Retrospective Study on the Effects of Accelerate Learning STEMscopes Science in Grade 5 on STAAR Science Achievement in 2018-19

Michael A. Cook, PhD
Elizabeth Humberstone, PhD
Steven M. Ross, PhD

December 2023

JOHNS HOPKINS
SCHOOL of EDUCATION

Center for Research and Reform in Education

Contents

Retrospective Study on the Effects of Accelerate Learning STEMscopes Science in Grade 5 on STAAR Science Achievement in 2018-19

The purpose of the present study was to evaluate the efficacy of Accelerate Learning's STEMscopes Science curriculum on science achievement, as measured by the State of Texas Assessments of Academic Readiness (STAAR) Science test. The present retrospective quasi-experimental design (QED) examined secondary data from the 2018-19 school year. STAAR Science proficiency percentages were compared between schools that used the STEMscopes Science curriculum in 2018-19 and propensity-matched comparison schools that did not use STEMscopes.

Research Questions

Research questions for this study include:

1. What is the impact of the STEMscopes Science curriculum on STAAR Science proficiency rates for Grade 5 students?
    a. Do impacts differ across subgroups (i.e., race/ethnicity, gender, EL status)?

Method

Research Design

This study was a retrospective quasi-experimental design (QED) that examined school-level proficiency rates of Grade 5 students on the STAAR Science assessment in the 2018-19 school year. Accelerate Learning provided the Center for Research and Reform in Education (CRRE) with lists of schools that used STEMscopes in the 2018-19 school year. Comparison schools were identified using propensity-matching procedures (described in more detail below). Data sources included publicly available data from the Texas Education Agency from the 2018-19 school year.

The main outcome variable of interest was the percentage of students in a school that met or exceeded the state benchmark of "approaching grade level." The main analyses in this study included school-level linear regression models comparing proficiency rates of students in treatment (STEMscopes) and matched comparison schools, controlling for a host of school-level demographic variables.

Participants

Participant schools included those who used the STEMscopes Science curriculum in Grade 5 in 2018-19, as well as matched comparison schools identified by propensity-matching procedures. STEMscopes Science was used by 2,379 schools across 427 districts in 2018-19.

To be considered a "treatment" school, the school had to have nonzero Grade 5 usage data, as measured by tracking of web page visits. As defined by Accelerate Learning for sufficient implementation, treatment schools must have used at least 80% of topics (scopes). There are a total of 18 scopes for Grade 5; thus, treatment schools were expected to use at least 14 scopes in Grade 5. Schools that used STEMscopes but did not reach the 14-scope cutoff were not classified as treatment schools and were excluded from the potential pool of matched comparison schools.

Initially, 1,905 schools with Grade 5 used STEMscopes in 2018-19. Of these schools, 1,070 schools with Grade 5 classrooms used at least 14 scopes; these were classified as treatment schools. Of these treatment schools, complete data were available for 846 schools with Grade 5 classrooms. Table A2 provides a comparison of the school demographic characteristics for the treatment and matched controlled schools, including percent of the student population that were White, female, Latino, and FARMS-eligible.

Measures

Data sources for the current study include:

STAAR Science. School-level STAAR Science rates from Grade 5 in the 2018-19 school year were used as the primary outcome variables in these analyses. STAAR Science is only administered in Grades 5 and 8, and students are classified as achieving at one of four levels: "Did Not Meet Grade Level" (Level 1), "Approaching Grade Level" (Level 2), "Meets Grade Level" (Level 3), and "Masters Grade Level" (Level 4). Accelerate Learning defined "passing" the STAAR Science assessment as achievement at "Approaching Grade Level" (Level 2) or higher. Thus, the main outcome variables in these analyses were the combined school-level percentages of students at "Approaching Grade Level" or higher on the STAAR Science assessment.

STAAR Math. As the STAAR Science assessment is only administered in Grades 5 and 8, no direct state-administered science pretest assessment was available. For this reason, school-level Grade 4 STAAR Math proficiency levels from 2017-18 were used as the pretest measures in all analyses. STAAR Math uses the same four proficiency levels as are used in STAAR Science (and in all STAAR assessments).

Demographics. The Texas Education Agency (TEA) provided publicly-available datasets of school- and district-level demographic data. A variety of files, some at school-level and some at district-level, were combined to obtain as much demographic

data as possible. Counts of students by gender and ethnicity subgroups were available and computed into percentages, based on total numbers of Grade 5 students at each time-point. Economically disadvantaged status was available in a school-level file examining student programs and special school sub-populations.

Analytical Approach

Data for all Texas schools were analyzed descriptively by examining patterns in school-level proficiency levels, as well as by examining school-level proportions of student subgroups of interest, such as race/ethnicity and special education students. All analyses were performed using R version 4.3.0. Originally, Hierarchical Linear Modeling (HLM) was proposed for the main analyses, with schools nested within districts. However, a large number of districts contained only one school with Grade 5 students. Due to the large number of "one school" districts, we lacked the necessary sample size to use models that account for clustering of schools within districts, and we necessarily restricted our analyses to school-level variables. The school-level demographic variables included as covariates in regression analyses include percentages of female, Latino and White students, ELL status, economically disadvantaged status, and prior math achievement. These demographic variables were initially included in all analytic models.

To adjust for prior achievement and demographic differences between treatment and comparison groups, propensity score matching (PSM) was used to create comparison groups of schools that were as similar as possible to treatment schools. Propensity scores were computed using the MatchThem package in R, with one-to-one matching using optimal pair matching. Matching variables included the school demographic variables detailed above, as well as percentages of at risk and special education students, school size, schools' regions, and school-level prior math achievement variables. This procedure created comparison groups that were of equal size to the treatment group and, based on prior achievement and demographic variables, as similar as possible to treatment schools. Thus, sample sizes throughout the report reflect half treatment schools and half control schools unless otherwise noted.

The data had a high degree of missingness: in the 2018-19 academic year, approximately 21% of observations were missing at least one variable. STAAR proficiency rates were the most missing variables for the 2018-19 school year. Specifically, Latino prior math achievement proficiency rate was the most missing variable, with 9% of cases missing for Grade 5.

To account for this missing data throughout the datasets, multiple imputation by chained equations was done using the 'mice' package in R (15 imputations, 10 iterations per imputation). Analyses were then run with the 'mice' package's pool function. The results reported in the main regression analyses are from the imputed data. Complete case analyses were also run and are reported in Appendix B.

Results

We begin by descriptively examining STEMscopes usage across Grade 5 and each cohort. This is followed by regression analyses that examine the impacts of the STEMscopes curriculum on STAAR Science proficiency rates in the 2018-19 school year.

STEMscopes usage. Table 1 shows average STEMscopes usage by cohort. We examine usage for schools that met the 14-scope criterion and were considered treatment schools, as well as usage across all schools, including those that did not meet the 14-scope criterion and were thus excluded from further analysis in this study. In Table 1 below, Treatment Schools are those using 14-scopes or more; All Schools indicate the treatment schools plus schools that used less than 14-scopes.

Table 1

Average STEMscopes Usage Across Schools Using Curriculum

	Grade 5 Scopes	Grade 5 *N*
Treatment Schools	17.64 (2.26)	1,070
All Schools	12.95 (6.21)	1,905

Note. SD in parentheses.

Grade 5 treatment schools averaged approximately 18 scopes completed. When considering all schools, Grade 5 treatment schools averaged approximately 13 scopes completed in 2018-19.

Average Proficiency Rates. Table 2 shows the average "Approaching Grade Level" proficiencies for all schools in the analytic sample including those that were not matched (Overall Mean), treatment schools (Treatment Mean), and matched control schools that did not use STEMscopes (Control Mean). The table includes means from one imputed data set and from only complete cases as a comparison.

Table 2

Average "Approaching Grade Level" Proficiency for Treatment and Matched Control Schools for the 2018-19 School Year

	Overall Mean	Treatment Mean	Control Mean
Imputed (*n* = 3,211)	72.62 (16.23)	74.45 (14.17)	73.26 (16.35)
Complete Cases (*n* = 2,528)	72.93 (15.01)	74.21 (13.90)	73.28 (15.43)

Note. SD in parentheses.

The averages from the imputed data set were similar to the dataset restricted to only complete cases for Grade 5. The standardized mean differences between the treatment schools and matched controls were only .08 SDs for imputed data and .06 SDs for complete cases.

Impact Analyses

Table 3 shows the regression impact estimates for the STEMscopes curriculum on school-wide STAAR Science proficiency rates for Grade 5 students in the 2018-19 school year for the 963 treatment schools and 963 matched control schools. Separate regression analyses were conducted to examine the impact of STEMscopes on each of the Approaching, Meets, and Masters STAAR Science proficiency levels. Note that each successive proficiency level assumes any prior level proficiency. That is, "Masters" aggregates frequencies from its level combined with the "Approaching" and "Meets" levels.

Table 3

Impact of STEMscopes in the 2018-19 School Year on Grade 5 School-wide STAAR Science Proficiency Rates (n = 1,926)

Level	Estimate	Standard Error	*p* value
Approaching	0.957^	0.535	.075
Meets	1.415*	0.609	.021
Masters	1.104*	0.459	.017

Note. ^ $p < .10$; * $p < .05$.

Impacts of the STEMscopes curriculum on school-wide proficiency rates were directionally positive and met statistical significance at the "Meets" and "Masters" STAAR Science proficiency levels. The regression estimates in these analyses can be interpreted as the average increase in STAAR Science proficiency rates for schools that used STEMscopes, in relation to comparison schools that did not use STEMscopes, while the constant can be interpreted as the regression-adjusted proficiency level for the comparison group. For example, STEMscopes schools had "Meets" STAAR Science proficiency rates that were approximately 1.41 points larger than those in comparison schools, after controlling for prior year mathematics proficiency rates and school-level demographic variables. Similar magnitude impacts were observed for "Approaches" and "Masters" STAAR Science proficiency levels.

Table 4 shows the results of similar analyses conducted on student subgroups of interest, which included female, Latino, and FARMS students. These regression analyses replicate the main regression analyses, but only for data reported for subgroups at the school level. Thus, these analyses are examining the impact of STEMscopes on school-wide proficiency rates for female, Latino, and FARMS students.

Table 4

Impact of STEMscopes in the 2018-19 School Year on Grade 5 STAAR Science Proficiency Rates, Student Subgroups (n = 1,926)

Level	Estimate	Standard Error	*p* value
Female			
Approaching	1.400*	0.618	.025
Meets	1.770*	0.691	.011
Masters	1.084*	0.507	.033
Latino			
Approaching	0.938	0.636	.142
Meets	1.145	0.721	.113
Masters	0.655	0.531	.218
FARMS			
Approaching	0.833	0.636	.192
Meets	0.950	0.694	.172
Masters	0.648	0.474	.172

Note. * $p < .05$.

The regression estimates in this table can be interpreted as the average increase in STAAR Science proficiency level for schools using STEMscopes, in relation to comparison schools. STEMscopes impacts were generally directionally positive, with impacts on female students reaching statistical significance. Proficiency rates for females in STEMscopes schools were between 1.08 and 1.77 points higher than those for female students in comparison schools. Similar patterns of directional impacts were generally observed for Latino and FARMS students, as well.

Table 5 examines the associations of STEMscopes usage, as defined by the number of completed scopes at a school, with proficiency rates.[1] Schools were categorized into STEMscopes usage groups of 1-7 scopes used, 8-14 scopes used, and 15+ scopes used, with 0 scopes used as the reference category.

Table 5

Usage Analysis of STEMscopes in the 2018-19 School Year on Grade 5 STAAR Science Proficiency Rates (n = 4,002)

Level	Estimate	Standard Error	*p* value
Approaching			
1-7 Scopes	-0.49	0.62	.43
8-14 Scopes	0.66	0.60	0.27

[1] Usage analyses used a separate imputation from the main analysis.

15+ Scopes	1.48***	0.46	<.001
Meets			
1-7 Scopes	-0.69	0.74	.35
8-14 Scopes	0.21	0.74	.77
15+ Scopes	2.21***	0.55	<.001
Masters			
1-7 Scopes	-0.23	0.58	.69
8-14 Scopes	0.09	0.56	.87
15+ Scopes	1.82***	0.43	<.001

Note. *** $p < .001$.

These dosage analyses showed that associations between using 15 or more scopes at a school and STAAR Science proficiency rates were directionally and significantly positive. Compared to schools using no scopes, schools using 15+ scopes were associated with between a 1.5 to 2.2-point increase in the science proficiency rates at all levels.

Summary and Discussion

The purpose of this evaluation was to examine the impacts of the STEMscopes Science curriculum on school-level STAAR Science proficiency rates across the state of Texas in the 2018-19 school year. Publicly available datasets from the Texas Education Agency (TEA) were examined as the main data sources of this evaluation. Outcome variables included school-level STAAR Science proficiency rates at the "Approaching," "Meets," and "Masters" levels, and impacts were estimated across all students, as well as for subgroups of interest. Linear regression modeling was used to examine STEMscopes' impact on STAAR Science proficiency rates, controlling for school-level demographics and prior year mathematics proficiency rates. Additional analyses examining the associations between counts of scopes completed and school-level proficiency rates were conducted, allowing for an examination of the impacts of different usage levels of the STEMscopes program on science proficiency rates. This report focuses on results for Grade 5 in the 2018-19 school year.

Across the present analyses, STEMscopes impacts were generally small but positive and significant. Specifically, Grade 5 students in STEMscopes schools were significantly more likely than those in comparison schools to achieve "Meets" and "Masters" proficiency levels. Similarly, impacts across subgroups were generally positive, but small in magnitude. One notable exception is Grade 5 female students, where a significant positive impact of STEMscopes on proficiency rates was evidenced across all three proficiency levels.

Analyses that examined the associations between the number of scopes completed and science proficiency rates showed varied patterns of associations. Notably, in relation to schools using no scopes, there was a significant and positive

association between using 15+ scopes and a 1.5 to 2.2-point increase in science proficiency for 5th grade schools in 2018-19.

Some important limitations of this evaluation should be noted. First, the unit of analysis was the school; while coverage of the entire state of Texas provided generalizability and representativeness in the analytic sample, this resulted in very coarse-grained analyses. Student-level analyses may be more sensitive to program impacts. In addition, no data were available regarding STEMscopes program implementation in schools or classrooms, outside of school-level counts of completed scopes. Future research may focus on a small number of districts and compare student-level achievement scores for science classrooms that use STEMscopes to science classrooms that do not use the program.

Appendix A: Baseline Equivalence Tables

Table A1

Baseline Equivalence of Prior Math Proficiency Level for the 2017-18 School Year

	Treatment Mean	Comparison Mean	Standardized Mean Diff.
Imputed	77.34 (12.71)	77.17 (14.96)	0.01
Complete Cases	77.17 (12.52)	76.69 (13.88)	0.04

Note. SDs in parentheses.

Table A2

School Demographic Variables, in Average Percentages, for the 2018-19 School Year

Variable	Grade 5 Complete Cases	Grade 5 Imputed
Female		
Treatment	48.75	48.75
Comparison	48.65	48.69
FARMS		
Treatment	66.63	65.58
Comparison	67.35	65.13
Latino		
Treatment	59.10	56.83
Comparison	59.54	56.16
White		
Treatment	24.59	26.14
Comparison	24.30	26.73

Appendix B: Complete Case Analysis Tables

Table B1

Impact of STEMscopes in 2018-19 School Year on Grade 5 School-wide STAAR Science Proficiency Rates (n = 1,692)

Level	Estimate	Standard Error	*p* value
Approaching	0.527	0.489	.281
Meets	0.718	0.589	.223
Masters	0.450	0.457	.325

Table B2

Impact of STEMscopes in 2018-19 School Year on Grade 5 STAAR Science Proficiency Rates, Student Subgroups (n = 1,692)

Level	Estimate	Standard Error	*p* value
Female			
Approaching	0.861	0.568	.130
Meets	0.754	0.658	.252
Masters	0.257	0.492	.601
Latino			
Approaching	0.458	0.589	.437
Meets	0.163	0.672	.808
Masters	-0.152	0.503	.762
FARMS			
Approaching	0.540	0.586	.357
Meets	-0.007	0.651	.992
Masters	0.238	0.450	.597