th grade								
Year 2 Implementation (2017-18)	new 5th grade 1 YR TRT	new 6th grade 1 YR TRT	new 7th grade 1 YR TRT	new 8th grade 1 YR TRT	baseline 4th grade	baseline 5th grade	baseline 6th grade	baseline 7th grade				
Year 3 Implementation (2018-19)	6th grade 2 YR TRT	7th grade 2 YR TRT	8th grade 2 YR TRT		new 5th graders 1 YR TRT	new 6th graders 1 YR TRT	new 7th graders 1 YR TRT	new 8th graders 1 YR TRT	baseline 4th grade	baseline 5th grade	baseline 6th grade	baseline 7th grade
Year 4 Implementation (2019-20)	7th grade 3 YR TRT	8th grade 3 YR TRT			6th grade 2 YR TRT	7th grade 2 YR TRT	8th grader 2 YR TRT		new 5th graders 1 YR TRT	new 6th graders 1 YR TRT	new 7th graders 1YR TRT	new 8th graders 1 YR TRT

Table 8: *G^{2}ROW STEM* Actual Implementation with Cohorts and Years of STEM Exposure

Study Year	Cohort 1	Cohort 1	Cohort 1	Cohort 1	Cohort 2	Cohort 2	Cohort 2	Cohort 2
Year 1 (2016-17) Initial Implementation	baseline 5th grade	baseline 6th grade	baseline 7th grade	baseline 8th grade				
Year 2 Implementation (2017-18)	6th grade 2 YR TRT	7th grade 2 YR TRT	8th grade 2 YR TRT		baseline 5th grade	baseline 6th grade	baseline 7th grade	baseline 8th grade
Year 3 Implementation (2018-19) Field test no TNReady Science	7th grade 3 YR TRT	8th grade 3 YR TRT			6th grade 2 YR TRT	7th grade 2 YR TRT	8th grade 2 YR TRT	
Year 4 Implementation (2019-20) No TNReady Math or Science, March school closures	8th grade 4 YR TRT				7th grade 3 YR TRT	8th grade 3 YR TRT		
Year 5 Implementation (2020-21): Full year remote programs					8th grade 4 YR TRT			

Teal cells are one-year math or science implementation
Gold cells are one-year math implementation
Gray cells are two-year math implementation

2.3 Baseline Equivalence

Prior to testing baseline equivalence, all *G²ROW STEM* students were matched to comparison students using a 1:1 nearest neighbor propensity score model for each impact analysis (both one-year and two-year impact analyses for math, one-year impact analysis for science). The propensity score models matched treatment and comparison students on baseline TNReady achievement test scores (math and science NCEs from the 2016-17 administration for Cohort 1 and the 2017-18 administration for Cohort 2), gender, minority, and English learner status.

Baseline Analytic Model. The linear model used for testing baseline equivalence is shown below.

$Y_i = \alpha + T_i\beta_1 + \varepsilon_i$
Where:
Y_i = the baseline measurement for student *i*
α = the intercept
$T_i\beta_1$ = the impact of the *G²ROW STEM* treatment condition (1 = treatment and 0 = comparison)
ε_i = a random error term for student *i*

Baseline Analytic Model Specifics. The baseline regression equations tested whether the treatment and comparison group students were equivalent at baseline on key variables, including test scores, gender, minority status, English learner status, grade, and cohort. Ordinary least squares regression was used for the continuous outcome of baseline test scores and the ordinal variable for grade level. Logistic regression was used for binary outcomes (gender, minority status, English learner status, and cohort). Baseline equivalence was met with all three impact analysis models. The testing revealed no significant differences between the treatment and comparison students. The characteristics of both the treatment and comparison samples at baseline for variables and covariates are presented in Tables 9-14. For the math one-year impact sample, effect sizes ranged from -0.03 to 0.10. For the math two-year impact samples, effect sizes ranged from -0.05 to 0.12. For the one-year science impact samples, effect sizes ranged from -0.02 to 0.05. In all cases, the standardized mean difference between treatment and comparison on baseline measures was less than 0.25. All variables were still included in the confirmatory analyses.

Table 9. Characteristics of *G2ROW STEM* Students at Baseline for the Math One-Year Impact Sample

Characteristic	Mean	N	Standard Deviation
Baseline Math NCE	43.46	542	19.04
Gender	0.58	542	0.49
Minority	0.69	542	0.46
English Learner	0.08	542	0.27
Grade	5.46	542	0.71
Cohort	1.36	542	0.48

Table 10. Characteristics of Comparison Students at Baseline for the Math One-Year Impact Sample

Characteristic	Mean	N	Standard Deviation
Baseline Math NCE	41.97	542	18.93
Gender	0.58	542	0.49
Minority	0.71	542	0.46
English Learner	0.08	542	0.28
Grade	5.45	542	0.71
Cohort	1.31	542	0.46

Table 11. Characteristics of *G2ROW STEM* Students at Baseline for the Math Two Year Impact Sample

Characteristic	Mean	N	Standard Deviation
Baseline Math NCE	41.99	249	17.54
Gender	0.57	249	0.50
Minority	0.64	249	0.48

Characteristic	Mean	N	Standard Deviation
English Learner	0.07	249	0.25
Grade	5.31	249	0.46

Table 12. Characteristics of Comparison Students at Baseline for the Math Two Year Impact Sample

Characteristic	Mean	N	Standard Deviation
Baseline Math NCE	39.83	249	17.50
Gender	0.59	249	0.49
Minority	0.67	249	0.47
English Learner	0.06	249	0.25
Grade	5.30	249	0.46

Table 13. Characteristics of *G2ROW STEM* Students at Baseline for the Science One-Year Impact Sample

Characteristic	Mean	N	Standard Deviation
Baseline Science NCE	41.90	330	19.04
Gender	0.58	330	0.49
Minority	0.69	330	0.46
English Learner	0.08	330	0.27
Grade	5.52	330	0.71

Table 14. Characteristics of Comparison Students at Baseline for the Science One-Year Impact Sample

Characteristic	Mean	N	Standard Deviation
Baseline Science NCE	42.34	330	19.30
Gender	0.56	330	0.50
Minority	0.69	330	0.46
English Learner	0.07	330	0.25
Grade	5.54	330	0.71

2.4 Data Elements

Table 15 summarizes the outcome data elements that were collected to answer confirmatory and exploratory research questions. Data was collected across four school years — 2016-17 through 2019-20. However, no statewide TNReady testing occurred in science in 2018-19 due to field testing for new science standards. Statewide testing did not occur in 2019-20 because of school closures due to COVID-19.

Table 15. *G²ROW STEM* Impact Study Outcome Measures

Impact Study	Measure	Type	Source
One-Year Math	TN Ready NCE scores for math	Continuous	District administrative records
Two-Year Math	TN Ready NCE scores for math	Continuous	District administrative records
One-Year Science	TN Ready NCE scores for science	Continuous	District administrative records

Table 16 describes the covariates included in each of the confirmatory analyses. A blocking variable for schools was used in the impact analysis models to account for differences in school characteristics. All seven schools were dummy coded with Croft Middle School serving as the reference group.

Table 16. *G²ROW STEM* Impact Study Covariates and Blocking Variable

Variable	Description	Type of Measure	Data Source
Baseline TNReady NCEs for Math and Science	One Year Math: TN Ready math NCE in spring 2017 (Cohort 1) and spring 2018 (Cohort 2) One Year Science: TN Ready science NCE in spring 2017	Continuous	District administrative records
Treatment	Identified whether a student was a *G²ROW STEM* student or a comparison student	Binary 0 = Comparison 1 = *G²ROW STEM*	*G²ROW STEM* program records
Gender	Identified the student's gender	Binary 0 = male 1 = female	District administrative records
Minority Status	Identified whether a student was a racial/ethnic minority	Binary 0 = white 1 = non-white	District administrative records
English Learner Status	Identified whether the student was an English learner	Binary 0 = not an EL 1 = EL	District administrative records
Grade	Identified the grade of the student at baseline	Discrete, ordinal	District administrative records
Cohort	Identified the cohort of the student	Discrete, nominal 1 = Cohort 1 2 = Cohort 2	*G²ROW STEM* program records
School	Identified the school of attendance for the student	Discrete, nominal	District administrative records

3. ANALYSIS AND RESULTS

3.1 Confirmatory Analytic Model

The linear model used for conducting the analyses for the confirmatory research questions is shown below.

$$Y_i = \alpha + \text{Baseline}_i\beta_1 + \text{Treatment}_i\beta_2 + \text{Gender}_i\beta_3 + \text{MinorityStatus}_i\beta_4 + \text{EnglishLearnerStatus}_i\beta_5 + \text{Grade}_i\beta_6 + \Sigma\text{School}_i\beta_7 + \text{Cohort}_i\beta_8 + \varepsilon_i$$

Where:

Y_i = the outcome for student i

α = the intercept

$\text{Baseline}_i\beta_1$ = parameter estimate for the effect of the student baseline score

$\text{Treatment}_i\beta_2$ = covariate adjusted difference in the mean student outcome for treatment group students minus the mean student outcome for comparison group students (1 = treatment and 0 = comparison)

$\text{Gender}_i\beta_3$ = effect of student gender (1 = female and 0 = male)

$\text{MinorityStatus}_i\beta_4$ = effect of student minority status (1 = minority and 0 = not a minority)

$\text{EnglishLearnerStatus}_i\beta_5$ = effect of whether or not a student is an English learner (1 = English learner and 0 = not an English learner)

$\text{Grade}_i\beta_6$ = effect of student grade level

$\Sigma\text{School}_i\beta_7$ = blocking variable for student's school

$\text{Cohort}_i\beta_8$ = effect of student cohort (students are either in cohort 1 or cohort 2)

ε_i = a random error term for student i

Confirmatory Analytic Model Specifics. Following *G²ROW STEM* implementation for five years, a one-year math impact, a one-year science impact, and a two-year math impact were calculated by comparing *G²ROW STEM* students' mean achievement scores to comparison students' mean achievement scores, controlling for baseline characteristics. It was not possible to assess more years of impact for math or science due to the lack of TNReady outcome data.

The unit of assignment and the unit of analysis were both student-level. Assignment to treatment and comparison groups occurred within each of the seven middle schools in the sample. An intent to treat approach was used in the analysis so that students assigned to the treatment group remained in the treatment group even if they dropped out of the *G²ROW STEM* program. There was no imputation of outcome or baseline data. Students who had missing baseline or outcome data were not included in the analyses. OLS fixed effect regression equations were used to analyze the effect of one or more years of *G²ROW STEM* participation on math and science achievement test outcomes.

3.2 Exploratory Analytic Model

The linear model used for conducting the analyses for the exploratory research questions is shown below.

$$\mathrm{Y_i} = \alpha + \mathrm{Baseline_i}\beta_1 + \mathrm{Treatment_i}\beta_2 + \mathrm{Gender_i}\beta_3 + \mathrm{MinorityStatus_i}\beta_4 + \mathrm{EnglishLearnerStatus_i}\beta_5 + \mathrm{Grade_i}\beta_6 + \Sigma\mathrm{School_i}\beta_7 + \mathrm{Cohort_i}\beta_8 + Treatment * \mathrm{Gender_i}\beta_9 + \varepsilon_i$$

Where:
Y_i = the outcome for student i
α =the intercept
$\mathrm{Baseline_i}\beta_1$= parameter estimate for the effect of the student baseline score
$\mathrm{Treatment_i}\beta_2$ = covariate adjusted difference in the mean student outcome for treatment group students minus the mean student outcome for comparison group students (1 = treatment and 0 = comparison)
$\mathrm{Gender_i}\beta_3$ = effect of student gender (1 = female and 0 = male)
$\mathrm{MinorityStatus_i}\beta_4$= effect of student minority status (1 = minority and 0 = not a minority)
$\mathrm{EnglishLearnerStatus_i}\beta_5$ = effect of whether or not a student is an English learner (1 = English learner and 0 = not an English learner)
$\mathrm{Grade_i}\beta_6$ = effect of student grade level
$\Sigma\mathrm{School_i}\beta_7$ = blocking variable for student's school
$\mathrm{Cohort_i}\beta_8$ = effect of student cohort (students are either in cohort 1 or cohort 2)
$\mathrm{Treatment} * \mathrm{Gender_i}\beta_9$ = interaction term to estimate program effects by gender
ε_i = a random error term for student i

Exploratory Analytic Model Specifics. Exploratory analytic models were analyzed for one-year math, one-year science, and two-year math outcomes to assess the differential impact of *G²ROW STEM* on female and male students. These exploratory models provide insight into the ability of the *G²ROW STEM* program to effectively engage female students in STEM instruction and activities.

3.3 Results for Confirmatory One-Year Math Analysis

Results indicated a statistically significant difference between the *G²ROW STEM* treatment group and the business-as-usual comparison group on the math one-year outcome. The average TNReady score at the 2017-18 administration for the treatment group was 45.47 NCEs, while the average for the comparison group was 35.87 NCEs. This difference is 9.60 NCEs which equates to roughly one-half of a standard deviation (SD). In addition to significant differences in outcome TNReady scores between treatment and comparison students, significant differences were found for student minority status. While whites had higher NCE scores than minorities in 2017-18 (a difference of 15.88 NCEs, nearly a full SD), the difference between treatment and comparison students was greatest for minorities (10.17 NCEs v. 7.49 NCEs for whites). This suggests that minority students benefited more from the *G²ROW STEM* program in terms of achievement scores in math. Grade level and gender also showed significant differences between baseline and outcome TNReady math assessment scores. As grade level increased the outcome score decreased. For gender, the difference in outcome TNReady math scores for girls and boys by treatment and comparison group are further detailed in Section 3.6: Results for Exploratory Analyses.

There were significant effects for some of the school blocks. This reflects a significant mean difference in 2017-18 TNReady math scores between that particular school and the reference school, Croft Middle, after controlling for the other covariates. This information does not indicate significant differences between baseline TNReady math scores and outcome 2017-18 TNReady math scores. What is shown from the school estimates is that some schools have lower mean scores than Croft Middle (Goodlettsville) and others have higher mean scores (Oliver, Madison).

In terms of the significant effects for the baseline TNReady math score, this reflects the fact that the best predictor of future academic performance is prior performance and is an expected result. Table 17 includes the regression model output. An R squared of 0.70 was found for the overall model.

Table 17. *G²ROW STEM* One Year Math Analysis

Variable	Estimate	Standard Error	*t*-value	*p*-value
Intercept	6.82	3.61	1.89	0.059
Math Baseline	0.81	0.02	37.02	< 0.001
Treatment	8.33	0.71	11.65	< 0.001
Gender	1.81	0.73	2.50	0.013
Minority Status	-4.06	0.86	-4.73	< 0.001
EL Status	-0.69	1.42	-0.48	0.628
Grade	-1.18	0.52	-2.26	0.024
Cohort	1.22	0.78	1.56	0.120
School (Goodlettsville)	-3.10	1.34	-2.32	0.021
School (Litton)	3.70	1.39	2.65	0.008
School (Madison)	4.74	1.49	3.17	0.002
School (McMurray)	3.31	1.37	2.41	0.016

Variable	Estimate	Standard Error	*t*-value	*p*-value
School (Oliver)	4.05	1.16	3.51	< 0.001
School (Stratford)	-2.58	1.50	-1.72	0.086

3.4 Results for Confirmatory Two-Year Math Analysis

Results indicated a statistically significant difference between the *G²ROW STEM* treatment group and the business-as-usual comparison group on the math two-year outcome as well. The average TNReady score at the 2018-19 administration for the treatment group was 43.90 NCEs, while the average for the comparison group was 39.24 NCEs. This difference is 4.66 NCEs which equates to roughly one-quarter of a SD. In addition to significant differences in outcome TNReady scores between treatment and comparison students, significant differences were found for student minority status. While whites had higher NCE scores than minorities in 2018-19 (a difference of 12.08 NCEs, over one-half of a SD), the difference between treatment and comparison students was greatest for whites (5.01 NCEs v. 4.01 NCEs for minorities). This suggests that white students benefited slightly from the *G²ROW STEM* program in terms of achievement scores in math when tracked over two years of achievement test data.

There were significant effects for some of the school blocks. This reflects a significant mean difference in 2018-19 TNReady math scores between that particular school and the reference school, Croft Middle, after controlling for the other covariates. This information does not indicate significant differences between baseline TNReady scores and outcome 2018-19 TNReady scores. What is shown from the school estimates is that some schools had higher mean scores over the two years when compared to Croft Middle's 2018-19 TNReady math scores (Madison, McMurray, and Oliver).

In terms of the significant effects for the baseline TNReady math score, this reflects the fact that the best predictor of future academic performance is prior performance and is an expected result. Table 18 includes the regression model output. An R squared of 0.54 was found for the overall model.

Table 18. *G²ROW STEM* Two Year Math Analysis

Variable	Estimate	Standard Error	*t*-value	*p*-value
Intercept	16.29	7.06	2.31	0.022
Math Baseline	0.69	0.04	18.42	< 0.001
Treatment	3.31	1.15	2.88	0.004
Gender	-0.53	1.16	-0.46	0.647
Minority Status	-1.17	1.34	-3.10	0.002
EL Status	-1.03	2.60	-0.40	0.698
Grade	-0.89	1.26	-0.71	0.481
School (Goodlettsville)	0.84	2.15	0.39	0.698
School (Litton)	2.21	2.19	1.01	0.312
School (Madison)	7.91	2.40	3.29	0.001
School (McMurray)	5.62	2.20	2.55	0.011
School (Oliver)	5.03	1.85	2.72	0.007
School (Stratford)	-0.95	2.49	-0.38	0.703

3.5 Results for Confirmatory One-Year Science Analysis

As with the other two impact models, significant differences in outcome TNReady science scores between treatment and comparison students were found. The average TNReady science score at the 2017-18 administration for the treatment group was 44.94 NCEs, while the average for the comparison group was 38.10 NCEs. This difference is 6.84 NCEs which equates to roughly one-third of a SD. In addition to significant differences in outcome TNReady scores between treatment and comparison students, significant differences were found for student minority status. While whites had higher NCE scores than minorities in 2017-18 (a difference of 20.53 NCEs, a full SD), the difference between treatment and comparison students was greatest for minorities (7.83 NCEs v. 4.64 NCEs for whites). This suggests that minority students benefited more from the *G²ROW STEM* program in terms of achievement scores in science.

There were significant effects for some of the school blocks. This reflects a significant mean difference in 2017-18 TNReady science scores between that particular school and the reference school, Croft Middle, after controlling for the other covariates. This information does not indicate significant differences between baseline TNReady science scores and outcome 2017-18 TNReady science scores. What is shown from the school estimates is that some schools have lower mean scores than Croft Middle (Stratford) and others have higher mean scores (Oliver, McMurray).

In terms of the significant effects for the baseline TNReady math score, this reflects the fact that the best predictor of future academic performance is prior performance and is an expected result. Table 19 includes the regression model output. An R squared of 0.76 was found for the overall model.

Table 19. *G²ROW STEM* One Year Science Analysis

Variable	Estimate	Standard Error	*t*-value	*p*-value
Intercept	4.82	4.08	1.18	0.238
Math Baseline	0.85	0.03	31.84	< 0.001
Treatment	6.85	0.86	8.00	< 0.001
Gender	1.23	0.86	1.44	0.152
Minority Status	-3.46	1.06	-3.27	0.001
EL Status	1.46	1.81	0.81	0.420
Grade	-0.64	0.62	-1.02	0.307
School (Goodlettsville)	1.31	1.60	0.82	0.412
School (Litton)	-0.51	1.59	0.32	0.747
School (Madison)	0.32	1.78	0.18	0.855
School (McMurray)	4.97	1.66	3.00	0.003
School (Oliver)	10.37	1.45	7.17	< 0.001
School (Stratford)	-3.54	1.78	-1.99	0.048

Effect sizes were calculated for mean differences between treatment and comparison students on TNReady math and science outcome data. The results of these analyses are shown in Table 20.

Table 20. Effect Size Calculations: Hedges *g*

Statistic	One-Year Math	Two-Year Math	One-Year Science
Treatment Adjusted Mean	44.2	42.6	45.0
Comparison Mean	35.9	39.2	38.1
Treatment SD	20.3	19.1	20.7
Comparison SD	21.1	17.5	22.2
N for Matched Group	542	249	330
Hedges *g*	.40	.19	.32

The greatest mean differences were found in the one-year math analysis, followed by the one-year science. Hedges *g* was used as a measure of the effect size or difference between the two groups in terms of standard deviation units. The thresholds to be used for interpreting Hedges *g* are that a value of 0.2 indicates a small effect, one that may not be seen by the naked eye. A medium effect is indicated with a *g* of 0.50 and a large effect with a *g* of 0.8 (Lovakov and Agadullina, 2017). Both the one-year math and one-year science analyses approach medium effect sizes. These results suggest that the students who participated in *G²ROW STEM* had greater gains on math and science achievement test scores when compared to students who did not participate in the program. The gains approached a half of a standard deviation difference or roughly 10 NCE points.

3.6 Results for Exploratory Analyses

Results indicated no statistically significant effect for the interaction of *G²ROW STEM* participation with gender in any of the three exploratory models (one-year math, two-year math, one-year science). However, a marginal effect was found for the one-year science exploratory model (p=.069). Table 16 details the results of the analyses.

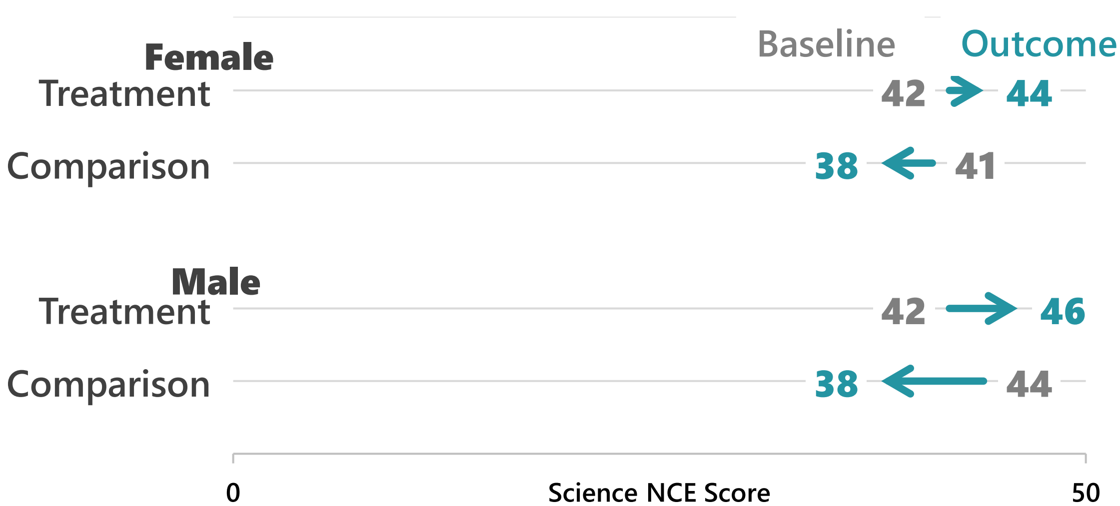
Table 16. *G²ROW STEM* Treatment x Gender Analysis

Variable	Estimate	Standard Error	*t*-value	*p*-value
One-Year Math Treatment x Gender	-0.12	1.44	-0.08	0.934
Two-Year Math Treatment x Gender	0.59	2.30	0.26	0.797
One-Year Science Treatment x Gender	-3.10	1.70	-1.82	0.069

Figure 4 clarifies the interaction found between treatment status and gender for science TNReady outcomes. Average outcome NCE scores increased more from baseline for males in *G²ROW STEM*. For this group there was a difference of 3.92 NCE points between baseline and outcome measures. For comparison male students, science achievement scores declined an average of 6.17 NCE's from baseline to outcome measure. Female students evidenced a similar pattern but to a lesser degree with gains from baseline to outcome for *G²ROW STEM* students of 2.39 NCE's and declines from baseline to outcome for comparison student of 2.74 NCE's. In summary, it appears that students, especially males, benefit from the program in terms of higher achievement test scores. Students that do not have *G²ROW STEM* program experience see a decline in science achievement test scores on the average.

Figure 4.

Differences in one-year science outcomes by gender approached significance, with scores for G²ROW STEM students increasing and scores for comparison students decreasing.



4. DISCUSSION

Through *G²ROW STEM*, MNPS extended the success of the career academy concept to the middle schools and developed STEM enhanced learning and career-focused applied experiences for students. Through STEM extended learning, applied PBL instruction, and mentoring experiences, *G²ROW STEM* students received up to five years of enhanced instruction and real-world experiences. In this impact study, these *G²ROW STEM* students were compared to students receiving business-as-usual instruction from within the same schools.

Fixed effect linear regression models with school blocks were analyzed to determine the effect of *G²ROW STEM* experiences on TNReady math and science outcome scores. Three different models were analyzed: one-year math, two-year math, and one-year science. The results of these analyses found significant differences between the treatment and comparison groups in all three models. These findings suggest that *G²ROW STEM* improved academic achievement in math and science for students in the program (treatment students). However, the strongest results were for the one-year math analysis, followed by the one-year science analysis. This may be due to the fact that tracking individual student participation over multiple years and using an index of saturation/intensity and consistency was not strictly applied. The two-year math impact analysis followed students who had been participants in either of two years (2016-17 and 2017-18). Future analyses should follow sub-groups of students who participated in the program consistently over the entire two-year period and those who were involved in multiple extended learning experiences.

Results for gender differences in *G²ROW STEM* students' achievement gains were not significant in math and only marginal in science. Results for science indicated that boys in the *G²ROW STEM* program had larger gains than girls. Boys in the comparison group, not receiving *G²ROW STEM,* had larger declines in achievement test scores than comparison group girls. The interaction of gender and treatment effects should be investigated further with multiple years of achievement test data.

There were significant results for minority status and treatment effects. However, for both the one-year analysis in math and science, minority students had greater gains in TNReady test scores from baseline to outcome measures. For the two-year math analysis, the white subgroup had the largest gain from baseline to outcome measures. More in-depth analyses of minority status interactions with *G²ROW STEM* program participation should be undertaken to tease apart the contributing factors underlying these results.

While there were differences in mean outcome TNReady scores between schools across all three analyses, they only consistently identified McMurray and Oliver as having higher mean achievement test scores than the reference school (Croft Middle). Further research should focus on examining gains from baseline to outcome for the schools and relating these results directly to implementation fidelity information for both teacher and student activities. This information could identify specific instructional difference or applied learning experiences that contributed to greater achievement gains. Continuous improvement efforts could focus on applying these successful strategies to other schools.

Further exploratory analyses should also be conducted to examine differences in achievement for students in the *G²ROW STEM* program and those students in the comparison groups for demographic subgroups (English language learners, and students who are economically disadvantaged).

REFERENCES

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

Bolshakova, V., Johnson, C., & Czerniak, C. (2011). It Depends on What Science Teacher You Got: Urban Science Self Efficacy from Teacher and Student Voices. *Cultural Studies of Science Education*, 6(4), 961-97. Retrieved from http://link.springer.com/article/10.1007%2Fs11422-011-9346-2

Bridgeland, J. M., Balfanz, R., Moore, L. A., & Friant, R. S. (2010). Raising their voices: Engaging students, teachers, and parents to help end the high school dropout epidemic. Retrieved from Civic Enterprises website: www.civicenterprises.net/MediaLibrary/Docs/ raising their voices.pdf

Charette, R. N. (2016). The STEM anxiety business, *Computer*, 49 (3), pp. 82-87.

Cunningham, C. M. (2009). Engineering is elementary. *The bridge*, 11-17.

Engineering is Elementary. (2020). *Engineering Everywhere*. Boston, MA: Museum of Science.

Gallagher, S., & Gallagher, J. (2013). Using Problem-based Learning to Explore Unseen Academic Potential. *Interdisciplinary Journal of Problem-Based Learning*, 7(1), 111-131. Retrieved from http://dx.doi.org/10.7771/1541-5015.1322

Hanover Research. (2011). K-12 STEM education overview. Retrieved from www.hanoverresearch.com/wp-content/uploads/2011/12/K-12-STEM-Education-Overview-Membership.pdf

Henderson, C., & Dancy, M. (2011). Increasing the impact and diffusion of STEM education innovations. Retrieved from www.nae.edu/File.aspx?id=36304

Howard-Brown, B., & Martinez, D. (2012). *Engaging Diverse Learners through the Provision of STEM Education Opportunities*. SEDL. Retrieved from http://secc.sedl.org/resources/briefs/diverse_learners_STEM/

Institute of Education Sciences. (2020). *What Works Clearinghouse*. Retrieved from National Center for Education Evaluation and Regional Assistance: https://ies.ed.gov/ncee/wwc/

Krishnamurthi, A., Bevan, B., Coulon, V.R., & Rinehart, J. (2013). *Defining Youth Outcomes for STEM Learning in Afterschool*. Afterschool Alliance. Retrieved from http://www.afterschoolalliance.org/STEM_Outcomes_2013.pdf

Liu, Y., Lou, S., & Shih, R. (2014). The Investigation of STEM Self-Efficacy and Professional Commitment to Engineering Among Female High School Students. *South African Journal of Education*, 34(2). Retrieved from http://www.ajol.info/index.php/saje/article/view/105536/95558

Museum of Science. (2015). *Engineering Everywhere: Final Report*. Boston, MA: Museum of Science.

Roth, W., & Van Eijck, M. (2010). *Fullness of life as minimal nit: Science, technology, engineering, and mathematics (STEM) learning across the life span*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/sce.20401/abstract

Sanders, M. (2009). STEM, STEM education, STEMmania. *Technology Teacher*, 68, 20-26. Retrieved from http://esdstem.pbworks.com/f/TTT%2BSTEM%2BArticle_1.pdf

Strobel, J., & van Barneveld, A. (2009). When is PBL more effective? A meta-synthesis of metaanalyses 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

Tennessee Department of Education (TN ED). (2020). State Report Card. Retrieved from https://www.tn.gov/education/topic/report-card

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