An Early Investigation of the Let's Go Learn Edge Program: Analyzing Program Impact after an Initial Implementation Year

---

James Marshall, Ph.D.
Kimberly Cavoto, Ph.D.

December 2020

Purpose:	To investigate Let's Go Learn Edge program efficacy one year post implementation, as measured by California's Smarter Balanced Summative Assessments.
Prepared for:	Let's Go Learn
Preferred citation:	Marshall, J. & Cavoto, K. (2020). *An early investigation of the Let's Go Learn Edge program: Analyzing program impact after an initial implementation year.* James Marshall Consulting, Inc.
Date of Publication:	December 18, 2020
Author Contact:	info@jamesmarshallconsulting.com

Table of Contents

Table of Figures

Table of Tables

An Early Investigation of the Let's Go Learn Edge Program: Analyzing Program Impact after an Initial Implementation Year

Abstract

Let's Go Learn Edge is a technology-delivered instructional program that provides a personalized learning experience to students of all ages in the subjects of reading and mathematics. Leveraging results from the publisher's diagnostic assessments, English/Language Arts Edge (ELA Edge) and Math Edge provide student-specific learning experiences that align with assessed ability levels.

This retrospective study was conducted with the support of a central California school district, and employed district-wide state achievement data to isolate program effects following the initial year of Edge implementation. Students who completed more than two Edge lessons within a subject area during the 2018-19 school year were compared to those who did not use the program, or used it for a minimal time. Baseline achievement was established using the spring 2018 state test data, and gain scores were calculated using spring 2019 testing data. Baseline scores were then used as a covariate to remove small differences that were observed between the Edge and comparison groups.

For the English Language Arts/Literacy (ELA) assessment, mean gain scores favored the LGL group (combination of all grade 4-6 students), and at each grade level analyzed (grades 4, 5, and 6). Gain score differences for all students and grade 6 students proved statistically significant (unlikely the result of random chance, $p < .006$ and $p < .001$ respectively). Effect sizes were small, with .003 for all students and .011 for grade 6. For the Mathematics (Math) assessment, mean gain scores favor the LGL group overall, and at each grade level analyzed (grades 4, 5, and 6). Gain score differences for all students, grade 4, and grade 6 students proved statistically significant (unlikely the result of random chance, $p=.000$). Effect sizes were small and ranged from .013 to .029.

Limitations to the investigation include inequalities between treatment and comparison groups, and possible bias regarding selection of students who used the LGL program. Following maturation of the program's implementation, future research should be conducted to optimize a rigorous plan, including purposeful assignment to treatment and comparison groups.

Introduction

From K-12, college, and far beyond, a student's ability to read, comprehend and communicate with varying levels of complexity is vital to understanding and interacting with every subject and discipline. This also applies to the ability to perform mathematical operations, which drive countless tasks and decisions. In a data-centric world, logic and reasoning cannot be understated.

Yet, students in the United States continue to face steep challenges in mastering these important gate-keeping skills.

The *National Assessment of Educational Progress*' "Nation's Report Card" provides the following statistics related to these skills from the most recent national testing data:

- Reading results from 2019 indicate just 34% of grade 8 students and 35% of grade 4 students performed at the Proficient level (or above); these percentages represent a *decrease* of two percentage points each, relative to 2017 results.
- Of the 2019 percentages reported above, just 4% of grade 8 and 9% of grade 4 students tested Advanced in reading achievement. The percentage of students testing Advanced has remained the same or has varied by one percentage point since 1992 (grade 8) and 2003 (grade 4).
- Mathematics results from 2019 indicate 34% of grade 8 students and 41% of grade 4 students performed at, or above, the established Proficient level; these percentages represent no change for grade 8 students, and a one percentage point increase for grade 4 students, relative to 2017 results.
- Of the 2019 percentages for mathematics, just 10% of grade 8 and 9% of grade 4 students tested Advanced in mathematics achievement. The percentage of students testing Advanced has remained the same or varied by one percentage point since 2017 (National Center for Educational Statistics, 2019).

These nationwide results vividly illustrate the challenges faced by young people across the country. Such challenges are often amplified for students who speak a language other than English, and/or have limited financial resources.

Technology-delivered instruction holds promise for turning these trends around. The increasing sophistication of technology-based interventions, including the ability to assess each student's achievement levels and then personalize instruction, stands to optimize learning for students. Whether students are beginning a program below grade level or seeking to accelerate their learning beyond their peers, personalized learning solutions can offer them a "right time, right level" solution for success.

While definitions of personalized learning (PL) vary, a national study conducted by the RAND Corporation and funded by the Bill and Melinda Gates Foundation offered the following global description:

> Personalized learning prioritizes a clear understanding of the needs and goals of each individual student and the tailoring of instruction to address those needs and goals. These needs and goals, and progress toward meeting them, are highly visible and easily accessible to teachers as well as students and their families, are frequently discussed among these parties, and are updated accordingly (Pane et al., 2017).

Pane et al. (2017) also assert that technology can facilitate personalized learning. Among its potential contributions is managing the complexity demanded by the personalization process. With seemingly infinite student abilities and learning needs, juxtaposed against diverse pathways through a curriculum, technology's promise to optimize such complexities is all but requisite. In a review of 71 studies that addressed personalized learning, Zhang et al. (2020) note that the majority of the reviewed studies, especially those with technology-supported personalized learning, were associated with positive results. However, these authors also caution about these early findings when they suggest "that while generally positive, the empirical research examining the effects of PL on PK-12 student learning outcomes is still in its initial stages of development. The intermix of educator and student activities with appropriate use of technology and the UDL framework have shown to be potential contributing supports for educators and researchers to design and implement PL" (p. 12).

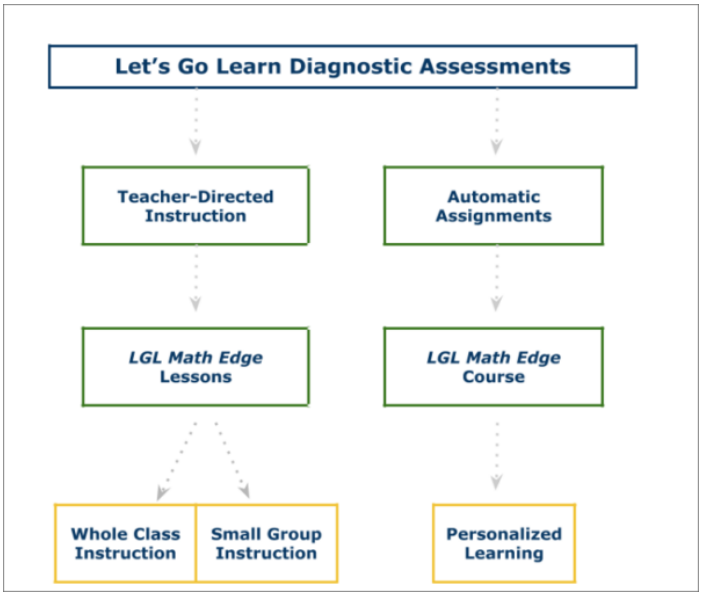
This brief introduction highlights how critical detailed and accurate student performance data is for expanding the personalized learning marketplace. Let's Go Learn (LGL), an online assessment and instructional technology publisher with two decades of experience developing diagnostic assessments in both English/Language Arts (ELA) and Mathematics, created LGL ELA Edge (ELA Edge) and LGL Math Edge (Math Edge), which each provide technology-based instruction through personalized learning.

This study investigated first-year results following the implementation of these Edge programs in a large California school district.

Let's Go Learn Solutions

Founded in 2000, Let's Go Learn offers a range of solutions designed to improve student achievement.

Figure 1: Let's Go Learn Solution Ecosystem



Initially, the company developed two diagnostic assessments: (1) Diagnostic Online Reading Assessment (DORA); and (2) Adaptive, Diagnostic Assessment of Mathematics (ADAM). These comprehensive assessment tools were designed to provide students, educators, and families with clear and actionable data related to student performance throughout an academic year. Each assessment is comprised of subtests that assess areas such as vocabulary and reading comprehension to develop a complete picture of student ability and opportunities for growth so that teachers could direct their instruction accordingly.

In addition to the diagnostic assessments, the LGL leadership sought to provide a technology-based instructional solution that could leverage the diagnostic data to build personalized learning experiences for young people.

The resulting learning solutions, ELA Edge and Math Edge, use assessment results to offer learners engaging, personalized instruction in over 300 gamified and interactive lessons. The programs allow teachers to immediately implement personalized learning and provide targeted activities to support their existing classroom learning initiatives. After students engage in the platform, educators can access reports by student, classroom, or site, to inform learning plans, determine support efforts, and elevate conversations with stakeholders.

Each personalized course provides explicit instruction and introduces concepts via animations, songs, and graphics. The instruction is intended to engage students as they learn skills and demonstrate knowledge during gamified instructional quizzes. As students work through the courses, the responsive platform employs direct instructional feedback. If, during a quiz, a student has an incorrect answer the platform helps them understand why and practice the right steps for mastery.

As students work with each EDGE program, the platform captures their progress and creates progress reports that are instantly available for stakeholders to review. Teachers can explore sharable reports for an individual student or for the whole class, and use the data to plan whole-class instruction or target skill gaps for scaffolding.

Let's Go Learn: Program Effectiveness

Case studies and data analysis investigating the efficacy of LGL programs have been conducted with participating school districts. The following two examples typify early investigations of program impact.

- Sussman Middle School, one of four middle schools within Downey Unified School District in California, implemented Math Edge in January 2018. After less than half a year, the Edge-implementing school outperformed all other middle schools in the district. That same school made significant gains in both 6th and 8th grades on the Smarter Balanced Summative Assessment, the state assessment test.
- In 2016-17, Jersey City Public School (JCPS) began a large-scale implementation of ELA Edge in grades 3 to 8. Data representing a grade 4 cohort indicated that ELA Edge helped JCPS: Elevate conversations across assessments and instruction; integrate individualized reading instruction into classrooms for improved learning; and demonstrate significant gains in the grade 4 Partnership for Assessment of Readiness for College and Careers (PARCC).

Complete reports are available from the publisher at www.letsgolearn.com.

Research Questions

This study investigated the following research questions to determine any relationship between the use of ELA Edge and Math EDGE, and student achievement:

1. Does the academic performance of students using ELA Edge differ from that of their non-using peers, as measured by California Assessment of Student Performance and Progress' (CAASPP) Smarter Balanced Summative Assessment-derived ELA gain scores following the program's initial, full implementation year?
2. Does the academic performance of students using Math Edge differ from that of their non-using peers, as measured by CAASPP Smarter Balanced Summative Assessment-derived Math gain scores following the program's initial, full implementation year?
3. To what extent do gain scores differ based on group membership defined by the student's CAASPP score-derived proficiency level category established at baseline?

Methods

This study was conducted retrospectively through the cooperation of the participating school district. It benefitted from achievement scores from the CAASPP Smarter Balanced Summative Assessments in ELA and Math.

Research Design

Univariate effects of intervention condition on gain score measures were examined using between-subjects analysis of covariance (ANCOVA), adjusting for pretest scores. The study employed a quasi-experimental, retrospective design with post-hoc assignment of students to the ELA Edge and comparison group, and the Math Edge and comparison group.

Measures

The study benefited from data provided by Edge that was used to quantify system use, which included metrics of intervention time, lessons attempted, and lessons completed. Additionally, the participating school district, following human subjects review, provided complete state testing records for each student in the district for the 2018 and 2019 testing administrations. The full complement of involved measures is described below.

Edge Use

ELA Edge and Math Edge was used throughout the district for the full 2018-19 school year. These data were used, retrospectively, to place students with valid Smarter Balanced Summative Assessment data into either the LGL or comparison group.

LGL's assessment and learning management system tracks student performance on LGL diagnostic assessments (DORA and ADAM) and records varied dimensions of ELA Edge and Math Edge use. Recorded data include the number of lessons attempted, the number of lessons successfully completed, and the related time intervals in which such activity occurred. Successful completion was defined as achieving a mastery score of 80% on a lesson-specific assessment, which is completed as a final task for each lesson.

Academic Achievement

Results from the CAASPP's Smarter Balanced Summative Assessments in ELA and math were employed to develop the independent variable for this study. These assessments, commonly referenced collectively as the "California state test," are conducted each spring and described by the state as an "academic check-up for students in grades 3-8 and grade 11" (California Department of Education, n.d.). The California Department of Education highlights the use of vertical scaling that makes Smarter Balanced scale scores comparable over time:

> Because of the vertical scaling of the Smarter Balanced assessments, scale scores for a test may be compared to scale scores for the same student or groups of students in different years for the same content area, as well as for between specific grade levels and content areas. This allows users to say that achievement for a given content area and grade was higher or lower one year as compared with another. Scale scores for the Smarter Balanced assessments may be compared across grades since the scales are vertically aligned across grades. Scores for the paper-pencil versions of the Smarter Balanced Summative Assessments are linear forms but have the same scale as the online tests (California Department of Education, 2019, p. 24).

In addition, and specific to each subject area and grade level, scale scores align with one of four proficiency categories used to classify a student's academic performance.

The research dataset included state assessment results from the 2017-18 school year, with testing conducted in spring 2018 as a baseline measures of academic performance. Results from the same state assessments from the 2018-19 school year, with testing conducted in spring 2019, were used as post-implementation measures of academic performance. Using these two data points in time, gain scores were calculated for both ELA and Math by subtracting the 2018 baseline scale score from the 2019 scale score. Gain scores were then used as the dependent variable for all analyses.

Attendance

The impact of school attendance on academic achievement has been historically documented across countless studies (see, for example, Romero & Lee, 2007; Ginsberg et al., 2014; Gottfried, 2015). Given the likelihood of attendance rates influencing analysis results, the research design attempted to use attendance as a covariate in support of the investigation of Research Question 3. Attendance data was provided by the school district based on figures reported to the State of California during the 2018-19 school year. The attendance rate was defined as the number of days students reported absent during the defined school year.

Participants

This study benefitted from the cooperation of an urban school district in Southern California. The K-12 district operates 19 schools, 12 of which are elementary level. All 12 schools were included in the study. The district serves almost 14,000 students, of whom 94% are Latino, almost 5% African American, 1% White, and the remaining small percentages Asian, Pacific Islander, Native American, and multi-racial. Student gender figures suggest approximately one-half female, and one-half male. Within the district, approximately 93% of students qualify for free or reduce lunch.

Recent academic performance figures for the district place approximately 35% of students proficient in Reading/Language Arts and 24% proficient in Mathematics. The average graduation rate is 83%.

The district reviewed the analysis plan, and then provided the researchers will complete testing records, as held by the California Department of Education, for years 2017, 2018, and 2019. In addition, attendance data and LGL-collected program usage data were provided.

Group Membership

Beginning with the full set of state test data for 2018 and 2019, several initial analyses were used to establish group membership. This involved designating each individual student record into one of the following three classifications: (1) excluded from study; (2) included as comparison student; or, (3) included as an Edge student. Designations for ELA Edge and Math Edge were made independently of one another. Thus, based on available data, a student could be designated as ELA Edge only, Math Edge only, or both ELA Edge and Math Edge. The analysis approach treated each subject area independent from the other.

The following decision points were used to place group members into one of the three categories, as defined in the previous paragraph.

Initial consideration for placement into the comparison or treatment group required that a student must have met each of the following criteria: (a) been enrolled in elementary school in the participating district during both 2017-2018 and 2018-2019; and (b) had valid CAASPP scores (as indicated within State of California testing records) for both 2018 and 2019. These criteria necessarily eliminated younger students in grades K through 3, as Smarter Balance tests are administered to students in grade 3 and above.

With usage metrics provided by the LGL management system, ELA Edge and Math Edge students were initially identified as those who used the corresponding program in 2018-2019 (either ELA or Math). This necessarily meant they had a recorded value for total number of lessons completed. To be classified into either Edge subject-based category (treatment), the student must have: (a) completed more than two lessons; or, (b) completed less than three lessons if their proficiency level met or exceeded the standard. These parameters were set by the researchers with the goal of including a wide range of LGL-using students, as measured by time on the system. The latter criterion ("b" stated above) was met for just 2.4% for ELA Edge

and 2% for Math Edge. Conversely, comparison students were defined as: not participating in any lessons (77.8% for Comparison ELA and 76.3% for Comparison Math); or, completing less than three lessons if their proficiency level was not met or nearly met (22.2% for Comparison ELA and 23.7% for Comparison Math). Table 1 summarizes the resulting group membership for the full sample, and by grade level.

Table 1: Sample Size by Group, before Weighting and based on Lessons Completed

Group	ELA		Math	
	LGL # Students (>2 lessons)	**Comparison # Students (<3 lessons)**	**LGL # Students (>2 lessons)**	**Comparison # Students (<3 lessons)**
All Students (grades 4-6)	2,211	602	2,119	702
Grade 4	836	107	815	128
Grade 5	816	108	748	179
Grade 6	559	387	556	395

In addition, results were analyzed based on proficiency categories. The following table provides the number of students, by proficiency category and grade level, prior to weighting the sample (see following section for weighting detail).

Table 2: Sample Size by Proficiency Group, before Weighting

Group	ELA		Math	
	LGL # Students (>2 lessons)	**Comparison # Students (<3 lessons)**	**LGL # Students (>2 lessons)**	**Comparison # Students (<3 lessons)**
All Students (grades 4-6)				
Standard not met	757	297	664	365
Standard nearly met	545	142	671	233
Standard met	554	115	516	66
Standard exceeded	355	48	268	38
Grade 4				
Standard not met	254	53	218	55
Standard nearly met	240	26	203	37
Standard met	196	17	258	21
Standard exceeded	146	11	136	15
Grade 5				
Standard not met	318	56	219	82
Standard nearly met	184	19	271	72
Standard met	181	17	177	18
Standard exceeded	133	16	81	7
Grade 6				
Standard not met	185	188	227	228
Standard nearly met	121	97	197	124
Standard met	177	81	81	27
Standard exceeded	76	21	51	16

LGL Group: Edge Usage

The analyzed LGL groups were comprised of students who had completed more than two lessons using either ELA Edge or Math Edge (or both, for each respective group). The following table details Edge usage by student group for both ELA and Math.

Table 3: Edge Program Use Description, Prior to Weighting

		Total Lessons				Unique Lessons			
Group	**N**	**Mean**	**SD**	**Low**	**High**	**Mean**	**SD**	**Low**	**High**
ELA									
All Students	2,211	24.3	23.1	0	211	17.7	13.7	0	99
Grade 4	836	28.7	26.0	0	211	21.2	15.5	0	90
Grade 5	816	21.0	20.2	0	188	15.8	12.2	0	72
Grade 6	559	22.5	21.1	0	180	15.5	12.0	0	99
Math									
All Students	2,119	23.3	25.3	0	238	12.7	10.5	0	89
Grade 4	815	25.3	24.8	0	201	13.7	10.4	0	89
Grade 5	748	19.0	20.2	0	211	11.0	9.0	0	88
Grade 6	556	26.1	30.8	0	238	13.30	12.0	0	60

Equating Groups

Following the initial group assignments, raw distributions between the comparison group and each subject-based Edge group was examined to ensure they did not differ in ways that would bias findings. Using Analysis of Variance (ANOVA) procedures, we identified significant differences between groups based on gender, Individual with Disabilities Act (IDEA) indicator, and economic disadvantage. Groups were not significantly different on demographic variables related to ethnicity, primary language, or parent education level, with two small exceptions cited below. Table 4 summarizes the initial demographic comparison which predicated the weighting scheme's development.

Table 4: Significant Group Differences, Prior to Weighting

	ELA		Math	
Group	**LGL**	**Comparison**	**LGL**	**Comparison**
All Students				
N	2,211	602	2,119	702
Gender = Male	51%	54%	51%	52%
IDEA Indicator = YES	13%	18%	13%	16%
Qualify for Free/Reduced Lunch	94%	96%	94%	94%
Grade 4				
n	836	107	815	128
Gender = Male	50%	60%	51%	54%
IDEA Indicator = YES	13%	24%	13%	17%
Qualify for Free/Reduced Lunch	94%	99%	94%	95%
Grade 5				
n	816	108	748	179
Gender = Male	52%	62%	52%	56%
IDEA Indicator = YES	13%	31%	13%	21%
Qualify for Free/Reduced Lunch	94%	95%	94%	93%
Grade 6				
n	559	387	556	395
Gender = Male	52%	50%	52%	50%
IDEA Indicator = YES	13%	13%	13%	13%
Qualify for Free/Reduced Lunch	94%	95%	94%	95%

To address these differences, a weighting scheme using gender, IDEA indicator (yes/no) and grade (based on 2019 grade level) was developed. Economic disadvantage was not used as a weighting factor because over 95% of the district's students are designated as economically disadvantaged (e.g., qualifying for free or reduced lunch).

The weighting scheme was developed by examining these three variables for the "sample" (e.g., identified students) and the "population" data. Population data was defined using the full

district level data for grades 4, 5 and 6. The population data included all students who were present in the district for the spring 2019 test, and with testing records that contained complete demographic data. Demographically, groups were similar after weighting.

Specific to the weighting procedure, two post-weighting exceptions merit notation. In ELA, grade 4 comparison students had more parents who did not graduate from high school (43%) compared to grade 4 LGL students (29%). In Math, grade 4 comparison students had slightly more students on Individualized Education Plans (IEP), relative to grade 4 LGL students (1% for comparison, 0% for LGL). These differences were acknowledged and accepted as limitations to the study.

While the weighted sample was used to produce each reported result, for reasons of simplicity, the authors have chosen to cite unweighted sample sizes (reported n) for all findings in this report.

Analysis Procedures

Using groups as previously described, the statistical analysis employed Univariate Analysis of Covariance (ANCOVA). Analyses were conducted separately, based on subject area, which produced independent results for ELA Edge and Math Edge. Gain scores, based on the difference between the 2019 and 2018 Smarter Balanced Summative Assessment scale scores, served as the dependent variable. Group affiliation (e.g., ELA Edge and comparison students; Math Edge and comparison students) was the single factor in each analysis, which was conducted with the full sample and independently based on grade level. This resulted in four analyses for each subject area to represent all students grades 4-6, and then students in grade 4, grade 5, and grade 6. Additionally, analysis within the four groups established by baseline proficiency level group membership was conducted.

Addressing Potential Bias using Covariates

With the following procedures and limitations acknowledged, two key issues that could potentially challenge accurate measurement remained: attendance and baseline score differences.

Regarding attendance, comparison group students were, on average, absent roughly three more days relative to LGL group students. Addressing this bias prior to analysis and through statistical adjustment measures was impossible due to significant amounts of missing attendance data for comparison group students. For example, in the ELA area, of the 602 comparison group students, just 381 had attendance data. In Math, these numbers reduced from 702 to 479. Therefore, the number of days absent was used as a covariate to statistically control for attendance rate differences. Thus, analyses to determine whether any initial gains remained after controlling for, or statistically removing, the difference for the number of absences were employed. Table 5 highlights attendance rate differences for the full group and by grade level, for both ELA and Math.

Table 5: 2018 Absence Mean Rate Comparison

Group	ELA Days Absent				Math Days Absent			
	LGL (n=2,211)		Comparison (n=381)		LGL (n=2,119)		Comparison (n=479)	
	M	SD	M	SD	M	SD	M	SD
Full Group	7.5	7.4	11.5	14.8	7.5	7.5	10.6	12.6
Grade 4	7.5*	7.6	12.7*	20.4	7.4*	7.7	12.5*	17.5
Grade 5	7.4*	7.1	11.8*	11.8	7.3*	7.2	9.9*	9.1
Grade 6	7.5*	7.3	9.8*	9.5	7.7*	7.6	9.2*	8.9

*indicates statistically significant difference ($p \leq .05$)

Initial analysis of baseline scores (spring 2018 assessments) for each group established the fact that the ELA Edge and Math Edge groups each started at different points relative to their comparison groups.

Table 6: 2018 (Baseline) Scale Score Mean Comparisons, by Subject Area

Group	ELA 2018 Scale Score				Math 2018 Scale Score			
	LGL (n=2,211)		Comparison (n=602)		LGL (n=2,119)		Comparison (n=702)	
	M	SD	M	SD	M	SD	M	SD
Full Group	2442.9*	93.6	2420.9*	95.6	2450.2*	79.7	2419.3*	76.7
Grade 4	2407.3*	84.4	2384.2*	90.2	2426.9*	77.6	2397.7*	78.0
Grade 5	2438.3	92.7	2434.9	97.9	2451.3*	76.4	2419.6*	69.3
Grade 6	2481.8*	88.0	2443.1	88.4	2471.7*	78.5	2439.8*	76.6

*indicates statistically significant difference ($p \leq .05$)

Because comparison students had lower baseline (2018) scale scores, the 2018 scores were used as a covariate to statistically control for this difference. Like attendance, analyses were performed to determine whether any gains initially present persisted after controlling for (statistically removing) differences in baseline scores. The same approach was applied within proficiency category analyses to determine any differences in gains based on baseline proficiency level (e.g., standard not met, etc.) group membership.

Results

The analysis investigated the impact of ELA Edge and Math Edge on student achievement following approximately seven-to-eight months of program use. It is important to note the point-in-time based on program's implementation. The resulting data align with the program's initial full-implementation year in the participating school district. The researchers characterize these results as formative and resulting from an analysis that leverages state test data as a dependent measure at the first point possible within the program's implementation timeline.

English Language Arts/Literacy

The initial analysis examined student academic growth in ELA as measured by the difference in scale scores between the spring 2018 and spring 2019 testing periods. The analysis utilized ANCOVA to eliminate the influence of slight differences in spring 2018 baseline scores. Table 7 presents results of this procedure, which demonstrate an advantage for students in the LGL group, after removing the influence of the baseline scores.

Table 7: ANCOVA Results with Baseline Score Covariate—ELA, All Students and by Grade

		Gain Score		Condition effect		
Group	**N**	**Mean (SD)**	**Adj Mean**	***F***	***p***	**Partial eta^2**
All Students				7.64	.006	.003
LGL	2,211	41.3 (58.1)	42.2			
Comparison	602	37.8 (65.4)	34.8			
Grade 4				1.39	.239	.001
LGL	836	38.3 (56.3)	39.3			
Comparison	107	37.2 (89.8)	33.7			
Grade 5				.47	.493	.001
LGL	816	48.7 (52.2)	48.8			
Comparison	108	46.2 (86.9)	45.7			
Grade 6				10.64	.001	.011
LGL	559	37.3 (67.5)	39.0			
Comparison	387	30.4 (47.9)	24.1			

The same analyses were then conducted for groups based on the students' spring 2019 grade level. Across all grade levels, means were higher for the LGL students. However, the mean difference was only significant for the grade six students.

An analysis of gain scores based on proficiency-category defined groups was also pursued. The spring 2018 baseline scores were used to organize students into groups as defined by the four Smarter Balanced assessment proficiency categories (standard not met, standard nearly met, standard met, standard exceeded). Gain scores were then compared within each established group. Table 8 summarizes ANCOVA results for each grade level, as organized by the four proficiency categories.

Table 8: ANCOVA Results with Baseline Score Covariate—Comparison by ELA Baseline Proficiency Category and Grade

		Gain Score		Condition effect		
Group	**N**	**Mean (SD)**	**Adj Mean**	***F***	***p***	**Partial eta²**
Standard Not Met						
Grade 4				1.70	.194	.006
LGL	254	58.4 (59.2)	58.7			
Comparison	53	48.8 (91.1)	48.2			
Grade 5				.59	.445	.002
LGL	318	63.7 (54.6)	63.9			
Comparison	56	59.2 (80.5)	58.6			
Grade 6				1.89	.170	.005
LGL	185	54.0 (77.0)	55.3			
Comparison	188	48.2 (52.3)	44.9			
Standard Nearly Met						
Grade 4				.001	.972	.000
LGL	240	39.2 (55.1)	39.2			
Comparison	26	39.5 (78.2)	39.5			
Grade 5				.22	.640	.001
LGL	184	47.1 (47.5)	47.2			
Comparison	19	43.0 (107.2)	42.5			
Grade 6				4.90	.028	.022
LGL	121	42.9 (63.9)	42.9			
Comparison	97	23.4 (41.3)	23.4			
Standard Met						
Grade 4				2.19	.141	.010
LGL	196	24.5 (55.6)	24.3			
Comparison	17	7.2 (109.0)	8.1			
Grade 5				.00	.998	.000
LGL	181	36.6 (53.5)	36.5			
Comparison	17	36.2 (90.8)	36.6			
Grade 6				3.45	.064	.013
LGL	177	24.8 (57.6)	24.7			
Comparison	81	8.2 (36.3)	8.8			
Standard Exceeded						
Grade 4				4.24	.041	.027
LGL	148	20.0 (42.8)	19.4			
Comparison	11	36.6 (59.3)	40.3			
Grade 5				.03	.867	.000
LGL	133	29.6 (40.7)	29.6			
Comparison	16	28.0 (73.5)	28.1			
Grade 6						
LGL	76	16.5 (54.1)	16.5	3.71	.057	.038
Comparison	21	-14.6 (36.1)	-14.5			

Of the 12 analyses, 10 of the mean comparisons favor the LGL groups, one slightly favors the comparison group (+0.3 for grade 4 standard nearly met), and one greatly favors the

comparison group (+16.6 for grade 4 standard exceeded). Neither of these differences was statistically significant. However, three of the twelve analyses that favored LGL groups did prove statistically significant, two of which were for grade 6 (standard nearly met, standard exceeded) and one for grade 4 (standard exceeded).

An earlier section of this report detailed differences between the LGL and comparison groups specific to number of school days missed (absences). To remove the potential bias introduced by absence rates, an ANCOVA was performed using absences as a covariate. It should be noted that this analysis does not control for the slight baseline score differences between LGL and comparison groups. Additionally, the comparison group is smaller due to a significant number of cases that were missing attendance data. Table 9 summarizes results of this analysis.

Table 9: ANCOVA Results with Absence Covariate—ELA, All Students and by Grade

		Gain Score		Condition effect		
Group	**N**	**Mean (SD)**	**Adj Mean**	***F***	***p***	**Partial eta^2**
All Students				0.00	.990	.000
LGL	2,211	41.3 (58.1)	41.3			
Comparison	381	40.9 (68.1)	41.3			
Grade 4				.268	.604	.000
LGL	836	38.3 (56.3)	38.4			
Comparison	75	35.5 (94.2)	35.4			
Grade 5				4.182	.041	.005
LGL	816	48.7 (52.2)	48.7			
Comparison	81	38.5 (88.1)	37.9			
Grade 6				5.078	.025	.006
LGL	559	37.3 (67.5)	37.1			
Comparison	225	50.0 (45.8)	50.9			

For the full group, gain scores did not differ significantly after removing the variance attributed to attendance. The same was true for grade 4. The LGL group outperformed the comparison group at grade 5, while the opposite was true for grade 6. In both cases, differences were significant.

While ANCOVA procedures with absence rate as the covariate were attempted within proficiency category-defined groups, the resulting small sample sizes prohibited completing the analysis.

The final attempted ANCOVA procedure analyzed ELA gain scores with both the baseline score and absence covariates. Due to missing attendance (explained earlier in this report), both sample size and statistical power were limited. The following table presents results of the analysis with both covariates entered.

Table 10: ANCOVA Results with Baseline Score and Absence Covariates—ELA, All Students and by Grade

		Gain Score		Condition effect		
Group	**N**	**Mean (SD)**	**Adj Mean**	***F***	***p***	**Partial eta^2**
All Students				6.84	.009	.003
LGL	2,211	40.9 (57.9)	41.8			
Comparison	381	38.4 (72.1)	33.7			
Grade 4				3.59	.059	.004
LGL	836	37.7 (55.4)	38.8			
Comparison	75	34.4 (98.1)	29.1			
Grade 5				10.93	.001	.012
LGL	816	48.6 (52.2)	49.1			
Comparison	81	34.6 (93.0)	32.6			
Grade 6				.113	.736	.000
LGL	559	36.9 (67.7)	38.3			
Comparison	225	48.7 (46.1)	40.2			

The previously reported analysis, conducted within groups established by baseline proficiency levels, were once again impossible due to small sample sizes.

Mathematics

As with ELA, the first analysis quantified student academic growth in Math using gains based on scale scores differences between the spring 2018 and spring 2019 testing periods. Here again, ANCOVA was used to eliminate the influence of slight differences in spring 2018 baseline scores. The results in Table 11 demonstrate consistent advantages for students in the LGL group after removing the influence of the baseline scores (pre-test).

Table 11: ANCOVA Results with Baseline Score Covariate—Math, All Students and by Grade

		Gain Score		Condition effect		
Group	**N**	**Mean (SD)**	**Adj Mean**	***F***	***p***	**Partial eta²**
All Students				38.164	.000	.013
LGL	2,119	31.6 (51.4)	32.3			
Comparison	702	20.5 (56.6)	18.1			
Grade 4				13.048	.000	.014
LGL	815	31.8 (43.8)	33.4			
Comparison	128	25.8 (67.6)	21.0			
Grade 5				2.186	.140	.002
LGL	748	30.8 (46.4)	31.4			
Comparison	179	27.6 (60.0)	25.8			
Grade 6				28.017	.000	.029
LGL	556	32.1 (66.3)	32.2			
Comparison	395	8.7 (49.6)	8.3			

Students in the LGL group outperformed their comparison group peers at each grade level. For grades 4 and 6, differences in gain scores proved significant.

As with ELA, a Math Edge analysis was performed based on proficiency categories established using baseline scores. The following tables present results, by grade level, within each proficiency category. ANCOVA procedures employed the baseline score as a covariate to remove differences from baseline scores.

Table 12: ANCOVA Results with Baseline Score Covariate—Comparison by Math Baseline Proficiency Category and Grade

		Gain Score		Condition effect		
Group	**N**	**Mean (SD)**	**Adj Mean**	***F***	***p***	**Partial eta²**
Standard Not Met						
Grade 4				8.91	.003	.032
LGL	218	56.9 (45.5)	57.4			
Comparison	55	40.9 (63.4)	39.9			
Grade 5				1.27	.261	.004
LGL	219	37.8 (49.5)	37.8			
Comparison	82	31.3 (60.0)	31.3			
Grade 6				15.44	.000	.033
LGL	227	33.1 (76.3)	33.4			
Comparison	228	7.4 (52.7)	6.9			
Standard Nearly Met						
Grade 4				1.91	.168	.008
LGL	203	30.5 (40.6)	30.5			
Comparison	37	21.6 (64.9)	21.5			
Grade 5				.09	.766	.000
LGL	271	27.8 (48.0)	27.9			
Comparison	72	26.5 (63.3)	26.1			
Grade 6				4.80	.029	.015
LGL	197	28.6 (60.0)	28.6			
Comparison	124	13.0 (46.6)	12.8			
Standard Met						
Grade 4				3.07	.081	.011
LGL	258	20.6 (38.8)	20.8			
Comparison	21	9.3 (63.7)	8.4			
Grade 5				.00	.992	.000
LGL	177	25.1 (42.0)	24.9			
Comparison	18	23.9 (65.0)	25.0			
Grade 6				2.14	.146	.020
LGL	81	35.0 (59.1)	34.9			
Comparison	27	12.7 (56.8)	13.0			
Standard Exceeded						
Grade 4				.62	.432	.004
LGL	136	14.9 (39.4)	15.0			
Comparison	15	7.8 (79.0)	7.5			
Grade 5				2.90	.092	.033
LGL	81	33.6 (38.7)	33.4			
Comparison	7	7.2 (64.6)	8.8			
Grade 6						
LGL	51	36.1 (51.1)	36.1	10.49	.002	.141
Comparison	16	-14.5 (21.7)	-14.9			

Four of the twelve analyses proved statistically significant. In each case, the gain score difference favored the LGL group. These differences were slightly more common at the lower

proficiency levels, with a single significant result at the highest proficiency level (standard exceeded) for grade 6.

The final analysis of gain scores involved again using an ANCOVA procedure with absences as a covariate. As was the case with ELA, this analysis does not control for the slight baseline score differences between LGL and comparison groups. Also, here too, the comparison group is smaller due to missing attendance data. Table 13 summarizes these results.

Table 13: ANCOVA Results with Absence Covariate—Math, All Students and by Grade

		Gain Score		Condition effect		
Group	**N**	**Mean (SD)**	**Adj Mean**	***F***	***p***	**Partial eta²**
All Students				9.068	.003	.003
LGL	2,119	31.6 (51.4)	31.3			
Comparison	479	22.6 (56.1)	23.5			
Grade 4				10.721	.001	.012
LGL	815	31.8 (43.8)	31.8			
Comparison	95	18.8 (66.7)	18.6			
Grade 5				1.061	.303	.001
LGL	748	30.8 (46.4)	30.4			
Comparison	152	24.8 (62.6)	26.2			
Grade 6				1.536	.216	.002
LGL	556	32.1 (66.3)	31.9			
Comparison	232	24.0 (46.1)	25.0			

Gain scores for the full group and grade 4 differed significantly, with LGL students demonstrating, on average, greater gain scores. While results from grades 5 and 6 both favor LGL, neither difference proved significant after removing the influence of attendance differences.

Small sample sizes prohibited ANCOVA analysis with absence rate as the covariate within proficiency category-defined groups.

As with ELA, a final ANCOVA procedure analyzed Math gain scores with both the baseline score and absence covariates. Due to missing attendance data (explained earlier in this report), both sample size and statistical power were limited. Table 14 provides results of this analysis which attempted to control for baseline score and attendance rate differences.

Table 14: ANCOVA Results with Baseline Score and Absence Covariates—Math, All Students and by Grade

		Gain Score		Condition effect		
Group	**N**	**Mean (SD)**	**Adj Mean**	***F***	***p***	**Partial eta^2**
All Students				34.805	.000	.013
LGL	2,119	31.4 (51.9)	32.2			
Comparison	479	20.8 (57.5)	18.1			
Grade 4				49.829	.000	.048
LGL	815	31.8 (43.8)	34.1			
Comparison	95	17.0 (66.4)	12.1			
Grade 5				3.886	.049	.004
LGL	748	30.8 (46.4)	31.1			
Comparison	152	24.4 (61.8)	23.3			
Grade 6				1.741	.187	.002
LGL	556	31.5 (67.0)	31.6			
Comparison	232	24.0 (46.1)	24.0			

As with ELA, attempts to conduct the above analysis, but within groups established by baseline proficiency levels, were impossible due to small sample sizes.

Conclusion

This analysis investigated early evidence of Edge program efficacy in ELA and Math. Using district-wide California state test scores, students were placed into LGL and comparison groups based on Edge program use. Scale scores from the spring 2018 Smarter Balanced Summative Assessment in both ELA and Math were used to establish baseline performance levels prior to the first full school year use of Edge across the district.

Scores from the spring 2019 state test administration were used to calculate gain scores (defined as the difference between spring 2019 and spring 2018). Weighting was used to make LGL and comparison groups equivalent on variables of gender and IDEA. Univariate effects of intervention condition on gain score measures were examined using between-subjects analysis of covariance (ANCOVA). Baseline scores and days absent were then used as covariates to remove the effects of differences in each area.

Results of the primary analysis using scale score gain figures indicate that, after controlling for baseline (2018) score differences:

- In English Language Arts/Literacy, mean gain scores favor the LGL group overall, and at each grade level analyzed (grades 4, 5, and 6).
    - Gain score differences for all students and grade 6 students proved statistically significant (unlikely the result of random chance, $p < .006$ and $p < .001$ respectively).
    - Effect sizes were small, with .003 for all students and .011 for grade 6.
- In Mathematics, mean gain scores favor the LGL group overall, and at each grade level analyzed (grades 4, 5, and 6).
    - Gain score differences for all students, grade 4, and grade 6 students proved statistically significant (unlikely the result of random chance, $p=.000$).
    - Effect sizes were small and ranged from .013 to .029.

The limitations of this analysis must be acknowledged.

- First, the analysis follows the initial full school year of LGL program implementation. It is reasonable to assume that teachers used this time to become familiar with the program. The sophistication of program use is likely to increase over time. Assuming such an increase occurs, it is possible that program effects will increase.
- Second, while this investigation benefitted from district-wide (census) student data, students were not assigned to treatment and comparison groups purposefully, nor through random selection. Instead, group membership was determined after program implementation and solely by LGL program usage levels, based on the 2018-2019 school year records. There are likely many reasons why students in the comparison group did not use LGL. Some or all of these reasons for non-use have likely influenced the results reported here.

- Finally, this investigation was limited to analysis. Consideration should be given to a formal, quasi-experimental program evaluation. Such an effort would begin with a detailed evaluation design, including the definition of LGL and comparison group membership. Additionally, a program evaluation could pursue complementary measures to define classroom implementation practices. Variables of interest might include frequency of use, method of lesson assignment, and integration into existing curricula and practice. More robust student-related measures that go beyond lessons attempted and completed and system use time could contribute to a more nuanced understanding of the program, its use, and its impact.

Results presented in this report provide highly formative, yet initially promising evidence of LGL program efficacy. While influenced by several key limitations, the small yet reliable and favorable differences observed for the LGL group, after controlling for initial group performance inequalities, are remarkable. Future investigation, ideally accomplished with a carefully defined and detailed research plan, should continue to develop an increasingly detailed picture of the LGL Edge program's efficacy.

Conflict of Interest Statement

The authors were compensated by Let's Go Learn to conduct the independent efficacy analysis reported herein. Apart from that project-specific engagement, the authors have no financial or non-financial interest in the organization, nor the programs it produces.

References

Ginsberg, A., Chang, H., & Jordan, P. (2014). *Absences add up: How school attendance influences student success*. http://www.attendanceworks.org/absences-add-up/

Gottfried, M.A. (2015). Chronic absenteeism in the classroom context: Effects on achievement. *Urban Education,* 1–32. doi: 10.1177/0042085915618709

National Center for Education Statistics. (2019). *National assessment of educational progress: An overview of NAEP*. National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education.

Pane, J.F., Steiner, E.D., Baird, M.D., Hamilton, L.S., & Pane, J.D. (2017). *How does personalized learning affect student achievement?* RAND Corporation. https://www.rand.org/pubs/research_briefs/RB9994.html

Romero, M., & Lee, Y.S. (2007). *A national portrait of chronic absenteeism in the early grades*. New York: National Center for Children in Poverty.