AccelerateLEARNING
THE LEADER in PreK-12 STEM EDUCATION

5177 Richmond Ave., Suite 800, Houston, TX 77056
1-800-531-0864
stemscopes.com
stemscopes@acceleratelearning.com

December 11, 2023 **Efficacy Research**

Comparing 2023 STAAR 5th Grade Science Achievement for STEMscopes and Non-STEMscopes Schools in Texas

Background

The current report focuses on the efficacy of the STEMscopes Science 5th grade curriculum. We used a post-facto quasi-experimental design (QED) that utilizes a matched control group to evaluate the potential associations between STEMscopes Science and science achievement for 5th graders across the state of Texas. QEDs with matched samples attempt to overcome the hurdle of "non-random" assignment.

In addition to examining the potential relationship between STEMscopes Science implementation and science achievement for all students, we also consider achievement in several sub-groups of students. Specifically, previous results in the field (Morgan, Farkas, Hillemeier, & Maczuga, 2016) suggest science achievement gaps among students who are considered minority groups relative to their non-minority peers, as well as achievement gaps related to socio-economic status. However, past STEMscopes reports suggest that STEMscopes may have a stronger association with standardized test passing scores among minorities (particularly black/African American students and Latino/Hispanic students) as well as students considered economically disadvantaged, English language learners (ELL), and students who receive special education services.

Thus, within the current report we consider: 1) potential group differences in science achievement for STEMscopes versus non-STEMscopes schools. Overall, we hypothesize that even with the more stringent matched control group QED design, schools that purchased and used STEMscopes during the 2022-2023 school year will include a higher percent of students, on average, that either approached, met, or mastered grade level expectations on the 2023 State of Texas Assessments of Academic Readiness (STAAR) than schools that did not purchase STEMscopes (i.e., "non-STEMscopes schools"). 2) We also anticipate significant associations between STEMscopes Science and school 5th grade STAAR passing rates among subgroups of students.

Results

To examine the effectiveness of STEMscopes Science to increase STAAR 5th grade school achievement rates, we conducted multiple regression analyses with 2,484 matched Texas schools. Our first set of analyses focused on predicting 2023 STAAR 5th grade school "approaches, meets, or masters expectations" rates with a binary variable indicating whether a school was a STEMscopes school or non-STEMscopes school, and covariates (see methods). Results were significant for "meets expectations: (non-STEMscopes M = 32.00, STEMscopes M = 33.04, b = 1.04, $p < 0.05$, ES = 0.08), see Figure 1 for comparison of rates between STEMscopes schools and non-STEMscopes schools and Table 1 for all model parameters. Analysis results also indicated a significant association between STEMscopes and STAAR 5th grade school "masters expectations" rates[1] (non-STEMscopes M = 13.35, STEMscopes M =

[1] The school "masters expectations" outcome was positively skewed (1.45) due to a high number of schools with zero or a very low percentage of students in this category. With this in mind, all analyses with "master expectations" school rate as the outcome were also ran with a zero-inflated poisson distribution

5177 Richmond Ave., Suite 800, Houston, TX 77056
1-800-531-0864
stemscopes.com
stemscopes@acceleratelearning.com

14.26, b = 0.91, $p < 0.05$, ES = 0.10), but not for the "approaches" level. In addition, the sub-group analyses indicated significant findings among Hispanic/Latinx students such that there was a positive association between STEMscopes Science and STAAR school "meets expectations rates" (non-STEMscopes M = 27.27, STEMscopes M = 28.44, b = 1.17, $p < .05$, ES = 0.08), as well as masters expectations rates non-STEMscopes M = 10.32, STEMscopes M = 11.06, b = 0., $p < .05$, ES = 0.08).

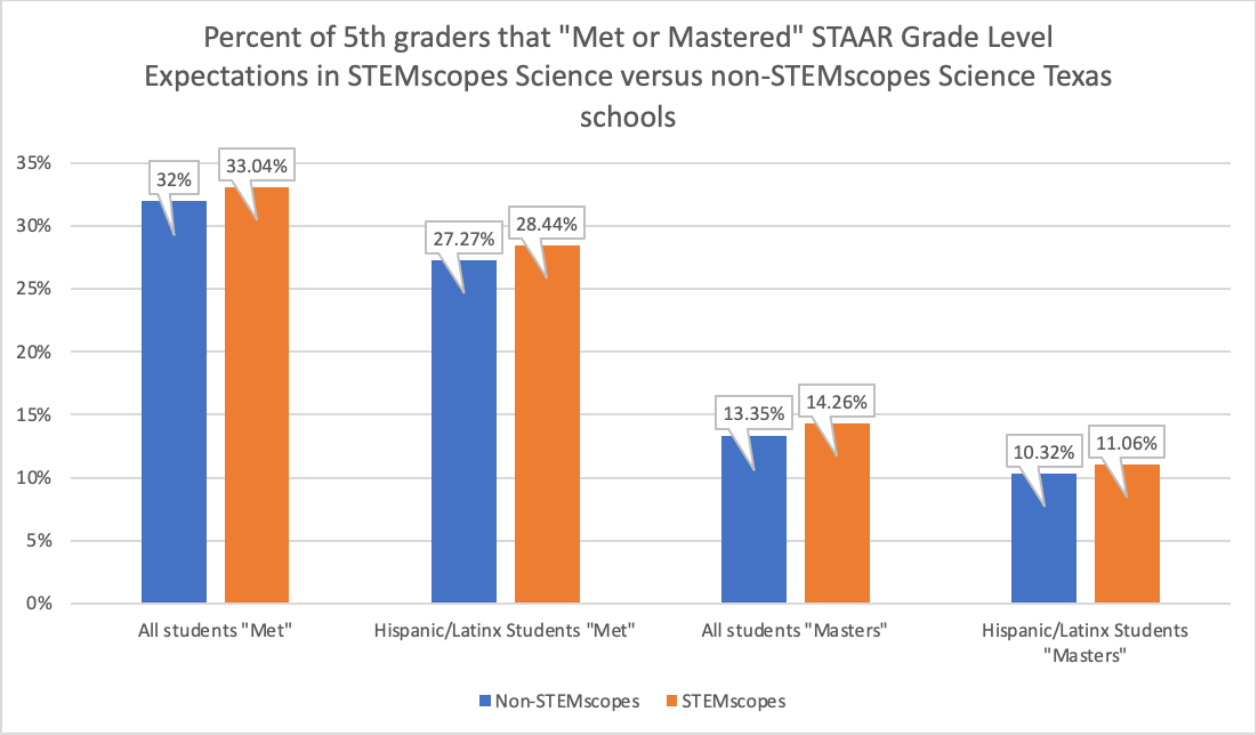

Figure 1. Comparison of 2023 school "met or mastered" rates between non-STEMscopes and STEMscopes schools.

Table 1. Full model comparing non-STEMscopes and STEMscopes school passing rate, all students

Model parameters	"Approaches" model estimates (SE)	"Meets" model estimates (SE)	"Masters" model estimates (SE)
Intercept	62.69 (0.36)**	32.00 (0.36)**	13.35 (0.25)**
STEMscopes curriculum district	0.28 (0.49)	**1.04 (0.49)***	**0.91 (0.34)****
Baseline school 4th grade math passing rate	8.68 (0.31)**	7.28 (0.30)**	3.83 (0.21)**
District Size	0.92 (0.28)**	0.76 (0.28)**	0.33 (0.20)
Percent economically disadvantaged students	-5.91 (0.52)**	-7.26 (0.51)**	-5.06 (0.36)**
Percent African American/Black students	-1.63 (0.33)**	-1.18 (0.33)**	-0.44 (0.24)
Percent Latino/Hispanic students	-0.01 (0.46)	0.07 (0.46)	0.26 (0.32)
Percent Asian students	0.25 (0.32)	1.28 (0.32)**	1.29 (0.22)**
Percent of ELL students	0.17 (0.38)	0.97 (0.38)*	0.88 (0.27)**

model. The pattern of findings matches those presented above. We chose to present the linear model findings here to ease interpretation.

5177 Richmond Ave., Suite 800, Houston, TX 77056
1-800-531-0864
stemscopes.com
stemscopes@acceleratelearning.com

Percent of special education students	0.00 (0.30)	-0.06 (0.30)	-0.16 (0.21)
Charter status	-3.16 (0.85)**	-3.28 (0.85)**	-2.19 (0.60)**

*= p < .05; **= p<.01.

Methods

In this section we provide details about study procedures including the data sources, variables used, and participating schools.

Data sources

Data for this study came from two sources. First, schools that purchased and used STEMscopes for 5th grade in the 2022 - 2023 school year were identified through the STEMscopes analytics and Sales Force systems. Within the analytics reports, we used the number of 5th grade scopes accessed as a metric of use, and then confirmed usage with our internal Sales Force reports (to rule out schools that had free trial usage). We defined a school as using the curriculum if the teacher accessed at least 8 of the available science units (called scopes) for 5th grade (~33%) at least once during the period from Aug 1 – May 15.

Second, school demographic data and school performance on the State of Texas Assessments of Academic Readiness (STAAR) were accessed through the Texas Education Agency website. We used the 2021 - 2022 STAAR campus data file for 4th grade and focused on the school level "approaches grade level percent" on the STAAR mathematics test as a baseline measure of academic achievement. Specifically, the state of Texas creates proficiency benchmarks in all academic content and identifies students as not proficient, approaching grade-level proficiency, meeting grade-level proficiency, and mastering grade-level proficiency. We use the 2021-2022 math school approaches grade level rate for 4th grade because a STAAR science test is not administered in 4th grade. We wanted to ensure it was (approximately) the same students contributing scores to a school's passing rate. Since the math and science components of the STAAR correlate highly (r = 0.85 in 2020-2021 5th grade), this is an appropriate way to ensure that there were no baseline differences across STEMscopes and non-STEMscopes matched schools in prior academic achievement.

We also downloaded 2022-2023 school year school enrollment data including race/ethnicity count data, student program data including count data for students considered economically disadvantaged, emergent English learners, students who were eligible to receive special education services, as well as gifted and talented program students. This file also includes information regarding which region of Texas a school was located in. We focused on these covariates as previous research has indicated they are associated with science achievement. All count data was then converted to percentage data by school (e.g., number of economically disadvantaged students/ total number of students in a school). These variables were used to match STEMscopes and non-STEMscopes schools (see participants section below for details on matching). Once matching was complete and baseline analyses were conducted (see baseline equivalence), we downloaded the 2023 STAAR campus data file. We analyzed school level data across "approaching, meets, and masters" grade level expectations. This file also includes sub-population percent meeting each benchmark score by school.

Missing data

If a variable (e.g., program category/sub-population) has less than 10 students, the data is removed by the state and "<10" is included instead to protect student privacy under the Family Educational Rights and Privacy Act (FERPA). This leads to missingness by design in both

5177 Richmond Ave., Suite 800, Houston, TX 77056
1-800-531-0864
stemscopes.com
stemscopes@acceleratelearning.com

covariate and outcome variables. We handled this missing data in two ways. For any covariate variable used to match schools, we used multiple imputation by chained equations via R-studio's "MICE" package with the "CART" imputation method (see Van Buuren and Groothuis-Oudshoorn, 2011). We use MICE procedures during this step to ensure complete data for matching procedures via R-Studio's "Match-it" package. Once data were matched, in all final analyses we used R-Studio's "Lavaan" package which uses full information maximum likelihood procedures to handle missing outcome data.

Participants

In the 2022-2023 school year, the overall number of regular public and public Charter Texas schools that purchased and used STEMscopes (in any capacity) for 5th grade was 2,809 out of 4,325 in-person schools (~65%). For the current report, virtual and alternative schools were excluded. Of the STEMscopes schools, 2,649 (94%) met the criterion of schoolwide usage of at least eight 5th grade science units (scopes) and were eligible for the study. Of note, 1,619 STEMscopes schools (58%) accessed *every* scope.

As mentioned previously, to match schools based on the data available from the TEA website we used R-Studio's "Match-it" package. Typically, we use the "optimal" matching method. This method matches every available school using propensity score matching. It circles through the school list to optimize the lowest propensity score between two matched schools. In the current data, there was a greater number of STEMscopes schools than non-STEMscopes schools. In this unusual situation, the "optimal" matching method will use every control school and leave out treatment schools that do not find a suitable partner. We included school level variables to match data: 4 race/ethnicity variables representing the percent of the school population that fit each category: White/Caucasian, Hispanic/Latinx, Black/African American, and Asian; as well as the total number of students enrolled in a school as well as the number of 5^{th} grade students, the percent of students that qualified for special education or gifted and talented programs, the percent of English language learners, the percent of economically disadvantaged students, whether the school was a public Charter or not, and a school's region in Texas. Finally, schools were matched based on 'baseline' academic achievement as indexed by the STAAR 2022 4th grade math "approaches" rate for a given school. The initial matching, with the optimal matching technique, included all 1,516 non-STEMscopes schools matched with 1,516 STEMscopes schools (excluding 1,133 treatment schools).

Unfortunately, this resulted in less-than-ideal matching results. Specifically, the average propensity score distance across schools was 0.34 (closer to zero is better), with some school pairs with distances greater than 1. Typical field standards suggest that propensity scores less than 0.50 between pairs is ideal (Huber, Lechner, & Wunsch, 2013). In addition, the standard mean difference between schools in relation to whether schools in the non-STEMscopes group and the STEMscopes group were Charters or not was outside of acceptable values based on the What Works Clearinghouse standards of baseline differences less than 0.25. The difference for Charter schools was 0.31, with double the number of Charter schools in the control group sample relative to the STEMscopes sample.

This represents a true difference in the Texas school population that chose to use STEMscopes versus not using STEMscopes. The control sample (all of whom were used with

5177 Richmond Ave., Suite 800, Houston, TX 77056
1-800-531-0864
stemscopes.com
stemscopes@acceleratelearning.com

optimal matching) does indeed include more Charter schools; thus, Matchit was unable to fully match on this category. This also lays bare the unusual situation wherein the treated (STEMscopes) population is larger than the control population. This suggests that optimal matching may be a poor technique as the control group may not be the best comparison when trying to capture a sample of "typical Texas" schools across the STEMscopes and non-STEMscopes comparison.

Bearing this in mind, we used an alternative approach. The results above represent an approach using the Match-it program, but with the Mahalanobis Distance matching method. Mahalanobis Distance is designed to consider the multivariate space between numerous covariates when matching. Specifically, rather than propensity scores, Mahalanobis Distance (Gu & Rosenbaum, 1993) is used as the distance metric, and it considered both more robust to multiple covariate usage and to correlated covariates (which these are). Finally, using Mahalanobis distance in Match-it has the added benefit wherein one can prespecify an allowable multivariate distance between matched school pairs. In this case, we used a conservative multivariate distance of 0.10. By setting a pre-specified distance, the program will not return school pairs that are too dissimilar to a degree outside this distance. Matching with the Mahalanobis Distance method resulted in 1242 non-STEMscopes schools (274 not matched) and 1242 STEMscopes schools (1370 not matched) for a total sample of 2484 matched schools, representing 57% of Texas Regular and Charter Elementary schools.

Baseline Equivalence

For all covariate variables (the variables used for matching) including baseline academic achievement, there were no significant differences between matched groups (see Table 2). As a reminder, we specified that the multivariate difference (that is, this distance considers *all* covariates at the same time, there still could be large differences in individual covariates) could not be greater than 0.10. However, the What Works Clearinghouse (WWC) standards require that baseline differences for a single variable with a standard mean difference greater than 0.05 must be controlled for statistically. Following the advice of Stuart, 2010, we include all covariates that were not collinear in the final analyses as a complementary approach to matching, and a more stringent test of effects. Several race/ethnicity variables were correlated above -0.70. Specifically, the percent of students that were White/Caucasian was highly negatively correlated with the percent of students who reported being Hispanic/Latinx. With this in mind, we present models above with only the variable representing the percent of Hispanic/Latinx students included in the model (however, we re-ran models that included White/Caucasian percent variable and the pattern of findings was the same). Inclusion of these covariates satisfies the WWC standard as several variables had standard mean differences greater than or equal to 0.05.

5177 Richmond Ave., Suite 800, Houston, TX 77056
1-800-531-0864
stemscopes.com
stemscopes@acceleratelearning.com

Table 2: Baseline comparison of matched STEMscopes and non-STEMscopes schools

Variables	State Total	Unmatched Sample Total	Matched Sample Total	Non-STEMscopes	STEMscopes	t-value	p-value	Effect Size
Baseline school 4th grade Math passing rate 2021	69%	66.91%	67.18%	67.00%	67.35%	0.50	0.62	0.02
School Size	-	530	516	508	524	1.82	0.07	0.07
Percent economically disadvantaged students	62.00%	66.78%	66.11%	66.40%	65.82%	0.54	0.59	0.02
Percent Black/African American students	12.62%	12.66%	12.56%	12.61%	12.50%	0.16	0.87	0.01
Percent Latino/Hispanic students	51.96%	54.96%	53.43%	53.13%	53.74%	0.52	0.60	0.02
Percent Asian students	5.47%	4.48%	4.77%	4.38%	5.16%	1.76	0.08	0.08
Percent White/Caucasian students	26.40%	27.04%	28.42%	28.90%	27.95%	0.88	0.38	0.04
Percent Charter Schools	11.03%	9.39%	10.55%	10.79%	10.31%	0.39	0.70	0.02
Percent of English Language Learners (ELLs)	23.02%	26.23%	26.84%	27.07%	26.61%	0.53	0.60	0.02
Percent of special education students	12.74%	14.22%	13.94%	13.99%	13.89%	0.43	0.66	0.02
Percent of gifted and talented students	8.22%	6.95%	7.11%	7.04%	7.18%	0.53	0.60	0.03

5177 Richmond Ave., Suite 800, Houston, TX 77056
1-800-531-0864
stemscopes.com
stemscopes@acceleratelearning.com

Planned analyses

Analyses were conducted with R-studio's Lavaan structural equation modeling package because this package includes estimation with full information maximum likelihood (FIML) to handle missing data. FIML procedures to handle missing data estimation ensure that in the final analyses the estimation is not biased. Our main variables of interest were the 2023 5th grade science "approaches, meets, and masters" school rates. We also include sub-population analyses for ELL and non-ELL students, economically disadvantaged students, and for two racial/ethnic categories (Black/African American and Hispanic/Latinx). All other racial/ethnicity percent passing variables included too much missingness to be considered (e.g., greater than 50% missing).

Conclusion

Findings provide evidence of the efficacy of the STEMscopes Science 5th grade curriculum. The STEMscopes Science curriculum was associated with a small but significant point increases in the average percent of schools' students "Meets or Masters" rates relative to schools that did not use STEMscopes and in consideration of different sub-groups of students. The effect sizes associated with these increases indicated small effects based on field standards for psychological research (e.g., Brydges, 2019). Although effect sizes were small, they are consistent with numerous years of evaluation in Texas regarding the effectiveness of STEMscopes. We also note that the format of the 2023 STAAR assessment was re-designed, and this may have impacted scores for this first year of use. The STEMscopes Science curriculum has since redesigned our entire platform to ensure that Texas students are receiving the best resources to succeed in the classroom. Nevertheless, as a practical measure of the current effect size, we multiplied the overall average estimated change in the percent of students who met or mastered STAAR grade level expectations for STEMscopes schools (versus non-STEMscopes schools) times the total number of students tested in STEMscopes schools. Given that schools were closely matched regarding enrollment, this provides a rough estimate of the number of students who more likely met or mastered their STAAR 2023 5th grade science assessment in relation to their school using STEMscopes Science. Specifically, we estimate 976 additional students met the 2023 STAAR science achievement test, and 851 additional students mastered grade level expectations in STEMscopes schools.

5177 Richmond Ave., Suite 800, Houston, TX 77056
1-800-531-0864
stemscopes.com
stemscopes@acceleratelearning.com

Works Cited

Brydges, C. R. (2019). Effect size guidelines, sample size calculations, and statistical power in gerontology. *Innovation in aging*, *3*(4).

Huber, M., Lechner, M., & Wunsch, C. (2013). The performance of estimators based on the propensity score. Journal of Econometrics, 175(1), 1-21.

Gu, X. S., & Rosenbaum, P. R. (1993). Comparison of multivariate matching methods: Structures, distances, and algorithms. Journal of Computational and Graphical Statistics, 2(4), 405-420.

Morgan, P.L., Farkas, G., Hillemeier, M. M., and Maczuga, S. (2016). Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educational Researcher*, 45, 18-35.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, *25*, 1.