The Association of Legends of Learning Usage and Science Achievement in a Large School District in the Southeastern United States

David McKinney, PhD, WestEd
Scott Strother, PhD, WestEd
Steven Schneider, PhD, WestEd

January 2023

Suggested citation: McKinney, D., Strother, S., & Schneider, S. (2023). The Association of Legends of Learning Usage and Science Achievement in a Large School District in the Southeastern United States. WestEd.

WestEd is a nonpartisan, nonprofit agency that conducts and applies research, develops evidence-based solutions, and provides services and resources in the realms of education, human development, and related fields, with the end goal of improving outcomes and ensuring equity for individuals from infancy through adulthood. For more information, visit WestEd.org. For regular updates on research, free resources, solutions, and job postings from WestEd, subscribe to the E-Bulletin, our semimonthly e-newsletter, at WestEd.org/subscribe.

Table of Contents

LIST OF FIGURES

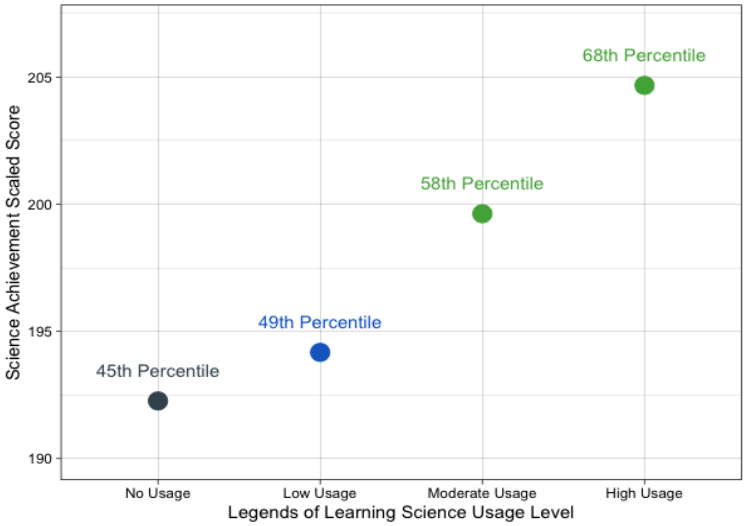
LIST OF TABLES

Executive Summary

Fifth- and eighth-grade students with higher levels of usage of the Legends of Learning science program in the 2020-21 school year had, on average, higher science achievement than their peers with low or no usage of the program in a large school district in the southeastern United States.

Legends of Learning is a web-based, standards-aligned, and game-based learning solution that engages students in short math and science games. This study focuses on the use of Legends of Learning in a large school district in the southeastern United States that implemented the program throughout the district in January 2020 for supplemental science instruction for grades 4-8. During the 2020-2021 school year, over 630 teachers in the district utilized Legends of Learning, translating to over 91,000 hours of engagement during which students played over 336,000 games and answered over 2.15 million assessment items.

Figure 1. Estimated Mean Science Achievement of Fifth- and Eighth-Grade Students by Legends of Learning Usage Level



WestEd, a nonpartisan education research agency, conducted a study aimed at understanding how students' use of the Legends of Learning science program in the 2020-21 school year was associated with their science achievement. The sample used in the study consisted of over 14,000 fifth- and eighth-grade students with diverse racial and ethnic backgrounds. Many students in this sample were English language learners (ELLs) or had an individualized education program (IEP). Students' scores on the state end-of-year science assessment were used as a measure of science achievement. Students' use of the Legends of Learning science program was categorized into four levels of use: 1) No Usage, 2) Low Usage, 3) Moderate Usage, and 4) High Usage. Multilevel linear regression was used to estimate the association between level of Legends of

Learning usage and science achievement, controlling for students' grade level, sex, race, ethnicity, ELL status, and IEP status.

Findings

Results indicated that there were **significant** and **educationally-meaningful** differences in the mean science achievement of students with moderate and high usage of the Legends of Learning science program in the 2020-21 school year compared to students who did not use the program. **As the level of Legends of Learning science program usage increased, science achievement also increased**. Figure 1 shows how student science achievement is related to levels of activity in the Legends of Learning science program. The mean science achievement is plotted for each usage level and is labeled with that score's percentile among all students' scores in the sample. Mean science achievement for students in the No Usage group was 192.3 which was in the 45th percentile among all students in the study. Students in the Low Usage group had a mean science achievement of 194.2 (49th percentile). This mean is 0.09 standard deviations (SD) above the mean of the No Usage group. The Moderate Usage group mean science achievement was 199.6 (58th percentile, 0.34 SD above the No Usage group). The High Usage group mean science achievement was 204.7 (68th percentile, 0.58 SD above the No Usage group).

*There were **significant** and **educationally-meaningful** differences in the mean science achievement of students with moderate and high usage of the Legends of Learning science program in the 2020-21 school year compared to students who did not use the program.*

Conclusion

The results of this study indicated a promising association between Legends of Learning science program usage and science achievement. Further research is needed in the form of a randomized controlled trial (RCT) or a carefully planned quasi-experimental design (QED) study to further control for differences between students who do and do not use Legends of Learning.

Report

The purpose of this study was to understand how students' use of the Legends of Learning science program in fifth and eighth grade is related to student science achievement in one large school district in the southeastern United States.

Legends of Learning

Legends of Learning was founded in 2017 with the belief that everyone learns best through engaging experiences. Legends of Learning's game-based learning solution is web-based and standards-aligned to provide the best educational content, in the simplest manner, to the widest audience possible.

The platform hosts more than 2,000 math and science games and over 100,000 assessment questions that are all aligned to national and state standards. Over the years, more than 10 million students, teachers and administrators have relied on Legends of Learning to help students attain content mastery through short, curricula-aligned "mini games." The primary mini game activity type on the Legends of Learning platform is a 10–15-minute standards-aligned instructional game. These mini games are used to both introduce and reinforce lesson content. Question games, simulations, and assessments are also prevalent features. The questions can be used as formative assessments at the student, group, class, school, and district levels, which can aid educators in evaluating and assigning individualized learning experiences to meet all students' needs.

There are three components to Legends of Learning: the Learning Universe, Math Basecamp, and Awakening. Each product focuses on a different aspect of learning: in-school math and science learning, math fact fluency, and at-home math and science learning, respectively. Legends of Learning game-based instruction is standards-aligned, research-based, and uses industry best practices to provide productive struggle and academic challenges to teach students science concepts while developing 21st century skills.

School District

This study was conducted in a large school district in the southeastern United States. The district enrolls over 100,000 students who come from a diverse set of racial and ethnic backgrounds

(about 40% Hispanic; 35% white; 20% Black; and <5% students identified as Asian, Native American, Alaskan Native, Pacific Islander, or multiracial). Additionally, about 10% of students in the district are English Language Learners (ELLs) and about 15% are students with individualized education programs (IEPs).

Legends of Learning in the District

In January 2020, the district implemented Legends of Learning in order to add supplemental science instruction for grades 4 - 8 throughout the district. With rostering and an LMS integration at the outset, teachers received professional learning and curriculum support from both Legends of Learning and the school district. While the COVID-19 pandemic closed schools in March 2020, students and teachers were able to continue instruction and learning virtually. The majority of schools in the district utilized Legends during spring 2020 in a pilot capacity.

During the 2020-21 school year, the district transitioned their use of the Legends of Learning science program from a pilot to a fully available, supplemental resource for all teachers and students. Legends of Learning mathematics was also added, reaching all students and teachers in fourth through eighth grade. Over 630 teachers used Legends of Learning during the school year for over 91,000 hours of engagement, where students played over 336,000 games and answered over 2.15 million assessment items. Teachers assigned mini games and assessments to be completed by learners in a variety of settings. Usage ranged from 0 to over 50 activities per student.

WestEd

WestEd was selected as an independent evaluator of the Legends of Learning science program in the school district. WestEd, a nonpartisan research, development, and service agency, works with education and other communities to promote excellence, achieve equity, and improve learning for children, youth, and adults. WestEd is a preeminent educational research, development, and service organization with over 800 employees and 13 offices nationwide. WestEd has been a leader in moving research into practice by conducting research and development (R&D) programs, projects, and evaluations; by providing training and technical assistance; and by working with policymakers and practitioners at state and local levels to carry out large-scale school improvement and innovative change efforts. The agency's mission is to promote excellence, achieve equity, and improve learning for children, youth, and adults.

Study

The study focused on all fifth- and eighth-grade students in the district in 2020-21 who had a state standardized science assessment scaled score. Students who did not have a valid score or who took an alternative assessment that is scored on a different scale were excluded from the sample. Table 1 describes the sample in the study, which consisted of over 14,000 students who

generally reflected the diversity observed in the district overall. The students in this sample were more likely to be an ELL or to have an IEP than students in the district as a whole (10% and 15%, respectively).

Table 1. Sample Demographics and Educational Status

Category	*N* (Percent)
Grade Level	
Fifth Grade	7459 (51.4)
Eighth Grade	7039 (48.6)
Sex	
Female	7143 (49.3)
Male	7355 (50.7)
Race/ethnicity	
Hispanic	5696 (39.3)
White	5286 (36.5)
Black	2853 (19.7)
Multiracial	404 (2.8)
Asian	206 (1.4)
Native American or Alaskan Native	30 (0.2)
Native Hawaiian or Other Pacific Islander	23 (0.2)
Has IEP	3160 (21.8)
ELL	3411 (23.5)

Measures

Students' scaled scores from the spring 2021 administration of the state standardized science assessment were used as the measure of science achievement. The science assessment is administered to all eligible fifth- and eighth-grade students in the district. The assessment consists of over 50 items which students have over two hours to answer. The state department of education uses Item Response Theory (IRT) to create a scaled score that better reflects a student's understanding than a simple raw score of the number of correct items would.[1] Students' scored responses are converted into a scaled score that ranges from 140 to 260[2]. The scaled

[1] Anonymous. (2021).
[2] Anonymous. (2022).

score is also categorized into five performance levels, as shown in Table 2. Notably, a performance level of three or higher is considered satisfactory performance by the state.

Table 2. State Science Assessment Performance Levels and Scaled Scores

Level	Performance	Fifth-Grade Range	Eighth-Grade Range
1	Inadequate	140 - 184	140 - 184
2	Below Satisfactory	185 - 199	185 - 202
3	Satisfactory	200 - 214	203 - 214
4	Proficient	215 - 224	215 - 224
5	Mastery	225 - 260	225 - 260

Students' activity in the Legends of Learning science program was captured throughout the 2020-21 school year. The type and duration of each activity was logged by Legends of Learning for each activity for each student. For each student, a total instructional activity count was created by counting all activities they "played" in the 2020-21 school year, including instructional games, question games, simulations, and assessments. Legends of Learning usage levels were then created from these activity counts by categorizing a student's total number of activities into four levels of usage, shown in Table 3.

Table 3. Students' Usage Levels as Determined by the Number of Legends of Learning Science Activities Played

Usage	Activity Count
No usage	0
Low usage	1 - 24
Moderate usage	25 - 49
High usage	50 or more

These categories were created to roughly reflect usage of none, less than once a week, about once a week, and two or more times a week. A school year in the United States is typically 36 weeks, but there are many weeks during a school year when instruction may differ from a typical week (e.g., the weeks before holiday breaks, the weeks when standardized testing occurs, etc.).

Thus, we chose 25 weeks as a rough approximation of the number of weeks during which it would be reasonable to expect a teacher to utilize the Legends of Learning science program.

Analysis

Multilevel linear regression (MLR), which accounts for the nesting of students within schools, was used to model the association between Legends of Learning usage level and science achievement. A series of models were estimated using the lme4 package[3] in the R statistical computing platform[4]. In each model, science achievement was the outcome. In the first model, student grade level, sex, race/ethnicity, and usage level were modeled as predictors of science achievement. In subsequent models, these predictors were retained, and additional interaction terms were added to test the interaction of usage level and student demographic and educational characteristics in predicting science achievement. In MLR, a reference group must be selected for each categorical variable (e.g., fifth-grade students as a reference group for the grade level variable). For this analysis, the reference groups for each predictor variable were fifth grade, female, Hispanic, no IEP, and not ELL. Thus, the resulting mean science achievement predicted by the estimated models reported below is for fifth-grade students identified as female and Hispanic who do not have an IEP and are who are not an ELL.

Results

In the sample, each student's number of logged science activities varied from a minimum of 0 to a maximum of 224. The majority of students were in the No Usage (n = 9,415) and Low Usage groups (*n* = 4,205, Table 4). The mean number of science activities completed by students in the sample was 4.9 (*SD* = 11.9). Table 4 shows the mean number of science activities for each usage level. While the typical usage level was low, over 800 students were in the Moderate Usage or High Usage groups, averaging over 30 science activities each over the course of the school year. Mean science achievement in the sample was 194.2 (*SD* = 21.51). Results of the multilevel linear regression indicated that there were **significant** and **educationally-meaningful** differences in the

Table 4. Frequency and Mean Number of Activities Overall and by Legends of Learning Usage Level

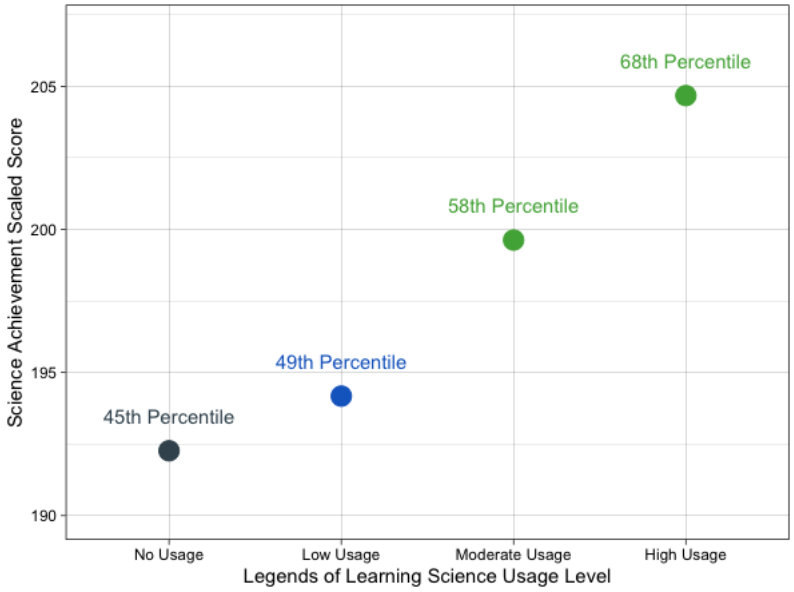
	Overall	No Usage	Low Usage	Moderate Usage	High Usage
Number of Students	14,498	9,415	4,204	675	204
# of Legends Science Activities, Mean (SD)	4.9 (11.9)	0.0 (0.0)	7.9 (6.5)	33.7 (7.0)	71.0 (25.7)

[3] Bates, D., Mächler, M., Bolker, B, & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software 67(1), 1-48*

[4] R Core Team (2022). R: A language and environment for statistical computing. R doundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

mean achievement of students with moderate usage (25 to 49 activities) or high usage (50 or more activities) of Legends of Learning science in the 2020-21 school year compared to students with no usage of Legends of Learning science, controlling for students' grade level, sex, race, ethnicity, ELL status, and IEP status. Additionally, mean science achievement for students with low usage was also significantly different and larger than the mean science achievement of students with no usage. Figure 2 shows how predicted student performance on the state standardized science assessment, based on the estimated regression model, is related to usage levels in the Legends of Learning science program. The mean scaled score is plotted for each level of student activity in the program. As can be seen in the figure, as the level of Legends of Learning

Figure 2. Estimated Mean Science Achievement of Fifth- and Eighth-Grade Students by Legends of Learning Usage Level



science activity increases, science achievement also increases. Mean science achievement for students in the No Usage group was 192.3 which was in the 45th percentile among all students in the study. Students in the low usage group had mean science achievement of 194.2 (49th percentile). This mean is 0.09 standard deviations above the mean of the no usage group. The moderate usage group mean science achievement was 199.6 (58th percentile, 0.34 SD above the no usage group). The high usage group mean science achievement was 204.7 (68th percentile, 0.58 SD above the no usage group).

Performance Levels

If the predicted mean science achievement were translated into the state's five performance levels (Table 2), then the mean for the High Usage group would be level 3, satisfactory, while the mean for the No Usage group would be level 2, below satisfactory.

Interactions

Interactions between usage level and each of student sex, race/ethnicity, ELL status, and IEP status in predicting science achievement were tested with a series of models, one for each student demographic or educational status variable. For each variable (e.g., sex) two models—the model presented above in the main analysis without any interaction terms and a model with usage level interacted with the variable—were estimated and compared to determine if including the interaction in the model significantly fit the data better than omitting the interaction.

Results indicated that there was no significant interaction between usage level and student sex, race/ethnicity, and IEP status. These results indicate that there is no significant difference in the relationship between science achievement and usage level between students of different sexes, different race/ethnicity groups, and different IEP statuses.

There was a significant interaction between ELL status and usage level in predicting science achievement. This indicates that the association between usage level and science achievement is significantly different for students who are ELLs compared to students who are not ELLs. Specifically, mean differences in science achievement scores for ELLs who had moderate usage compared to those with no and low usage was smaller than the same contrast in usage levels for students who are not ELLS. Nonetheless, for students who are ELLs, as with students who are not ELLs, as Legends of Learning usage increases, mean science achievement also increases.

Limitations

This was an observational study. Usage of Legends of Learning was not randomly assigned, where some students would be assigned, at random, to receive instruction with the Legends of Learning science program and some would not. As a result, this study can only describe the association between Legends of Learning science usage levels and science achievement and cannot make any causal claims that the differences in science achievement are caused by use of the Legends of Learning science program. Additionally, this study took place in a district where a relatively small percentage of teachers had adopted the Legends of Learning platform as a substantial part

of their science instruction. As a result, only a small proportion of the sample was observed with high levels of usage.

Conclusion and Future Directions

The results of this study indicated a promising association between Legends of Learning science program usage and science achievement. Further research is needed in the form of a randomized controlled trial (RCT) or a carefully planned quasi-experimental design (QED) research study to further reduce differences between students who do and do not use Legends of Learning in the study. Furthermore, as Legends of Learning is adopted by more and more science teachers, studies can include a larger number of students with moderate and high usage to determine if the associations observed in this study hold across a larger sample.