LGL Edge

Technical Manual

Published by Let's Go Learn, Inc.

Introduction: *LGL Edge* Series

The Nation's Report Card indicates that only 33% of 4th graders and 31% of 8th graders read at or above proficient levels. Similarly, 33% of 4th graders and 26% of 8th graders are at or above proficient levels in mathematics (NAEP, 2022). A majority of students in schools struggle with the reading and math skills required at their grade levels. It's critical to our success as a nation, as a diverse society, and as individuals that educators determine how to lead students to proficiency at an accelerated rate.

When *LGL Edge* is implemented according to best practices, students' academic achievement is accelerated by filling gaps and continuing to move forward to reach proficiency in grade-level content in reading and math. In alignment with the recommendations of the National Council of Learning Disabilities (2021), Let's Go Learn's design has incorporated these characteristics to accelerate learning:

- Reduce cognitive load to focus on grade-level content with scaffolding to fill in gaps
- Provide context for cultural relevance
- Drive engagement by aligning learning to student interests (music, narration, delivery media)
- Leverage multiple modalities to support learning styles and reinforcement
- Develop executive function and critical thinking skills through gamification features to support intrinsic and extrinsic motivation, focus, real time scoring, and optimal performance

To ensure ongoing differentiation and academic progress and gains, the *LGL Edge* series requires that educators have students take the front-end diagnostic assessments three times a year, at regular intervals, which should include beginning of the year (within the first four weeks of the school year), mid-point, and end of the year (within four weeks of the conclusion of the school year). In addition, students must use the online lessons 3 to 5 times a week for 45 minutes a session to move steadily toward accelerated proficiency.

LGL Edge Research Base

A key part of developing an effective and equitable K-12 educational program is creating a research base that is both deep and wide and maintaining the research base for the life of the program to ensure that the program serves the needs of educators and students. During the needs analysis literature review, design, and revision phases of program design, a thorough research base ensures an optimal foundation of pedagogy and best practices: “key decisions regarding the design of instruction are based on research and experience related to human learning, instruction, and general systems theory” (Hirumi, 2022).

LGL Edge: Diagnostic Assessments and Reduced Cognitive Load

To provide the most effective learning for all students, the design of *LGL Edge* uses individual student diagnostic data to focus instruction on grade-level content while filling gaps at appropriate learning points (Levin, 1988). We do this by requiring that implementation begins with our adaptive diagnostic assessments, which drive the creation of individual learning paths. The result is that all students, whether performing above, at, or below grade level, experience optimal learning interactions. According to Olenchak (2009) and Moon and Reis (2004), “Rather than approaching instruction from a deficit model, efforts should focus on student strengths, simultaneously providing compensatory strategies and additional instruction to address gaps in learning and needed areas of growth.”

Because our learning paths are unique to individual students’ strengths and weaknesses, we are able to provide a reduced cognitive load. Cognitive load theory (CLT), developed out of the study of problem solving by John Sweller in 1988, supports the theory that when cognitive load is reduced by instructional design, learning increases in effectiveness. LGL has incorporated three types of cognitive load theory into our program design. According to CLT, the process of construction and automatization of cognitive schemas constitutes learning (van Merriënboer & Sweller, 2010). Therefore, efficient and successful learning requires an ease in the process of creating and modifying cognitive schemas to optimize intrinsic and extraneous cognitive load for upcoming learning to levels that do not exceed the learner’s cognitive capacity and do not impede the learning performance of an individual (Reif, 2010; Tracey et al., 2022).

Intrinsic: Initially, all students complete the full diagnostic *ADAM* assessment. This assessment breaks math into 44 subtests and finds students' instructional points within the linear scope and sequence of each of the subtests. The same is true for *DORA*, which breaks the teaching of reading into seven subtests. This means that *LGL Edge* uses the identification of these multiple instructional points as the basis for how lessons are assigned and delivered to each student.

Thus, the intrinsic cognitive load is reduced by not presenting topics or skills that are too hard or complex for a student to learn.

Extraneous: This type of cognitive load refers to the observation that some learning is more easily achieved with certain teaching methods. For instance, it is usually more effective to teach learners the concept of a triangle by showing them a picture of a triangle, rather than by trying to describe it in words. In the case of *LGL Edge*, lessons use music, graphics, and audio to present each lesson. These factors reduce the extraneous cognitive load significantly for each learner by presenting and teaching skills in a format that is easier for students to process and absorb. In addition, the gaming-interface design of *LGL Edge* allows students to learn more easily.

Germane: *LGL Edge* was designed with long-term retention in mind, which is the focus of germane cognitive load theory. Students are required to demonstrate the highest level of mastery by repeating lessons until they get to the Gold level or to 95% accuracy. Students are not allowed to repeat a lesson until two days after their first time completing it. These factors support the deeper learning of students across the *LGL Edge* content areas of reading and mathematics.

Driven by our adaptive diagnostic assessments, *LGL Edge* math and reading programs provide individualized instruction that is based on student skill gaps rather than on grade level. Implementation of the *LGL Edge* series requires that teachers have students take the diagnostics three times a year so that differentiated instruction is ongoing throughout the year. *LGL Edge* uses the identification of multiple instructional points as the basis for how lessons are assigned and delivered to each student. *LGL Edge* instructional paths reduce cognitive load by using diagnostic data to identify gaps and strengths aligned to national and state standards.

The most essential element of the *LGL Edge* series is its foundation in Let's Go Learn's adaptive diagnostic assessments in reading and mathematics. The assessments are criterion-referenced, valid, and reliable. If students miss a chunk of skills or subskills in reading or math, or if they miss the opportunity to practice the cognitive and metacognitive processes that guide critical thinking in these areas, their progress in these content areas is seriously stymied.

Given that our diagnostics are criterion-referenced adaptive measures, student data is not tied to a single grade level. For example, a sixth grader may need an "instructional" level that is below their grade level. Our diagnostics are able to identify gaps, even if they exist over multiple years. Thus, student instructional paths are unique to their performance on each tested standard. The same is true for students working above grade level. This design feature allows the diagnostic-to-learning path system to control for a student's prior academic history, exposure (i.e. poverty), language ability, and disabilities.

The validity of an instructional program refers to its ability to support valid instructional inferences. That is, when implementation is followed according to best practices, do results support a valid conclusion about student learning? Building a valid program begins with accurate definitions of the content (i.e., the knowledge domains and skills). If the instructional activities correlate to the constructs that the program is designed to teach, then the program has content validity. Content validity is the basic logical bedrock of any instructional program. The content validity of *LGL Edge* programs is driven by the valid and reliable diagnostic assessment data that organize instructional content.

Content validity is the basic bedrock of any instructional program. Building a valid program begins with precise identification of discrete knowledge domains and skills necessary to bring each learner towards expertise in a specific content area. The content validity of *LGL Edge* programs is based on a hierarchical task analysis (HTA) conducted by experts in mathematics and reading instruction (Stanton, 2006). Each HTA is translated into a list of key skills to be taught for each course in the *LGL Edge* program. The content covered, the sequence of the activities delivered, and the specific items of feedback given to each learner are driven by expert knowledge from each field.

Let's Go Learn's Reading Diagnostic: *DORA*

DORA (Diagnostic Online Reading Assessment) is criterion-referenced, adaptive, and delivered online. It is diagnostic in nature and can be used as a measure of student growth. After assessment, comprehensive reports are provided to teachers and administrators to help with SLO creations and monitoring. *DORA* diagnostically evaluates each student's reading abilities while providing the highest level of accuracy through assessments with high overall coefficient alphas. In addition, test-retest consistency is high, from 0.69 to 0.84.

Sections that make up individual subtests are items written to test specific skills within the scope and sequence of the subtest. These CBM-level sections acquire their reliability in part from the test design that aggregates specific skill items together while maintaining p-values that range from 0.25 to 0.75. Individual field testing of each CBM-level section requires a mastery versus non-mastery score of 0.75 or higher, which was the lowest threshold requirement for decision consistency by pools of students with previously established skills mastered.

DORA was created to paint a picture of an individual's reading strategies more accurately across multiple measures which follow a constructivist perspective (Flores et al., 1991). The most effective way to characterize students' reading ability is to assess their reading skills across a set of criterion-referenced categories that are important to the reading process. The eight reading skills measured by Let's Go Learn are: 1) High Frequency Words, 2) Phonemic Awareness, 3) Phonics, 4) Word Recognition, 5) Vocabulary, 6) Spelling, 7) Silent Reading Comprehension, and 8) Fluency.

High Frequency Words subtest

This subtest assesses children's ability to automatically recognize words that have been identified as frequently occurring in books, newspapers, and other texts. This subtest uses words from Edward B. Fry's 300 sight words as test items which have been broken down into three general levels of difficulty (Fry, Kress, & Fountoukidis, 2004). A child's response time in identifying these sight words is recorded and factored into the scoring of the child's performance on the assessment.

Phonemic Awareness subtest

According to Ruddell (1998), by the time children are between three and four years old, they have learned most of the approximately 40 phonemes (discrete sounds in words) which comprise the English language. The ability to hear and manipulate these discrete sounds in spoken words is referred to as "phonemic awareness." Children demonstrate their phonemic awareness by segmenting words into individual sounds (i.e., /fish/ into /f/-/i/-/sh/), deleting sounds in words, blending sounds, adding sounds, or substituting sounds within a word to make a new word. Some researchers have indicated that phonemic awareness is one of the best predictors of reading success (Stanovich, 1993-1994). Others further argue that phonemic awareness is both the prerequisite and consequence of learning to read (Yopp, 1992). As such, it is especially important to determine children's level of phonemic awareness in the primary grades to ensure that they get any necessary intervention as early readers, lest they struggle with reading as young adults. Specific phonemic awareness categories tested include: 1) addition, 2) deletion, 3) substitution, 4) identification, 5) categorization, 6) blending, 7) segmenting, 8) isolation, and 9) rhyming.

Phonics subtest

In addition to having an awareness of the discrete sounds in words, children need to master how sounds and words are represented in English. This is important because children need to be able to effortlessly decode and recognize familiar and unfamiliar words to help facilitate the process of negotiating the meaning behind the text (Adams, 1990; Snow, Burns, & Griffin, 1998). The phonics subtest assesses a child's ability to recognize basic English phonetic principles of high utility (Pressley & Woloshyn, 1995). These phonetic principles include: 1) beginning sounds, 2) short vowel sounds, 3) blends, 4) the silent E rule, 5) consonant digraphs, 6) vowel digraphs, 7) r-controlled vowels, 8) diphthongs, and 9) syllabification.

Word Recognition subtest

As in many informal reading inventories such as the Qualitative Reading Inventory (Leslie & Caldwell, 1994), the Basic Reading Inventory (Johns, 2001), and the Diagnostic Assessment of Reading (Roswell & Chall, 1992), *DORA*'s Word Recognition subtest assesses a learner's ability to recognize leveled lists of words. In this subtest, children are presented with a number of increasingly difficult words until they reach a level at which they "frustrate" or stop recognizing the words presented to them. The final outcome of the assessment gives teachers an idea of the grade-level ability of a child to recognize words out of context. This assessment is important in identifying how well individuals can use what they know about text to recognize words outside the context of a sentence and of increasing difficulty.

Vocabulary subtest

A learner's knowledge of words and what they mean is an important part of the reading process, as knowledge of word meanings affects the extent to which learners comprehend what they read (National Reading Panel, 2000). The vocabulary subtest assesses a child's understanding of words. The words from this subtest were selected by teachers and reading specialists to reflect the types of words children learn in various disciplines at different grade levels and in various stages of their lives. Similar to the Peabody Picture Vocabulary Test (Dunn, 1959), in the vocabulary subtest children are asked to select the picture that correctly corresponds to a word they hear. The program continues to present children with increasingly difficult words until they make a certain number of errors. This subtest provides information about a child's level of oral vocabulary.

Spelling subtest

The process of spelling involves many cognitive processes. While each person uses different strategies for spelling words, these strategies usually have in common a familiarity with a particular word (i.e., familiarity with its meaning and visual exposure to the word), letter-sound matching, and confirmation of how the word "looks" (Bear et al., 2000; Ruddell, 1999; Gillet & Temple, 1994). Because spelling is also a generative process (as opposed to a decoding and meaning-making process in reading), it is natural for young readers' spelling abilities to lag a few months behind their reading abilities. *DORA*'s spelling subtest tries to capture the nuances of the different processes that children use to spell words by employing target words with increasing difficulty in different domains. In the process of creating the items for the *DORA* spelling subtest, reading specialists created a list of recommended target spelling words by examining words commonly encountered in or taught at specific grade levels. The program stops administering words when a child consistently spells words incorrectly. Items from this subtest were chosen by reading specialists and classroom teachers to approximate the kinds of words children of a particular age would see in their classroom instruction.

Silent Reading Comprehension subtest

The silent reading comprehension subtest forms the crux of *DORA*, which attempts to provide a window into the semantic domain of a learner's reading abilities. The content of each silent reading passage is expository and written to reflect the subject areas that students of a particular grade level would encounter. In a variation on protocols for some informal reading inventories (Gillet & Temple, 1994; Leslie & Caldwell, 1994), children silently read passages of increasing difficulty and answer questions about each passage immediately after they read it. The questions for each passage are broken up into three factual questions, two inferential questions, and one contextual vocabulary question. The program stops administering passages and questions once a student misses a certain number of questions on a passage. It provides teachers with information about a child's comprehension level.

Fluency subtest

Fluency is included as a teacher-administered measure. In this subtest, children read aloud short leveled passages with increasing syntactic complexity. Teachers time children's reading of these passages and record their errors and prosody using the National Assessment of Educational Progress (NAEP) Oral Reading Fluency Scale (1995).

Let's Go Learn's Math Diagnostics

Let's Go Learn has three math diagnostics: *ADAM*, *DOMA Pre-Algebra*, and *DOMA Algebra*. Their content validity comes from best practices in math curricula. *ADAM* and *DOMA Pre-Algebra* employ a gains score, or trajectory, model for student growth. Our gains score model captures grade-level progress on a particular scale or subscale between time 1 and time 2. The model is represented as: GL(s)2 – GL(s)1, where GL = grade level and where (s) denotes the particular scale or subscale. The combination of an interval scale design with a K-7 set-item range allows *ADAM* to measure the growth of students' ability either within a single school year or across students' entire K-7 experience. Likewise, the combination of an interval scale design with a grade 4 to 7 set-item range allows *DOMA Pre-Algebra* to measure the growth of students' ability either within a single school year or across students' grade 4 to 7 experience. *ADAM* and *DOMA Pre-Algebra* scores can be used both to diagnose student needs and to track student growth over time.

The development of these cutting-edge math products has been spear-headed by math specialist and teacher-trainer Paul Giganti of UC Berkeley and CalState Hayward. Prior to his work in professional development, Paul Giganti taught math in public schools for over 15 years. He has directed federally funded professional development programs in California under the auspices of the California Post-Secondary Educational Commission. Currently he is the coordinator of the California Mathematics Council Festivals Programs and Professional Development. In addition to his classroom teaching and professional development career, Giganti has published several children's picture books about mathematics. Supplementing the expertise of Giganti, LGL derives construct validity for the *ADAM* & *DOMA* series of tests by its alignment to both Common Core State Standards (CCSS) and state standards. *DOMA: Basic Math Skills* was originally aligned to California state mathematics standards in the Numbers and Measurement strands, as well as NCTM National Standards for Mathematics. *ADAM K-7*, the sequel to the *DOMA Basic Math Skills* assessment, was redesigned fundamentally and expanded to cover all 5 NCTM major math strands and nearly all of the CCSS. *ADAM* is aligned to CCSS and state standards in all 50 states. Further, *DOMA: Pre-Algebra* and *DOMA: Algebra* are aligned to NCTM standards, CCSS, and all 50 state standards.

ADAM

ADAM is a K-7 assessment that is multiple measured, criterion referenced, adaptive, and delivered online. It is diagnostic in nature and designed to identify each student's Zone of Proximal Development. Post assessment, comprehensive reports are provided to teachers and administrators to help with SLO creations and monitoring. *ADAM* diagnostically evaluates each student's math abilities while providing the highest level of reliability and accuracy and high overall coefficient alphas. In addition, test-retest consistency is high–from 0.69 to 0.84. Sections that make up individual subtests are written to test specific skills within the scope and sequence of the subtest. These CBM-level sections acquire their reliability in part from a test design that aggregates specific skills items together while maintaining p-values that range from 0.25 to 0.75.

ADAM assesses across five major math strands that span 44 subtests of K-7/8 mathematics. The grade score range for all strands is K to 7. *ADAM* is used for grades K-7/8 for assessment of foundational math skills.

- Numbers and Operations: 14 subtests; 661 criterion-referenced test items in 105 constructs
- Measurement: 7 subtests; 133 criterion-referenced test items in 34 constructs
- Geometry: 11 subtests; 203 criterion-referenced test items in 53 constructs
- Data Analysis: 8 subtests; 106 criterion-referenced test items in 36 constructs
- Algebraic Thinking: 4 subtests; 305 criterion-referenced test items in 43 constructs

DOMA Pre-Algebra

DOMA Pre-Algebra is a grade 4-7 multiple-measured criterion-referenced assessment. It consists of 14 subtests that address key foundational skills in mathematics. These subtests employ scope and sequence math skills organized in the order in which they would be taught to students across each of these subtests. These leveled skills are also aligned with instructional grade-level content standards. *DOMA Pre-Algebra*, by design, uses an interval scale, given that it is aligned to grade-level skills that span grades 4-7. *DOMA Pre-Algebra* scores are reported as grade-level scores with partial-year growth also noted. A single adaptive *DOMA Pre-Algebra* assessment is used for all grade-level students who are learning their grade 4 to 7 foundational math skills. The adaptive nature of *DOMA Pre-Algebra* was designed so that the assessment identifies the zone of proximal development (ZPD) of each student regardless of the student's actual grade level. The grade 4 to 7 focus of *DOMA Pre-Algebra* allows teachers and administrators to identify gaps in students' learning (previous years' standards that have not been met) as well as identify students who are working above their grade level.

DOMA Pre-Algebra uses test items that are criterion-referenced to pre-requisite knowledge expectations:

- Pre-Screening: 14 criterion-referenced test items, one from each subtest of the full assessment
- Integer Operations: 11 criterion-referenced test items
- Fraction Operations: 12 criterion-referenced test items
- Decimal Operations: 9 criterion-referenced test items
- Comparing and Converting: 10 criterion-referenced test items
- Estimating and Rounding: 6 criterion-referenced test items
- Evaluating Exponents: 6 criterion-referenced test items
- Ratios and Proportions: 5 criterion-referenced test items
- Simplifying Expressions: 6 criterion-referenced test items
- Coordinate Graphing: 8 criterion-referenced test items
- Linear Functions and Extending Patterns: 8 criterion-referenced test items
- Simple Equations: 6 criterion-referenced test items
- Geometry: 11 criterion-referenced test items
- Interpreting Data: 10 criterion-referenced test items
- Simple Probability: 7 criterion-referenced test items

<u>DOMA Algebra</u>

DOMA Algebra, a course-specific diagnostic assessment, consists of 11 Algebra I-specific constructs, as well as a pre-screening section much like the *DOMA Pre-Algebra* assessment.

- Pre-Screening: 22 criterion-referenced test items, representing two questions from each subtest

- Evaluating Advanced Exponents: 7 criterion-referenced test items
- Solving Linear Equations: 6 criterion-referenced test items
- Graphing and Analyzing Linear Equations: 9 criterion-referenced test items
- Relations and Functions: 7 criterion-referenced test items
- Solving and Graphing Inequalities: 5 criterion-referenced test items
- Solving and Graphing Systems: 8 criterion-referenced test items
- Polynomial Operations: 8 criterion-referenced test items
- Factoring Polynomials: 7 criterion-referenced test items
- Radical Expressions and Equations: 7 criterion-referenced test items
- Quadratic Equations: 7 criterion-referenced test items
- Rational Expressions and Equations: 8 criterion-referenced test items

LGL Edge: Culturally Relevant Context

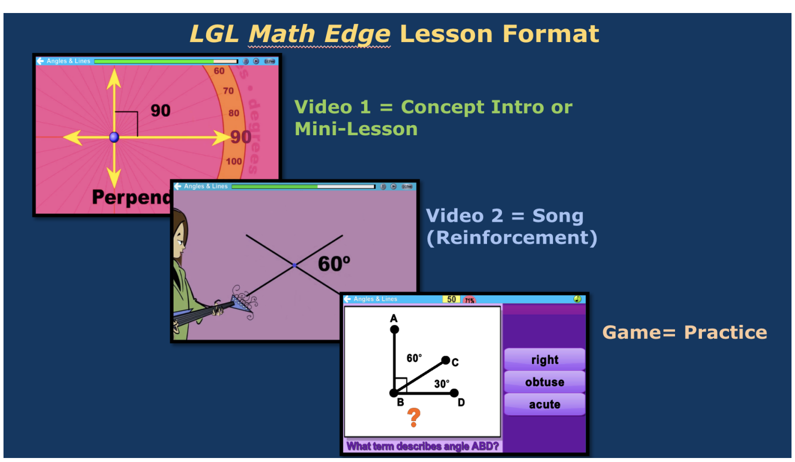
The more that research reveals how we learn and remember, the wiser we become about the elements necessary for an optimal environment for each unique learner. Let's Go Learn's instructional designers recognized early that equity and diversity could best be served by creating a pop culture environment with diverse characters, edgy art, bold colors, and a wide range of environments. If learners aren't fully engaged in the learning experience, content doesn't stick. Research by Eppart et al. (2021) found "that cultural, methodological and pedagogical barriers can significantly affect the use of educational technology in face-to-face and online classes and can consequently impact student learning." In other words, context matters: "Culturally responsive education that recognizes and affirms students' cultural and racial identity also leads to better academic outcomes" (Aceves & Orosco, 2014).

In Stembridge's book on Culturally Responsive Education (CRE), she recommends CRE as a model for including "the awareness of culture, race, ethnicity, gender, ability, and other social identity markers" that drive learning experiences that include all students. She offers that the most engaging instruction is the most personal (Stembridge, 2020).

LGL Edge creates a meaningful context carefully designed to support student learning. Each lesson features characters, music, animation, songs, contextual art, color, narration, and video. Each student learns in a unique way based on individual abilities and interests, preferred learning style, cultural and social background, and family and personal experience. We can take a page from Hollywood as it comes to the conclusion, finally, that more diverse casts perform far better at the box office. A spokesperson for the Academy stated: "Our values at the Academy are based on the belief that arts and sciences, including the arts and sciences of filmmaking, thrive from diversity" (as cited in Wilson, 2022). It is no wonder that LGL's *Edge* is so effective in supporting student gains.

LGL Edge: Music and the Brain

LGL Edge is one of the very few supplemental instructional programs that leverage the power of music to deepen and accelerate student learning. The music that drives *LGL Edge* lessons offers unique benefits, contributing to engagement, memory, recall, and comprehension (Hoeckner & Nusbaum, 2013). In fact, every lesson includes music to reinforce reading and mathematical instructional concepts. Music awakens the brain and "enables the left and right hemispheres to communicate, allowing for coordinated body movement as well as complex thoughts that require logic (left side)" (Pegasus, n.d.).



Patel (2010) in an article on music, evolution, and the brain puts forward the premise that "music is biologically powerful, meaning that it can have lasting effects on nonmusical abilities (such as language and attention) during the lifetime of individual humans." He goes on to say, "[M]usic often provides an important mnemonic device for storing long sequences of linguistic information" (Patel, 2010). As an example, he refers to the alphabet song that children in English-speaking countries learn to concretize the order of the letters and that adults still refer to when categorizing or organizing information according to order.

Both research and experience inform us that music impacts emotion. According to Ahmad and Rana (2015), "Music has the potential to influence mood, feelings, and thoughts; it has the ability to change the emotional and physical status of people, whether they are in bad, good, or sad moods." A growing body of neurological research provides evidence that when it comes to learning, emotions matter: "[T]he aspects of cognition that are recruited most heavily in education, including learning, attention, memory, decision making, motivation, and social functioning, are both profoundly affected by emotion and in fact subsumed within the process of emotion" (Immordino-Yang & Damasio, 2007, p. 7).

Another positive characteristic of the integration of music and learning is that it brings a contemporary form of technology and entertainment into the classroom. In a recent ATD blog, Meacham (2022) reminds us of the following: "Music makes learning more fun, which makes us want to learn more. Music increases dopamine levels in the brain's reward center, stimulating a desire to learn more. This reward cycle can increase memory performance for nonmusical information that is associated with the music."

LGL Edge: Narrative Rate and Prosody

Narration of instructional content in *LGL Edge* lessons deliberately adjusts the rate of speech and prosody to increase learner oral comprehension and fluency. The reason for adding these features to each lesson is to ensure optimal learning for all students. Often students miss content when oral instruction is delivered at a fast rate. Not only does this reduce listening and content comprehension, but it impacts fluency. Narrative prosody models not just oral fluency but also reading fluency. Research proposes "providing a narrated text to a visual source (multimodality) instead of combining the visual source with an explanatory text in writing (unimodal)" (Tracey, 2022).

Adjusting the rate of speech of the content-rich narration increases the comprehension, particularly of English learners and struggling readers: "For people who lack proficient comprehension…slowing speech rate can provide a substantial advantage" (King & East, 2011; Hux et al., 2020). Research by McBride (2011) also found "a slower rate of speech yielded higher scores on comprehension questions."

Prosody and reading and oral fluency go hand in hand. Prosody "encompasses a variety of phenomena: emphasis, pitch accenting, intonational breaks, rhythm, and intonation" (Wagner & Watson, 2010). Anyone who has sat through a lecture delivered in a monotone or worked with a digital program that uses AI narration will not be surprised that prosody improves student engagement (Servan et al., 2017). Research on infants provides evidence that prosody has an impact on language development: "[P]rosody influences how infants remember linguistic stimuli and even helps with extracting groups of words from continuous speech" (Hawthorne, 2014). Researchers have also found that prosody "can convey extra information beyond just words. This powerful form of communication can be used to improve students' recall" (Parr, 2020).

LGL Edge: Multiple Modes of Learning and Thinking

LGL Edge lessons present instruction in multiple modes to take advantage of the different ways that each person learns and thinks. Using multiple modes reduces the extraneous cognitive load significantly for each learner by teaching skills in a format that is easier for individual students to process and absorb.

Each student learns in a unique way based on individual abilities and interests, preferred learning style, cultural and social background, and family and personal experience. *LGL Edge* offers a diverse blend of multimedia experiences so that every student is engaged and motivated by the learning activities: "[A]ll students are capable of learning, provided the learning environment attends to a variety of learning styles" (Irvine and York, 1995; Guild, 2001).

In addition to keeping the difficulty of learning activities within the learner's instructional level, we also know that learning happens best when it speaks to the affective dimensions of the learner's profile. That includes (a) how the instruction is tailored to the learner's interests and socioemotional level, (b) how the instruction provides feedback, and (c) how it challenges the learner to apply previously learned skills to new concepts (Ambrose et al., 2010; Miller, 2014).

Instructional modules were built to help students apply foundational concepts they already know about the topic and to present the material in an engaging context that is relevant to their age group. The *LGL Edge* series uses multimedia that is carefully designed to support student learning. Each lesson features characters, music, animation, songs, contextual art, color, narration, and video. Students learn in a unique way based on individual abilities and interests, preferred learning style, cultural and social background, and family and personal experiences.

Every *LGL Edge* lesson begins with a direct instruction segment, presented with animation and/or music. This segment serves to reteach and review concepts, strategies, and processes. Research shows that direct instruction is effective: “The findings of a recent longitudinal follow-up study of over 1,000 low-income minority students in compensatory education are illuminating. In both rural and urban areas, we found positive long-term effects, with students achieving higher reading, language, and mathematics scores on standardized tests than students who either had not participated in direct instruction or who had participated in other programs. Participation in direct instruction also lowered dropout rates and raised the proportion of students applying to college” (Gersten & Keating, 1987, p. 29).

LGL Edge: Gamification Supports Interactivity, Repeatability, Scoring, and Rewards

LGL Edge is designed with a game-based paradigm that intrinsically motivates students to accomplish activities without the limitations of time or previous failures. Game-based design benefits the learner by lowering the threat of failure, fostering a sense of engagement through immersion, sequencing tasks to allow early success, linking learning to goals, and creating a social context (Jenkins, 2005). In the context of a game, students can experiment and practice in a virtual environment without fear of reprisal. According to Jenkins (2005), “At their best, games put kids in charge of their own learning and, at the same time, make them conscious of the learning process itself by presenting challenges they need to work through or around.” In an engaging game-like environment, students can experiment and practice in a virtual world without fear of reprisal.

Let’s Go Learn’s *LGL Edge* is designed to help students achieve grade-level proficiency in reading and math while earning points in a motivating game. This is a design that resonates with today's students, who enjoy computer-based games and entertainment. *LGL Edge* was designed with long-term retention in mind, which is the focus of germane cognitive load theory. Students are required to demonstrate the highest level of mastery by repeating lessons until they get to the Gold level or to 95% accuracy. Students are not allowed to repeat a lesson until two days after their first time completing it. These factors support the deeper learning of students across the *LGL Edge* content areas of reading and mathematics.

References for Research Base

LGL Edge References

Aceves, T. C., & Orosco, M. J. (2014). Culturally responsive teaching. (Document No. IC-2). University of Florida, Collaboration for Effective Educator, Development, Accountability, and Reform Center. Retrieved from http://ceedar.education.ufl.edu/tools/innovation-configurations/

Ahmad, N., & Rana, A. (2015). Impact of music on mood: Empirical investigation. *Research on Humanities and Social Sciences*. https://www.researchgate.net/publication/285055978_Impact_of_Music_on_Mood_Empirical_Investigation

Conley, D. T. (2007). *Redefining college readiness*. Eugene, OR: Education Policy Improvement Center.

Dahan, D. (2015). P*rosody and language comprehension*. Wiley Interdiscip Rev Cogn Sci., 441-52. doi: 10.1002/wcs.1355. Epub 2015 Jun 25. PMID: 26267554.

Eppard, A. K., Bowles, M., & Johnson, J. (2021). EdTech culturation: Integrating a culturally relevant pedagogy into educational technology. *Electronic Journal of e-Learning*, *19*(6), 516–530. https://doi.org/10.34190/ejel.19.6.2065

Gee, J. P. (2004). *Language, learning, and gaming: A critique of traditional schooling*. New York: Routledge.

Gersten, R., & Keating, T. (1987). Long-term benefits from direct instruction. Association for Supervision and Curriculum Development. Retrieved from http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_198703_gersten.pdf

Guild, P.B. (2001). *Diversity, learning style and culture*. Retrieved from http://education.jhu.edu/PD/newhorizons/strategies/topics/Learning%20Styles/diversity.html

Hawthorne, K., & Gerken, L. (2014, November). From pauses to clauses: Prosody facilitates learning of syntactic constituency. *Cognition 133* (2):420-8. doi: 10.1016/j.cognition.2014.07.013. Epub 2014 Aug 23. PMID: 25151251; PMCID: PMC4163511

Hirumi, A. (2022). Instructional systems design. University of Central Florida. https://pressbooks.online.ucf.edu/eme6613ah/chapter/overview-systematic-design-process/#:~:text=pseudo%2Dinstruction)%2C%20key%20decisions,in%20unsubstantiated%20fads%20or%20opinions.

Hoeckner, B., & Nusbaum, H.C. (2013). Music and memory in film and other multimedia: The *Casablanca* effect. In S.L. Tan, A.J. Cohen, S.D. Lipscomb, & R.A. Kendall (Eds.), *The psychology of music in multimedia.* Retrieved from http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199608157.001.0001/acprof-9780199608157-chapter-11

Hux, K., Brown, J.A., Wallace, S., Knollman-Porter, K., Saylor, A., & Lapp, E. (2020, February 7). Effect of text-to-speech rate on reading comprehension by adults With aphasia. *American Journal of Speech and Language Pathology 29* (1), 168-184. doi: 10.1044/2019_AJSLP-19-00047. Epub 2019 Nov 5. PMID: 31689376; PMCID: PMC7231913

Immordino-Yang, M.H., & Damasio, A. (2007) We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education*. Retrieved from https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1751-228X.2007.00004.x

Irvine, J.J., & York, D.E. (1995). Learning styles and culturally diverse students: A literature review. In J.A. Banks & C.A. Banks (Eds.), *Handbook of research on multicultural education*. New York: Macmillan.

Jenkins, H. (2004a, June). Make meaning, not war. Independent Schools, *63*(4), 38.

Jenkins, H. (2004b). Reality bytes: Eight myths about video games debunked. *The video game revolution*. Retrieved from http://www.pbs.org/kcts/videogamerevolution/impact/myths.html

Jones, S. (2003). Let the games begin: Gaming technologies and entertainment among college students. Washington, DC: Pew Internet and American Life Project.

Kress, G. (2003). *Literacy in the new media age*. London: Routledge.

Levin, H. M. (1988). Accelerating elementary education for disadvantaged students. In *School*

Success for Students at Risk. Harcourt, Brace, Jovanovich, 209–225. Retrieved from https://files.eric.ed.gov/fulltext/ED305150. pdf#page=211

Mathematics and Reading Report Cards at Grades 4 and 8. (2022). The nation's report card. https://www.nationsreportcard.gov/

McBride, K. (2011). The effect of rate of speech and distributed practice on the development of listening comprehension. *Computer-assisted language learning*. https://www.researchgate.net/publication/232943547_The_effect_of_rate_of_speech_and_distributed_practice_on_the_development_of_listening_comprehension

McCaulley, M.H. (1980). *Psychological (Myers-Briggs) type differences in education*. Gainesville, FL: Center for Applications of Psychological Type, Inc.

Meacham, M. (2022). Strategic use of music to support learning. *ATD Blog*. Association for Talent Development. Retrieved from https://www.td.org/atd-blog/strategic-use-of-music-to-support-learning

Mitchell, A., & Savill-Smith, C. (2004). *The use of computer and video games for learning: A review of the literature*. London: Learning & Skills Development Agency.

Moon, S. M., & Reis, S. M. (2004). Acceleration and twice-exceptional students. In *A nation deceived: How schools hold back America's brightest students*, *2*, 109–119. Retrieved from https://gagc.org/resources/Documents/Schwartz/ A%20Nation%20Deceived-How%20Schools%20Hold%20Back%20America's%20Brightest %20 Students.pdf#page=118

National Center for Learning Disabilities. (2021). *Promising practices to accelerate learning for students with disabilities during COVID-19 and beyond*. Retrieved from https://www.ncld.org/wp-content/uploads/2021/02/2021-NCLD-Promising-Practices-to-Accelerate-Learning_FINAL.pdf

Olenchak, F. R. (2009). Effects of talents unlimited counseling on gifted/learning disabled students. *Gifted Education International*, *25*(2), 144–164. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?-doi=10.1.1.879.159&rep=rep1&type=pdf

Papert, S. (1998, June). Does easy do it? Children, games, and learning. Retrieved from www.papert.org/articles/Doeseasydoit.html

Parr, C. (2020). Why prosody is a power tool for teachers. *TES Magazine*. https://www.tes.com/magazine/archived/why-prosody-powerful-tool-teachers

Patel, A.D. (2010). Music, biological evolution, and the brain. In C. Levander & C. Henry (Eds.) *Emerging disciplines*. Rice University Press. Retrieved from https://pages.ucsd.edu/~rbelew/courses/cogs260_s10/readings/Patel10_music_evolution.pdf

Pegasus. (n.d.) Your brain on music. University of Central Florida. Retrieved from https://www.ucf.edu/pegasus/your-brain-on-music/

Popham, W. J. (2008). *Transformative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

Pratt, M. L. (1991). Arts of the contact zone. *Profession*, *91*, 33–40.

Rideout, V. J., Vandewater, E. A., & Wartella, E.A. (2003). *Zero to six: Electronic media in the lives of infants, toddlers, and preschoolers*. Retrieved from https://www.dcmp.org/caai/nadh169.pdf

Serban, O., Barange, M., Zojaji, S., Pauchet, A., Richard, A., & Chanoni, E. (2017). Interactive narration with a child: Impact of prosody and facial expressions. *ICMI '17: Proceedings of the 19th ACM International Conference on Multimodal Interaction*. Retrieved from https://dl.acm.org/doi/10.1145/3136755.3136797

Squire, K.D. (2004). Replaying history: Learning world history through playing *Civilization III* (Doctoral dissertation). Retrieved from http://www.academia.edu/1317076/Replaying_history_Learning_world_history_through_playing_Civilization_III

Stembridge, A. (2020). *Culturally Responsive Education in the classroom : An equity framework for pedagogy*. Routledge.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 257-285. Retrieved from http://onlinelibrary.wiley.com/doi/10.1207/s15516709cog1202_4/pdf

Thibodeaux, T. (2001). Gaming: The first CE contact. *Dealerscope: The Business of CE Retailing*, *3* (43), 42.

Taylor, T.A.H., Kamel-El Sayed, S., Grogan, J.F., Hajj Hussein, I., Lerchenfeldt, S., & Mohiyeddini, C. (2022, June 24). Teaching in uncertain times: Expanding the scope of extraneous cognitive load in the Cognitive Load Theory. *Front Psychol*, *13*. doi: 10.3389/fpsyg.2022.665835. PMID: 35814084; PMCID: PMC9263515

Wagner, M., Watson, D. G. (2010, January 1). Experimental and theoretical advances in prosody: A review. *Lang Cogn Process*, *25*, (7-9), 905-945. doi: 10.1080/01690961003589492. PMID: 22096264; PMCID: PMC3216045

Wilson, Josh. (2022). How is Hollywood's performance in diversity and inclusion in 2022? *Forbes*. Retrieved from https://www.forbes.com/sites/joshwilson/2022/04/22/how-is-hollywoods-performance-in-diversity-and-inclusion-in-2022/?sh=4bd30c95b514

LGL Math References

Center for Public Education. (2012). *High school rigor and good advice: Setting up students for success*. Retrieved from http://www.centerforpubliceducation.org/Main Menu/Staffingstudents/High-school-rigor-and-good-advice-Setting-up-students-to succeed

Garland, S. (2013). The math standards: Content and controversy. *The Hechinger Report*. Retrieved from http://hechingerreport.org/content/the-math-standards content-and-controversy_13325/

Hein, V., et al. (2013, November). Predictors of post-secondary success. *College & Career Readiness & Success Center at American Institutes for Research*. Retrieved from https://ccrscenter.org/sites/default/files/CCRS%20Center_Predictors%20of%20Postsecondary%20Success_final_0.pdf

Hirsch, C.R., & Coxford, A.F. (1997). Mathematics for all: Perspectives and promising practices. *School Science and Mathematics*, *97* (5), 232-241.

Holton, D., and Clarke, D.(2006, August). Scaffolding and metacognition. *International Journal of Mathematical Education in Science and Technology*, 37 (2), 127–43.

Jungjohann, K., Clarke, B., & Cary, M.S. (2009, October 2). *Early learning in mathematics: A formula for success.* Retrieved from http://www.slideserve.com/mulan/early-learning-in-mathematics-a-formula-for success

Liljedahl, P. (2021). *Building thinking classrooms in mathematics, grades K–12: 14 teaching practices for enhancing learning*. Thousand Oaks, CA: Corwin.

Musen, L. (2010). *Pre-algebra and algebra: Enrollment and achievement*. Retrieved from http://annenberginstitute.org/pdf/LeadingIndicator_Math.pdf

National Council of Teachers of Mathematics (NCTM). (2014). *Principles to actions: Ensuring mathematical success for all*. Reston, VA: NCTM. https://www.nctm.org/PtA/

NCTM, NCSM, & ASSM. (2021). *Continuing the journey: Mathematics learning 2021 and beyond*. Reston, VA: National Council of Teachers of Mathematics. https://www.nctm.org/uploadedFiles/Research_and_Advocacy/collections/Continuing_the_Journey/NCTM_NCSM_Continuing_the_Journey_Report-Fnl2.pdf

TNTP & Zearn. (2021). *Accelerate, don't remediate: New evidence from elementary math classrooms (report)*. Retrieved from https://tntp.org/publications/view/teacher-training-and-classroom-practice/accelerate-dont-remediate

Webb, N.L., & Coxford, A. F. (Eds.) (1993). *Assessment in the mathematics classroom: 1993 yearbook.* Reston, VA: National Council of Teachers of Mathematics.

William, D. (1999, Spring). Formative assessment in mathematics, part 1: Rich questioning. *Equals: Mathematics and Special Educational Needs*, *5* (2), 15-18.

LGL ELA/Reading References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Urbana Champaign, IL: University of Illinois, Reading Research and Education Center.

Adams, M. J. (1998). The three-cueing system. In F. Lehr & J. Osborn (Eds.), *Literacy for all issues in teaching and learning* (pp. 73-99). New York: Guilford Press.

Allen, M. & Yen, W. (1979). *Introduction to measurement theory*. Long Grove: Waveland Press.

August, D. & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth*. Mathway, NJ: Lawrence Erlbaum Associates, Inc.

Barshay, J. (2020). Four things you need to know about the new reading wars. *The Hechinger report.* Retrieved from https://hechingerreport.org/four-things-you-need-to-know-about-the-new-reading-wars/

Bear, D., Invernizzi, M., Templeton, S., & Johnston, F. (2000). *Words their way: Word study for phonics, vocabulary, and spelling instruction*. Upper Saddle River, NJ: Prentice-Hall, Inc.

Dunn, L.M. (1959). *Peabody Picture Vocabulary Test*. Bloomington, MN: Pearson Assessments.

Farrell, L., Hunter, M., Davidson, M., & Osenga, T. (2019). The simple view of learning. *Reading rockets*. Retrieved from https://www.readingrockets.org/article/simple-view-reading

Flores, B., Tefft, C., & Esteban, D. (1991). Transforming deficit myths about learning, language, and culture. *Language Arts*, *68*, 369-379.

Fry, E., Kress, J., & Fountoukidis, D. (2004). *The reading teacher's book of lists*. Paramus, NJ: Prentice-Hall, Inc.

Gillet, J.W., & Temple, C. (1994). *Understanding reading problems: Assessment and instruction* (4th ed.). New York: Harper Collins College Press.

Green, R. (2022). Evidence-based reading system outweighs state's "balanced literacy" approach. CalMatters. Retrieved from https://calmatters.org/commentary/2022/07/evidence-based-reading-system-outweighs-states-balanced-literacy-approach/

Johns, J. (2005). *Basic Reading Inventory: Pre-primer through grade twelve and early literacy assessments*. Dubuque, IA: Kendall/Hunt Publishing.

Juel, C. (1996). What makes literacy tutoring effective? *Reading Research Quarterly*, 31(3), 268-289.

Krashen, S. (2010). The Goodman/Smith Hypothesis, the Input Hypothesis, the Comprehension Hypothesis, and the (even stronger) case for free voluntary reading. In *Defying convention, inventing the future in literacy research and practice: Essays in tribute to Ken and Yetta Goodman*. P. Anders (Ed.) New York: Routledge. Retrieved from http://www.sdkrashen.com/content/articles/many_hypothesis.pdf

Leslie, L., & Caldwell, J. (1994). *The Qualitative Reading Inventory* (2nd ed.) New York: Harper Collins College Press.

Lyon, G. R. & Chhabra, V. (2004). The science of reading research. *ASCD*. Retrieved from https://www.ascd.org/el/articles/the-science-of-reading-research

International Literacy Association Board Board. (2020). ILA launches landmark literacy research resource. Retrieved from https://www.literacyworldwide.org/docs/default-source/resource-documents/ila-board-statement-on-science-of-reading.pdf?sfvrsn=c762bc8e_8

McCallum, R., & Wasik, B. A. (1997). Volunteer tutoring programs: A review of research on achievement outcomes. Johns Hopkins University & Howard University, Center for Research on the Education of Students Placed at Risk.

National Reading Panel. (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Retrieved from https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

Ordetx, K.(2021). What is the science of reading? Institute for Multi-Sensory Education. Retrieved from https://journal.imse.com/what-is-the-science-of-reading/

Pinell, G.S., Pikulski, J.J., Wixson, K.K., Campbell, J.R., Gough, P.B., & Beatty, A.S. (1995). National Assessment of Educational Progress (NAEP) Oral Reading Fluency Scale: Listening to children read aloud. Washington, D.C.: Office of Educational Research and Improvement, U.S. Department of Education.

Pressley, M., & Woloshyn, V. (1995). *Cognitive strategy instruction that really improves* *children's academic performance*. Cambridge, MA: Brookline Books.

Ruddell, R. (1999). *Teaching children to read and write: Becoming an influential teacher* (2nd ed.). Needham Heights, MA: Allyn & Bacon.

Schwartz, S. (2022). Which states have passed science of reading laws? What's in them? *EdWeek*. Retrieved from https://www.edweek.org/teaching-learning/which-states-have-passed-science-of-reading-laws-whats-in-them/2022/07

Shanahan, T., & Barr, R. (1995). Reading recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly*, *3* (4), 958-996.

Snow, C., Burns, S., & Griffin, P. (Eds). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Stanovich, K. E. (1993-94). Romance and reality (Distinguished Educator Series). *Reading Teacher*, *47* (4), 280-91.

Townsend, D. (2021). What is the "science of reading" and why does it matter? *Nevada Today.* Retrieved from https://www.unr.edu/nevada-today/news/2021/atp-science-of-reading

Wasik, B. A. (1998). Using volunteers as reading tutors: Guidelines for successful practices. *The Reading Teacher*, *51* (7), 562-570.

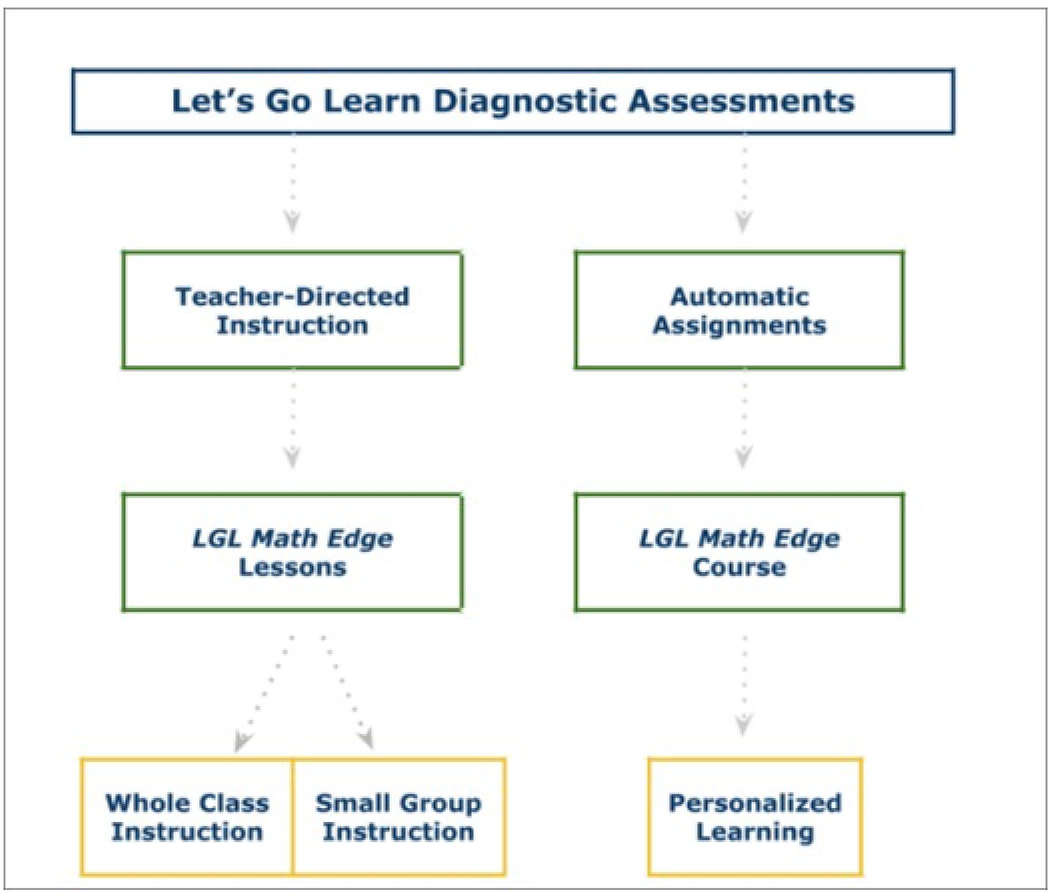
Woodward, J., & Baxter, J.A. (1997). The effects of an innovative approach to mathematics on academically low-achieving students in inclusive settings. *Exceptional Children*, *63* (3), 373-388.

Yopp, H. K. (1992). Developing phonemic awareness in young children. *Reading Teacher*, *45* (9), 696-703.

LGL Edge Evidence Base

Introduction

Founded in 2000, Let's Go Learn offers a range of solutions designed to improve student achievement. Initially, the company developed two diagnostic assessments: (1) The *Diagnostic Online Reading Assessment* (*DORA*) and (2) The *Adaptive, Diagnostic Assessment of Mathematics* (*ADAM*). These comprehensive assessment tools were designed to provide students, educators, and families with clear, actionable data related to student performance throughout an academic year. Each assessment consists of subtests that evaluate areas such as vocabulary and reading comprehension to develop a complete picture of student ability and opportunity for growth so that teachers can direct their instruction accordingly.



In addition to creating 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. During a quiz, if students give 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 shareable 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.

Third Party Studies

Let's Go Learn is committed to documenting case studies to ensure product effectiveness and provide feedback for ongoing revisions. This section documents recent case studies of Edge use at schools across the United States.

Study 1: California School District, 2018-2019

<u>Introduction</u>

This retrospective study was conducted with the support of a central California school district; it 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 (a 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.

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 the 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's 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 as 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 benefited 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 groups, and to the *Math Edge* and comparison groups.

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* were 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 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 measure 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 were reported absent during the defined school year.

Participants

This study benefited 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 reduced lunch.

Recent academic performance figures for the district identify approximately 35% of students as proficient in reading/Language Arts and 24% as proficient in mathematics. The average graduation rate is 83%.

The district reviewed the analysis plan and then provided the researchers with 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 as 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 had met each of the following criteria: (a) had 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 grades 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 the total number of lessons completed. To be classified into either *Edge* subject-based category (treatment), the student had to 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," as stated above) was met for just 2.4% of students 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 is broken down by grade level.

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

	ELA		**Math**	
Group	**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).

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 the following section for weighting details).

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 composed 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 groups 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 were 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, Individuals with Disabilities Education 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, using 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 included slightly more students on Individualized Education Plans (IEPs) 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 differences 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 than 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 were 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 ($p \leq .05$)

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. As with 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 the 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 eta²**
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 levels. Across all grade levels, means were higher for the LGL students. However, the mean difference was only significant for students in grade six.

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 eta^2**
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 the 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 completion of the analysis.

The final attempted ANCOVA procedure analyzed ELA gain scores with both the baseline scores 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, was 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 (pretest).

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 eta^2**
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 eta^2**
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 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 benefited 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 but 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 in the programs it produces.

Study 2: Jersey City 2015 - 2017

Purpose of the Study

The purpose of this study was to examine the effect of adaptive supplemental lessons on elementary students' mathematics achievement in a large urban school district serving a diverse population. The district is diverse in race, ethnicity, and socioeconomic status and has a large population of English language learners and students with disabilities. This study examined the impact adaptive learning had on mathematics achievement in the overall population and among individual subgroups that must be reported on as per NCLB. This study explored the *ADAM* results of students in third through sixth grades. Although *LGL Math Edge* can cover more levels, these grades were selected because algebraic skills are built in grades three through six (Knuth, Stephens, Blanton, & Gardiner, 2016; Napaphun, 2012). Research shows that students are more likely to take advanced math courses if they achieve success in Algebra I (Byun, Irvin, & Bell, 2015).

It is important to evaluate adaptive learning systems based on learner outcomes. Understanding the effectiveness of *LGL Math Edge* on student mathematics achievement may guide educators' future decisions regarding classroom practices. For example, administrators may make it mandatory for teachers in their building to incorporate adaptive supplemental lessons in a blended learning environment. It is necessary to understand how adaptive learning systems impact the achievement of all learners to determine if there is a correlation between adaptive learning and student outcomes.

Theoretical Framework

The theoretical framework for this study is rooted in cognitive load theory (CLT; see also Chapter 2). Cognitive load theory suggests that learning is at its best when occurring in conditions aligned with human cognitive architecture (Sweller, Ayres, & Kalyuga, 2011). Working memory is limited to material that it can accommodate simultaneously and is limited in capacity and duration for new information. Cognitive load refers to the amount of mental effort used in working memory (Kalyuga, 2011). Limitations of working memory impact achievement in complex tasks.

Sweller’s (1988) cognitive load theory argues that the type and amount of cognitive load that learners experience is a critical component in successful learning. In order to be effective, the cognitive load must be minimized during the learning process. A way to reduce cognitive load is to practice skills until they become automatic.

Technology-based instruction specifically tailored to conform to existing knowledge of human cognition is likely to be effective in reducing cognitive load. Adaptive learning technology aids learners who require support in building new knowledge bases without hindering knowledgeable learners (Kalyuga, 2011).

Whenever learners acquire more knowledge in a specific domain, instructional design methods must be adjusted. The expertise reversal effect occurs when successful instructional procedures for novice learners have no effect, or an adverse effect, as learners acquire more expertise (Kalyuga, Ayres, Chandler, & Sweller, 2003). A major instructional implication of the expertise reversal effect is the need to tailor instruction to different learner expertise levels. As a learner develops more experience in a task domain, the level of expertise may change. Adaptive learning considers levels of expertise in real time and adapts as the levels change. Learners get less guided instruction as their expertise increases, which may assist in acquiring advanced knowledge in a domain.

LGL Math Edge, the adaptive learning system used in this study, was designed using cognitive load theory as a framework (McCallum, 2016). Initially, all students complete the fully diagnostic *ADAM* assessment. This assessment breaks math into 44 subtests and finds students' instructional points within the linear scope and sequence of each subtest. This means that target skills are selected exactly at each student's instructional point. Thus the intrinsic cognitive load is reduced by not presenting topics that are too hard or complex for a student to learn. Regarding extraneous cognitive load, the program uses music, graphics, and audio to present each lesson (McCallum, 2016). Students are not required to read lessons. These factors reduce the cognitive load significantly (Kalyuga & Liu, 2015; Mayer & Moreno, 2003; Murray & Perez, 2015). As far as germane cognitive load, *LGL Math Edge* was designed with long-term retention in mind (McCallum, 2016). Thus, students must demonstrate the highest level of mastery by repeating lessons until they get to 95% accuracy. Students are not allowed to repeat lessons until two days after their first time completing the lesson. When students complete the adaptive mathematics lessons in *LGL Math Edge*, they are expected to show achievement gains on the *ADAM* assessment.

Research Questions

The current study utilized a quasi-experimental research design. The study examined the results achieved using a treatment in which all third through sixth-grade students were to receive at least 0.5 hours per week of adaptive math lessons in a blended learning environment during the regular math block, and these results were compared with the results achieved without the treatment. As described in the previous sections, the primary research goal was to determine whether there was a significant difference in the results for the students receiving the treatment compared to the control group (those not receiving the treatment). Mathematics achievement was defined as the difference between pretest and posttest outcomes on the *ADAM* assessment. The study was guided by the following research questions:

- RQ1. What effect do adaptive mathematics lessons have on elementary students' mathematics achievement?
- H1. There is a statistically significant difference in elementary students' mathematics achievement when adaptive mathematics lessons are implemented.
- Ho. There is no statistically significant difference in elementary students' mathematics achievement when adaptive mathematics lessons are implemented.
- RQ2. What effect do adaptive mathematics lessons have on student achievement based on the subgroups of gender, race/ethnicity, SWD classification, ELL classification, and socioeconomic status?
- H1. There is a statistically significant difference in student mathematics achievement based on the subgroups of gender, race/ethnicity, SWD classification, ELL classification, and socioeconomic status when adaptive mathematics lessons are implemented.
- Ho. There is no statistically significant difference in student mathematics achievement based on the subgroups of gender, race/ethnicity, SWD classification, ELL classification, and socioeconomic status when adaptive mathematics lessons are implemented.
- RQ3. Does a significant relationship exist between time spent on *LGL Math Edge* and mathematics achievement from pretest to posttest?
- H1. There is a statistically significant relationship between time spent on *LGL Math Edge* and mathematics achievement from pretest to posttest.
- Ho. There is no statistically significant relationship between time spent on *LGL Math Edge* and mathematics achievement from pretest to posttest.

Significance of the Study

Students who are active learners tend to experience greater academic success (Como & Randi, 1999). This type of learning creates the opportunity for differentiated, personalized, and independent learning. The personalized nature of adaptive learning allows the program to meet students at their academic level. This permits the scaffolding of learning, thus facilitating improvement in weaker areas. Previous studies (Johnson & Samora, 2016; Murray & Perez, 2015) have examined the effect of adaptive lessons on student achievement, however, there is limited research in which data is disaggregated by race, socioeconomic status, SWD classification, and home language (Marchand-Martella, 2014). Much of the research on the effect of adaptive learning on student achievement is based on small-scale studies, which means that researchers have relied on results from single classes or schools in order to reach conclusions about their interventions (Blackwell, Trzesniewski, & Dweck, 2007). This study used district-wide data and had a large sample size of approximately 15,000 participants for analysis. The study may provide insight into the effects of adaptive learning on mathematics achievement for various subgroups. Additionally, this study provides a meaningful contribution to the literature on adaptive learning's effect on student achievement.

Research Design

To fulfill the purpose of this research, the experimental study utilized a quantitative methodology. Experimental research attempts to ascertain whether a specific treatment influences an outcome (Creswell, 2014). In this study, the specific treatment was the use of adaptive mathematics lessons, and the outcome was student achievement in mathematics. Quasi-experiments, which use nonrandomized assignments, were conducted (Keppel, 1991). The researcher received IRB approval from the school district that participated in this study, Jersey City Public Schools, and IRB exemption from the researcher's university, New Jersey City University. Secondary data was used for this study. Let's Go Learn provided the researcher with data for the 2015-2016 and 2016-2017 school years that was stored on the LGL platform. The control group consisted of all third through sixth grade students during the 2015-2016 school year. The control group did not receive adaptive learning lessons. The treatment group consisted of all third through sixth grade students during the 2016-2017 school year. The treatment group was to receive 0.5 hours per week of supplemental adaptive mathematics lessons. The adaptive lessons were completed over the 30-week timespan between the pretest and posttest, meaning every student should have a total of at least 15 hours in *LGL Math Edge*. Although there were approximately 32 weeks between the pretest and posttest, schools were closed during the weeks of winter and spring break. Therefore, students were able to receive 30 weeks of adaptive instruction. The data for both groups included student demographic information and *ADAM* results.

ADAM was administered in October and May, and it served as both the pretest and posttest. The mean gain score was determined by finding the average difference between the pretest and posttest scores of each group. The data for each group was analyzed as a whole, meaning the total number of third through sixth grade students, and was not broken down by grade level. Student demographic information included gender, race/ethnicity, SWD classification, ELL status, and socioeconomic status. The data for the treatment group also included the amount of time each student spent in *LGL Math Edge*. The target amount of time for each student to spend in *LGL Math Edge* was 15 hours.

The study's methodology was threefold. First, the overall achievement gains of third through sixth graders on *ADAM* for the treatment group were analyzed. *ADAM* was administered in October and May, and it served as both the pretest and posttest. The achievement gains were the difference in total scores between the pretest and posttest. In the 2016–2017 school year, the district mandated that all elementary and middle school students receive at least 0.5 hours per week of adaptive math lessons in a blended learning environment during the regular math block. This examination demonstrated the overall impact of the adaptive learning intervention on the treatment group. Second, the achievement gains of student subgroups in the treatment group were analyzed and a mean gain score was determined. The analysis consisted of distinct subgroup data, including overall student population, SWD population, ELL population, population by race and ethnicity, population by gender, and population by socioeconomic status. Although holistic data highlights the overarching trends in a large-scale implementation, it is essential to look at how adaptive learning impacts various student subgroups. Disaggregating the results helped to remove any potential masking of data trends and ensured equity amongst all students by determining if outcomes varied by subpopulation. Finally, the mean gain scores of the control and treatment groups were compared. Data comparisons consisted of all the aforementioned subgroups. This comparison determined if the adaptive supplemental lessons had had a significant impact on elementary student achievement in mathematics.

This study also investigated if a significant relationship existed between time spent in *LGL Math Edge* and mathematics achievement from pretest to posttest. Data from the treatment group were analyzed. A random effect model was used to test the significance of the continuous variable of math time in relation to other variables, including gender, race/ethnicity, SWD classification, ELL status, and socioeconomic status.

Definition of Terms

Achievement gap: A significant disparity in academic performance between different groups of students (“Achievement Gap,” 2013).
Adaptive learning: An instructional method that uses technology as a means of lesson delivery, allowing teachers to spend more time with students and offer personalized learning (Lishon-Savarino, 2016).
Blended learning: A teaching method that combines online learning with traditional face-to-face instruction (Chen & Jones, 2007).
Conventional instruction: Instructional delivery that occurred without adaptive lessons during the 2015–2016 school year.
Intelligent tutoring systems: Educational applications of artificial intelligence and machine-learning technologies that provide customized instruction and immediate feedback without teacher intervention (Oliveira & Nascimento, 2012).
Personalized learning: A process that uses observation to tailor interventions for individual students to increase the likelihood of success and may or may not involve technology (Newman, 2013).

Limitations

It is critical to recognize that all studies have methodological limitations. This study had a large sample size and ample data points, and it followed current educational data collection practices to minimize any potential limitations. The data were collected using a valid and reliable measure that is aligned to mathematics standards. The study components, *ADAM* (diagnostic test) and *LGL Math Edge* (treatment), are both by Let’s Go Learn. Chapter 3 details the validity and reliability of the diagnostic test and how diagnostic test performance can predict standardized assessment success.

One limitation of the study was the implementation of the adaptive mathematics lessons. During the 2015–2016 school year, it was expected that teachers follow the curriculum and prepare students for the standardized test. The expectations for the 2016–2017 school year were the same, except that the district mandated 0.5 hours per week of adaptive lessons during the regular math block. That meant that over the 30-week span between the pretest and posttest, every student should have a total of at least 15 hours in *LGL Math Edge*. Teachers were not given a specific way to implement this mandate, nor were there consequences for teachers whose students did not meet the target of 0.5 hours of adaptive lessons per week.

Another study limitation was that the data that were employed were secondary and came from a quasi-experimental method. The data were collected without direct involvement or control over the variables among individual classrooms. The large sample ensured that extraneous variables were minimized. During the presentation of the findings and recommendations, this limitation was addressed.

Controlling the study as a collection of secondary quantitative data allowed for an objective analysis. The evaluation of a technology implementation, such as adaptive learning, is often based upon changes in assessment outcomes. Standardized test scores are utilized in a similar nature by almost every educational institution so that administrators can make informed decisions about instructional practices. Additionally, large-scale educational research studies often use state and district data to conduct secondary research and make comparisons (Center for Research on Education Outcomes, 2013). This type of data has been used to draw comparisons between districts, states, and countries. The researcher was not personally invested in the creation of the assessment or treatment, which facilitated an impartial evaluation of the adaptive learning system.

The purpose of this experimental study was to examine the effect of adaptive supplemental lessons on elementary students' mathematics achievement. All third through sixth grade students in the treatment group were to receive at least 0.5 hours per week of adaptive math lessons in a blended learning environment within their regular math block. The study examined the differences in the mean grade score gains on *ADAM* between the control and treatment groups. In this study, grade score gains are the differences in score from pretest to posttest on *ADAM*. The pretest was administered in October and the posttest in May of each year. In addition, the study examined the mean grade score gains of the control and treatment groups using the subgroups of gender, race/ethnicity, SWD classification, ELL classification, and socioeconomic status. Furthermore, this study examined the relationship between time spent in *LGL Math Edge* and changes in mathematics achievement from pretest to posttest. The data consisted of *ADAM* grade score gains from a pretest to posttest for third through sixth-grade students in the 2015– 2016 and 2016–2017 school years from a large urban school district in New Jersey.

Demographics

The participants in this study consisted of 7,114 students in grades 3–6 in the control group and 7,733 students in grades 3–6 in the treatment group. The participants were enrolled in 27 Title I schools within a large urban school district in New Jersey. There were 3,446 females and 3,668 males in the control group. There were 3,789 females and 3,944 males in the treatment group.

The subgroup of race/ethnicity was divided into four categories: African American, Asian, Caucasian, and Hispanic. The control group consisted of 2,192 African-American students, 1,085

Asian students, 791 Caucasian students, and 2,934 Hispanic students. The treatment group consisted of 2,307 African-American students, 1,254 Asian students, 931 Caucasian students, and 3,088 Hispanic students. These subgroups were reported according to NLCB (2001) requirements using the OMB (1997b) standards of Asian, Black, Hispanic, and White.

The subgroup of classification referred to whether the student was classified as a student with a disability. The Individuals with Disabilities Education Act (IDEA; 2004) identifies 13 conditions by which a student can be classified as a SWD: specific learning disability, other health impairment, Autism spectrum disorder, emotional disturbance, speech or language impairment, visual impairment, deafness, hearing impairment, deaf or blindness, orthopedic impairment, intellectual disability, traumatic brain injury, and multiple disabilities. There were 1,046 students classified as SWD and 6,068 students not classified as SWD in the control group. The treatment group consisted of 1,200 students classified as SWD and 6,533 students not classified as SWD.

ELL students have limited English proficiency and are unable to learn cogently in English. ELL students are typically participants in bilingual, dual language, or ESL programs. There were 511 students with an ELL classification in the control group and 613 students with an ELL classification in the treatment group.

Economically disadvantaged students are students who receive a free or reduced lunch. The National School Lunch Program (2008) considers students from low-income families earning below 185% of the federal poverty line eligible to receive a free or reduced lunch. There were 5,345 students classified as economically disadvantaged in the control group and 5,809 students classified as economically disadvantaged in the treatment group. The demographics of the control and treatment groups are shown in Table 3.

Table 1: Demographic Information for the Control and Treatment Groups

Characteristic	Control (2016)	Treatment (2017)
Male	3668	3944
Female	3446	3789
African American	2192	2307
Asian	1085	1254

Caucasian	791	931
Hispanic	2934	3088
SWD	1046	1200
Non-SWD	6068	6533
ELL	511	613
Non-ELL	6603	7317
Economically Disadvantaged	5345	5809
Non-economically Disadvantaged	1768	1924
Total Students	7,114	7,733

Findings

Research Question 1

In Research Question 1, the researcher asked: What effect do adaptive mathematics lessons have on elementary students' mathematics achievement? A Welch two-sample, one-tailed t-test was used to determine if a statistically significant difference existed between the mean gain score of students in adaptive math learning conditions and the mean gain score of students in conventional learning conditions. Welch's t-test is designed for unequal variances, but the assumption of normality is maintained. A Welch two-sample, one-tailed t-test was conducted to analyze the first null hypothesis, H_o1: There is no statistically significant difference in elementary students' mathematics achievement when adaptive mathematics lessons are implemented.

Hypothesis testing. Table 2 represents the descriptive statistics for the control and treatment groups. There were 7,114 students in the control group and 7,733 students in the treatment group. The mean grade score gain for the control group was .49 with a standard deviation of .41. The mean grade score gain for the experimental group was .72 with a standard deviation of .48.

Table 2: ***Descriptive Statistics for the Control and Treatment Groups***

Group	n	M	SD
Control	7114	.48	.41
Treatment	7733	.71	.47
Total	14847	.61	.46

Analysis: A Welch two-sample, one-tailed t-test was used to determine if a significant difference in student achievement existed between adaptive math learning conditions and conventional learning conditions according to *ADAM* mean grade score gains. The grade score gains on *ADAM* functioned as the dependent variable, with the control group taught by conventional instruction and the treatment group taught by conventional instruction and adaptive lessons. The year tested represented the independent variable. There was a statistically significant difference between the control and treatment groups at $\alpha = .05$ level, $t\ (14775.89) = -31.001$, $p < .001$, CI.95 -.24, -.21. Furthermore, Cohen's effect size was large ($d = .52$). Students in the treatment group ($M = .72$, $SD = .48$) scored significantly higher than students in the control group ($M = .49$, $SD = .41$), with a mean difference of .23. Therefore, null hypothesis one was rejected.

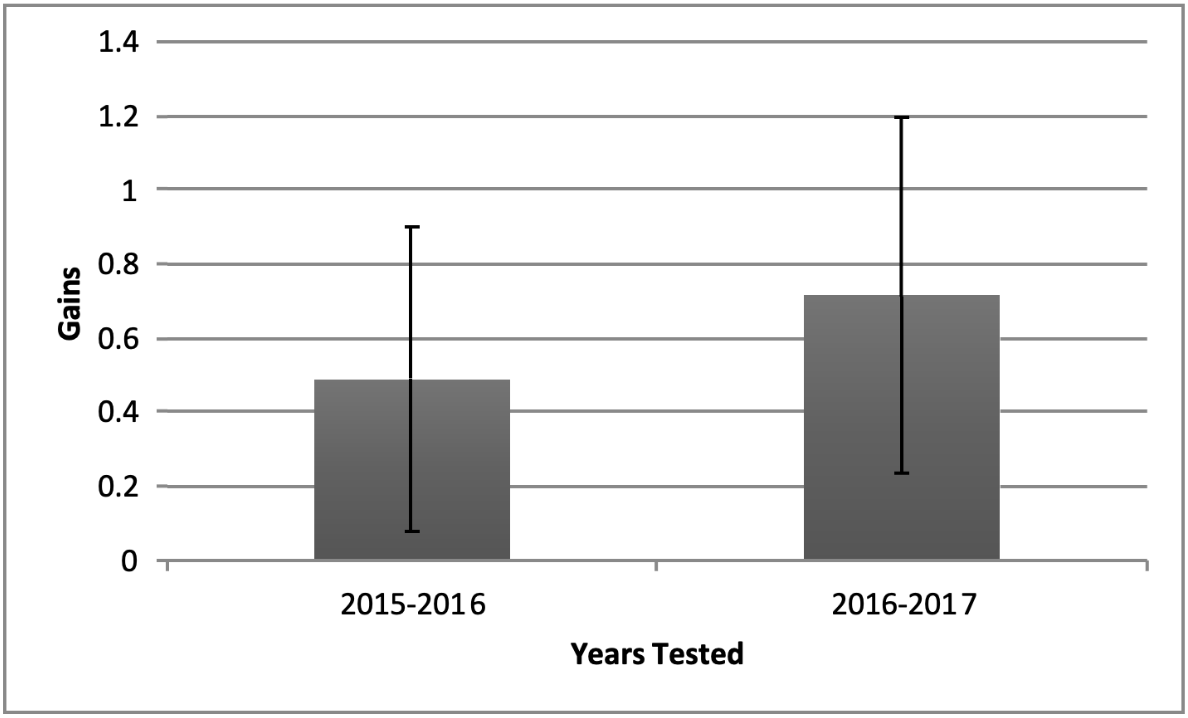


Figure 1 Mean grade score gains in the control and treatment groups

Hypothesis results: The first null hypothesis stated that there is no statistically significant difference in elementary students' mathematics achievement when adaptive mathematics lessons are implemented. Therefore, the researcher rejected the first null hypothesis. The mean grade score gains for the control group (*M* = .49, *SD* = .41) and the treatment group (*M* = .72, *SD* = .48) revealed a statistically significant difference in mathematics achievement scores. The treatment group scored statistically significantly higher than the control group, with a mean difference of .23.

Research Question 2

In the second research question, the researcher asked if there was a statistically significant difference in student mathematics achievement based on the subgroups of gender, race/ethnicity, SWD classification, ELL classification, and socioeconomic status when students participated in adaptive learning conditions. Given that the subgroups had four or more distinct subgroups within them to test, a one-way analysis of variance (ANOVA) was used to investigate whether there was any evidence that the means differed. Each sample size was large enough to assume normality of means based on the central limit theorem, except for some smaller ethnicity subgroups. Testing on these smaller groups was eliminated. All one-way ANOVA tests violated the assumption of homogeneity of variance; therefore, the Welch statistic was used. If the Welch ANOVA determined there was evidence that there was a statistically significant difference in means, it was followed up by a Games-Howell post-hoc test to find significant differences within the groups. An alpha level of .05 was used for all statistical tests. A Welch ANOVA was conducted to analyze the second null hypothesis, H_o2: There is no statistically significant difference in student mathematics achievement based on the subgroups of gender, race/ethnicity, SWD classification, ELL classification, and socioeconomic status when adaptive mathematics lessons are implemented.

Gender

Hypothesis testing: Table 5 represents the descriptive statistics for females and males. There were 14,845 participants in the study. The control group consisted of 7,114 participants, composed of female (n = 3,445) and male (n = 3,668) students. The treatment group consisted of 7,732 participants, composed of female (n = 3,788) and male (n = 3,944) students. The mean grade score gain for females in the control group was .49 with a standard deviation of .39. The mean grade score gain for females in the treatment group was .73 with a standard deviation of .43. The mean grade score gain for males in the control group was .47 with a standard deviation of .42. The mean grade score gain for males in the treatment group was .70 with a standard deviation of .51.

Table 3: Descriptive Statistics for Gender

Group	n	M	SD
Control Female	3445	.49	.39
Control Male	3668	.47	.42
Treatment Female	3788	.73	.43
Treatment Male	3944	.70	.51
Total	14845		

Analysis: A Welch ANOVA was administered to determine the differences between the mean grade score gains of the control and treatment groups based on gender. The grade score gains functioned as the dependent variable and gender as the independent variable. Results of the

Welch ANOVA showed an overall statistically significant difference in group means at α = .05 level, Welch's F (3, 8242) = 334.79, p <.001, $\eta^{\#}$ = .06. The magnitude of the difference in the means and the effect size was medium ($\eta^{\#}$= .06). A Games-Howell post-hoc test was conducted given the statistically significant omnibus Welch ANOVA test. The Games-Howell test indicated the female treatment group scored statistically significantly higher than the female control group, with a mean difference of .24. The male treatment group scored statistically significantly higher than the male control group, with a mean difference of .23. Therefore, Null Hypothesis 2 was rejected. There was a statistically significant difference in student mathematics achievement based on the subgroup of gender when adaptive mathematics lessons were implemented. Figure 2 shows the mean grade score gains by gender in the control and treatment groups.

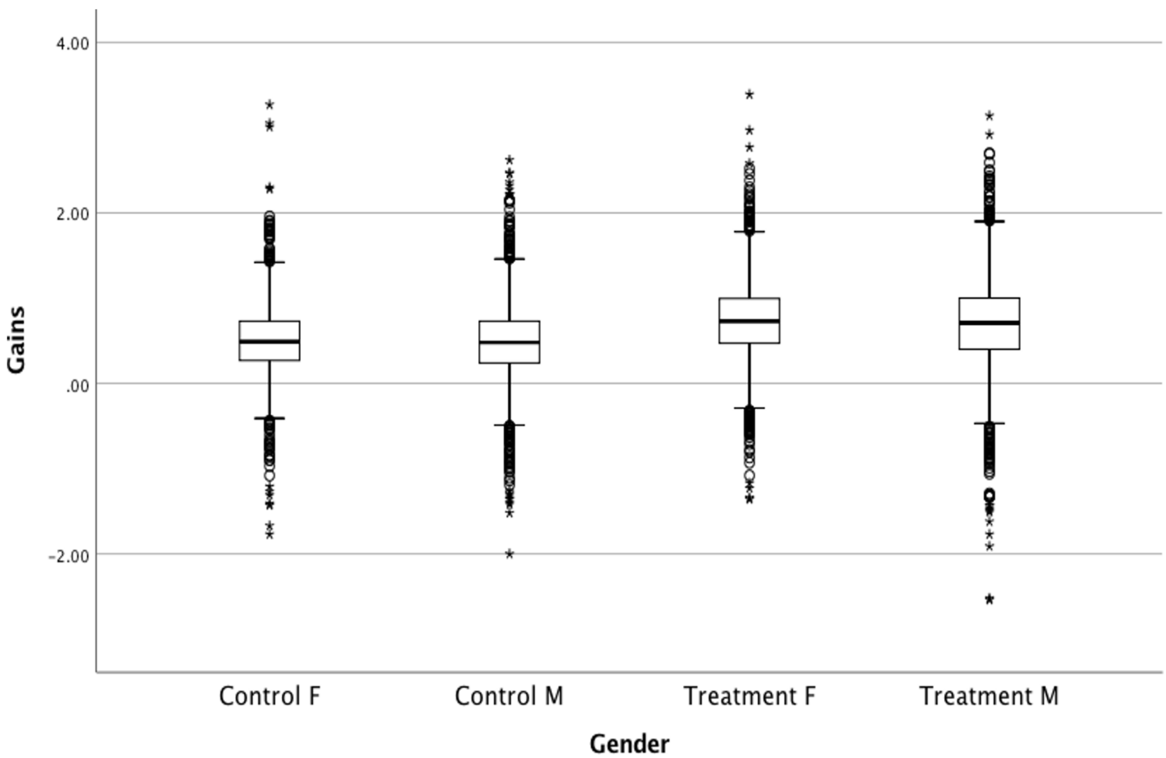


Figure 2. **Gains by gender.**

Hypothesis results: The second null hypothesis stated that there is no statistically significant difference in student mathematics achievement based on gender when adaptive mathematics lessons are implemented. The researcher rejected the second null hypothesis. There was a statistically significant difference between the mean grade score gains for the female control group (M = .49, SD = .39) and female treatment group (M =.73, SD = .43). The difference between the mean grade score gains for females in the control and treatment groups was .24, with the treatment group scoring higher. There was a statistically significant difference between the scores for the male control group (M =.47, SD = .42) and male treatment group (M = .70, SD = .51). The difference between the mean grade score gains for males in the control and treatment groups was .23, with the treatment group scoring higher. There was a statistically significant difference in student mathematics achievement based on gender when adaptive mathematics lessons were implemented. In other words, students in the treatment group scored statistically significantly higher than students in the control group, regardless of gender.

Students with Disabilities classification.

Hypothesis testing: Table 4 represents the descriptive statistics for students classified as students with disabilities (SWD) and students not classified as SWD. The control group consisted of students who were classified as SWD (n = 1046) and students not classified as SWD (n = 6067). The treatment group consisted of students who were classified as SWD (n = 1200) and students not classified as SWD (n = 6532). The mean grade score gain for students classified as SWD in the control group was .39 with a standard deviation of .48. The mean grade score gain for students classified as SWD in the treatment group was .55 with a standard deviation of .55. The mean grade score gain for students not classified as SWD in the control group was .50 with a standard deviation of .39. The mean grade score gain for students not classified as SWD in the treatment group was .75 with a standard deviation of .45.

Table 4: Descriptive Statistics for SWD Classification

Group	n	M	SD
Control Non-SWD	6067	.50	.39
Control SWD	1046	.39	.48
Treatment Non-SWD	6532	.75	.45
Treatment SWD	1200	.51	.55
Total	14845		

Analysis: A Welch ANOVA was administered to determine the differences between the mean grade score gains of the control and treatment groups based on the subgroup of SWD classification. The grade score gains functioned as the dependent variable and SWD classification as the independent variable. Results of the Welch ANOVA showed an overall statistically significant difference in group means at α = .05 level, Welch's F (3, 2935) = 431.36, $p < .001$, $\eta^{\#}$= .081. The magnitude of the difference in the means and the effect size was large ($\eta^{\#}$= .081).

Post-hoc analyses using the Games-Howell test indicated that the mean grade score gain for the SWD treatment group was statistically significantly higher than for the SWD control group, with a mean difference of .12. The non-SWD treatment group scored statistically significantly higher than the non-SWD control group, with a mean difference of .25; the SWD treatment group, with a mean difference of .24 mean; and the SWD control group, with a mean difference of .35. Therefore, the second null hypothesis was rejected. There was a statistically significant difference in student mathematics achievement based on the subgroup of SWD classification when adaptive mathematics lessons were implemented. Figure 3 shows the mean grade score gains by SWD classification in the control and treatment groups.

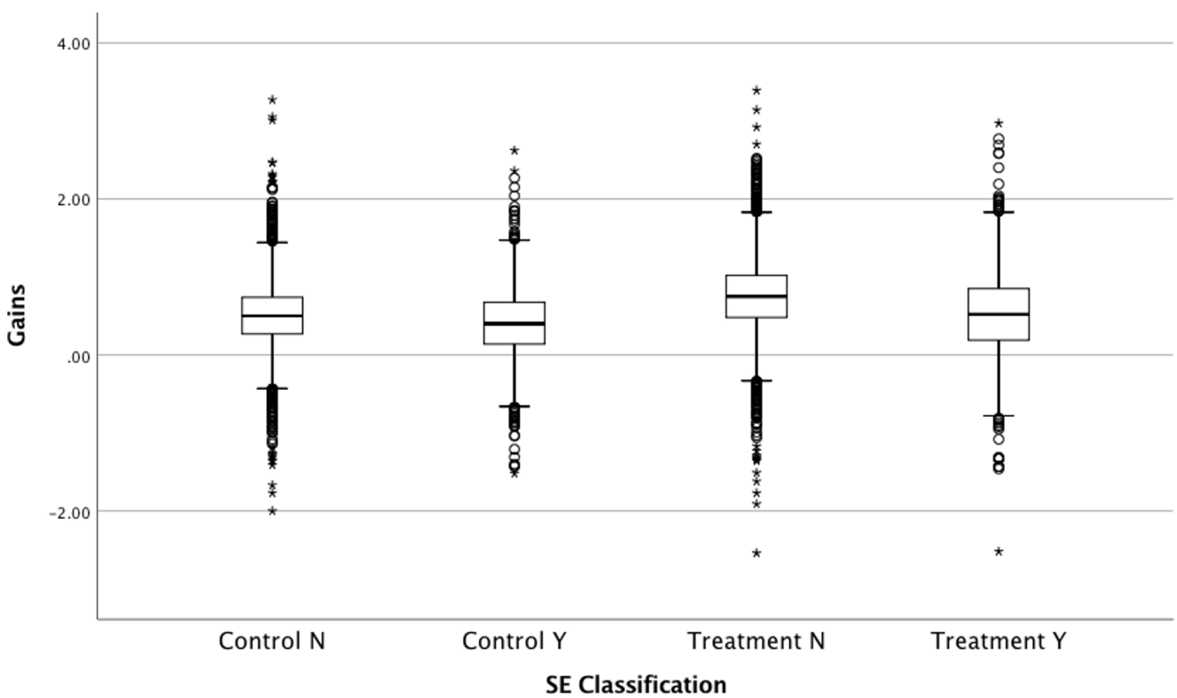


Figure 3. **Gains by SWD classification.**

Hypothesis results: The second null hypothesis stated that there is no statistically significant difference in student mathematics achievement based on the subgroup of SWD classification when adaptive mathematics lessons are implemented. The researcher rejected the second null hypothesis. There was a statistically significant difference between the mean grade score gains of the SWD control group (M = .39, SD = .48) and SWD treatment group (M = .51, SD = .55). The difference between the mean grade score gains for SWD control and treatment groups was .12, with the treatment group scoring statistically significantly higher. There was a statistically significant difference between the mean grade score gains of the non-SWD control group (M = .50, SD = .39) and the non-SWD treatment group (M = .75, SD = .45). The difference between the mean grade score gains for non-SWD in the control and treatment groups was .25, with the treatment group scoring statistically significantly higher. The non-SWD treatment group (M = .75, SD = .45) scored statistically significantly higher than the SWD treatment group (M =.51, SD = .55), with a difference of .24. The non-SWD treatment group (M = .75, SD =.45) scored statistically significantly higher than the SWD control group (M = .39, SD =.48) with a difference of .35. There was a statistically significant difference in student mathematics achievement based on the subgroup of SWD classification when adaptive mathematics lessons were implemented.

ELL classification.

Hypothesis testing: Table 7 represents the descriptive statistics for students classified as ELL and students not classified as ELL. The control group consisted of students who were classified as ELL (n = 511) and students not classified as ELL (n = 6602). The treatment group consisted of students who were classified as ELL (n = 612) and students not classified as ELL (n = 7120). The mean grade score gain for students classified as ELL in the control group was .67 with a standard deviation of .49. The mean grade score gain for students classified as ELL in the treatment group was .79 with a standard deviation of .57. The mean grade score gain for students not classified as ELL in the control group was .47 with a standard deviation of .39. The mean grade score gain for students not classified as ELL in the treatment group was .70 with a standard deviation of .46.

Table 5: ***Descriptive Statistics for ELL Classification***

Group	n	M	SD
Control non-ELL	6603	.47	.39
Control ELL	511	.67	.49

Treatment non-ELL	7120	.70	.46
Treatment ELL	613	.79	.57
Total	14847		

Analysis: A Welch ANOVA was administered to determine the differences between the mean grade score gains of the control and treatment groups based on ELL classification. The grade score gains functioned as the dependent variable and ELL classification as the independent variable. Results of the Welch ANOVA showed an overall statistically significant difference in group means at α = .05 level, Welch's F (3, 1428) = 366.34, p < .001, $\eta^{\#}$ = .068. The magnitude of the differences in the means and the effect size was medium ($\eta^{\#}$ = .068). A Games-Howell test was conducted and indicated that the ELL treatment group scored statistically significantly higher than the ELL control group, with a mean difference of .12; the non-ELL students in the treatment group, with a mean difference of .08; and the non-ELL students in the control group, with a mean difference of .32. The non-ELL treatment group scored statistically significantly higher than the non-ELL control group, with a mean difference of .23. Therefore, the second null hypothesis was rejected. There was a statistically significant difference in student mathematics achievement based on the subgroup of ELL classification when adaptive mathematics lessons were implemented. Figure 4 shows the mean grade score gains by English language learner classification in the control and treatment groups.

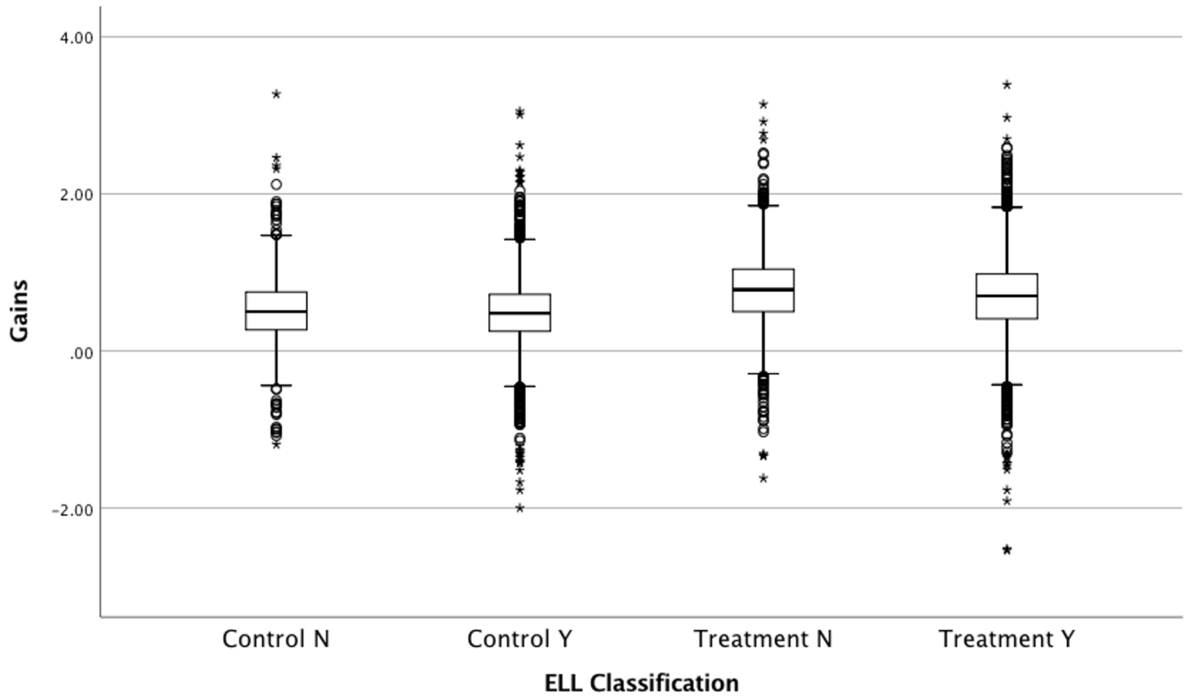


Figure 4. **Gains by English language learner classification.**

Hypothesis results: The second null hypothesis stated that there is no statistically significant difference in student mathematics achievement based on ELL classification when adaptive mathematics lessons are implemented. The researcher rejected the second null hypothesis. There was a statistically significant difference between the mean grade score gains for the ELL control group (M = .67, SD = .49) and the ELL treatment group (M = .79, SD = .57). The mean difference between the ELL control and ELL treatment groups was .12, with the treatment group scoring higher. There was a statistically significant difference between the mean grade score gains for the non-ELL control group (M = .47, SD = .39) and the non-ELL treatment group (M = .70, SD = .46). The mean difference between the non-ELL control and treatment groups was .23, with the treatment group scoring statistically significantly higher. The ELL treatment group (M = .79, SD = .57) showed statistically significant higher grade score gains than the non-ELL treatment group (M = .70, SD = .46) with a mean difference of .08. The ELL treatment group (M = .79, SD = .57) showed statistically significant higher grade score gains than the non-ELL control group (M = .47, SD = .39) with a mean difference of .32. There was a statistically significant difference in student mathematics achievement based on the subgroup of ELL classification when adaptive mathematics lessons were implemented.

Socioeconomic status.

Hypothesis testing: Table 8 represents the descriptive statistics for students classified as economically disadvantaged and students not classified as economically disadvantaged. The control group consisted of students who were classified as economically disadvantaged (n = 5345) and students not classified as economically disadvantaged (n = 1768). The treatment group consisted of students who were classified as economically disadvantaged (n = 5809) and students not classified as economically disadvantaged (n = 1923). The mean grade score gain for students classified as economically disadvantaged in the control group was .48 with a standard deviation of .41. The mean grade score gain for students classified as economically disadvantaged in the treatment group was .69 with a standard deviation of .47. The mean grade score gain for students not classified as economically disadvantaged in the control group was .51 with a standard deviation of .40. The mean grade score gain for students not classified as economically disadvantaged in the treatment group was .77 with a standard deviation of .47.

Table 6: Descriptive Statistics for Economically Disadvantaged

Group	n	M	SD
Control Non-economically Disadvantaged	1768	.51	.40
Control Economically Disadvantaged	5345	.48	.41
Treatment Non-economically Disadvantaged	1923	.77	.47
Treatment Economically Disadvantaged	5809	.69	.47
Total	14845		

Analysis: A Welch ANOVA was administered to determine the differences between the mean grade score gains of the control and treatment groups based on socioeconomic status. The grade score gains functioned as the dependent variable and socioeconomic status as the independent variable. Results of the Welch ANOVA showed an overall statistically significant difference in group means at $\alpha = .05$ level, Welch's $F(3, 14841) = 335.122$, $p < .001$, $\eta^{\#} = .063$. The magnitude of the difference in the means and the effect size was medium ($\eta^{\#} = .063$).

Post-hoc analyses using the Games-Howell test indicated that the mean grade score gain for the economically disadvantaged treatment group was statistically significantly higher than for the economically disadvantaged control group, with a mean difference of .21, and for the non-economically disadvantaged control group, with a mean difference of .18.

The non-economically disadvantaged treatment group scored statistically significantly higher than the non-economically disadvantaged control group, with a mean difference of .25; the economically disadvantaged treatment group, with a mean difference of .07; and the economically disadvantaged control group, with a mean difference of .29. Therefore, the second null hypothesis was rejected. There was a statistically significant difference in student mathematics achievement based on socioeconomic status when adaptive mathematics lessons were implemented. Figure 5 shows the mean grade score gains by socioeconomic status in the control and treatment groups.

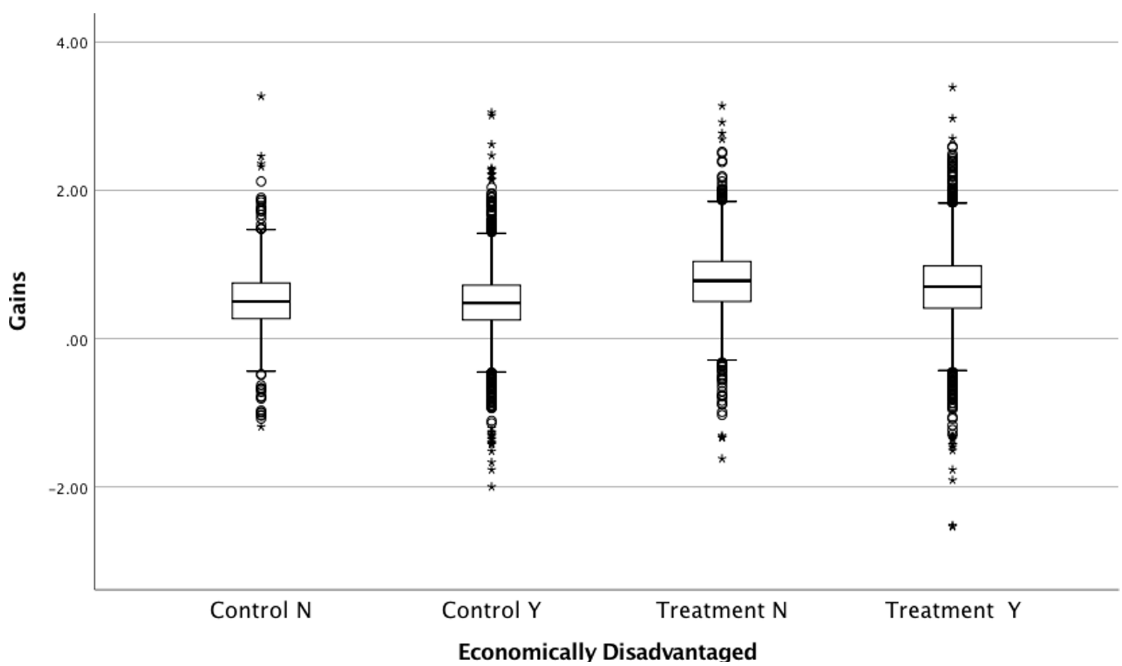


Figure 5. **Gains by socioeconomic status.**

Hypothesis results: The second null hypothesis stated that there is no statistically significant difference in student mathematics achievement based on socioeconomic status when adaptive mathematics lessons are implemented. The researcher rejected the second null hypothesis. There was a statistically significant difference between the mean grade score gains for the economically disadvantaged control group (M = .48, SD = .41) and the economically disadvantaged treatment group (M = .69, SD = .47). The mean difference was .21, with the treatment group scoring higher. There was a statistically significant difference between the mean grade score gains for the economically disadvantaged treatment group (M = .69, SD = .47) and the non-economically disadvantaged control group (M = .51, SD = .40). The mean difference was .18, with the treatment group scoring higher. There was a statistically significant difference between the mean grade score gains for the non-economically disadvantaged control group (M = .51, SD = .40) and the non-economically disadvantaged treatment group (M = .77, SD = .47). The mean difference was .25, with the treatment group scoring higher. The non-economically disadvantaged treatment group (M = .77, SD = .47) scored statistically significantly higher than the economically disadvantaged treatment group (M = .69, SD = .47) with a mean difference of .07. The non-economically disadvantaged treatment group (M = .77, SD = .47) scored statistically significantly higher than the economically disadvantaged control group (M = .48, SD = .41), with a mean difference of .29. There was a statistically significant difference in student mathematics achievement based on the subgroup of socioeconomic status when adaptive mathematics lessons were implemented.

Race/ethnicity.

Hypothesis testing: Table 9 represents the descriptive statistics for various races and ethnicities. The control group consisted of Asian (n = 1085), Black (n = 2192), Hispanic (n = 2934), and White (n = 791) students. The mean grade score gain for Asian students in the control group was .61 with a standard deviation of .38; the mean grade score gain for Asian students in the treatment group was .91 with a standard deviation of .43. The mean grade score gain for Black students in the control group was .41 with a standard deviation of .42; the mean grade score gain for Black students in the treatment group was .59 with a standard deviation of .47. The mean grade score gain for Hispanic students in the control group was .48 with a standard deviation of .39; the mean grade score gain for Hispanic students in the treatment group was .69 with a standard deviation of .46. The mean grade score gain for White students in the control group was .54 with a standard deviation of .43; the mean grade score gain for White students in the treatment group was .82 with a standard deviation of .47.

Table 7: Descriptive Statistics for Race/Ethnicity

Group	n	M	SD
Control Asian	1085	.60	.38
Control Black	2192	.40	.42
Control Hispanic	2934	.48	.39
Control White	791	.54	.43
Treatment Asian	1254	.90	.43
Treatment Black	2307	.59	.47
Treatment Hispanic	3088	.69	.46
Treatment White	931	.82	.47
Total	14582		

Analysis: A Welch ANOVA was administered to determine the differences between the mean grade score gains of the control and treatment groups based on race/ethnicity. The races/ethnicities tested were Asian, Black, Hispanic, and White. The grade score gains functioned as the dependent variable and race/ethnicity as the independent variable. Results of the Welch ANOVA showed an overall statistically significant difference in group means, Welch's F (7, 4578) = 236.44 = $p < .001$, $\eta^{\#} = .10$. The magnitude of the difference in the means and the effect size was large ($\eta^{\#} = .10$).

Post-hoc analyses using the Games-Howell test indicated that each race/ethnicity treatment group scored significantly statistically higher than their control-group counterpart. Post-hoc tests indicated that statistically significant differences in mean grade score gains were obtained between Asian and Black students, between Asian and Hispanic students, and between Asian and White students. Significant differences in mean grade score gains were obtained between

White and Black students and White and Hispanic students. Significant differences in mean grade score gains were obtained between Hispanic and Black students. Asian students had the highest improvement scores, followed by Caucasian, Hispanic, and Black students. There was a statistically significant difference in student mathematics achievement based on the subgroup of race/ethnicity when adaptive mathematics lessons were implemented. Figure 6 shows the mean grade score gains by race/ethnicity in the control and treatment groups.

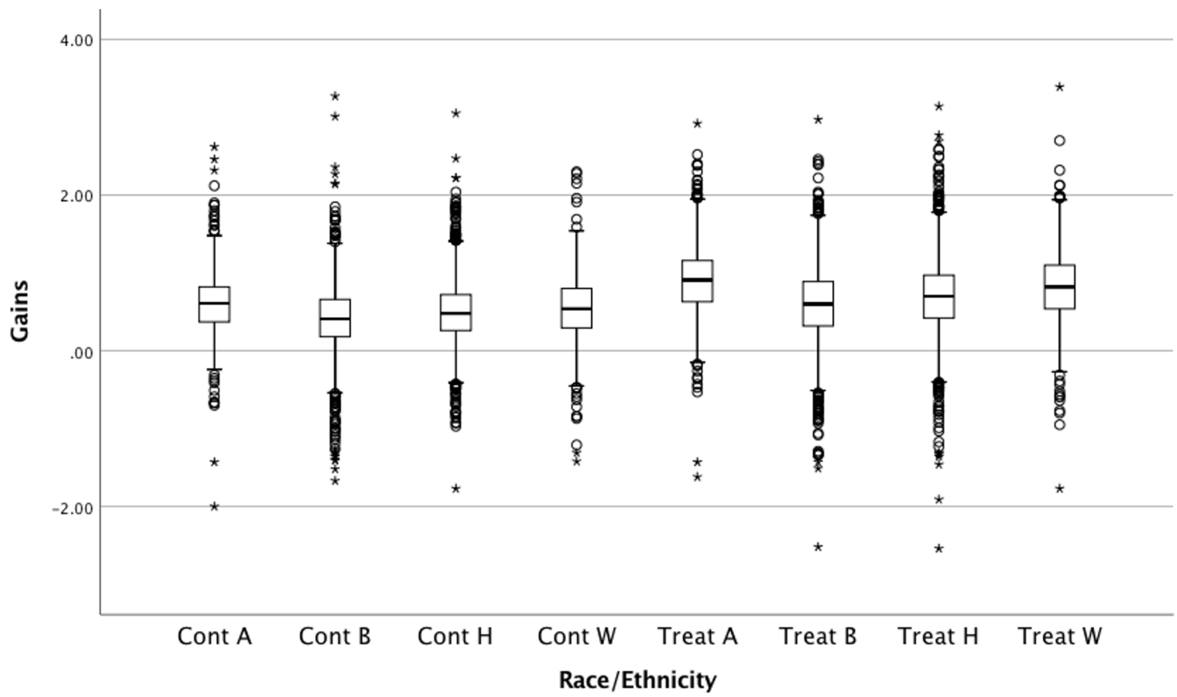


Figure 6. **Gains by race/ethnicity.**

Hypothesis results: The second null hypothesis stated that there is no statistically significant difference in student mathematics achievement based on race/ethnicity when adaptive mathematics lessons are implemented. The researcher rejected the second null hypothesis. Each race/ethnicity treatment group scored statistically significantly higher than their control-group counterpart. The Asian treatment group (M = .91, SD = .43) scored statistically significantly higher than the Asian control group (M = .61, SD = .38) with a mean difference of .3. The Black treatment group (M = .59, SD = .47) scored statistically significantly higher than the Black control group (M = .41, SD = .42240) with a mean difference of .18. The Hispanic treatment group (M = .69, SD = .46) scored statistically significantly higher than the Hispanic control group (M = .69, SD = .46) with a mean difference of .20. The White treatment group (M = .82, SD = .47) scored statistically significantly higher than the White control group (M = .82, SD = .47) with a mean difference of .28. There was a statistically significant difference in student mathematics achievement based on the subgroup of race/ethnicity when adaptive mathematics lessons were implemented.

Research Question 3

Research Question 3 asked if a significant relationship existed between time spent in *LGL Math Edge* and mathematics achievement from pretest to posttest. A simple linear regression was used to investigate the belief that the initial correlation and regression will determine if there are any predictive qualities between the variables. A simple linear regression was conducted to analyze the third null hypothesis, H_o3: There is no statistically significant relationship between time spent in *LGL Math Edge* and mathematics achievement from pretest to posttest. The significance of the continuous variable of math time in relation to other variables was tested using a random effect model.

Hypothesis testing: Table 10 represents the descriptive statistics for time spent in *LGL Math Edge* and mean grade score gains. There were 7,733 students who received the treatment of *LGL Math Edge*. The mean gain for the treatment group was 0.61 with a standard deviation of .45. The average amount of time each student spent in *LGL Math Edge* was 8.34 hours with a standard deviation of 11.33.

Table 8: ***Descriptive Statistics for Time Spent in LGL Math Edge and Grade Score Gains***

Category	n	M	SD
Grade Score Gains	7733	.61	.45
Math Time	7733	8.34	11.33

Analysis: The researcher performed a series of analyses to check assumptions for the regression. The data indicated that none of the major assumptions for simple linear regression had been violated in the analysis. Specifically, the assumptions of normality, multicollinearity, singularity, linearity, and homoscedasticity were all within the range of acceptable outcomes for this analysis. A simple linear regression was calculated to predict student achievement gains in mathematics based on the time students spent in *LGL Math Edge*. The results indicated a moderately strong, significant, and positive relationship between the variables of time spent in *LGL Math Edge* and gains in mathematics achievement from pretest to posttest: $r = .31$, $p <$

.001. The data indicated a significant model within the results: F (1, 14651) = 1597.30, $p < .001$. The model explained nearly 10%, adjusted R^2, of the variance in mathematics achievement from pretest to posttest.

A significant regression equation was found: (F (1, 14651) = 1597.30, $p < .001$) with an R^2 of .09. Participants' gains were equal to .50 + .01 (time on *LGL Edge)* points when time on *LGL Edge* is measured in hours. Participants' gains increased .01 for each hour of time spent on *LGL Edge*. In addition to a significant predictive model, time spent on *LGL Math Edge* significantly contributed to changes in mathematics achievement from pretest to posttest: $\beta = .31$, $p < .001$. The $\beta = .31$ was quite large in terms of a single contribution of a single variable. These results suggest that time spent on *LGL Math Edge* does indeed make a difference in mathematics achievement from pretest to posttest. The more time students spend in the program, the more likely they are to increase their scores from the pretest to the posttest. There is a statistically significant relationship between time spent on *LGL Math Edge* and mathematics achievement from pretest to posttest.

The researcher tested the significance of the continuous variable of math time in relation to other variables. The researcher performed a series of analyses to check assumptions for the regression. The Breusch-Pagan test (adjusted $R^2 = .16$, F (1, 12) = 100.28 $p < .001$) found heteroscedasticity when trying to apply a linear model; therefore, a random effect model was used. The data indicated that none of the major assumptions for simple linear regression had been violated in the analysis. Specifically, the assumptions of normality, multicollinearity, singularity, and linearity were all within the range of acceptable outcomes for this analysis. The researcher constructed a random effect linear model using R to test the variable of math time and its significance in relation to other variables. The formula entered into R was: Grade score gain = gender + SWD classification + ELL classification + socioeconomic status + ethnicity + math time + (1 | SiteId) + (1 | GradeLevel).

The fixed effects of gender, SWD classification, ELL classification, socioeconomic status, race/ethnicity, and math time were entered into the model (see Table 11). The random effects of site ID and grade level were entered into the model (see Table 12). The t values of the fixed effect variables showed that there were only two significant variables to predict the grade score: SWD classification (est = - .14, $t = 10.15$) and math time spent in adaptive learning conditions (est = .01, $t = 19.24$) when $t > 2.6$. The formula to construct a predictive model of math time by subgroup was: Grade score gain = - 0.146*SWD (Yes) + 0.01 * Adaptive Math Hours.

Visual inspection of the residual plot for the math time model (see Figure 7) showed almost no bias in the model, which inspires confidence that it was correct in estimating increased grade scores by the math time variable. The residual plot also showed that variance was well established and consistent except for the extreme ends.

Table 9: ***Random Effects for Time Spent in LGL Math Edge***

Groups	Name	Variance	SD
Site ID	(Intercept)	.01	.11
Grade	(Intercept)	.00	.00
Residuals		.19	.44

Table 10: ***Fixed Effects***

Group	Estimate	Std. Error	t- value
Intercept	1.16	.44	2.60
Gender	-.01	.01	- .75
ELL	- .01	.01	- .77
Disadvantaged	.06	.06	.97
SWD	- .14	.0	- 10.15
Asian	- .47	.44	1.05
Black	- .63	.44	- 1.43
Hispanic	- .65	.44	- 1. 16
White	- .52	.44	- 1. 17
Math Time	.01	.000	19.24

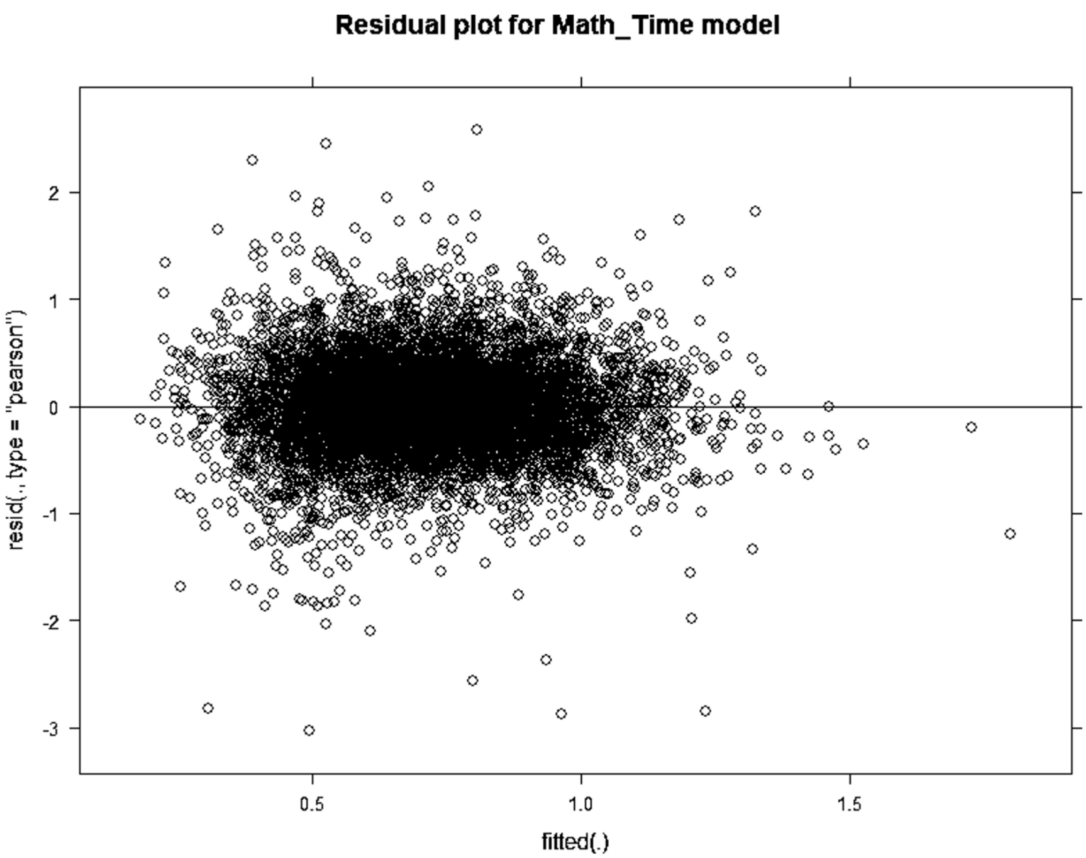


Figure 7. **Residual plot for math time.**

Results: The third null hypothesis stated that there would be no statistically significant relationship between time spent in *LGL Math Edge* and mathematics achievement from pretest to posttest. The researcher rejected the third null hypothesis. The results of a simple linear regression indicated a moderately strong, significant, and positive relationship between the variables of time spent in *LGL Math Edge* and gains in mathematics achievement from pretest to posttest: $r = .31$, $p < .001$. The data indicated a significant model within the results: $F (1, 14651) = 1597.304$, $p < .001$. A random effect linear model was constructed to test the variable of math time and its significance in relation to other variables. The t values of the fixed effect variables showed that there were only two significant variables that predicted the grade score: SWD classification (est = - .14, t = 10.15) and math time spent in adaptive learning conditions (est = .01, t = 19.24) when $t > 2.6$. Except for being classified as SWD, the amount of time spent in adaptive learning was the only key indicator of improved success. The formula to construct a predictive model of math time by subgroup was: Grade score gain = -0.146*SWD (Yes) + 0.01 * Adaptive Math Hours.

Summary of Findings

Research Question 1 focused on the impact of adaptive mathematics lessons on elementary students' mathematics achievement. A Welch two-sample, one-tailed t-test was used to determine if a statistically significant difference existed between the mean grade score gains of students in the control and treatment groups. Students in the treatment group scored statistically significantly higher than students in the control group according to the p-value of < .001. As a result, the first null hypothesis, Ho1–There is no statistically significant difference in elementary students' mathematics achievement when adaptive mathematics lessons are implemented–was rejected.

Research Question 2 focused on the difference in student mathematics achievement between the control and treatment groups based on the subgroups of gender, race/ethnicity, SWD classification, ELL classification, and socioeconomic status. A Welch ANOVA was administered to determine the differences between the mean grade score gains of the control and treatment groups based on the subgroups of gender, race/ethnicity, SWD classification, ELL classification, and socioeconomic status. Students in the treatment group of each subgroup scored statistically significantly higher than students in the control group of each subgroup in all subgroups according to the p-value of < .001. As a result, the second null hypothesis, H_o2–There is no statistically significant difference in student mathematics achievement based on the subgroups of gender, race/ethnicity, SWD classification, ELL classification, and socioeconomic status when adaptive mathematics lessons are implemented–was rejected.

Research Question 3 focused on determining if a significant relationship existed between time spent in *LGL Math Edge* and mathematics achievement from pretest to posttest. During the 30-week time span between pretest and posttest, the treatment group was to receive .5 hours per week of adaptive supplemental mathematics lessons. Each student should have received a minimum of 15 hours of adaptive supplemental mathematics lessons between the pretest and posttest. The results of a simple linear regression indicated a moderately strong, significant, and positive relationship between the variables of time spent in *LGL Math Edge* and gains in mathematics achievement from pretest to posttest: $r = .314$, $p < .001$. As a result, H_o3– There is no statistically significant relationship between time spent in *LGL Math Edge* and mathematics achievement from pretest to posttest–was rejected.

The purpose of this study was to examine the effects of adaptive supplemental lessons on elementary students' mathematics achievement in a large urban school district serving a diverse population. Data for three research questions were analyzed. A Welch two-sample, one-tailed t-test used in research question one determined that a statistically significant difference existed between the mean grade score gains of students in the control and treatment groups. A Welch ANOVA administered in Research Question 2 determined that the differences between the mean grade score gains of the control and treatment groups based on the subgroups of gender, race/ethnicity, SWD classification, ELL classification, and socioeconomic status were statistically significant. Games-Howell post-hoc analyses determined that statistically significant differences existed between the groups. The results of a simple linear regression in Research Question 3 indicated a moderately strong, significant, and positive relationship between the variables of time spent in *LGL Math Edge* and gains in mathematics achievement from pretest to posttest. A random effect linear model was constructed to test the variable of math time and its significance in relation to other variables, determining that SWD classification and math time spent in *LGL Math Edge* were the only significant variables when predicting grade score gains.

Implications

At the individual level, the results of this study have implications for positive change in student performance in mathematics. Historically, mathematics has been delivered through direct instruction, with the usual practice of waiting until the teacher feels every student understands the concept before moving on. A challenge students face is having too much down time listening to the lecture and not enough application time to practice the skills (Brabeck, Jeffrey, & Fry, 2011; Rohrer, 2009). The results of this study indicate that when students engage with personalized learning technology, they perform better than when they are taught through conventional instruction alone.

This study has significant implications at the organizational level. The results show that time spent in *LGL Math Edge* has a significant effect on student achievement, and aside from SWD classification, this time was the only predictive factor in determining student achievement. This is significant in terms of closing the achievement gap in the subgroup of race/ethnicity. It was evident that different race/ethnic groups were exposed to adaptive learning at different levels. This finding may be helpful to teachers and administrators for many reasons. Teachers may not be aware of the amount of time students spend in adaptive math. Even though teachers may believe all students receive equitable access to adaptive learning, in reality, they might not. This can help schools specifically target the students who make up their achievement gap and perhaps offer all students more time in adaptive learning conditions, in addition to their math block. Teachers could also allow more time for adaptive learning within the math block. The results thus encourage districts to further investigate why a particular group is logging less time in adaptive math. Perhaps the groups with lower usage make up a section of the district whose schools do not have one-to-one access to electronic devices. If this is a cause, administrators can recommend that principals use their budgets to purchase additional technology in order to give all students access and the opportunity to be successful.

This study has several other implications for teachers. Teachers may be motivated by the results to combine adaptive learning with their own best practices to increase student achievement in mathematics. They may be able to spend more time providing interventions or enrichment to students on a more personalized level. Teachers will be able to use this study as a benchmark to measure success when implementing adaptive lessons.

Additional implications relate to the theoretical framework of the study. The results of this study indicate that approaching the problem of reducing cognitive load solely with conventional instruction is not beneficial to solving the problem of declining mathematics scores. Being that a major instructional implication of the expertise reversal effect is the need to tailor instruction to different learner expertise levels, the results of this study validate that *LGL Math Edge* lessons provide an option to assist learners in acquiring advanced knowledge in a domain.

Recommendations for Further Research

The primary question that needs to be investigated in future research is whether the adaptive mathematics lessons will have significant effects on state assessments similar to those on *ADAM*. As this study shows, all subgroups made significant gains on *ADAM* when lessons were delivered through adaptive lessons in conjunction with conventional instruction. However, whether the success will be replicated on state assessments is unknown. Future research should include replicating this study and analyzing the score gains by grade level. The results may indicate whether adapting learning is more effective at a particular grade level. Additionally, this study should be replicated using various adaptive software programs to determine if adaptive learning programs are effective, or if it was specifically *LGL Math Edge* that was effective in increasing student achievement.

Another goal for future research should be to address the effects of adaptive mathematics lessons on ELL achievement when the lessons are delivered in the student's home language. As it is impossible to conduct a study in every home language that is not English, a good place to start would be conducting this study using a Spanish version of the treatment. Results can be compared to this study to determine if adaptive learning had a more significant impact on students when delivered in their first language.

Future research should also be conducted regarding the effect of adaptive mathematics lessons on SWD achievement. In this study, students classified as SE in the treatment group improved nearly to the point of being indistinguishable from students not classified as SWD without adaptive math learning conditions. Replicating this study and strictly enforcing fidelity among students with a SWD classification would give insight into whether the SWD results from this study are as promising as they appear to be.

Finally, a study should be conducted to determine the effect of adaptive lessons on student achievement in mathematics when a strict time requirement is enforced. Although this study determined that there was a significant gain in student achievement when using adaptive lessons, the time spent on *LGL Math Edge* varied greatly. It would be beneficial to see the impact adaptive lessons have on student achievement in mathematics if all participants have a strictly enforced minimum of 30 hours of adaptive lessons.

Case Studies

Introduction

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

Case Study 1: Downey Unified School District, Sussman Middle School

Downey Unified School District is home to over 30,000 students. Nearly 1,000 dedicated educators strive every day to develop each student to be a self-motivated learner and a productive, responsible, and compassionate member of an ever-changing global society. With over twenty individual schools, the district prides itself on fostering meaningful relationships with students, parents, and the community while providing a relevant and rigorous curriculum in facilities that advance teaching and learning.

Sussman Middle School, one of four middle schools within the district, implemented Let's Go Learn's personalized instruction solution, *LGL Math Edge*, in January 2018. After less than half a year, educators witnessed noticeable growth. The school was outperforming all other middle schools in the district, and they made significant gains in both 6th and 8th grade on the CAASPP state assessment, also known as the SBAC assessment.

Students with higher *LGL Math Edge* usage of about 1+ hour a week outperformed students with lower usage by 85%, according to the "Total" math score. Within each of the five major mathematics strands, they had consistently greater gains, ranging from 30% to 86%. Gains in grades 6 and 8 on the CAASPP correlated with a further breakdown of student usage by grade level. Seventh-grade students fell within the lower usage group, which was consistent with the lack of CAASPP gains. Overall, these results are valid and significant, given the large sample sizes of 216, 522, and 329 in each of the three leveled usage groups examined.

Students using *LGL ELA Edge* significantly outperformed students who were not using it. This is clear in the subtests of High-Frequency Words, Word Recognition, and Phonics. These subtests are focused on decoding. In addition, students made greater gains in reading comprehension.

Case Study 2: Montebello ELA

Montebello Unified School District (MUSD) in Southern California is home to nearly 30,000 diverse students. Across over thirty individual schools, educators strive to empower students to achieve academic excellence as model citizens. MUSD's commitment to students is immediately evident in the countless programs and resources dedicated to improving student learning.

Among many efforts to support student learning, Montebello Unified School District ran short 14-day summer school sessions in 2017 and 2018. The initiative was organized and headed by the Federal Programs Department, and it targeted RtI Tier 2-qualified students in grades 2 to 8 (students who were below 25% proficient on the state test) with a focus on either math or reading. The district adopted Let's Go Learn's personalized learning platform, the *LGL Edge* series, to accelerate remediation with data-driven personalized instruction and achieve intensive intervention efforts over two weeks.

Prior to summer school, students completed the *DORA* assessment at the end of the year (spring testing window). These scores served as a pretest for summer sessions, and on days 13 and 14 of summer school, students were given post-assessments. To ensure that students made the most of their time on the platform, teachers were trained in the use of *LGL ELA Edge* and knew how to assist students in using the program successfully.

RESULTS: In both the 2017 and 2018 summer sessions, student gains in reading were significant. In 2017, 279 students using *LGL ELA Edge* for more than 6.4 hours had extremely high gains as measured by the six subtests of *DORA*. These results are statistically significant. Similarly, in 2018, there were measurable gains.

Case Study 3: Montebello Math

Among many efforts to support student learning, Montebello Unified School District ran short 14-day summer school sessions in 2017 and 2018. The initiative was organized and headed by the Federal Programs Department, and it targeted RtI Tier 2-qualified students in grades 2 to 8 (students who were below 25% proficient on the state test) with a focus on either math or reading. The district adopted Let's Go Learn's personalized learning platform, the *LGL Edge* series, to accelerate remediation with data-driven personalized instruction and achieve intensive intervention efforts over two weeks.

Prior to summer school, students completed the *ADAM* assessment at the end of the year (spring testing window). These scores served as a pretest for summer sessions, and on days 13 and 14 of summer school, students were given post-assessments. To ensure that students made the most of their time on the platform, teachers were trained in the use of *LGL Math Edge* and knew how to assist students in using the program successfully.

RESULTS: In both the 2017 and 2018 summer sessions, student growth in math was significant as measured by the *ADAM* assessment, and the data indicate that overall, students who had more time on task with *LGL Math Edge* realized greater gains. Given the sample sizes and the usage of *LGL Math Edge* as the summer intervention content, these results are statistically significant.

Case Study 4: Jersey City Public School

Jersey City Public Schools (JCPS) is made up of 42 unique schools, including fourteen elementary schools (Pre-K through 5) and thirteen grammar schools (Pre-K through 8). The dedicated team at JCPS works to fulfill the district's promise to give every student E.A.R.S: Enrichment, Acceleration, Resources, and Support. Driven by this vision, the district adopted Let's Go Learn diagnostic assessments and personalized instruction to guide their learning initiatives.

In 2016-17, Jersey City Public Schools began a large-scale implementation of *LGL Math Edge* in grades 3 to 8. While this was their primary district initiative, a number of schools implemented *LGL ELA Edge* in addition to *LGL Math Edge*. The greatest number of students enrolled in *LGL ELA Edge* was a group of 103 students in fourth grade. The data from this case study examines the growth of this cohort.

References for Evidence Base

Crowley, K. (2018). The impact of adaptive learning on mathematics achievement. New Jersey City University.

Ginsberg, A., Chang, H., & Jordan, P. (2014). *Absences add up: How school attendance influences student success*. Retrieved from 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

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.

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. Retrieved from 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.