

Strong Promise: Moving from Interest and Engagement in STEM to Real Academic Results

Evaluating Girlstart After School Participants' Academic Performance

Summary

Girlstart's mission is to increase girls' interest and engagement in STEM through innovative, nationally-recognized informal STEM education programs. Girlstart examines in this report how STEM education directed toward elementary school girls influences long-term readiness and participation to math and science learning. This report compares Girlstart After School participants' academic performance to nonparticipant performance. Specifically, it examines how Girlstart After School influences science STAAR performance and course enrollment in subsequent elementary and early middle school grades. This quasi-experimental study uses a posttest intact group design to examine the STAAR science test performance of Girlstart After School participants in Central Texas. This study also evaluates girls' subsequent enrollment in 6th and 7th grade math and science. Eight independent school districts in Texas are analyzed using OLS regression and propensity-score matching techniques. This analysis enables Girlstart to identify the effect of program participation on test performance both immediately after program attendance and over time. In addition, qualitative findings from student and teacher surveys from 2010-2016 complement study results. The core findings of the study are that **Girlstart girls are more likely to achieve high-level STAAR science performance and enter advanced science and math courses over time than nonparticipant girls**. The qualitative results show that Girlstart girls have high rates of confidence in both STEM performance and knowledge of STEM careers. *We believe that this data demonstrates that Girlstart After School provides girls with the support and knowledge necessary to explore STEM topics at advanced levels of access and understanding.*

Sarah Harris, M.A., Evaluator

Special Thanks

We would like to thank A.J. Angus, Christina Caramanis and Mike Findley, who offered invaluable statistical support and feedback. We would also like to thank Cinda Christian, Reetu Naik, Chris Howard, Irma Guerra-Scott, Dan Cox, Vicki Barbosa, Joy Philpott, Garry Hardcastle, Dorian Jones, David Greiner, Crystal Hernandez, Cathy Malerba, and Megan Olshavsky, regional school district assessment staff, and Austin, Connally, Del Valle, Georgetown, Hays, North East, Pflugerville, and Round Rock ISDs.

Introduction

Girlstart, an Austin, Texas based science, technology, engineering, and math (STEM) education nonprofit organization, was founded in 1997. In 2010, Girlstart began to dramatically expand one of its core programs, Girlstart After School, to reach more girls and schools. In this report, we compare Girlstart After School participants' academic performance and subsequent advanced science and math course enrollment in math and science with nonparticipants in 8 school districts with the goal of understanding program impacts.

Girlstart's primary goal is to inspire girls' interest and aptitude in STEM as well as their confidence in STEM pursuits. As a partner to schools and districts generally, Girlstart is also committed to providing services that address the most immediate and pressing needs of partner schools and the students in the program. Girlstart understands that for the majority of partner schools, increasing students' passing rates in state-mandated tests is a very important goal. For example, Austin ISD's 2015-2020 Strategic Plan aims to support all students to be college, career, and life-ready. To measure success toward this goal,



Austin ISD will measure both the percentage of students achieving STAAR passing standards for math and the percentage of students in grades 3-8 participating in problem-based learning activities.ⁱ If Girlstart After School programs can help schools achieve higher passing rates, in addition to increasing girls' knowledge of STEM careers, these programs are achieving a significant impact for participants and partner schools and districts. For this reason, Girlstart expanded partnerships with multiple school districts in 2015 to collect and analyze data related to participants' performance on state-mandated tests and enrollment in advanced courses in math and science for grades 6-7

As an organization, Girlstart strives to expand its understanding of program impacts by increasing quantitative analysis of metrics that help to assess short and medium-term learning outcomes of program participation. Specifically, Girlstart aims to evaluate participants' educational achievement and attainment. To better facilitate this, Girlstart increased partnerships for data sharing from 2 districts in 2014 (Georgetown and Austin ISD) to 8 districts in 2015. Data sharing agreements were established with the following Texas school districts: Austin, Del Valle, and North East ISDs. In addition, Girlstart is currently developing data partnerships with Georgetown, Connally, and Hays ISDs.

Measuring educational achievement presents many challenges. For example, models that evaluate standardized assessment tests are vulnerable to numerous analytical sensitivities and biases such as scale interpretation. Recent research approaches to overcoming these biases include assessing non-arbitrary measures like adult earnings or educational attainment.ⁱⁱ Therefore in this study, Girlstart aims to measure programmatic impacts not only in achievement –as seen in STAAR passing rates in science – but also in educational attainment, as evaluated through advanced course enrollment of participants over time. Girlstart's study is guided by the following two research questions. 1) *Does Girlstart help girls perform better on standardized science tests?* 2) *Does Girlstart program participation increase the number of girls taking advanced math and science courses in subsequent years?*

About Girlstart After School

Girlstart is a weekly afterschool program in which clubs of 20+ girls meet for an hour of informal instruction in a variety of STEM topics. Each club is led by a member of Girlstart's STEM CREW (Creative, Resourceful, Empowered Women). The STEM CREW, college students majoring in STEM or education, receive classroom experience and training in STEM teaching while they work as Girlstart program leaders. Girlstart minimizes the influences of instructional variation through standardized training and through clear instructor and programmatic expectations. Each week, at each partner school, girls engage in hands-on problem-solving challenges. Participants collect and record data and discuss their findings with the group. In addition, at each session girls are introduced to STEM careers related to the activity they complete. Girlstart After School curriculums are aligned with the TEKS (Texas Essential Knowledge & Skills) learning standards as well as Next Generation Science Standards.

Girlstart retains a campus coordinator, a teacher or staff member whose task is to recruit girls to participate in the program, at each partner school. Although each campus coordinator conducts recruitment based on his or her understanding of the needs of the school and its students, campus coordinators receive the following general recruitment guidelines from Girlstart. First, Girlstart After School is not a tutoring program; girls who perform far below basic levels for their grade are not ideal candidates. In addition, Girlstart After School is not a program for only high-performing girls or girls who already express a strong interest or aptitude in STEM. Girlstart participants are girls who might value additional time in STEM subjects and/or a supportive environment in the afterschool space. Girlstart's program model is based on the belief that because all girls are underserved in STEM, all girls can benefit



from an engaging and challenging STEM afterschool program. One of the goals of Girlstart After School is to keep these promising students, particularly those who lack STEM learning resources outside of school, engaged in STEM learning, so that they may continue to do well in science and enter advanced science tracks in middle school.

Program Reach and Demographics

Girlstart After School currently reaches more than 1,400 girls through 55 weekly programs in 18 Texas districts. In 2014-15, Girlstart After School reached a total of 1,343 girls. Girlstart After School participants are girls in 4th through 6th grade, with 4th and 5th graders representing the largest number of participants. Girlstart purposefully seeks partnerships with schools that serve a high proportion of economically disadvantaged students, non-white students, and/or students classified as “at-risk” by the Texas Education Agency. Significant research shows that groups such as African-American and Hispanic students are less likely to have access to advanced science and math courses in high school, which negatively affects their ability to enter and complete STEM majors in college.ⁱⁱⁱ Among the 2014-15 Girlstart participants, 58% were Latina, 16% were Caucasian, 13% were African-American, 4% were Asian-American, and 9% identified as Multiethnic. Seventy percent of 2014-15 participants were considered economically disadvantaged (recipients of free/reduced-price lunch at school). It should be noted that Girlstart does not inquire about eligibility for these reduced lunch programs, so the percentage of economically disadvantaged participants may actually be higher.

Evaluation and Research

To monitor program outputs and evaluate progress toward educational and empowerment-based outcomes, Girlstart uses a mixed-methods approach that combines annual survey data with periodic quasi-experimental research. In this study, Girlstart focuses on the impact of Girlstart After School programs on academic performance in science. Girlstart After School curriculum predominantly focuses on science topics and is oriented toward learning outcomes in that area. However, this study also collects data on STAAR math performance in order to observe potential program effects in math knowledge for Girlstart participants. Therefore, primary data is collected on the 4th and 5th grade STAAR test in both math and science. STAAR, the State of Texas Assessments of Academic Readiness, replaced the prior TAKS (Texas Assessment of Knowledge and Skills) test in 2011-12, and is mandatory for all students at public schools in the state of Texas. The science STAAR test is administered in 5th and 8th grade; the math STAAR is administered each year starting in 4th grade. Fifth grade exams are important to individual students because it is the first time that students are tested in science (they take science tests again in 8th grade and high school).

Most importantly, student scores on 5th grade standardized tests play a crucial role in determining whether the students will take advanced or Pre-AP math and science courses in middle school. Students who enroll in advanced math and science classes in middle school are able to take advanced courses in high school, thereby increasing their readiness for STEM subjects at college. High performance on 5th grade math and science tests directly impact a student’s likelihood of accessing Pre-AP classes at middle school, while low scores discourage students from attempting to select into these courses. In 2015, the Greater Texas Foundation and E3 Alliance found that math and science readiness is key to the success of long term educational outcomes and recommends establishing a mastery of math in elementary school in order to increase the number and diversity of 8th grade students completing Algebra I, thereby helping to keep STEM career pathways, high wage and high demand jobs, open to more students.^{iv}



Girlstart annual pre- and post program surveys measure progress toward goals through participant reflection and instructor feedback; goals evaluated consider both competency in STEM knowledge as well as girls' confidence and interest in STEM subjects and careers. Girlstart After School surveys were developed by SEDL (formerly the Southwestern Educational Development Laboratory, now part of the American Institutes of Research) in 2011 and 2012. Girlstart's system of surveys, designed by SEDL, assess girls' skills, knowledge, interest, and confidence in STEM throughout their time in Girlstart After School. Complementing the National Science Foundation's suggestions for informal science education, the surveys measure program impact on awareness of, interest in, and attitude toward STEM concepts, processes or careers.⁹ The surveys provide an opportunity for both instructor feedback and for participants to self-report their perceptions toward Girlstart After School and toward STEM in general. Girlstart program leaders administer these surveys to the girls at the beginning and end of each semester. The results of this robust internal survey process consistently show that Girlstart After School achieves positive results for these indicators. Figure 1 shows stable, high-level percentages of both teacher reported and participants' self-reported data regarding personal development outcomes from 2010-16 Girlstart end-of-year surveys.

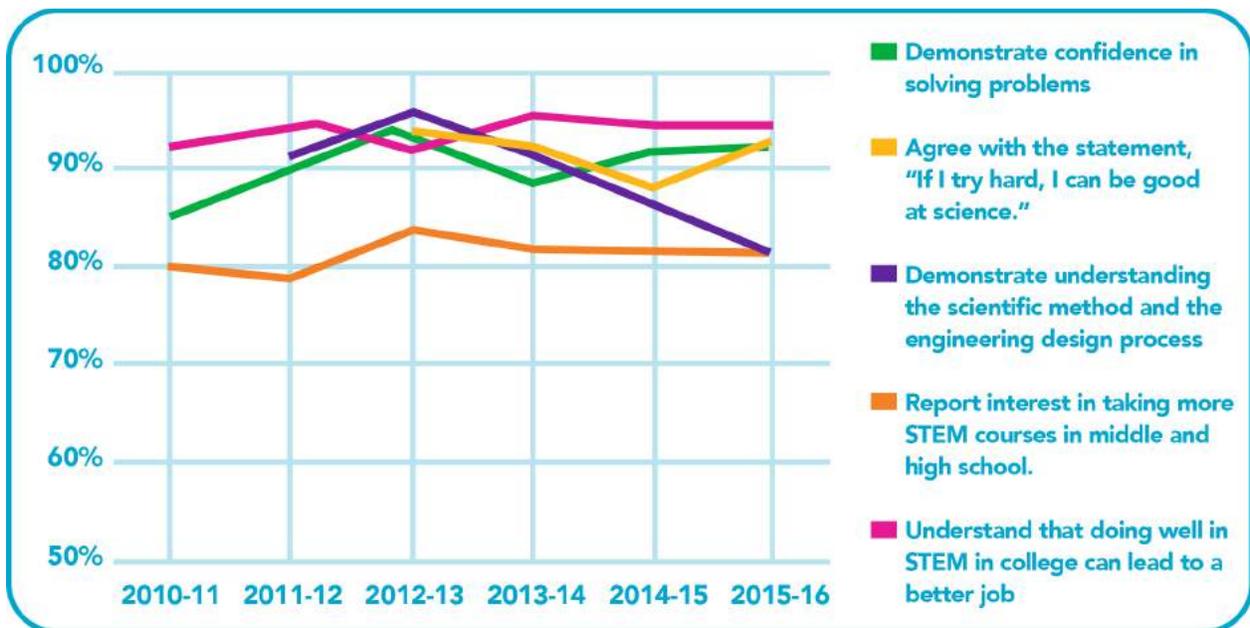


Figure 1 Selected Results from Girlstart Self-Reported Participant Surveys.

Overall, since 2010, participants report increasingly positive results in personal and career related metrics. While there is a slight decrease for specific STEM knowledge in scientific and engineering processes, more than 80% of participants progressed in this area. This result reflects survey adjustments made in 2012 toward reinforced positive self-perception, and resilience after failure. Similar metrics for nonparticipants would need to be collected in order to comparatively evaluate Girlstart After School's impact in these areas.

In this study, Girlstart collected STAAR test scores, STEM subject grades and course enrollment information to measure academic performance and progress toward short- and medium- term educational outcomes that relate to demonstrated knowledge in STEM subjects. As noted above, Girlstart's evaluation into academic performance is guided by two main research questions: 1) Does

Girlstart help girls perform better on standardized science tests? 2) Does program participation increase the number of girls taking advanced math and science courses in subsequent years? In order to collect data to help answer these questions, Girlstart designed a quasi-experimental study using a posttest only intact group design.

Girlstart After School is provided at the school level and is based on 3 criteria: collaborative leadership at the district and school level, Title 1 status, and availability of funding. Not all elementary schools in a given district receive Girlstart After School. If Girlstart After School has operated a year or less in a school, the school is not included in this study. Girlstart believes it takes at least a year to establish a program at a new school.

For research, treatment is at the school level and based on individual yearlong attendance in the Girlstart After School program. To ensure complete treatment, participants in this study attended no less than 50% of both fall and spring semester programming. Girlstart believes this level of commitment by students is necessary to achieve program learning objectives. The Greater Texas Foundation and E3 Alliance mirror this reasoning when researching mathematics course “mastery” rather than “completion” in order to capture more accurate levels of a student’s knowledge in a given subject.^{vi} The E3 Alliance study refers to “mastery” as a students’ passing of both semesters of a course rather than their overall average for the year. In future research, Girlstart will collect data on all participants and compare program impacts on academic performance for girls that attend less than half of each semester in the yearlong program.

To evaluate the impact of Girlstart After School programs on educational achievement, the primary measurement of treatment is the 5th grade STAAR test in science. Participants do not take a pretest for the STAAR math or science test, though pre-program subject grades for both Girlstart girls and nonparticipants were available in Austin, Georgetown, and Round Rock ISD. A difference in proportion OLS regression compares program means within districts and evaluates influences of program participation as well as free/reduced-price lunch status, race, and pre-program grades for Austin, Georgetown, and Round Rock ISD. Within district analysis is based on end of year test results and aims to measure immediate program impacts. In addition, advanced course enrollment rates for the 2013-14 Girlstart and nonparticipant groups are evaluated up to three years after program participation. For instance, if a girl attended Girlstart After School in the 4th grade, course enrollment data from the 7th grade provides information on program impacts three years post program when compared to nonparticipants. Small sample size made propensity-score matching techniques inadequate for all districts at the district level. A propensity-score matching technique is used in district analysis where demographic data or pre-program course grades in math and science are available. STAAR performance analysis is based on test results from the spring of the year in which participants attended Girlstart After School. All statistical analysis is conducted at the 95% confidence interval.

Nonparticipant data was requested from each district for girls who did not participate in Girlstart After School as comparison data is used to evaluate program impacts. The analysis attempts to control for differences in basic background characteristics and pre-program educational inputs by using the following data: free/reduced-price lunch status, race, and/or pre-program grades. Specific data and analysis limitations are addressed by district in the findings sections of this report. Differences in innate ability are potentially muted owing to Girlstart program selection process, which identifies girls who perform near the class mean. In addition, the study attempts to address one noneducational input by using free/reduced-price lunch status as a proxy measure of household income. For Austin, Del Valle,



Georgetown, and Round Rock ISD comparison data is from the same schools where Girlstart After School operates. At Pflugerville ISD, a limited sample size required some comparison data to come from other schools within the district. Comparison group selection varied slightly by district. Comparison data was not available for Connally, Hays, or North East ISD. Austin, Connally, Del Valle, Georgetown, North East, and Round Rock ISD provided course grades and enrollment information. Table 1 lists the number of Girlstart After School programs included in this report. Figure 2 shows the number of study participants by district.

YEAR	Austin ISD	Connally ISD	Del Valle ISD	Georgetown ISD	Hays ISD	North East ISD	Pflugerville ISD	Round Rock ISD	TOTAL
2013-14	8	1	2	4	3	1	5	3	27
2014-15	8	1	2	4	4	1	5	3	28
TOTAL	16	2	4	8	7	2	10	6	55

Table 1 Number and Distribution of Girlstart Programs in Study

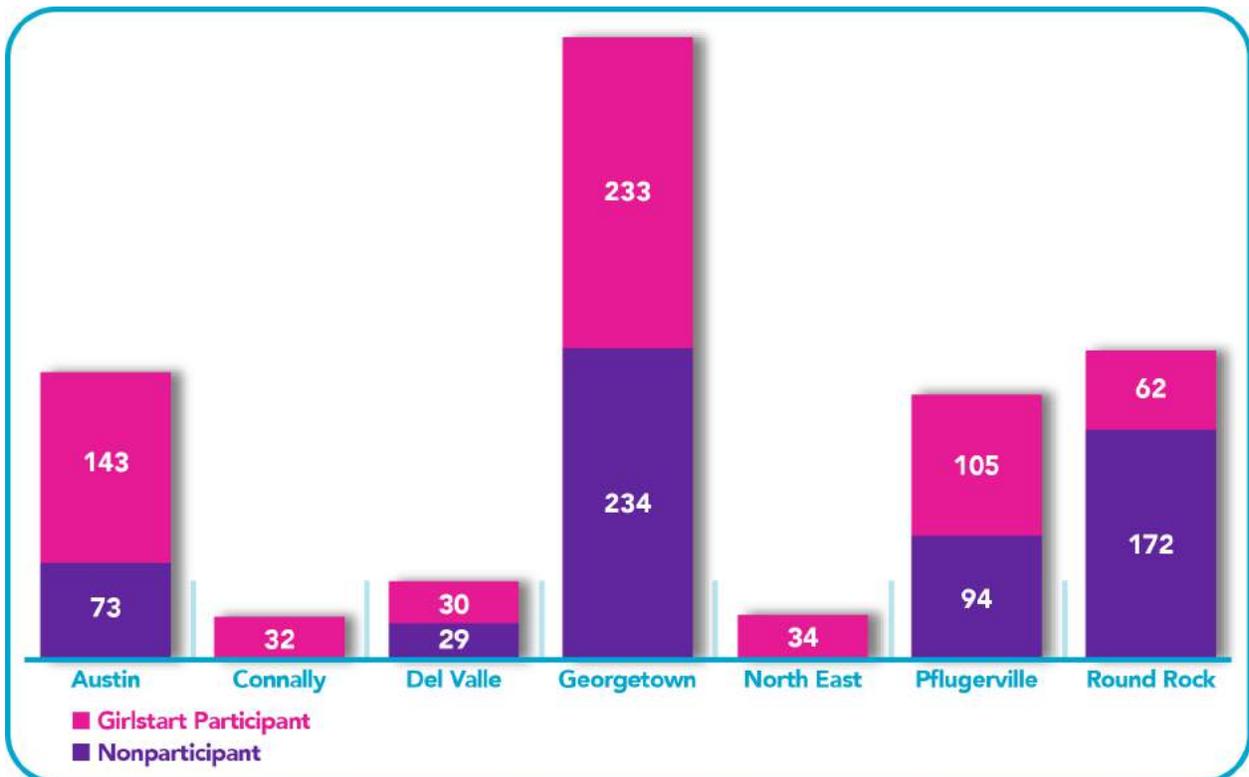


Figure 2 Number and Distribution of Study Participants by District

Overall Findings

Overall, girls who participated in Girlstart After School between 2013 and 2015 show consistently positive outcomes toward educational achievement and attainment in math and science subjects. Study



results reveal that Girlstart participants more often achieve standard and commended performance on STAAR math and science tests. In particular, Girlstart participants outperform nonparticipants in STAAR testing related to advanced science topics. Girlstart participants also achieve higher rates of educational attainment over time as a far greater percentage of Girlstart participants entering middle school enroll in advanced math and science courses up to three years after program participation.

Major Findings by District: Austin ISD

- When Girlstart participant and nonparticipant educational achievement in math and science is compared based on background characteristics and prior educational performance, Girlstart participants outperform nonparticipants in math and science across all levels.
- When Girlstart participant and nonparticipant educational attainment in math and science is compared based on background characteristics and prior educational performance, Girlstart participants attain advanced enrollment in math and science subjects in the 6th and 7th grade.
- In Austin ISD, girls who participated in 2013-14 Girlstart After School show consistently higher educational achievement and attainment over time in math and science subjects.

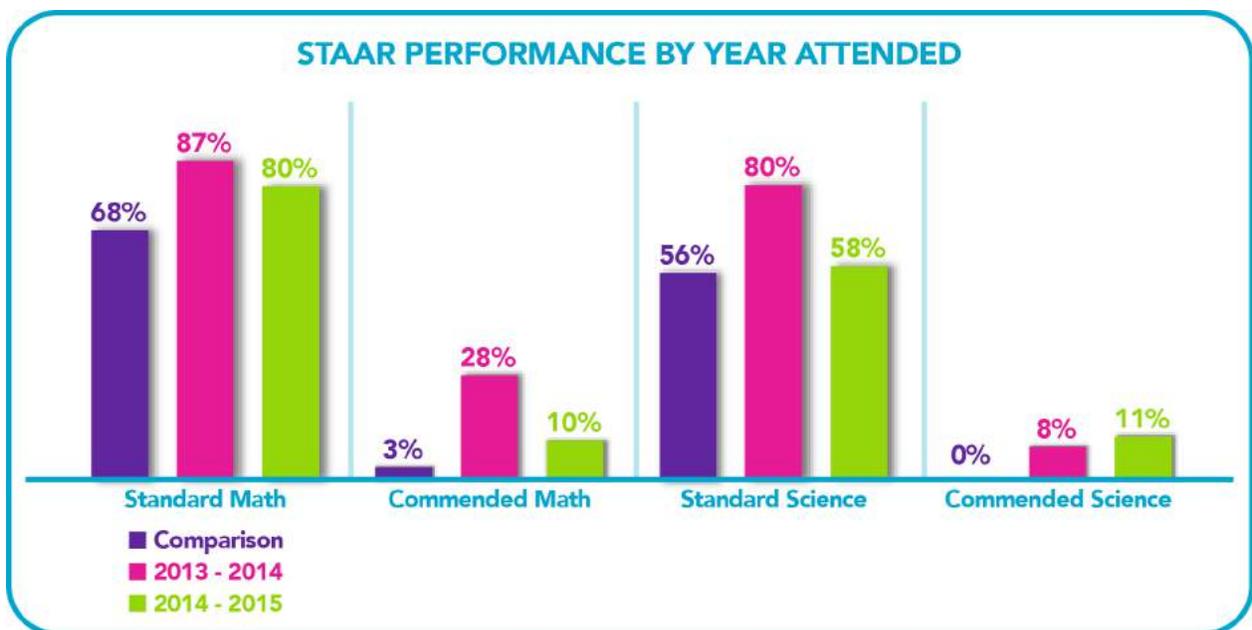


Figure 3 Austin ISD Average STAAR Performance Across Subject for 2013-14 and 2014-15 Study Groups

Figure 3 illustrates performance on STAAR math and science tests for nonparticipants and Girlstart participants from 2013-14 and 2014-15. The percentages in Figure 3 represent the portion of each group passing the associated STAAR test level. In Austin ISD, Girlstart participants consistently achieve higher passing rates on STAAR math and science tests than nonparticipants. The most noticeable difference in average passing rates occurs for standard and commended performance of the STAAR math test. 87% (2013-14) and 80% (2014-15) of Girlstart participants achieved standard performance in math compared to only 68% of the nonparticipant group. In addition, 28% (2013-14) and 10% (2014-15)



of Girlstart participants attained commended performance in math while only 3% of nonparticipants achieved this level. Furthermore, 8% (2013-14) and 11% (2014-15) of Girlstart participants attained commended performance in science, whereas 0% of nonparticipants achieved this. STAAR performance differences between nonparticipants and Girlstart participants is statistically significant for both standard and commended math achievement.

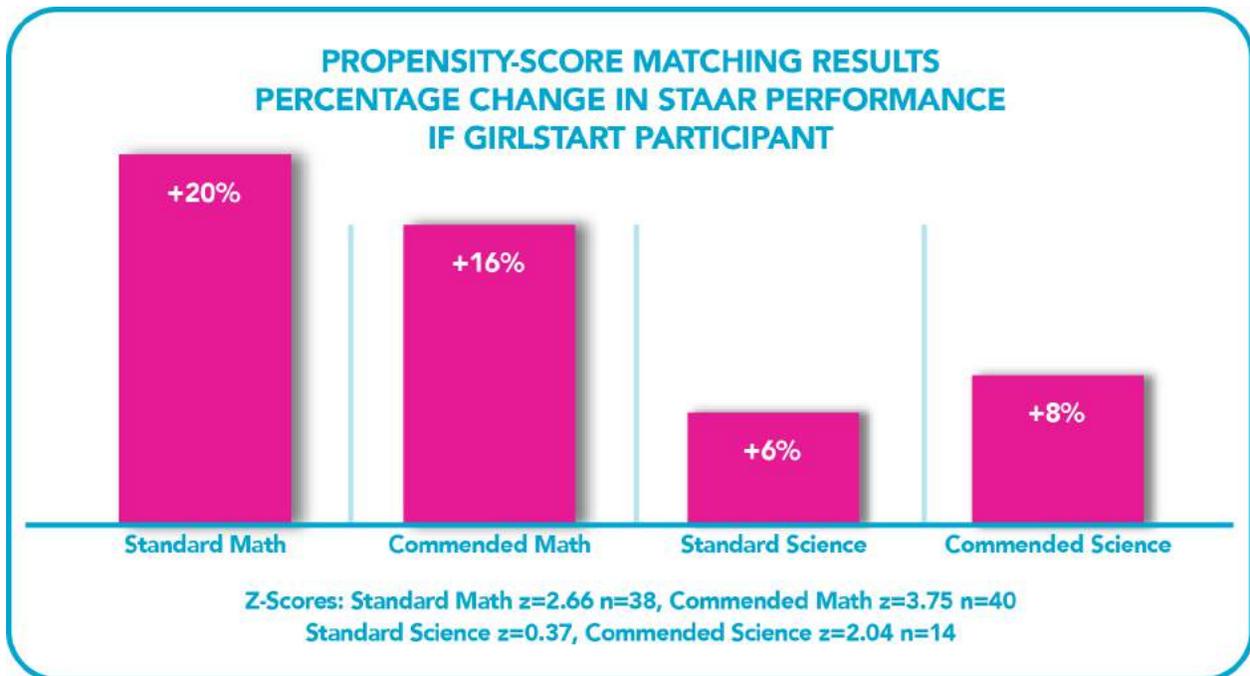


Figure 4 Austin ISD Propensity-Score Matching Results: STAAR Performance

Figure 4 illustrates STAAR achievement gains of girls who attended Girlstart After School from 2013-15 over their nonparticipant peers. The analysis attempts to control for educational and noneducational differences by taking into account free/reduced-price lunch status, race, and pre-program grades. Results for math and science performance are strongly positive toward Girlstart participation. Analysis revealed a highly significant difference in STAAR math performance between those attending Girlstart After School and nonparticipants. Compared to their nonparticipant peers in Austin ISD, 20% more Girlstart After School participants achieved STAAR standard math performance, and 16% more achieved STAAR commended math performance. In addition, analysis indicates that Girlstart girls achieve commended STAAR science performance at significantly higher rates than nonparticipants.

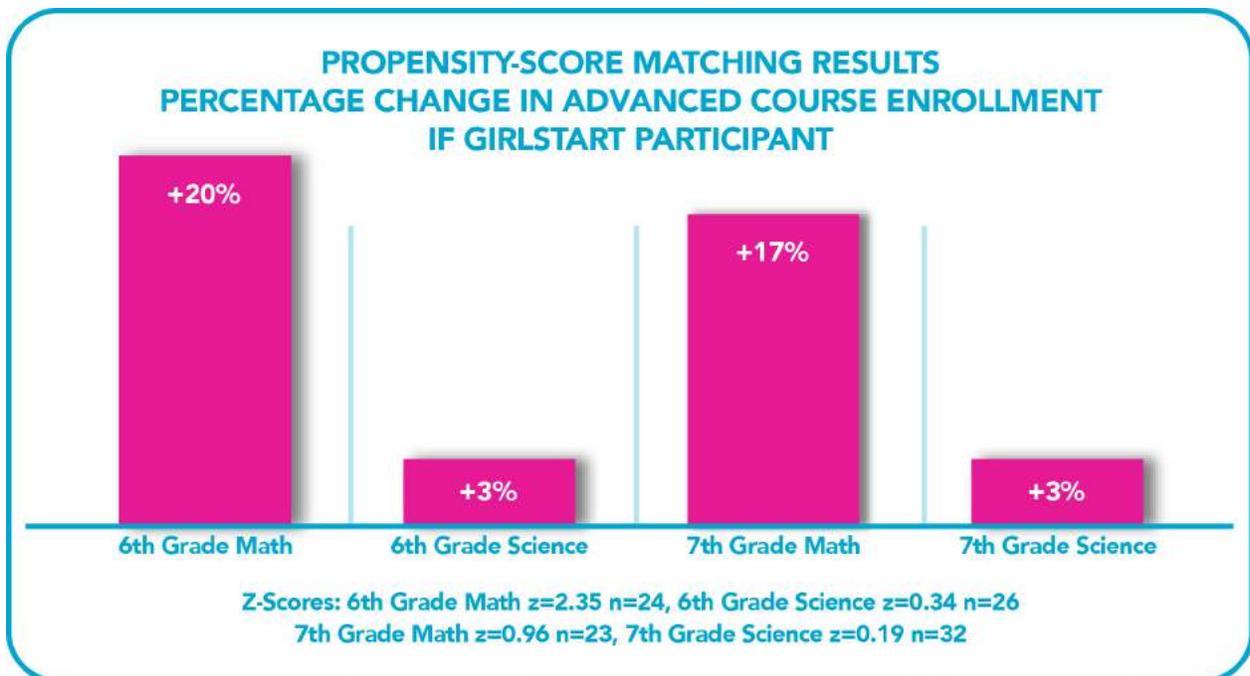


Figure 5 Austin ISD Propensity-Score Matching Results: Advanced Course Enrollment

Figure 5 illustrates percent differences in Girlstart After School participants advanced math and science course enrollment for up to three years after program participation as compared to nonparticipant advanced course enrollment. Accounting for the educational and non-educational inputs stated above, analysis suggests that up to 3 years after program participation, Girlstart participants consistently enroll in advanced math and science courses at higher rates than nonparticipants. Also, results indicate that Girlstart girls entering the 6th grade enroll in advanced math courses at significantly higher rates than girls who did not attend Girlstart After School. In Austin ISD, 20% more 6th grade girls enrolled in an advanced math course if they attended Girlstart After School in the 4th or 5th grade. Though not statistically significant, 17% more 7th grade girls enrolled in an advanced math course if they attended Girlstart After School in the 4th or 5th grade. In addition, enrollment in 7th grade advanced science courses is positive toward Girlstart participation.

Major Findings by District: Connally ISD (Waco)

- In Connally ISD, girls who participate in Girlstart After School show consistently higher educational attainment over time in math and science subjects than nonparticipants.
- High rates of advanced math and science course enrollment for Girlstart girls in Connally ISD is especially meaningful because these girls are compared to statewide nonparticipant rates where a large portion of girls come from historically higher performing districts.^{vii}

The number of participants from Connally ISD ranged from 9 to 30 girls depending on the year of program attendance, grade in program, and STAAR test subject. Girlstart was unable to obtain nonparticipant data for Connally ISD and, therefore, could not compare program impacts within the district. Owing to the small participant group size and lack of comparison data, program impacts are analyzed at the state level. Average results for educational achievement and attainment outcomes of interest for Girlstart participants in Connally ISD are listed below in Table 2.

LEVEL	MATH	PARTICIPANTS	SCIENCE	PARTICIPANTS
Standard	83%	30	73%	15
Commended	7%	30	0%	15
Scale Score	1566	30	3738	15
Advanced 6th Grade Enrollment	54%	24	46%	24
Advanced 7th Grade Enrollment	78%	9	78%	9

Table 2 Connally ISD Average Performance for 2013-14 and 2014-15 Girlstart Participants

In Connally ISD, Girlstart participants achieve the statewide Girlstart average of 83% for standard performance on the math STAAR, exceeding the statewide nonparticipant average of 76%. In advanced STAAR performance, Girlstart participants perform at lower rates than the statewide nonparticipant average in the math and science STAAR; however, as noted above, statewide comparison data comes from higher performing districts. In Connally ISD, advanced enrollment in math and science courses is well above the statewide nonparticipant average. For example, 54% of Girlstart participants in Connally ISD enrolled in an advanced 6th grade math course compared to only 39% of nonparticipants statewide. Furthermore, 46% of Girlstart participants enrolled in an advanced 6th grade science course while only 14% of statewide nonparticipants enrolled at this level. In addition, 78% of Girlstart participants enrolled in an advanced 7th grade math or science course as compared to a statewide nonparticipant enrollment of 31% in advanced math and 11% in advanced science. However, the small number of participants for 7th grade course enrollment data in Connally ISD should be noted.

Major Findings by District: Del Valle ISD

- Among all districts, Del Valle ISD Girlstart girls performed at the highest rates in educational achievement across subjects in STAAR performance as compared to district nonparticipants.
- Girls who participated in Girlstart After School show consistently higher educational achievement and attainment over time in math and science subjects than nonparticipants.
- Over time, girls who participated in Girlstart After School show significantly higher rates of educational achievement than nonparticipants, as seen through STAAR performance in math and science one year after program participation.



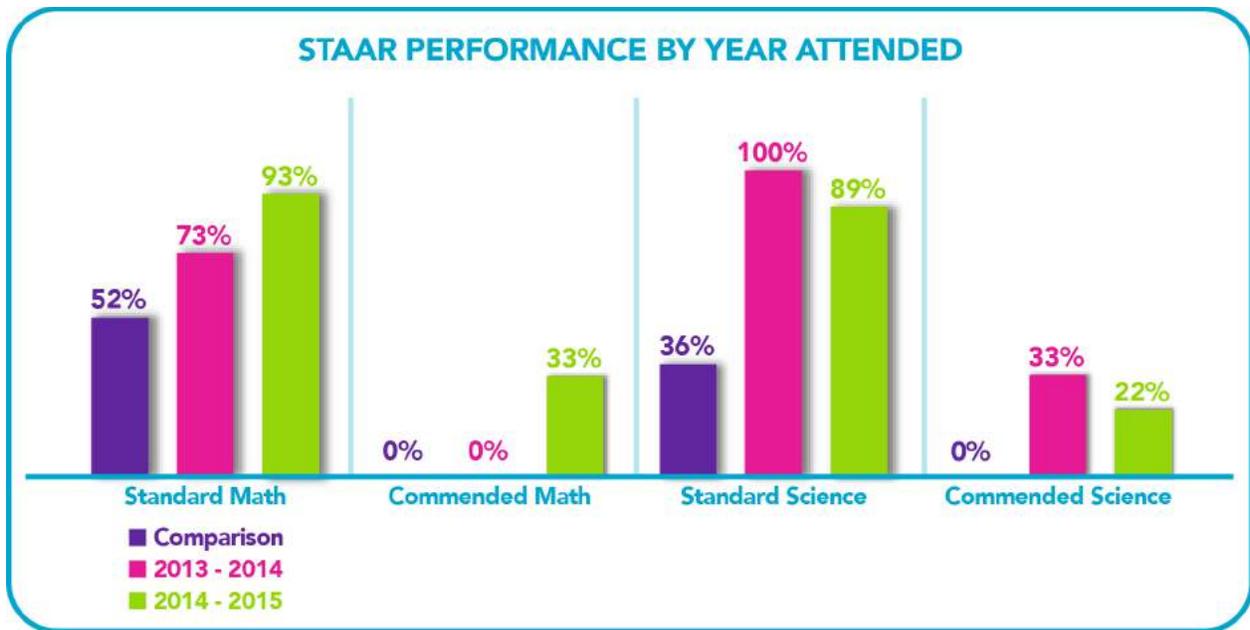


Figure 6 Del Valle ISD Average STAAR Performance Across Subject for 2013-14 and 2014-15 Study Groups

In Del Valle ISD, Girlstart was not able to obtain pre-program grades and therefore analysis did not control for pre-program educational differences. Girlstart obtained STAAR results for the year following program participation for the 2013-14 cohort and thus evaluated program impacts on Girlstart girls one year after program attendance compared to nonparticipant performance over the same period. Course enrollment data was not available for nonparticipants and the Girlstart participant group was small – ranging from 9 to 19. Therefore, educational attainment outcomes are analyzed only at the statewide level. However, analysis showed that advanced course enrollment rates for Girlstart girls are as high as, or exceeding, the statewide nonparticipant average in 6th and 7th grade math and 6th grade science.

Figure 6 compares performance on STAAR math and science tests for nonparticipants and Girlstart participants from 2013-14 and 2014-15. In Del Valle ISD, Girlstart participants consistently achieve higher passing rates on STAAR math and science tests. The most noticeable difference in average passing rates occurs for standard and commended performance on the STAAR science test. 100% (2013-14) and 89% (2014-15) of Girlstart participants attained standard science performance, while only 36% of nonparticipants reached this level of achievement. In addition, 33% (2013-14) and 22% (2014-15) of Girlstart participants attained commended performance in science compared to 0% of nonparticipants.

Figure 7 illustrates significant differences in educational achievement outcomes between participants and nonparticipants. In all four evaluated STAAR categories, there are significant performance differences between nonparticipants and Girlstart participants. For example, in math STAAR testing 52% of nonparticipants met standard performance while 84% of Girlstart participants achieved standard math performance. Results for commended STAAR performance in math are also strongly positive toward Girlstart participation; 19% of participants achieved commended math performance, while 0% of nonparticipants achieved this. These results are significant at the 99% confidence level.

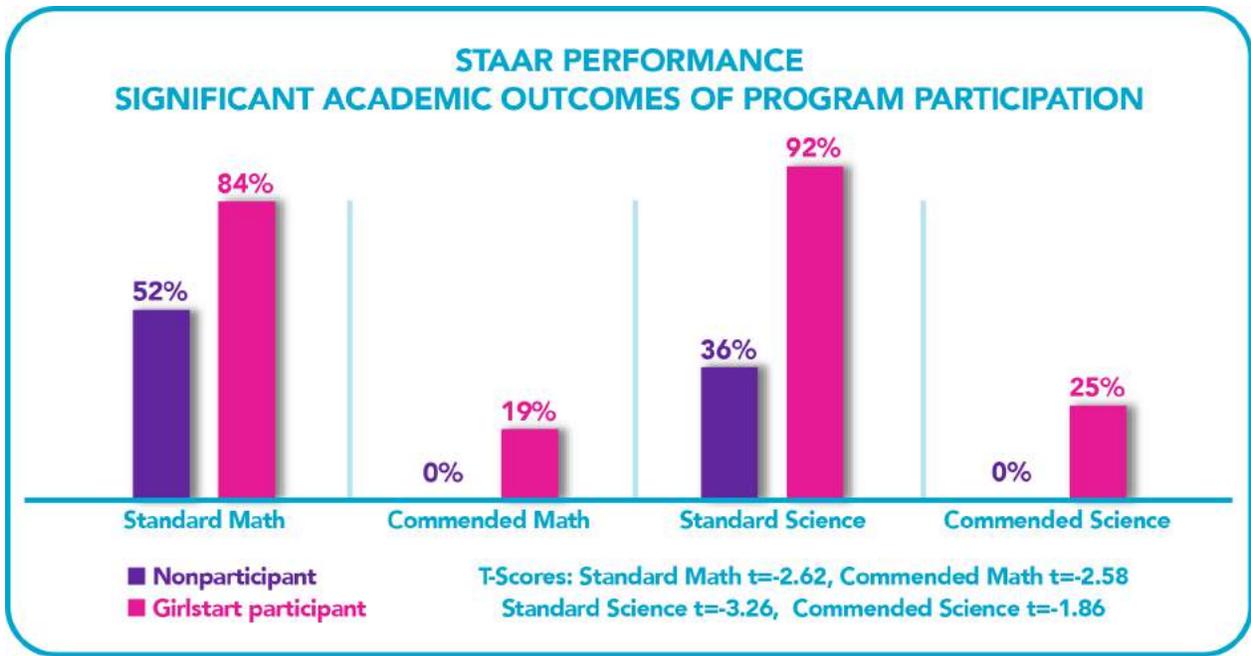


Figure 7 Del Valle ISD Percent of Study Groups Meeting Standard and Commended STAAR Performance

In Del Valle ISD, analysis indicates significant differences for standard and commended performance between study groups on the science STAAR test. 92% of Girlstart participants attained standard performance in science, whereas only 36% of nonparticipants reached this level of achievement. Results for commended performance in science are also strongly positive toward Girlstart participation; 25% of Girlstart girls attained commended science performance, compared to 0% of nonparticipants.

Educational Achievement Over Time: Post Program

Figure 8 illustrates that not only a larger percentage of Girlstart girls meet the STAAR standard math level than nonparticipants, but also Girlstart girls remain at higher passing rates one year after program participation. The difference in standard passing rates is statistically significant for Spring 2014 performance. Passing rates for math commended performance were equal for both groups at 8% for Spring 2014 and 0% for Spring 2015. Owing to a lack of pre-program grades, analysis does not control for differences in innate ability or educational inputs.

The sample size for STAAR science test data over time in Del Valle ISD is very small, ranging from 3-10 participants, and is therefore not included. However, the 2013-14 Girlstart girls achieved higher passing rates on both levels of the STAAR science test over time than nonparticipants, whose performance dropped dramatically for both science test levels over the same time period.



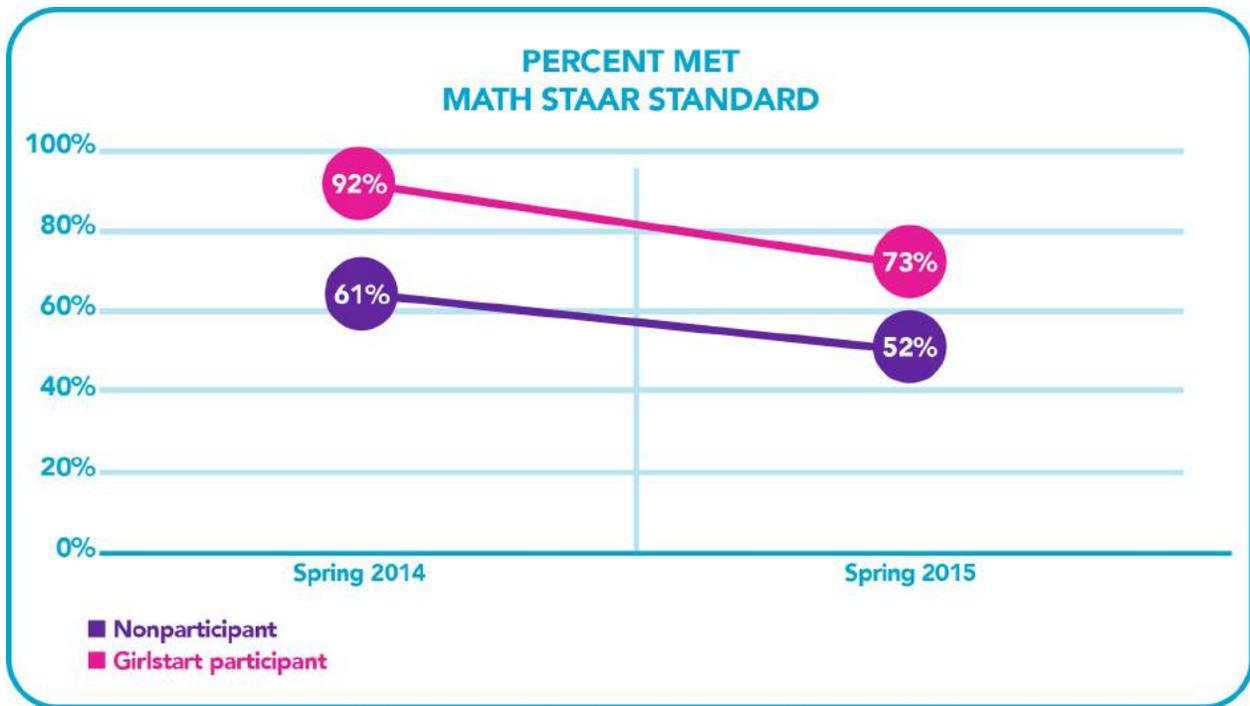


Figure 8 Del Valle ISD: 2013-14 Study Groups Standard Math Performance Over Time

Major Findings by District: Georgetown ISD

- In Georgetown ISD, girls who participated in Girlstart After School show consistently higher educational achievement in math and science subject areas than nonparticipants.
- When educational achievements are compared based on background characteristics and prior educational performance, Girlstart participants outperform nonparticipants in math and science across all levels.
- Girls who participated in Girlstart After School show significantly higher educational achievement than nonparticipants over time, as seen through STAAR performance in math and science one year after program participation.

Figure 9 compares performance on STAAR math and science tests for nonparticipants and Girlstart participants from 2013-14 and 2014-15 in Georgetown ISD. The most noticeable difference in average passing rates occurs for STAAR commended science performance. Girlstart participants reached commended performance in science at rates of 18% (2013-14) and 8% (2014-15) compared to only 3% of nonparticipants.

Figure 10 shows STAAR achievement differences for girls who attended Girlstart After School when accounting for free/reduced-price lunch status, race, and pre-program grades. Results for performance in math and science are strongly positive toward Girlstart participation. Analysis based on these criteria revealed significant differences for performance on commended STAAR achievement in math and science between those attending Girlstart After School and nonparticipants. Compared to nonparticipants in Georgetown ISD, 8% more Girlstart After School participants achieved commended STAAR performance in math, and 9% more Girlstart participants achieved commended STAAR

performance in science. In addition, results for performance in standard science and math are strongly positive toward Girlstart participation and almost significant at the 95% confidence level.

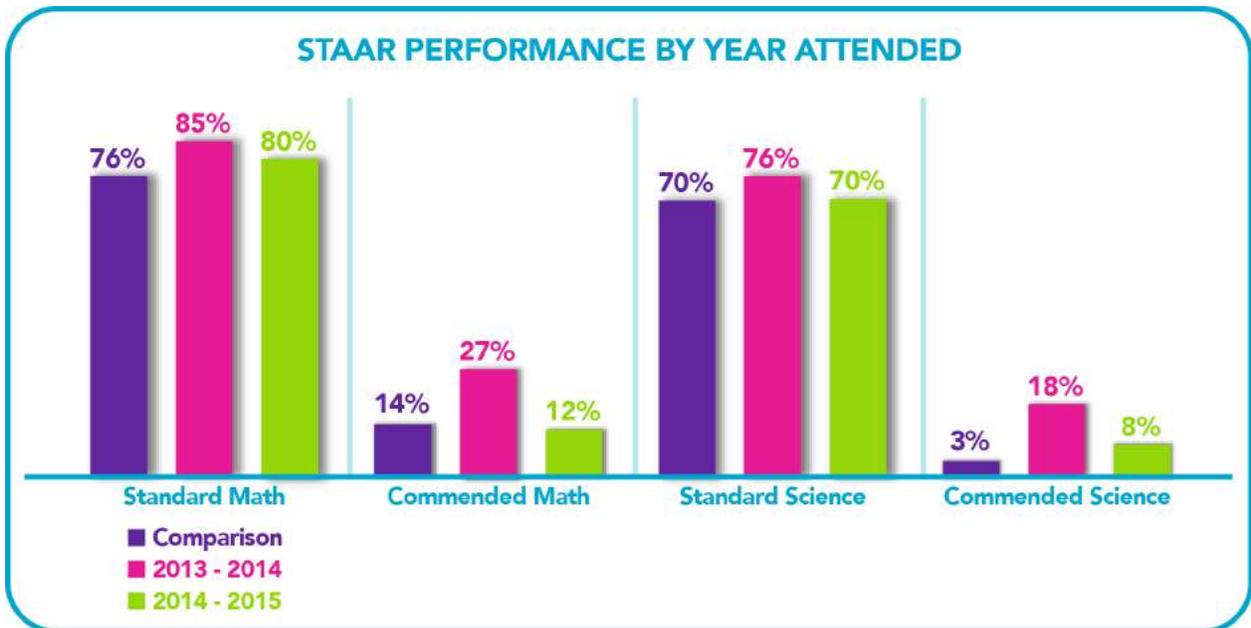


Figure 9 Georgetown ISD Average STAAR Performance Across Subject for 2013-14 and 2014-15

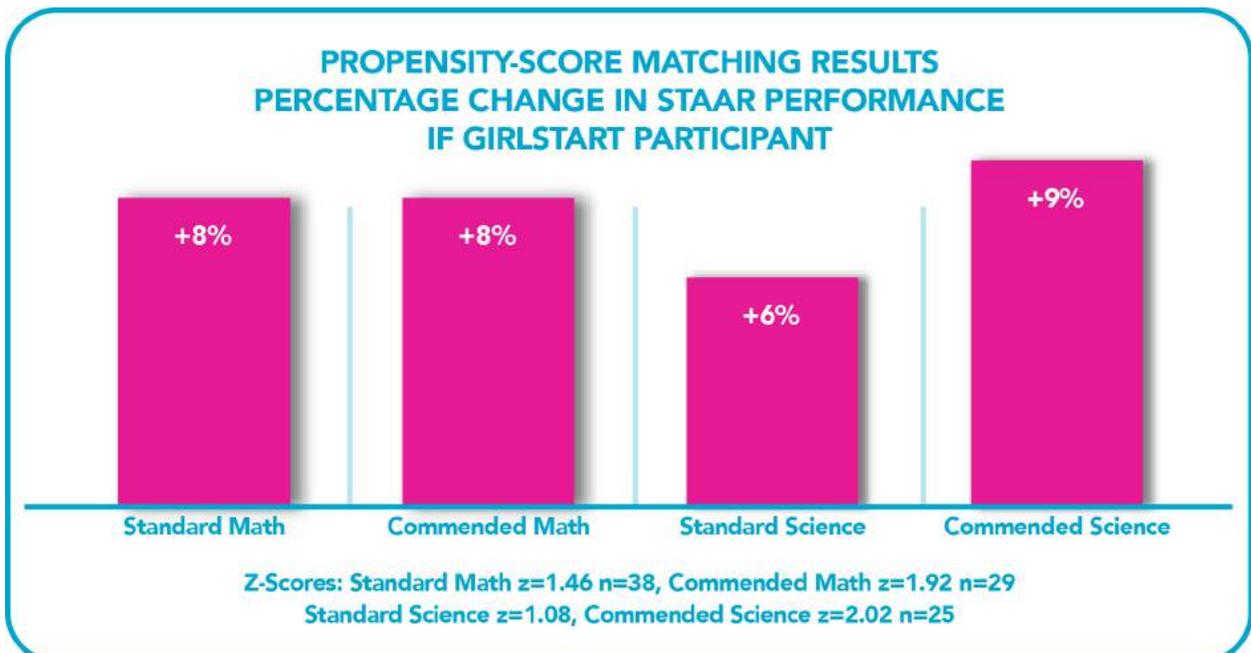


Figure 10 Georgetown ISD Propensity-Score Matching Results: STAAR Performance



Georgetown ISD Educational Achievement Over Time: Post Program

Results in Figure 11 show that larger percentages of Girlstart girls achieve commended STAAR performance in math and science immediately after program participation. The difference between Girlstart participants and nonparticipants in commended performance is statistically significant for math in Spring 2014. In addition, one year after program participation, Girlstart girls remain at higher performance rates for commended-level STAAR tests. Furthermore, performance in standard math and science over time is positive toward Girlstart participation and statistically significant for math immediately after program participation. It should be noted that analysis did not control for pre-program educational achievement differences using prior semester course grades.

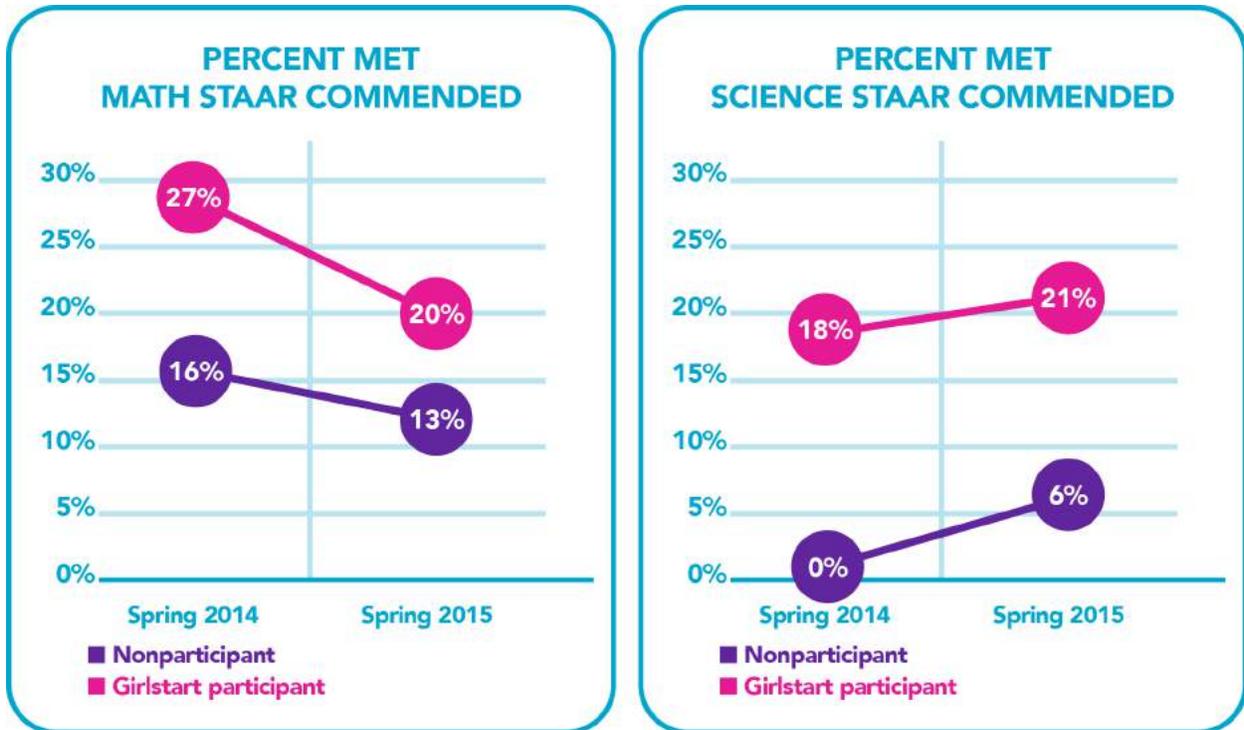


Figure 11 2013-14 Study Groups STAAR Performance Over Time

Major Findings by District: Hays CISD

Hays CISD data contained aggregated scale scores for STAAR math and science test for Girlstart participants and nonparticipants. Disaggregated nonparticipant STAAR scores were not available. It is unknown how the nonparticipant group was selected or how many students comprised this group. Since Girlstart did not receive STAAR math and science scores, advanced course enrollment attainment, or comparison data, Hays CISD was not included in statewide analysis. Table 3 lists average STAAR math and science scale scores for Hays ISD Girlstart girls and nonparticipants.

GROUP	MATH	PARTICIPANTS	SCIENCE	PARTICIPANTS
Girlstart Participants	1595	54	3935	31
Nonparticipants	1547	—	3702	—

Table 3 Hays ISD Study Group Scale Scores

Major Findings by District: North East ISD (San Antonio)

In North East ISD, Girlstart After School participants achieve higher rates of standard educational achievement in math and science compared to statewide nonparticipants.

- Girlstart participants achieve significantly higher rates of commended educational achievement in science compared to statewide nonparticipants.
- Girlstart participants achieve significantly higher rates of advanced course enrollment in math compared to statewide nonparticipants.
- A high rate of advanced math course enrollment is especially meaningful in North East ISD because participants are compared to statewide nonparticipants where a large percentage are from historically higher performing districts.^{viii}

METRIC	MATH	PARTICIPANTS	SCIENCE	PARTICIPANTS
Standard STAAR	84%	25	77%	22
Commended STAAR	12%	25	5%	22
Scale Score	1576	25	3705	22
Advanced 6th Grade Enrollment	62%	21	5%	21

Table 4 North East ISD Average Performance for 2013-14 and 2014-15 Girlstart Participants



The number of participants from North East ISD ranged from 5 to 25 girls depending on the year of program attendance, grade in program, and STAAR test subject. Girlstart was unable to obtain nonparticipant data for North East ISD and therefore could not compare performance within the district. Owing to the small participant group size and lack of nonparticipant data, analysis is performed at the statewide level. However, Table 4 lists average results for achievement outcomes of interest for Girlstart participants in North East ISD.

In North East ISD, Girlstart participants exceed the statewide nonparticipant average in both standard math and science STAAR performance. 84% of Girlstart girls achieve standard STAAR math performance as compared to 76% of statewide nonparticipants. In addition, 77% of Girlstart participants achieve standard STAAR science performance as compared to 73% of statewide nonparticipants. Analysis of advanced math performance shows that 12% of Girlstart girls achieve commended STAAR math performance while only 5% of statewide nonparticipants reach this level of achievement. In advanced performance for the science STAAR test, Girlstart participants perform at lower rates than the statewide nonparticipant average of 15%; however, as noted above, Girlstart girls in North East ISD are compared to a nonparticipant statewide average from historically higher performing districts.

In North East ISD, advanced enrollment in math courses is well above the statewide average for nonparticipants. 62% of Girlstart participants in North East ISD enrolled in advanced 6th grade math courses compared to only 39% of nonparticipants statewide. Girlstart girls enrolled at lower rates than the statewide nonparticipant average of 15% in advanced 6th grade science courses; as previously noted, Girlstart girls in North East ISD are compared to a nonparticipant statewide average from historically higher performing districts.

Major Findings by District: Pflugerville ISD

In Pflugerville ISD, girls who participate in Girlstart After School show consistently higher educational achievement in math and science. In particular, Girlstart participants achieve advanced science testing results at significantly higher rates than nonparticipants. Girlstart participants show higher rates of educational achievement in standard math performance than nonparticipants over time, as evaluated through STAAR performance one year after program participation.

Figure 12 compares academic performance on STAAR math and science tests for nonparticipants and Girlstart participants from 2013-14 and 2014-15 in Pflugerville ISD schools. In general, Girlstart girls consistently achieve equal or higher passing rates on STAAR math and science tests. The most noticeable difference in average passing rates occurs for commended performance on the STAAR science test. 14% (2013-14) and 24% (2014-15) of Girlstart girls achieved commended science performance compared to only 4% of nonparticipant girls.

Figure 13 shows statistically significant results of educational achievement outcomes based on program participation. Among the four categories of evaluated STAAR tests, the differences between study groups is significant for STAAR performance on standard math and commended science. 84% of Girlstart participants achieved standard math performance compared to 73% of nonparticipants. 19% of Girlstart participants achieved the commended level of science performance compared to only 3% of nonparticipants.



STAAR PERFORMANCE BY YEAR ATTENDED

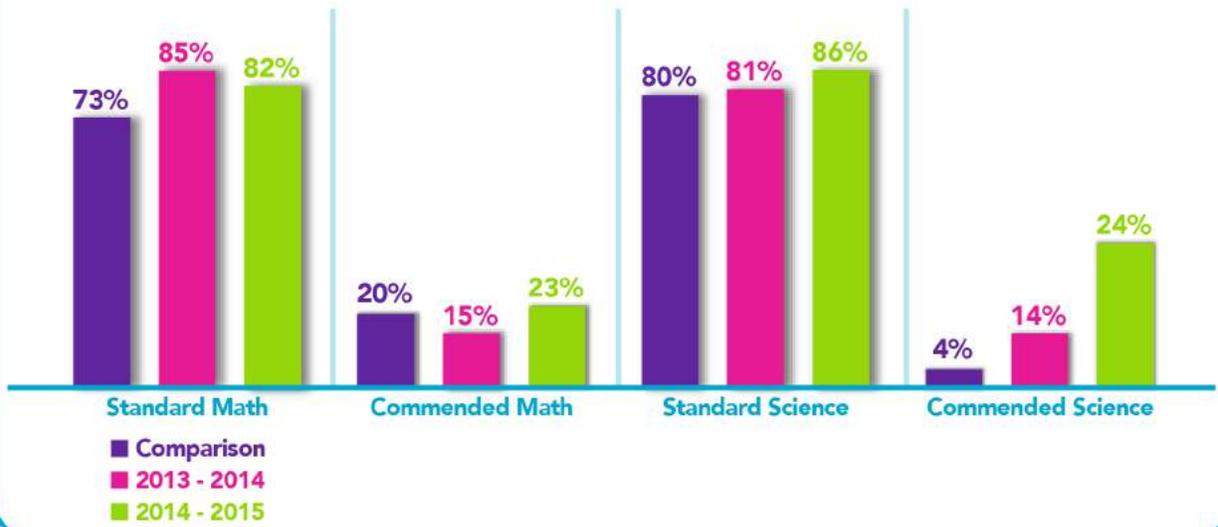


Figure 12 Pflugerville ISD Average STAAR Performance Across Subject for 2013-14 and 2014-15

STAAR PERFORMANCE SIGNIFICANT ACADEMIC OUTCOMES OF PROGRAM PARTICIPATION

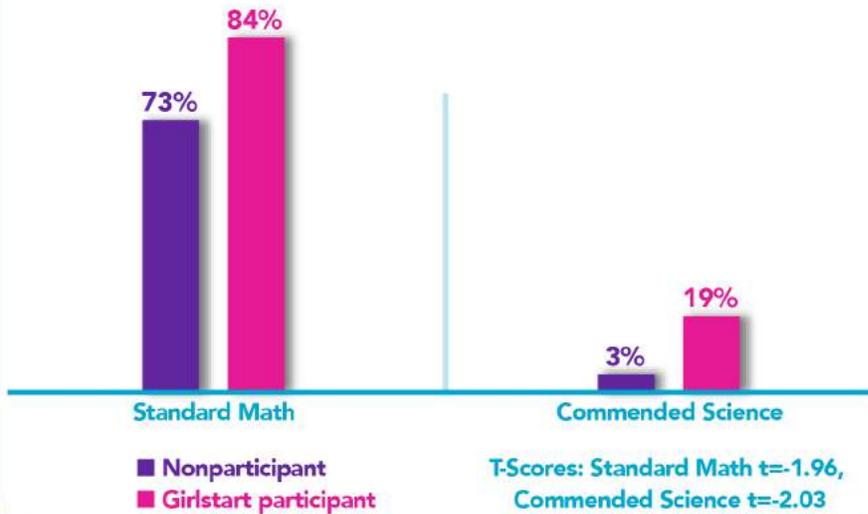


Figure 13 Pflugerville ISD Percent of Study Group Meeting STAAR Performance Levels



Results for commended STAAR math performance and standard science performance are positive toward Girlstart participation. Analysis of STAAR math performance reveals that 19% of nonparticipants and 20% of Girlstart girls achieved commended math performance. 79% of Girlstart participants achieved standard science performance compared to 74% of nonparticipants. These results are not statistically significant.

Educational Achievement Over Time: Post Program (2013-14 Cohort)

Figure 14 shows that a larger percentage of Girlstart girls meet standard math performance on the STAAR test immediately after program participation when compared to nonparticipants. In addition, Girlstart girls remain at higher passing rates on the math standard achievement level one year after the program. The difference in standard STAAR math passing rates is statistically significant over time. Among the other three categories of testing achievement, no significant differences were seen over time. Owing to a lack of pre-program grades, analysis does not control for educational differences prior to 2014.

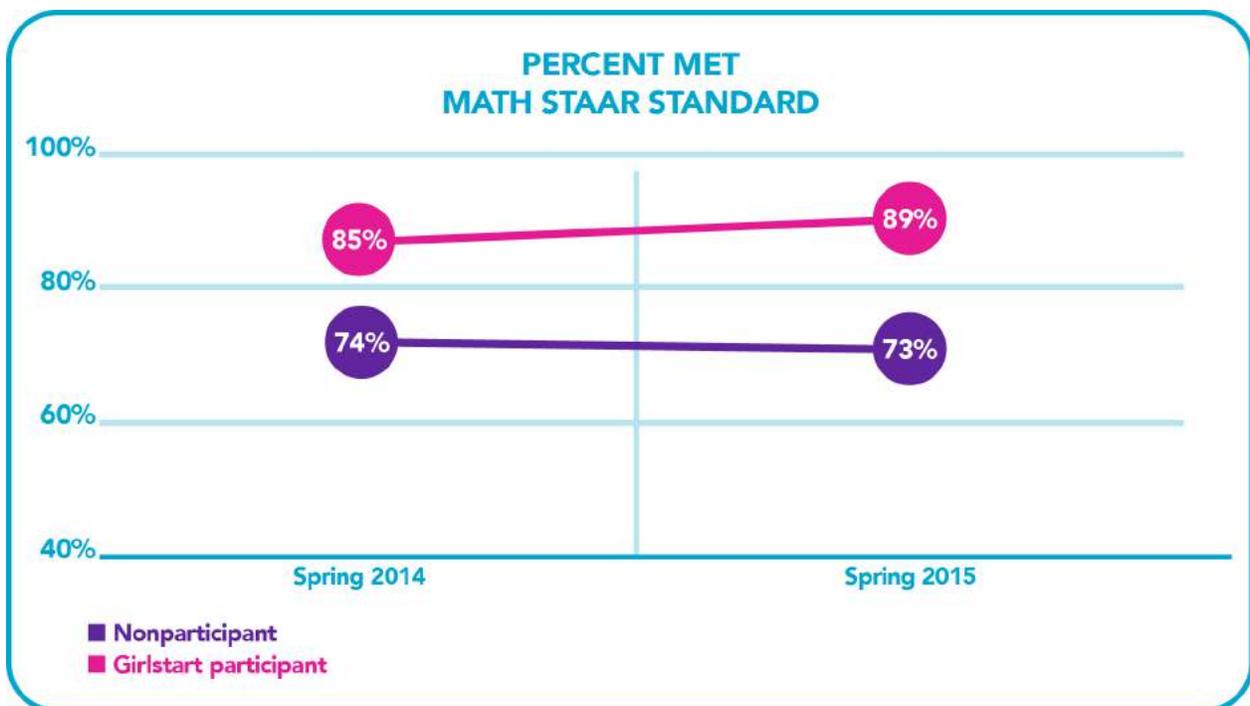


Figure 14 Pflugerville ISD 2013-14 Girlstart Participant STAAR Performance Over Time

Major Findings by District: Round Rock ISD

- When Girlstart After School participant and nonparticipant educational achievements in math and science are compared based on background characteristics and prior educational performance, Girlstart girls outperform nonparticipants in advanced math and science subject areas.
- In Round Rock ISD, Girlstart girls reach significantly higher rates of academic achievement in advanced science topics than nonparticipants.



- Girlstart girls show significantly higher achievement in advanced-level topics than nonparticipants over time, as seen through STAAR math and science passing rates one year after program participation.
- In 2014-15, program selection may have varied from average academic achievers and/or been limited by natural class composition.

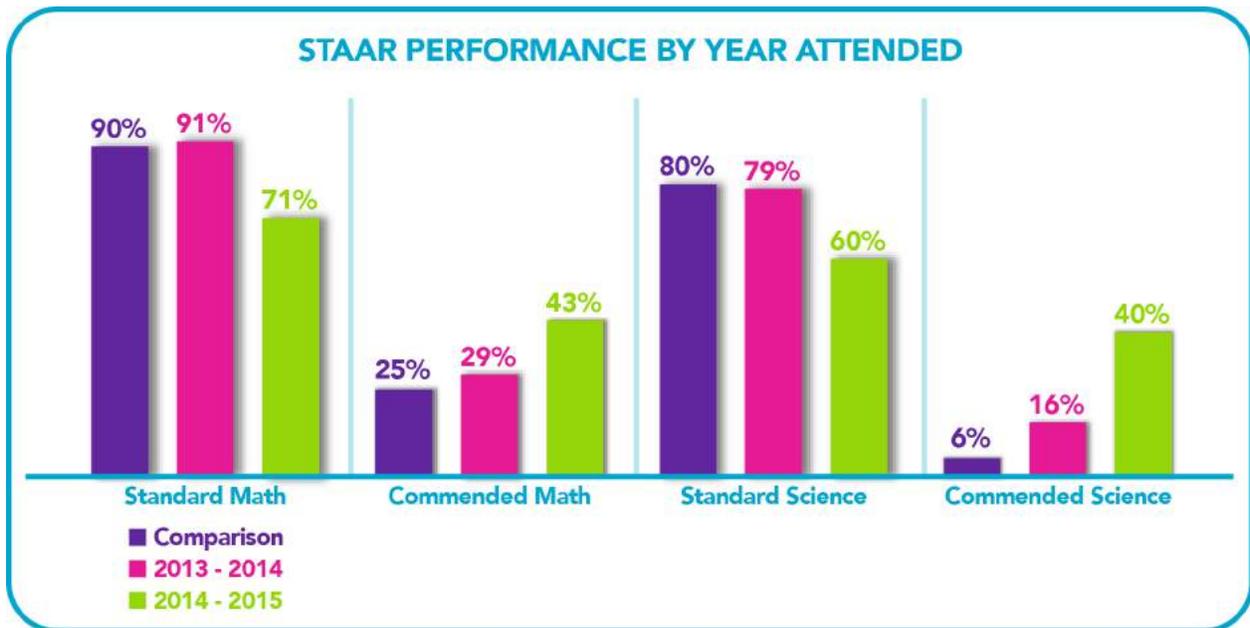


Figure 15 Round Rock ISD STAAR Performance Across Subject for 2013-14 and 2014-15 Study Groups

It should be noted that the 2014-15 Girlstart cohort sample size contains only five participants for STAAR science results and is therefore analyzed at district level. However, the 2014-15 Girlstart group’s unusually low and high performances influence district results and may indicate an atypical Girlstart After School selection process or some other unknown programmatic factor.

Figure 15 shows performance on STAAR math tests for nonparticipants and Girlstart participants from 2013-14 and 2014-15 in Round Rock ISD schools. Girlstart girls consistently achieve higher passing rates on commended STAAR math performance. 25% of nonparticipants achieved commended math performance, whereas 29% (2013-14) and 43% (2014-15) of Girlstart participants achieved commended performance in math. However, the 2014-15 Girlstart group performed lower than nonparticipants on standard STAAR math achievement, while the same group’s advanced-level math performance far exceeded passing rates of both the 2013-14 Girlstart girls and nonparticipants in math and science. The 2014-15 Girlstart group may have been influenced by program selection or some other unknown programmatic factor.

Results in Figure 16 illustrate STAAR achievement differences for Girlstart After School participants when accounting for free/reduced-price lunch status, race, and pre-program grades. Analysis shows that math and science performance is positive toward Girlstart participation and reveals significant differences for high-level science STAAR achievement between groups. In Round Rock ISD, 21% more girls achieve commended performance on the science STAAR test if they attended Girlstart After School.

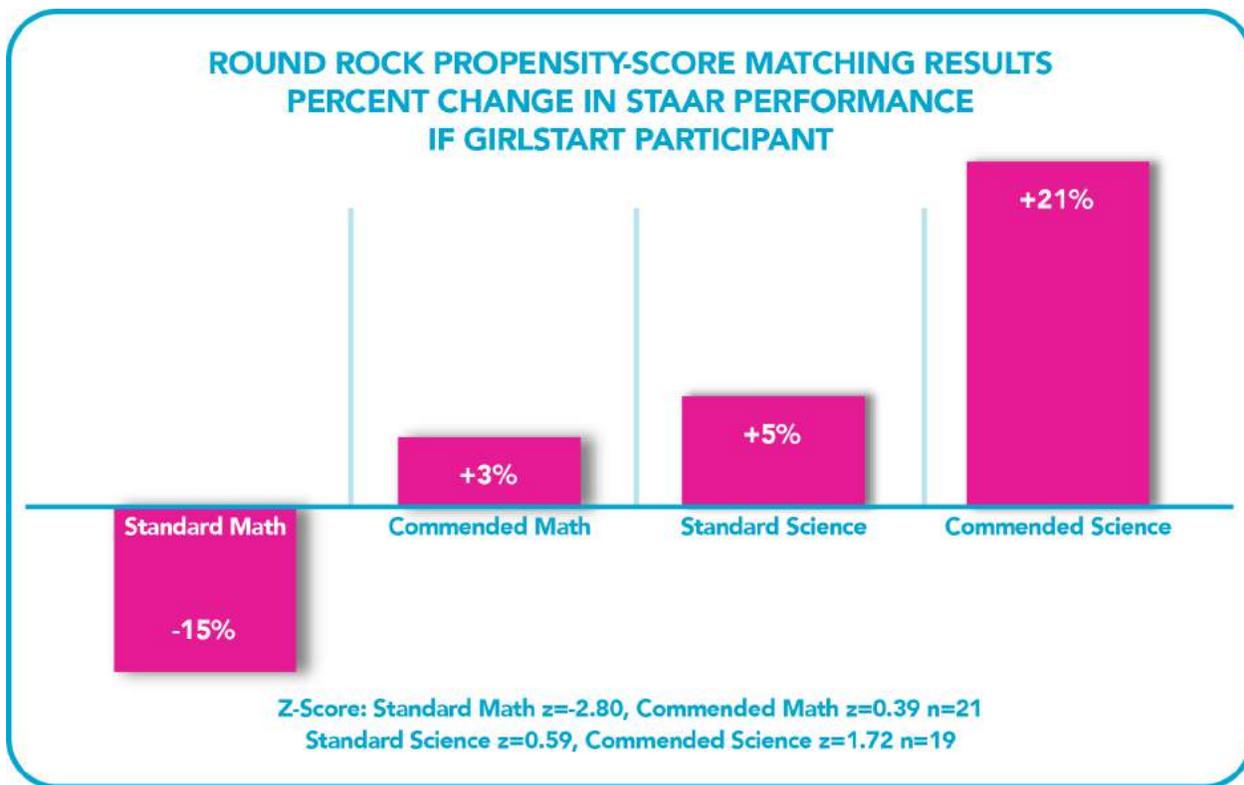


Figure 16 Round Rock ISD Propensity-Score Matching Results: STAAR Performance

Girlstart After School aims to impact science achievement and skills attainment. This report evaluated math to understand potential impacts on girls' math performance. However, math impacts are secondary outcomes that may not always mirror program impacts. Two factors may influence the STAAR standard math results in the Round Rock ISD. First, there is a large difference between the sample size of Girlstart participant and nonparticipant groups available for statistical matching. Round Rock ISD is a historically high-performing district, especially in meeting standard-level STAAR performance.^{ix} It is possible that a larger nonparticipant group – matched on background characteristics and prior academic performance – could hold a higher pre-performance average, while a smaller Girlstart group would not have this influence. In addition, a high variation in the 2014-15 participants' performance should be noted, as the standard deviation and standard error are 50 to 100 times greater than nonparticipant performance variation. This performance variation in the Girlstart group further indicates the previously mentioned potential divergence in Girlstart's After School selection process or class composition for 2014-15 in Round Rock ISD.

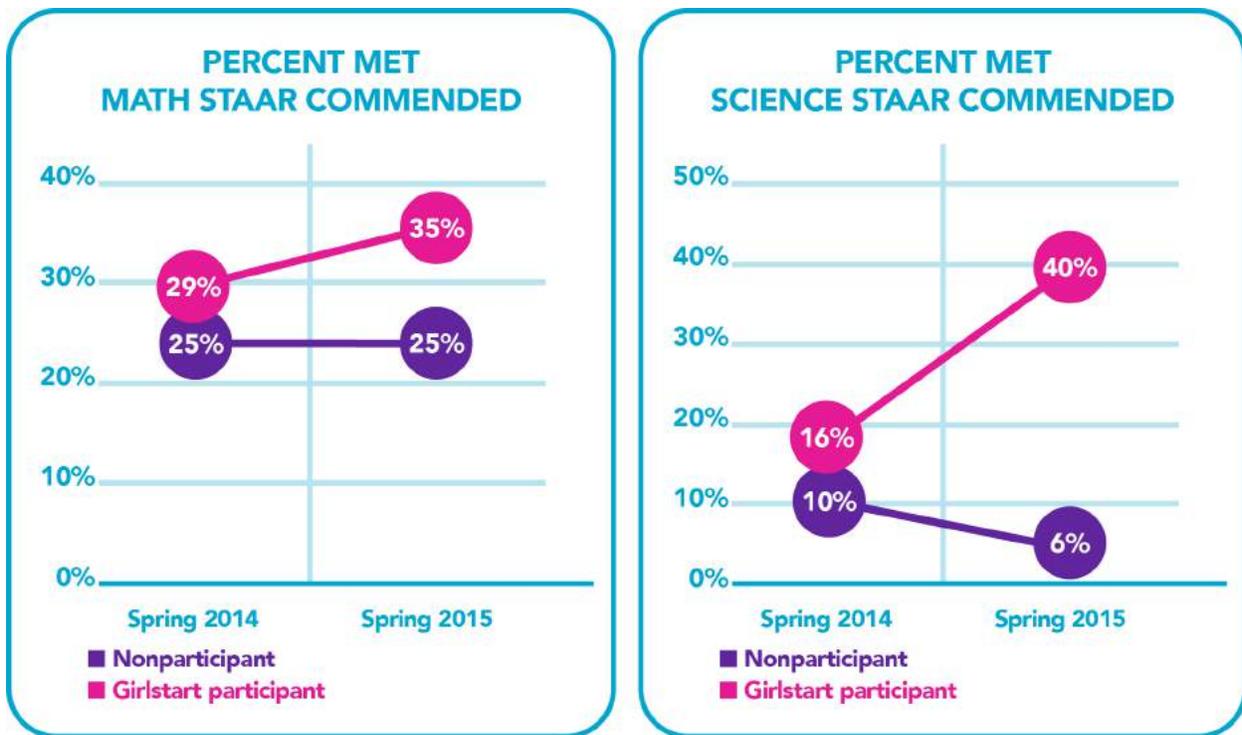


Figure 17 Round Rock ISD 2013-14 Study Groups STAAR Performance Over Time

Round Rock ISD Educational Achievement Over Time: Post Program (2013-14 Cohort)

Analysis reported in Figure 17 shows that a larger percentage of Girlstart girls met STAAR commended math and science performance immediately after participating in Girlstart After School than nonparticipants. In addition, Girlstart girls remain at higher passing rates for these performance levels over time. Performance differences between Girlstart girls and nonparticipants is highly significant for science STAAR commended performance over time. Performance over time is positive toward Girlstart participation in science STAAR standard performance and slightly negative in math STAAR standard performance. Neither of these differences are statistically significant. Analysis did not control for pre-program achievement differences using prior semester course grades, and, as noted above, the 2015 science cohort was very small.

Statewide: STAAR Performance and Subsequent Advanced STEM Course Enrollment

Statewide analysis includes data from Austin, Connally, Del Valle, Georgetown, North East, Pflugerville, and Round Rock ISDs. Hays CISD is not included in statewide analysis owing to a lack of data on math and science STAAR performance and course enrollment. Program impacts are compared between study groups controlling for influences of free/reduced-price lunch status and race.

Major Findings: Statewide

- When Girlstart participant and nonparticipant educational achievements are compared based on background characteristics, Girlstart girls outperform nonparticipants in math and science across all levels.
- Girls who participate in Girlstart After School consistently achieve higher rates of educational achievement in advanced science subjects.
- When Girlstart participant and nonparticipant educational attainment level is compared based on background characteristics, Girlstart participants attain increased access to advanced math and science courses in the 6th and 7th grade.

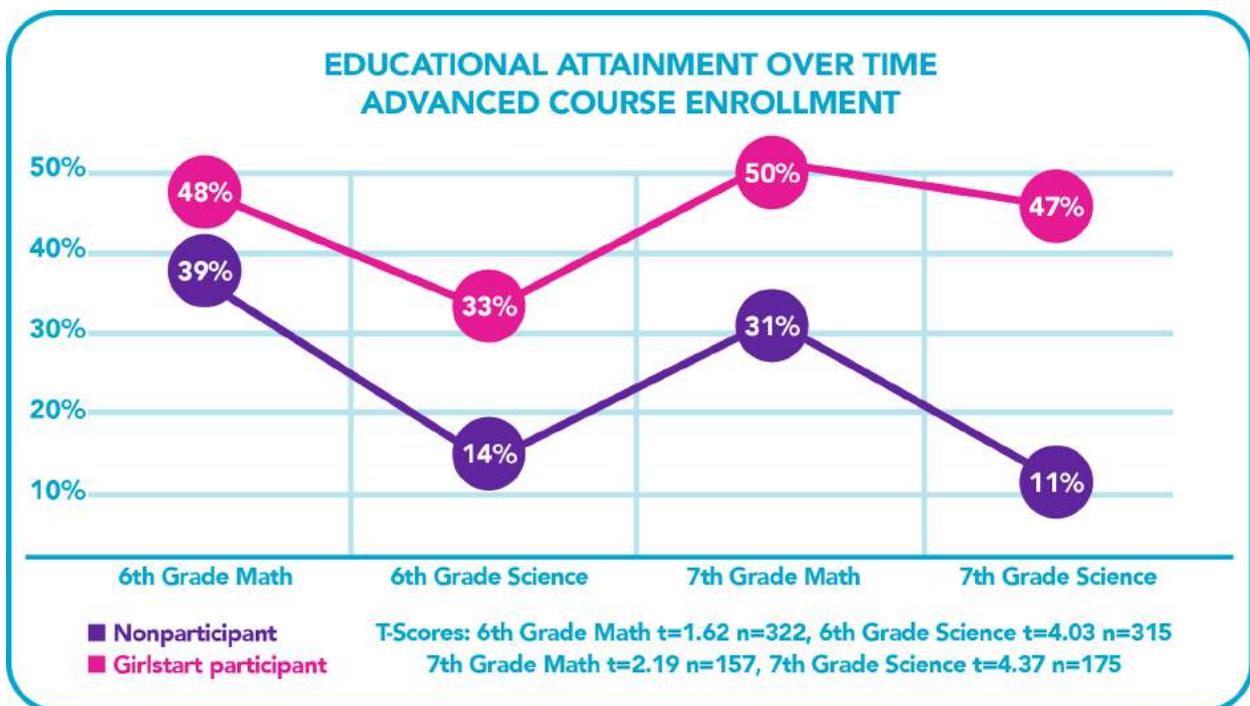


Figure 18 Statewide Percent of Study Group Enrolled in Advanced Courses Over Time

Statewide results in Figure 18 illustrate large differences between advanced course enrollment rates for Girlstart girls and nonparticipants; significantly higher percentages of Girlstart girls enroll in advanced math and science courses in the 6th and 7th grade when compared to nonparticipant course enrollment. Girlstart girls enroll in advanced math and science courses at significantly higher rates for one to three years after program participation. In addition, Girlstart girls appear to enroll in advanced math and science courses at increasingly higher rates compared to nonparticipants over time.

The differences between program participants and nonparticipants in advanced course enrollment rates are highly significant for 6th grade science and 7th grade math and science. Data for each district did not include pre-program grades, and therefore analysis did not control for pre-program grades. However, an analysis of Austin, Georgetown and Round Rock ISD controlled for pre-program differences in performance and background characteristics and confirmed statewide results and their significance.

It should be noted that parent perception of student capability may play a role in a girl's course enrollment. Parents' beliefs about their child's academic capability or readiness, or parental exposure to course selection processes, may influence a student's course enrollment. It is unclear in what way these potential influences may affect course enrollment. However, parents of low-income and/or first-generation college aspirant households may have less knowledge of the influence of course enrollment on future educational access or readiness, or be less able to obtain appropriate course enrollment knowledge for their child. If this is true, a higher percentage of Girlstart girls would be more likely to have parents under-assign advanced course enrollment, as Girlstart After School serves a high proportion of low-income and/or first generation college-aspirant households. As a result, Girlstart After School may have underestimated its influence on educational attainment for girls as observed in middle school advanced course enrollment in math and science.

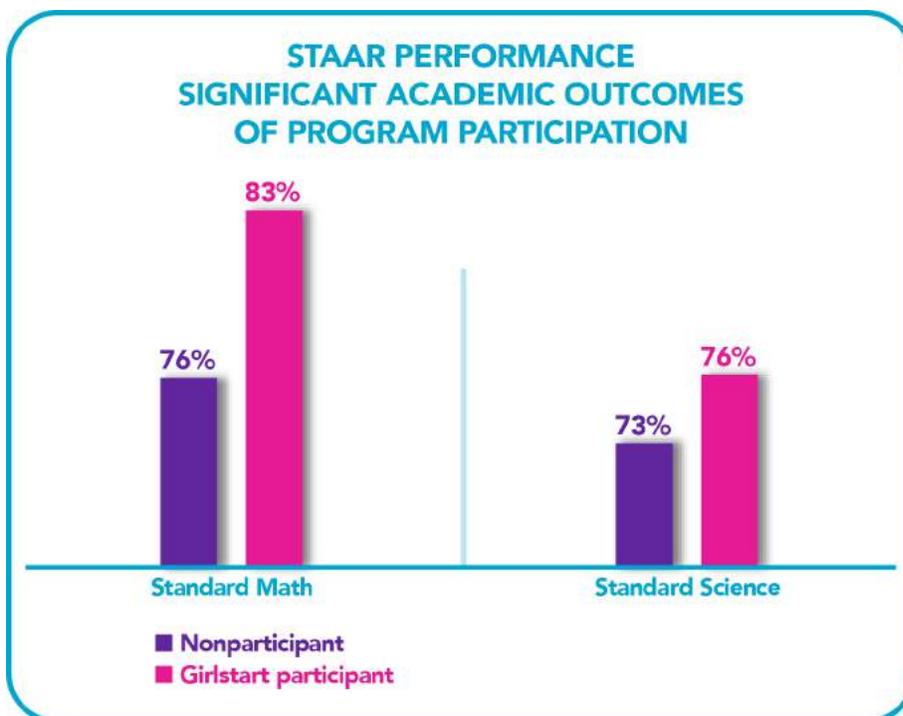


Figure 19 Statewide Percent of Study Group Meeting STAAR Performance

Throughout Central Texas, Girlstart participants outperformed nonparticipants on standardized math and science STAAR tests at all levels. For both the STAAR math and science test, Girlstart participants achieved standard and commended performance at higher rates than nonparticipants. Among these, the percent difference of Girlstart girls achieving advanced-level STAAR science performance was statistically significant when compared to nonparticipants; 13% of Girlstart participants achieved advanced performance in science topics while only 5% of nonparticipants achieved advanced science performance.

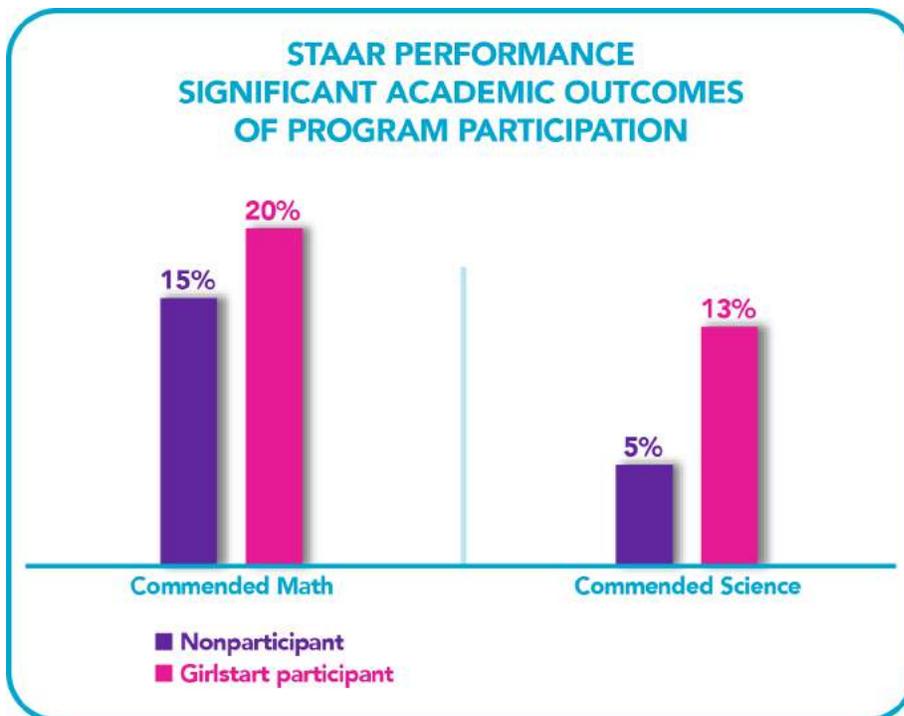


Figure 20 Statewide Percent of Study Group Meeting STAAR Performance

Major Findings: Overall

- Girlstart participants perform at significantly advanced levels on the STAAR science test compared to nonparticipants.
- Girlstart participants enroll in advanced science and math courses at significantly higher rates than nonparticipants in the 6th and 7th grade.
- Girlstart girls attain enrollment in advanced science and math courses at increasingly higher rates than nonparticipants over time.
- Girlstart After School appears to advance girls' short- and medium-term knowledge, skills, and abilities in math and science.

Research results of this report show that Girlstart After School academic impact is greatest in two areas: educational achievement and attainment. Firstly, Girlstart After School has a positive influence on educational achievement in math and science as measured by girls' performance on the State of Texas STAAR test. In particular, Girlstart participants consistently pass advanced levels of the STAAR science test more often than nonparticipants. Secondly, Girlstart After School has a positive influence on educational attainment in math and science subjects over time. Longitudinal analysis shows that Girlstart girls enroll in advanced science and math courses at significantly higher rates than nonparticipants in the 6th and 7th grade. This finding is especially meaningful because educational attainment may measure girls' agency and access to STEM career pathways in more meaningful ways than standardized test performance can do alone.

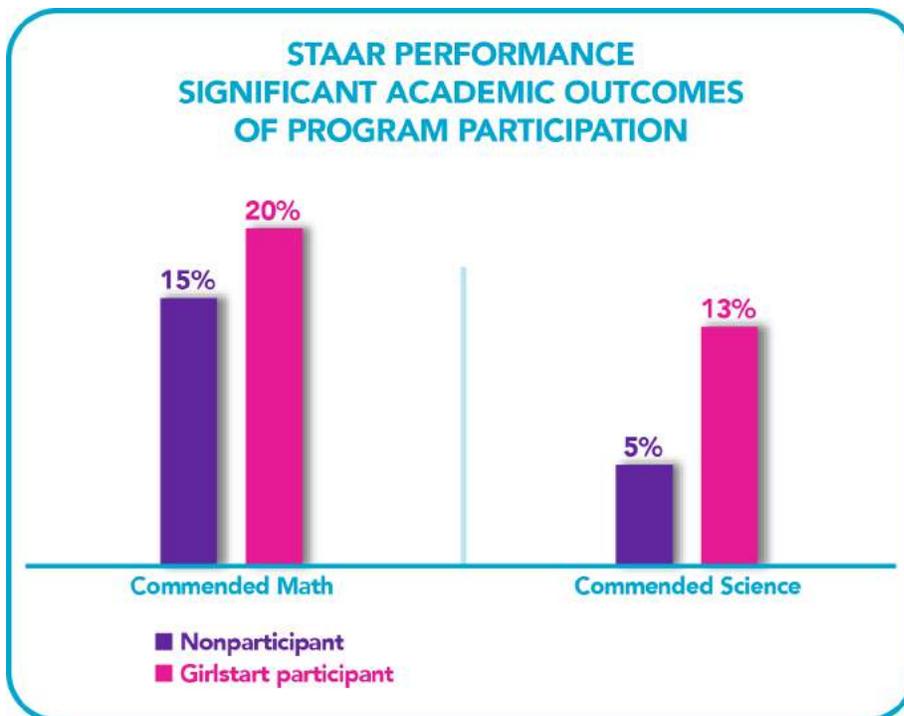


Figure 20 Statewide Percent of Study Group Meeting STAAR Performance

Although data on the math and science STAAR tests are included in this analysis, the 5th grade science STAAR test provides a stronger point of comparison for achievement outcomes of students’ skills and knowledge than the math test. Primarily, this is because Girlstart’s mission is to focus on science academic readiness and career exposure. In addition, science STAAR performance is a more valid statistical measure of program impacts because while students may take the math exam up to three times—and are given mandatory tutoring and remediation over several months if they do not pass — students are allowed only one attempt at the science test. Thus, science STAAR passing rates more closely reflect nonparticipant baseline abilities and academic program impacts. Therefore, it is especially meaningful to see that Girlstart participants consistently achieve advanced performance on the STAAR science test.

Also, research in this report found that Girlstart participants were more likely on average to enroll in advanced math and science courses after program participation. This finding indicates that Girlstart After School not only improves short- and medium-term academic achievement, but may also positively influence outcomes related to interest and beliefs surrounding STEM subjects and associated career opportunities. This is because major components of the Girlstart After School curriculum relate STEM topics and activities to careers, and instructors discuss the strengths and opportunities these jobs may provide. A 2010 report by the American Association of University Women (AAUW) reported that research shows girls perceive certain occupations as inappropriate for their gender and that occupational choice is influenced by a person’s values and expectant success.* Consequently, research also shows that girls and women are more likely to choose jobs where they understand or perceive a direct benefit of their work either to individuals or society as a whole. Girlstart After School appears to advance not only girls’ short- and medium-term knowledge, skills, and abilities in math and science, but may also positively influence girls’ understanding of opportunities in high demand, high paid jobs in math and science in ways that were previously unknown to participants.

Study Limitations

Limited Access to Robust Statewide Comparison Data

This study includes eight districts and, of those eight, five districts submitted comparison data. A large portion of nonparticipants, 234 out of 602, roughly 39% –are students from Round Rock ISD. Historical statewide reports show that Round Rock ISD achieves higher average STAAR performance than other districts where Girlstart After School works.^{xi} For example, in this study Round Rock ISD nonparticipants outperform other district nonparticipant groups on STAAR performance for both levels of math and science. A large portion of comparison data from a high performing district may result in underestimating the educational achievement and attainment impacts of Girlstart After School in statewide analysis.

In addition, not all comparison group data contained demographic and background characteristic information, or educational input differences regarding free/reduced-price lunch status, race, and/or pre-program course grades. This data limitation does not allow for statewide analysis to rely on more robust propensity-score matching techniques to compare girls' performance. However, when data is available, statistical methods that incorporate pre-program grades are compared with OLS regression results that do not account for pre-program performance differences.

Unobservable Influences

Based on free/reduced-price lunch status, race, and pre-program grades, the comparison data aimed to create an artificial control group to balance influences of educational and noneducational differences. However, other observable or unobservable influences may remain. Potential factors that may influence results include correlation between girls in the program and parental availability or support, and variation in teacher-based assignment, class composition, and/or girls' motivation. To increase understanding of long-term academic impacts, Girlstart aims to collect panel data on participant performance for 8th grade STAAR science performance and course enrollment. Girlstart attempts to control for influences of innate ability and instructional variability by working closely with teachers in the schools that select students to ensure that average academic performers are selected for the program, and by standardizing training for Girlstart After School instructors. However, analysis was unable to attempt to control for multiple noneducational inputs and other testing process factors like testing conditions or mood.^{xii}

Sample Size Limitations on Statistical or Substantive Analysis

The robustness of Round Rock ISD analysis may be vulnerable to variance in comparison group sample size; for math comparison, the sample size was 100 for nonparticipants and 62 for Girlstart participants, while science comparison data samples sizes were 62 and 20, respectively. Propensity-score matching aimed to mediate potential biases in study group sample sizes. Additionally, in Georgetown and Connally ISD the sample size for participants was as small as 5-6 participants for some measures of interest. Thus, substantive significance and analysis is omitted or limited in those particular results.

ⁱ Austin ISD, *AISD Strategic Plan 2015-2020 Reinventing the Urban School Experience Together*. (Austin: Austin ISD, 2015)

ⁱⁱ Jacob, B., & Rothstein, J. "The measurement of student ability in modern assessment systems." *Journal of Economic Perspectives*, 30(3), 85–108. doi:10.1257/jep.30.3.85

ⁱⁱⁱ Hill, C., Corbett, C., Rose, A. *Why So Few? Women in Science, Technology, Engineering, and Mathematics*. (Washington, DC: AAUW, 2010)

^{iv} Wiseman, A. Bailie, C., Gorgey, H., Hamrock, C. *Pathways of Promise: Statewide Mathematics Analysis*. (Austin: E3 Alliance, 2015)

^v Friedman, A. (Ed.). *Framework for Evaluating Impacts of Informal Science Education Projects* [On-line]. National Science Foundation. (Available at: http://insci.org/resources/Eval_Framework.pdf)



-
- ^{vi} Wiseman, A. Bailie, C., Gourgey, H., Hamrock, C. Pathways of Promise: Statewide Mathematics Analysis.
- ^{vii} STAAR Aggregate Data. Texas Education Agency, <http://tea.texas.gov/student.assessment/staar/aggregate/>.
- ^{viii} STAAR Aggregate Data. Texas Education Agency, <http://tea.texas.gov/student.assessment/staar/aggregate/>.
- ^{ix} STAAR Aggregate Data. Texas Education Agency, <http://tea.texas.gov/student.assessment/staar/aggregate/>.
- ^x Hill, Corbett, and Rose, *Why So Few?*
- ^{xi} STAAR Aggregate Data. Texas Education Agency, <http://tea.texas.gov/student.assessment/staar/aggregate/>.
- ^{xii} Jacob, B., & Rothstein, J. "The measurement of student ability in modern assessment systems." *Journal of Economic Perspectives*, 30(3), 85–108. doi:10.1257/jep.30.3.85