



Efficacy of an Adaptive Game-Based Math Learning App to Support Personalized Learning and Improve Early Elementary School Students' Learning

Hee Jin Bang¹ · Linlin Li² · Kylie Flynn²

Accepted: 25 February 2022 / Published online: 1 April 2022
© The Author(s) 2022

Abstract

This study investigates whether using My Math Academy, which provides personalized content and adaptive embedded assessments to support existing curricula, can improve learning outcomes and engagement for kindergarten and first grade students ($N=505$ treatment, 481 control). Findings indicate that students using My Math Academy made significant learning gains in math relative to children who did not. More skills mastered in My Math Academy was associated with greater learning gains on the external assessment, with the greatest impacts among students with lower levels of math knowledge, where there was more room for growth and on the most difficult skills. Teachers surveyed found My Math Academy easy to use in their classrooms and recognized it as a valuable learning resource that supplemented their existing curricula to improve students' engagement, motivation, and confidence in learning math.

Keywords Math learning · Game-based learning · Personalized learning · Adaptive instructional systems · Early childhood education

Introduction

Developmental and cognitive theories highlight the crucial role of early mathematics skills and knowledge in predicting later academic success and preparedness for twenty-first century STEM careers (Chu et al., 2016; Geary et al., 2013; Watts et al., 2014, 2018). The latest results from the National Assessment of Education Progress (NAEP), however, show that only 40% of fourth graders are proficient in math (NCES, 2019). The situation is worse for children from low socioeconomic (SES) backgrounds who begin school with significantly less math knowledge (Nores & Barnett,

2014; Starkey et al., 2004) and demonstrate gaps in the development of early math skills relative to their middle-class peers (Denton & West, 2002; Hecht et al., 2000; Nores & Barnett, 2014; Starkey et al., 2004). Many low-SES children are almost a full year behind their middle-class peers in math knowledge by the time they enter school, and this gap often persists and increases over time (Entwisle & Alexander, 1989, 1990; Entwisle et al., 2005; Morgan et al., 2009; Rathbun & West, 2004; Reardon, 2013). Moreover, with varying levels of math knowledge in the classroom, teachers face challenges of appropriately personalizing and individualizing learning for each student in their class (Dixon et al., 2014; Goddard et al., 2015).

Given the importance of early math for students' long-term success and the challenges of providing instruction appropriate for children with different skill levels, this study focuses on an adaptive digital learning resource, My Math Academy, designed for children in prekindergarten through 2nd grade. A primary goal is to investigate whether the digital learning environment, which provides personalized content and adaptive embedded assessments, can improve learning outcomes for kindergarten and first grade students when used in authentic classroom settings to complement existing math curricula and instruction. A secondary goal

✉ Hee Jin Bang
heejin.bang@aofl.com

Linlin Li
lli@wested.org

Kylie Flynn
kflynn2@wested.org

¹ Age of Learning, Inc., 101 N. Brand Blvd., 8th floor, Glendale, CA 91203, USA

² WestEd, 2470 Mariner Square Loop, Alameda, CA 94501, USA

is to determine the extent to which teachers value My Math Academy as a tool that supports their instruction and their students' learning and engagement. This study also builds on a prior study (Thai et al., 2021) which investigated the effectiveness of an earlier version of the My Math Academy with students in pre-K and kindergarten.

Literature

Technology and Young Learners

In recognition of the importance of building an early foundation to support lifelong learning and success, schools and districts seek tools and approaches that engage active learners and enhance personalized learning (Gross & DeArmond, 2018). While there are concerns about the appropriateness of technology use with young children, professionals have also recognized that technology and media can support children's learning and relationships especially when used intentionally and in developmentally appropriate ways (NAEYC, 2012, 2022; Paciga & Donohue, 2017). Moreover, the technology revolution has made it essential for all children to understand STEM, and high-quality STEM experiences early in life can support children's development across literacy and even executive functioning (McClure et al., 2017). A number of proven instructional STEM programs are available, but teachers face challenges in providing consistent personalized and individualized learning for each student in their class (Dixon et al., 2014; Goddard et al., 2015). Technology-based programs that are empirically validated and based on research about how children learn specific content can reduce these challenges for teachers and support students' personalized learning (Clements & Wright, 2022).

As technology makes it possible to learn anywhere, at any time, educational technology programs increasingly take into account the number and diversity of learners, the multiple choices, paths, and ways in which they learn (Levine, 2018), and the contexts in which they learn (Friedman, Masterson, & Wright, 2022). It must also account for the variability of time and space for learning opportunities, the resulting data about learning engagement and performance, the mass personalization of learning (Kucirkova, 2017), and the ways in which pedagogy must adapt to these needs (Pane et al., 2015). Furthermore, as many communities implement a combination of in-person and remote instruction during the current pandemic, teachers need tools that offer high-quality, adaptive, developmentally appropriate, engaging, personalized learning experiences, as well as data on student progress and instructional supports to effectively conduct instruction (NAEYC, 2012, 2022). Early childhood educators too often lack access to high-quality resources, especially in STEM, and guidance about what characterizes

high-quality resources and support for using these resources effectively is essential to providing instruction that will foster children's interest and self-efficacy in these subjects (Early Childhood STEM Working Group, 2017).

Game-Based Learning

Digital educational games leverage play, which is characterized as fun, freely chosen, governed by internal rules, and having a purpose of its own (Eberle, 2014). Play is important in learning and cognitive development (Dietze & Kashin, 2011; Fisher et al., 2012; Piaget, 1962), as it fosters social engagement and collaboration with others that can transform children's thinking (Gestwicki, 2017). Since play has such a significant influence on children's development and thought processes, a key role of play in learning is the constantly evolving (Slutsky & DeShetler, 2017) zone of proximal development (ZPD, Vygotsky, 1978), the "sweet spot" where the learners are ready to learn, with tasks that are not too easy nor too hard, in which they can succeed with some struggle (Plass et al., 2015). In digital environments, children learn from exploring open-ended play contexts that offer them the freedom to adapt and modify their experiences and opportunities to make sense of those changes in dynamic ways (Hatzigianni, 2018).

From an ecological perspective, proximal processes are the "engines of development" (Krebs, 2009). Effective proximal processes occur when children engage in activities with other people, objects, or symbols in their immediate external environment on a regular basis over extended periods of time (Bronfenbrenner, 1995; Tudge et al., 2017). In recent years, digital technology and interactive media have become increasingly popular tools that shape young children's lives in homes, schools, and communities (Brody, 2015; Hopkins et al., 2013). Digital educational games with sufficient content for children to interact with regularly for months will have the potential for effective learning and development. Therefore, it is important to understand the impact that various technologies may have on children's learning and development (Arnott, 2016). As young children increasingly engage with a wide range of digital technologies, research has shown how children integrate technology into play (e.g., interacting with voice-activated devices, asking it to count to 10 while playing hide-and-seek) and how digital play may be educational if designed with specific learning outcomes in mind (Scott, 2021). When technologies are used as tools to support learning, children's ability to interact with devices; adult support available to facilitate interactions with devices (Flynn & Richert, 2015); teachers' beliefs about using technologies with young children (Edwards, 2016); and the social interactions children engage in while using

the technology (Moore & Adair, 2015) also shape children's learning and development.

Effective learning games foster cognitive engagement and motivation by providing interactivity, adaptive challenges, and ongoing feedback (Gee, 2003; Rupp et al., 2010; Shute, 2008). They provide meaningful, socially interactive learning experiences that are guided by specific goals (Hirsh-Pasek et al., 2015). They also offer safe environments for failure, which encourage learners to take risks, explore, and try new things (Hoffman & Nadelson, 2010). Such discovery-based approach to learning provides opportunities for self-regulated learning, in which learners set goals, monitor their progress and achievement, and assess the effectiveness of strategies used to achieve specific goals (Plass et al., 2015). Consequently, digital games with high-quality, engaging content made available at scale can effectively help diverse learners master the targeted skills while supporting teachers in delivering instruction appropriate for each learner's ZPD. Furthermore, digital games, with their capacity to assess and generate data on student learning, can provide information for teachers to plan their instruction and understand how their students are learning.

Research syntheses indicate that carefully designed digital instructional solutions have positive benefits on learning (e.g., Higgins et al., 2012), and studies focusing on use of technology to develop early math skills have generally reported positive experiences and gains in achievement (e.g., Hubber et al., 2016; Kosko & Ferdig, 2016; Outhwaite et al., 2017). A recent student-level randomized control trial of interactive math apps designed for children ages 4 and 5 in the United Kingdom further affirmed the positive impact of math learning apps (Outhwaite et al., 2019). Children who used math apps in addition to standard math practice demonstrated the greatest math learning gains, outperforming their peers in the control group who engaged in standard math practice only (effect size = 0.31, ~3–4 months). Those who used the math apps only without the standard math practice also outperformed their peers who received the standard math practice only (effect size = 0.21, ~2 months). However, many studies report minimal or mixed effects for educational technology interventions in mathematics (Campuzano et al., 2009; Pane et al., 2014; Rutherford et al., 2014; Schenke et al., 2014; Slavin et al., 2017; Steenbergen-Hu & Cooper, 2013). Given the mixed findings, studies of potentially effective technologies targeting core mathematics outcomes are essential to help practitioners and policymakers select learning tools.

My Math Academy Intervention

As a highly engaging educational technology innovation targeting the early elementary grades, My Math Academy was designed to close the achievement gap and prepare students

for success in mathematics. As a supplemental curriculum, My Math Academy aims to improve student learning outcomes aligned to the Common Core State Standards in Mathematics (CCSS-M) in kindergarten, first, and second grade mathematics. It consists of game-based activities with adaptive learning trajectories, performance dashboards that help teachers support students' learning, and offline activities that parents can use to extend in-game learning experiences. With the adaptive learning activities, students engage with personalized learning experiences that enable them to master math skills and concepts. With the teacher dashboards, educators can monitor individual and group progress and assign learning activities to be completed in class or at home for additional reinforcement of specific concepts. My Math Academy, based in the learning sciences, leverages current approaches in [game-based learning](#), and uses evidence-centered design (Mislevy et al., 2003) to enable learners to master math concepts through playful experiences.

Theoretical Foundations Underlying My Math Academy

My Math Academy is a game-based, adaptive Personalized Mastery Learning System™ designed to help elementary age children build a strong understanding of fundamental number sense and operations (Dohring et al., 2019). Topics range from counting to 10 to adding and subtracting three-digit numbers using the standard algorithm. My Math Academy features more than 130 game-based activities that address 96 concepts and skills for pre-kindergarten through second grade.

Learning Trajectories

My Math Academy's highly detailed scope and sequence were informed by current research on hypothetical learning trajectories (Simon, 1995); the Learning Trajectories approach for early mathematics (Clements & Sarama, 2004, 2014; Sarama & Clements, 2004); literature on math interventions (e.g., the internationally recognized Math Recovery program); as well as state and national standards frameworks such as the CCSS-M and the National Council of Teachers of Mathematics' Standards and Principles for School Mathematics. Learning trajectories are defined as the learner's pathway through a hierarchy of goals and activities where each successive objective and interaction is designed to build on the understanding and mastery of previous objectives (Clements & Sarama, 2004; Sarama & Clements, 2013). Since all learners are different, multiple pathways are possible, and instruction is best when it is individualized. The Learning Trajectories approach depends on the learner's success with prior learning and uses that as a foundation for subsequent learning that is tailored to the individual child's needs (Clements & Sarama, 2004; Sarama & Clements,

2013). My Math Academy embodies the Learning Trajectories approach as it uses a sequence of learning objectives and adaptive algorithms to determine what the child knows, does not know, and is ready to learn next. The system places the child in game-based activities appropriate for the child’s anticipated zone of proximal development, adjusting to the child’s needs based on ongoing interactions and gameplay (Vygotsky, 1978). Consequently, each child experiences a unique learning trajectory, based on their prior knowledge, experience, learning ability, and agency within the game.

The scope and sequence capture the entirety of the concepts, principles, and skills involved in early number sense and operations, kindergarten through second grade, especially the most challenging areas of early math, and the hidden concepts and procedures / skills that drive children’s misunderstandings. A team of curriculum specialists unpacked each carefully curated standard into learning objectives that articulate fundamental concepts and procedures/skills that underpin the standard. They then collaborated with learning scientists and game designers to create a knowledge map, or a blueprint of fine-grained, measurable learning objectives, and pathways toward the development of early number sense that articulates the precursor and successor relationships between each learning objective.

Mastery-Based Learning and Evidence-Centered Design

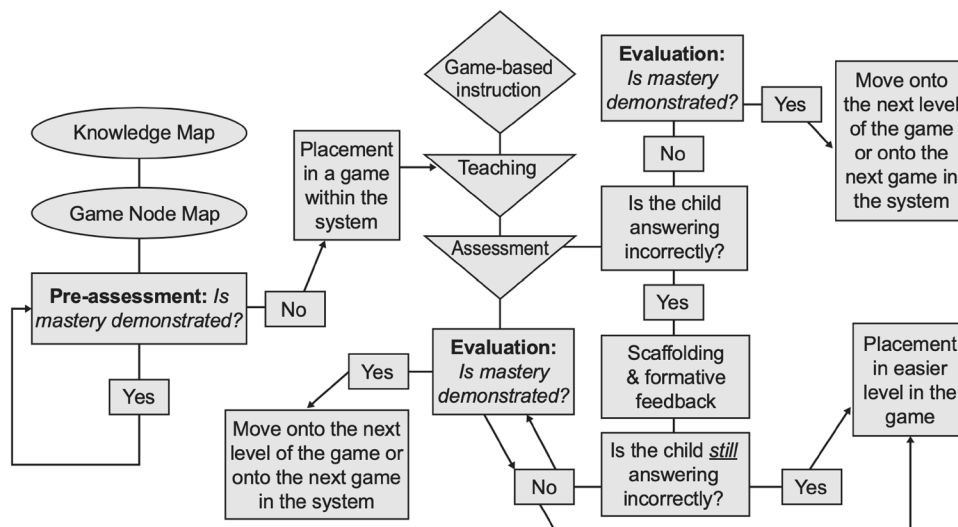
The program is grounded in Bloom’s (1968) Mastery Learning theory, which posits that all students can learn given needed time and appropriate instruction and advocates for a mastery-based personalized learning approach (see also Bingham et al., 2018; Plass & Pawar, 2020). This approach works because it respects learner variability by differentiating instruction and feedback, ensuring that children master each topic before moving on. The program also integrates

evidence-centered design – an assessment framework that enables the estimation of students’ competency levels via in-game learning evidence (Mislevy et al., 2003; Shute, 2011). The game operates on logical relationships between (1) learning objectives that constitute the constructs to be measured; (2) evidence in the form of game play performances that should reveal the target constructs; and (3) features of tasks that should elicit the desired evidence.

Figure 1 shows My Math Academy’s personalized mastery-based learning model, whose key components are aligned with Bloom’s Mastery Learning model (1968), including preassessments; instruction; feedback and correctives; evaluation; and alignment to a hierarchy of learning goals and objectives. It is based on the evidence-centered design framework which provides a principled alignment of the concepts, skills, and abilities a game is designed to teach with evidence of learning and task design (Clarke-Midura et al., 2012; DiCerbo et al., 2015; Shute, 2011). My Math Academy differs from other products in its ability to determine gaps in each child’s learning through engaging, interactive embedded formative and summative assessments (Shute, 2011).

When students use My Math Academy for the first time, they complete in-game preassessments that determine their prior knowledge and place them into proficiency-appropriate games. Each game begins with a teaching activity that provides an overview of the game, the problem-scenario, and audio instructions on the math content needed to successfully complete the task. After completing the teaching activity, the student moves onto other activity levels within the game, which include easy, medium, hard, and the “boss” level. A student demonstrates mastery on a learning objective by passing the “boss” activities within that game. Furthermore, each activity level includes four to six rounds (i.e., questions), and based on each learner’s performance,

Fig. 1 Personalized mastery-based learning model of My Math Academy



adaptive algorithms determine in real time which game, at which difficulty level, to recommend. Within each activity, adaptivity functions dynamically to modify instruction and provide scaffolding, feedback, and content to guide learners through each question. As students use the program, the data collected fuel in-system adaptive support and/or in-classroom educator intervention via other My Math Academy learning activities or small group instruction. A detailed example of a game in My Math Academy is included in “Appendix A”.

Engagement and Learning in Context

My Math Academy leverages games as a vehicle for playful engagement, learning in context, and formative assessments. It engages children in play, and an important engagement tool in My Math Academy is the story context for the game-based learning activities. The context of a math problem makes concepts and operations more meaningful to students and provides students with a scaffold and framework for understanding what they are expected to do, and why (Hirsh-Pasek et al., 2015; Sullivan et al., 2003). Storylines in [game-based learning](#) activities help all students, including struggling readers, gain access to the math and make sense of math problems in a story context. As Fisch’s Capacity Model (2000) posits, learning content integrated within narrative contexts creates mediated environments in which the narratives do not compete for limited cognitive resources. Story contexts can also help students transfer skills learned in games into the real world. Lastly, games enable integrated, ongoing formative assessments to provide useful feedback during the learning process (Shute & Kim, 2014), which enables ongoing feedback cycles and customized learner difficulty levels (Shute & Kim, 2014). The just-in-time feedback may change behaviors that are fed into the next round of formative feedback (e.g., Ke et al., 2019).

Developmentally Appropriate Tasks

Research in human cognition and development informed various aspects of the interaction design of My Math Academy. Frequent user testing and application of design research practices (Design Based Research Collective, 2003; Laurel, 2003) informed insights that drove design iterations. For example, with the user interface design, developers considered cognitive load, i.e., the amount of working memory used in a given context. Research on cognitive load (e.g., Sweller, 1988) informs how to implement visual and audio elements so as not to overwhelm the learner nor to distract them from the core learning interactions. Designers made deliberate decisions to include a minimal number of elements on screen required for each learning game. Furthermore, motor skill considerations shaped specific interaction

design, iterating with user testing data to support tap, drag, and drop interactions calibrated to the target age group for each game. Executive function capability was also a consideration in designing the complexity of games, to keep the layers of instruction, visual and verbal cues, and problem-solving steps at appropriate levels for early learners.

Methods

Design and Participants

This study took place between February and June 2019 and used a blocked cluster randomized design. A total of 41 classrooms were recruited from 11 high needs elementary schools in two school districts in Southern California. The percentage of English Learners across the eleven schools ranged from 35 to 75%; each school had a large proportion of Latinx students (77.9–95.5%), and most students in the schools (70–90% of school population) were eligible for free or reduced-price lunch. Blocking on school and grade, teachers were randomly assigned to one of two conditions: (1) a treatment condition where teachers would implement My Math Academy (21 classrooms), or (2) a control condition where teachers would implement whatever math curricula and interventions the district had in place (20 classrooms).

The 41 participating classrooms had a total of 988 students (507 in treatment, 481 in control), among whom 34 students opted out from the study. Among the 954 remaining students, 55% ($n = 523$) were in kindergarten, 43% ($n = 406$) were in first grade, and only 3% ($n = 25$) were in second grade (one treatment and one control classroom were 1st and 2nd grade combination classrooms). Given the focus of the study on kindergarten and first grade students, 918 students in these two grades completed the pre-assessment (baseline sample); 897 students in the same grades completed the post-assessment (98% retention); and 886 students completed both pre- and post-assessments (analytic sample). There were no significant differences between the treatment and control group teachers in their years of teaching their grades (treatment = 3.45 years [SD = 1.43]; control = 3.90 [SD = 1.37]). Additionally, the two groups of teachers did not differ significantly in their levels of confidence in their math knowledge, both groups reporting that they are somewhat confident (3) or extremely confident (4) on a scale of 1 to 4: (treatment pre = 3.33 [SD = 0.66], post = 3.10 [SD = 0.55]; control pre = 3.11 [SD = 0.68], post = 3.15 [SD = 0.75]). The two groups of teachers also did not differ significantly in their levels of confidence that they had the tools and resources they need to effectively teach math to their students, both groups reporting that they are somewhat confident (3) or not very confident (2) on a scale of 1–4: (treatment pre = 2.86 [SD = 0.57]; post = 2.95 [SD = 0.61];

Table 1 Analytic sample

	Treatment	Control	Total
Kindergarten	233	252	485
1st & 2nd grades	219	182	401
Total	452	434	886

control pre = 2.90 [SD = 0.58], post = 2.85 [SD = 0.67]). Table 1 shows the breakdown of the analytic sample by grade and condition. No classroom dropped from the study, which resulted in no attrition at the classroom level.

All participating teachers received a two-hour professional development training on My Math Academy. Treatment teachers received the training prior to the start of the study, while control teachers received the training after the study ended, in preparation for the following school year. The training provided an overview of the skills covered in My Math Academy, how it was designed to support mastery learning for individual students, and ideas about how to integrate it into classroom instruction as a supplement to the core math curriculum. Teachers also received a user guide for My Math Academy; treatment teachers attended a 1-h webinar (a month after the study began) on using the teacher dashboard that displayed student progress and usage.

Measures and Data Sources

Student Math Assessment

The researchers developed an assessment to measure student mathematics knowledge using high-quality, standards-based items selected from the Certica assessment item bank and administered on an electronic platform managed by one of Certica's partners, LinkIt! The selected assessment items were validated by Certica using point biserial coefficients (average = 0.42, range 0.39–0.48), and each targeted specific learning objectives students encountered while using My Math Academy. The full suite of assessment items administered can be found in "Appendix C." All assessment items were aligned to the California Common Core State Standards for kindergarten, first, or second grade (See "Appendix B" for the alignment of the California Common Core State Standards, the learning objectives, and the associated assessment items). A Spanish version of the assessment was available, but teachers expressed reluctance for having their students tested in Spanish, resulting in only 12 students (three treatment, nine control) taking the Spanish version of the pretest, and all students taking the assessment in English on posttest.

The pre-assessment included a total of 31 multiple-choice items, and the students had 30–45 min to complete it. Observational visits conducted near the end of the school year

indicated that many students had made substantial progress within My Math Academy, well into second grade materials; therefore, the post-assessment was modified to include seven additional items targeting second grade standards. Students had about 45 min to complete the posttest, but kindergarten students were allowed to stop after 31 items if they were struggling to finish.

The assessments were administered at the beginning of the study (February 2019) and after the classroom implementation had ended (May–June 2019). Researchers administered the assessment in small groups of 5–6 children using tablets and provided individual support to those who needed help with login or other technology-related issues. Students were provided with headphones to hear audio instructions associated with each assessment item; they responded to questions by tapping on the screen to select an answer choice. Researchers monitored the students to help them feel comfortable taking the assessment and to ensure that they had answered all questions before submitting the assessment.

Students' responses were coded as 0 (incorrect) or 1 (correct) for the subsequent analyses. The Kuder-Richardson (KR-20) reliability coefficient for measure of internal consistency was 0.89 for the pre-assessment and 0.90 for the post-assessment. The exploratory factor analysis indicated that there is one dominant factor (math knowledge) for each assessment.

My Math Academy Usage Data

Treatment teachers were asked to use My Math Academy for a minimum of 60 min per week (20-min sessions, three times/week) as a supplement to their core math curricula. The total amount of usage and total activities completed were the primary variables examined in this study. The usage data also includes the students' performance on the in-game preassessment; the specific games accessed, each of which targets one or more granular skills within the target learning objectives (e.g., counting from 1 to 20, adding and subtracting 3-digit numbers); all user interactions within the app; and the performance of each child at different difficulty levels within each game. The skills covered in the games can be found in "Appendix B."

Teacher Surveys

Participating teachers in treatment and control conditions completed both a pre-study survey (21 treatment, 19 control) and an end-of-study survey (21 treatment, 20 control). The pre-study survey asked teachers to give information regarding their regular classroom practices and activities. The end-of-study survey captured what occurred in the classroom during implementation. Treatment teachers gave feedback on

the feasibility and educational value of My Math Academy, while control teachers provided information regarding their use of educational technology tools and mathematic content taught over the course of the semester. They rated the impact of My Math Academy or the program they used on their students' math skills (e.g., counting skills, addition/ subtraction skills, interest in learning math, focus and attention during math lessons) using a Likert-type scale, with 1 = "very negative impact" and 5 = "very positive impact." They also responded to statements such as "My Math Academy was easy for me to use as a teacher" and "I find My Math Academy to be a valuable math learning resource." These were coded on a 4-point Likert-type scale, with 1 = "strongly disagree" and 4 = "strongly agree." Quantitative survey data were analyzed using SPSS.

Teacher Interviews

A subsample of teachers from both treatment (8) and control (4) conditions participated in an interview near the end of the study. Treatment teachers provided insight into their general math curriculum and typical classroom practices/resources, implementation of My Math Academy, and recommendations for the app. Control teachers reflected on their general math curriculum and classroom practices, as well as broader topics regarding the resources or interventions used during their math instruction. Qualitative analyses were conducted to better understand the fidelity and quality of classroom implementation of My Math Academy as well as the user experience of participants. Researchers developed codes based on research questions and emergent themes utilizing NVivo software. Data from teacher interview transcripts and pre/post teacher survey responses were triangulated to identify major themes and to create qualitative narratives of how teachers utilized My Math Academy and its potential impact on the student learning experience. Sample teacher interview questions can be found in "Appendix D."

Classroom Observations

A subsample of treatment (8) and control classrooms (4) were observed once by a pair of researchers. Observations in the treatment classrooms captured the extent to which My Math Academy was implemented without barriers, how engaged students were with My Math Academy, other teacher-led instruction, and the overall classroom environment. Observations in the control classrooms focused on the implementation of teacher-led math instruction and the overall classroom environment, along with the use of any digital math apps that were in use during the observations. Researchers used an observation protocol to maintain a running record ("Appendix E"), described the environment and technology, provided ratings for generalized fidelity of implementation (using three overarching domains: classroom management, quality of math instruction, quality of game play, if applicable), and provided a rationale regarding the rating. An inter-rater reliability score of 0.89 was calculated regarding the generalized fidelity of each classroom.

Analysis

Student Math Assessment

Baseline equivalence test was conducted using pre- and post-matched sample and the baseline sample. As shown in Table 2, there were no statistical differences between the treatment and control groups at baseline.

To estimate the impact of My Math Academy on student outcomes, Stata 16 was used to specify a two-level Hierarchical Linear Model. In addition to accounting for the nested structure of the data (students nested within classrooms), HLM allowed researchers to estimate the treatment effects and incorporate relevant variables from the different levels (i.e., student,

Table 2 Baseline differences on the pretest scores with 31 items

Adjusted means						
Study sample	Treatment (SD)	Control (SD)	Difference (SE)	<i>p</i> -value	95% confidence interval	Unweighted student sample size
Pre- and post-matched sample	15.38 (6.97)	15.53 (6.80)	-0.16 (0.60)	0.794	-1.34 to 1.02	886
Baseline sample for random assignment	15.46 (6.98)	15.54 (6.79)	-0.09 (0.59)	0.875	-1.25 to 1.07	918

A two-level regression model that accounted for study design characteristics (blocks used for random assignment purpose) was used to test the baseline equivalence between the treatment group and business-as-usual group. The standard error was estimated using the Huber-White procedure (Greene, 2003)

teacher) to examine potential moderators. Data were regression-adjusted to account for differences in baseline measure, demographics available through school records (i.e., grade level, gender, ethnicity, and free or reduced-price lunch status). In the model below, subscripts i and j denote student and teachers, respectively; Math represents student achievement in math; PreMath represents the baseline measure of math performance; ETH, GR, FRL, and GEN are dummy variables corresponding to the students' ethnicity, grade level, free and reduced lunch program, and female status, respectively. TREAT is a dichotomous variable indicating student enrollment in a classroom that has been assigned to the treatment condition; the intervention effect is represented by β_1 , which captures treatment/control differences in changes in the outcome variable between pretest and posttest. r_{0j} is a random effect term, the variance component of which accounts for the nesting of students within classes; e_{ij} represent a student-level random error term.

$$\text{Math}_{ij} = \beta_0 + \beta_1(\text{TREAT})_j + \beta_2(\text{PreMath})_{ij} + \beta_2(\text{ETH})_{ij} + \beta_3(\text{GR})_{ij} + \beta_4(\text{FRL})_{ij} + \beta_5(\text{GEN})_{ij} + r_{0j} + e_{ij}$$

Students who did not take the post-assessments were excluded from the impact sample. To account for missing values on the covariates, the missing-indicator method (White & Thompson, 2005) was used, which appears to refine the precision of impact estimates and standard errors. Additionally, to address concerns about omitted variables, the Konfound-it app (Frank et al., 2013) was used to quantify how much bias there would have to be due to omitted variables or any other source to invalidate any inferences made.

Qualitative Data Analysis

Researchers reviewed the classroom observation data, teacher survey data, and interview transcripts to identify major themes and patterns across classrooms. Guided by Miles and Huberman's (2014) qualitative data analysis process, researchers developed codes based on research questions and emergent themes and then utilized NVivo software to assist with the analysis of the data.

Teacher interview data contained four main topics: teachers' general math curriculum, use of math interventions, features of the intervention, and student experiences of the intervention. Each of these topics had specific codes. For example, both treatment and control teachers' responses to questions about their general math curriculum were documented using codes such as: (a) alignment to Common Core and/or District Standards, (b) implementation features, (c) curriculum resources, and

(d) progress monitoring. Treatment teachers were asked about My Math Academy specifically, while control teachers were asked about their use of math interventions more broadly. Some of the codes related to My Math Academy included experiences using the app, student appeal, student successes, student struggles, teacher resources, supporting student math learning, and improvements needed/recommendations.

Observations were conducted by two researchers at a time using a researcher developed protocol with three overarching domains comprised of sub-domains. Classroom management included (1) noise level, (2) organization/displays, (3) quality of transitions, (4) time management, and (5) on-task behaviors. Quality of math instruction included: (1) whole group interactions, (2) small group interactions, (3) individual or pair work, (4) teacher-child interactions, (5) quality of student learning, and (6) quality of math content and instruction. The game play domain (if applicable) included: (1) functionality, (2) child attention, (3) child posture, (4) child engagement. Observers followed a rubric to rate each subdomain as high (3), moderate (2) or low (1). For each pair of observers, their scores were compared, and interrater reliability was established at 0.89. Sub-domain scores were averaged to calculate overall scores for each overarching domain by classroom. The overarching domain scores were averaged by treatment condition to determine whether there were any meaningful differences between the groups.

Based on a triangulation of the qualitative data sources (teacher surveys, teacher interviews and classroom observation), three major themes emerged around the potential benefits of implementing My Math Academy in classrooms, including (1) the use of embedded assessments, (2) adapting to students' zone of proximal development, and (3) engaging or motivating students through new and/or challenging content. Each of these themes will be discussed in the results section below and help to contextualize how My Math Academy may have supported growth in [student math performance](#). Finally, the qualitative data also provided a snapshot of the overall classroom context and some information about other math instruction taking place.

Results

General Math Instruction

Information from surveys, interviews, and observations revealed that teachers in both conditions used a combination of math curricula and supplemental applications during their general math instruction and intervention time. They reported teaching math at least 4 times a week, usually for 45–60 min per day. Observation data indicated that teachers in both

groups had strong classroom management with respective average ratings of 2.74 and 2.85 (on a 3.00 scale). The overall quality of math instruction was also rated similarly at 2.65 and 2.85 for the treatment and control classrooms respectively. The overall quality of the My Math Academy game play for the treatment group was 2.67 suggesting that teachers were able to effectively implement My Math Academy as part of their math instruction.

Prior to the study, treatment teachers received training on implementation of My Math Academy, during which they were asked to implement the program as a supplemental intervention in addition to their general mathematics curriculum. They were also asked to not implement any other supplemental math intervention for duration of the study period, but survey results showed that they did not comply with this instruction. Specifically, they used Starfall (12), ST Math (9), My Math (5), Prodigy (2), as well as others. Control group teachers used some of these same interventions: Starfall (7), ST Math (7), My Math (6), IXL (5), among others.

Analytic data from My Math Academy indicated that on average, students spent 14.79 h (SD = 5.05 h), or between 74 and 81 min per week on My Math Academy, completing an average of 163 learning activities (SD = 67.12 activities). Observation data suggested that teachers often implemented My Math Academy with the whole class (individual students working on their tablets) during their designated math lessons. Another frequently observed implementation was having a small group of 5–6 students use My Math Academy as part of a rotating learning station equipped with tablets, while other groups of students worked on different activities such as reading, puzzles, drawing, etc. In interviews, teachers expressed that when My Math Academy was implemented with the whole class with individual headphones, teachers could roam and intervene when students requested teacher assistance. Teachers also felt comfortable using My Math Academy in small groups because the platform was easy for students to use independently. Students could confidently navigate through the app, which led to a seamless transition between activities. Survey and observation data also revealed that teachers were able to implement My Math Academy with very few hinderances related to technology (e.g., internet problems, login issues, glitches in the app, use of headphones, students using an incorrect account).

Student Math Performance

The results of the HLM showed that My Math Academy does help improve students’ math knowledge as measured by the researcher-developed math assessment. Specifically, treatment students who used My Math Academy for 12–13 weeks outperformed their control group peers as shown in Table 3. The difference between treatment and control groups is small but statistically significant (effect size = 0.11, $p = 0.026$).

Further analyses showed that 22.3% of the estimated effect of My Math Academy on students’ mathematics performance would have to be due to bias to invalidate the inference of an effect of the program on mathematics performance (Frank et al., 2013). To invalidate the inference, one would have to replace 198 of the observed data with null hypothesis cases of no effect of the program. This provides evidence of the practical significance of the difference between the treatment and control groups. Even though the difference in the raw scores of the two groups may be just one point, the difference is a meaningful one, given the substantial proportion of the estimated effect that would have to be due to bias to nullify the claim that My Math Academy had a positive impact on students’ math skills.

Subgroup analyses showed that the effects of the program varied by grade. In kindergarten, the treatment group scored about 1 point higher than the control group, and this difference is statistically significant (20.14 vs. 19.23, $p = 0.01$, effect size = 0.16). For first grade, the difference decreased

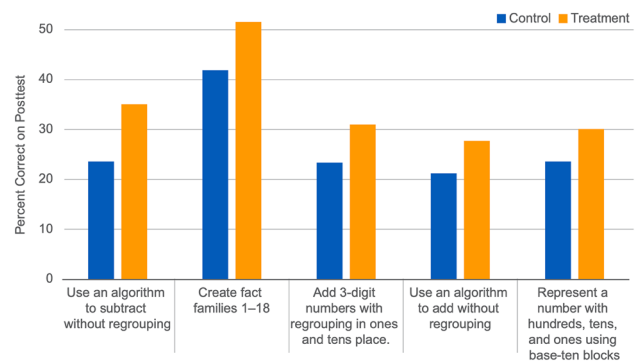


Fig. 2 Effect of My Math Academy on the most advanced skills assessed

Table 3 The effect of My Math Academy on student math achievement

Adjusted means						
Outcome measure	Treatment (SD)	Control (SD)	Difference (SE)	p-value	Effect size	Unweighted student sample size
Post-test (38 items)	20.97 (8.01)	20.08 (7.84)	0.89* (0.40)	0.026*	0.11	922

* $p < .05$

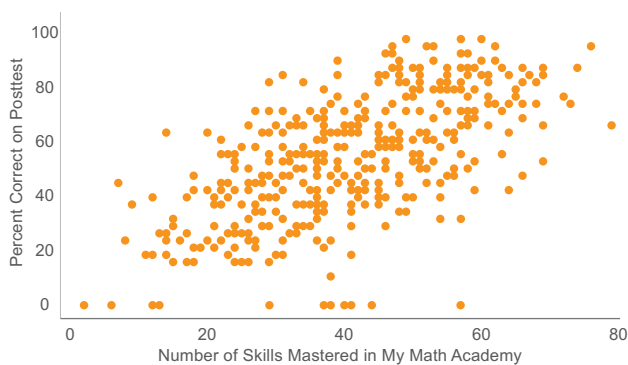


Fig. 3 Relationship between number of skills mastered in My Math Academy and posttest performance

to about 0.6 points and is not statistically significant (21.37 vs. 20.80, $p = 0.36$, effect size = 0.09). The program appears to have benefited most the students at kindergarten who had more room to grow.

Among the skills addressed in the assessment, My Math Academy appeared to have the greatest impact on the most difficult skills—the ones addressed in the additional items added to the posttest after observations that revealed many kindergartners accessing second grade-level games (see Fig. 2). On the posttest, 34% of treatment versus 22% of control group students were able to use an algorithm to subtract without regrouping ($X^2 [1, N = 886] = 17.49, p < 0.001$); 50% of treatment versus 42% of control group students could create fact families using numbers between 1 and 18 ($X^2 [1, N = 886] = 6.45, p < 0.05$); 31% vs. 23% of control group students could add three-digit numbers with regrouping in ones and tens places ($X^2 [1, N = 886] = 6.73, p < 0.01$); 27% of treatment versus 19% of control students could use an algorithm to add without regrouping ($X^2 [1, N = 886] = 6.91, p < 0.01$); and 29% of treatment versus 21% of control students could represent three-digit numbers using base-ten blocks ($X^2 [1, N = 886] = 6.37, p < 0.05$).

To determine whether progress within My Math Academy was related to growth in math skills as measured by the assessment, we examined the number of skills students mastered within the game as well as time spent using the program in relation to performance on the posttest. Overall, students mastered on average 61 skills ($SD = 21.4$ skills), and 68 students (15%) completed the entire game, demonstrating mastery on all 96 skills in the program. Among kindergartners, 40.4% of the students mastered at least 80% of grade-level skills, and 54.2% of the students mastered at least 50% of 1st grade level skills. Among 1st graders, 55.3% of the students mastered at least 80% of grade-level skills, and 42.5% of the students mastered at least 80% of 2nd grade skills. Although the correlation between post-test score and average use per week in minutes was weak at 0.18,

the correlation between post-test score and cumulative number of skills mastered was strong at 0.73 (see Fig. 3 below).

Student Learning and Engagement

The qualitative data sources were analyzed to understand the extent to which My Math Academy supported students' learning and engagement based on teacher feedback and researcher observations. The findings indicated that My Math Academy was a valuable learning resource for students that had a positive impact on student learning. Interviewed teachers identified three key components of My Math Academy that led to their students' success in mastering mathematical concepts, including embedded assessments, adaptiveness to students' zone of proximal development, and motivating students through new and/or challenging concepts. In the end-of-study survey, teachers in both treatment ($n = 20$) and control ($n = 20$) groups responded to questions about the extent to which the educational technology they used had a positive impact on their students' math skills. As shown in Fig. 4, teachers reported that in comparison to other technology used in control classrooms, My Math Academy had a significantly more positive impact on students' math skills. All differences between treatment and control group averages were statistically significant ($p < 0.05$) except "Enjoyment in learning math" ($p < 0.10$) and "Focus and attention during math" ($p = 0.26$). Effect sizes ranged from Cohen's $d = 0.6$ to 1.05. While these survey responses include subjective evaluations of teachers who used a new resource, they are meaningful given the implementation of the program in real classroom settings and the teachers' recognition of My Math Academy as a tool that provides personalized experiences for each learner.

Interview data showed that one of the components of My Math Academy that contributed to teachers' positive perception of the program was its ability to differentiate

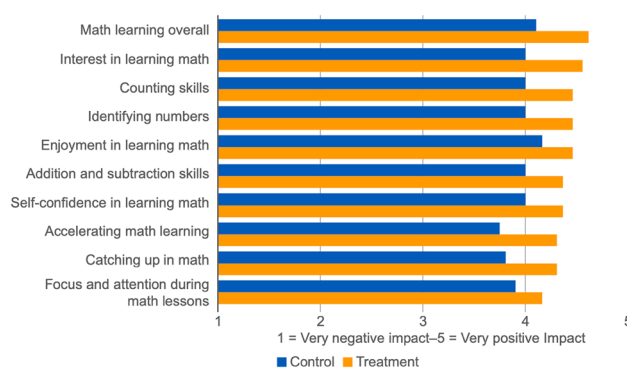
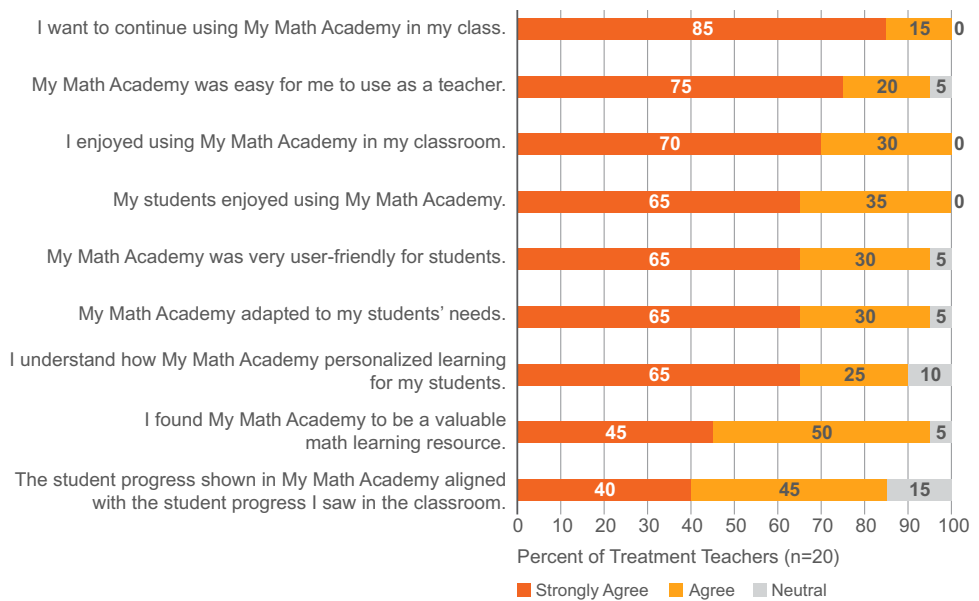


Fig. 4 Teachers' perception of educational technology on their students' math skills

Fig. 5 Treatment teachers' experiences of using My Math Academy in their classrooms



and meet students at their current level of understanding. Teachers were impressed that the program itself was able to adapt based upon each student’s zone of proximal development, and they witnessed several instances of student successes as a result. One teacher elaborated:

I have a kindergartener in my class who is very advanced, and so it’s great that [the student] was able to test in higher and he's working at a higher level.... I love it that My Math Academy is giving him more than what I can give him in the classroom setting.... I also have people who are working below grade level, and I feel it's going in and helping them fill in those gaps, ... stuff that they didn't catch towards the beginning of the year.

In general, students at all levels of math proficiency were able to steadily progress at their own pace, and teachers identified this as a valuable feature of My Math Academy.

Another common theme was how My Math Academy helped motivate students when they faced new, challenging concepts. Specifically, teachers reported that the engaging nature of My Math Academy enabled students to progress steadily through more difficult material without showing signs of fatigue or frustration. Teachers emphasized that when using My Math Academy, they saw improvement in both student understanding of material and persistence in working through with difficult concepts. Furthermore, teachers found that My Math Academy gave students the chance to understand math at a deeper level and described it as a constructive medium for their students to reinforce their math learning and continue developing their skills. The embedded assessments and reviews of fundamental math skills in My Math Academy helped students engage with

math content in multiple ways, strengthening their foundational knowledge.

One of the most notable findings was how much students enjoyed using My Math Academy. In interviews and end-of-study surveys (see Fig. 5), all teachers reported that their students enjoyed using the app, and one teacher elaborated:

[Students] look forward to going to the computers and the iPads. [When on My Math Academy] I can hear them singing, I can hear them counting, I can hear them really engaging in whatever it is they're doing. ... They're really into it and then I hear a lot of conversations with each other. "Oh yeah. I did that game too. Did you do the one with the [x]?" They're having these awesome conversations about their math experiences.

Researchers conducting classroom observations also saw students counting, singing, and dancing along with My Math Academy. Exclamations of excitement (“yes!”) from students or showing teachers and other students their screens when they got an answer correct were common behaviors among students who were engaged with My Math Academy.

Discussion

This study provides evidence to support the principles outlined in a position statement adopted by the National Association for the Education of Young Children (2012, 2022), which describe the need to implement developmentally appropriate instructional practices to ensure the effectiveness of early childhood education. The principles are built upon well-documented research on the developmental sequence of

children's learning; they highlight the variability in the rates at which learners mature and develop skills; and they emphasize the importance of play, social interactions, and contexts in helping young children construct their understanding of the world. My Math Academy is a mastery-based personalized learning system with a design that incorporates key principles such as mathematical learning trajectories (Clements & Sarama, 2004, 2014) and personalized, adaptive challenges that are generated through formative assessments (Gee, 2003), set in [game-based learning](#) contexts that are developmentally appropriate for young learners. The study results indicated that overall, those who used My Math Academy for about 15 h over the course of 12–13 weeks in spring 2019 outperformed their control group peers on the math assessment. This finding is especially noteworthy considering that treatment students were engaged in self-guided play within My Math Academy, without active instruction from their teachers.

Additionally, the features of My Math Academy that teachers identified as helpful in supporting their students' progress in mathematics – embedded assessments, adaptiveness to each student's ZPD, and feedback that motivate students through challenging concepts – are the essential elements of the personalized mastery-based learning model. Teachers' recognition of these features of My Math Academy suggests that the program can indeed alleviate the challenges that many teachers face when working with a classroom of children with a wide range of abilities or prior knowledge, as the program can appropriately challenge learners who are more advanced and provide the support for those who need it.

Another noteworthy finding of the study is that My Math Academy appeared to have the greatest impact on the most challenging math skills for young learners, which are also the skills that are most likely to be overlooked by teachers. Research indicates (e.g., Engel et al., 2013, 2016) that kindergarten teachers tend to spend most of the time on basic skills such as shape recognition and simple forward counting and forego some of the more advanced skills such as skip counting or counting backwards, even though most children enter kindergarten not having mastered these skills. This points to the value of My Math Academy as a resource that can provide personalized instruction to children who are ready for the challenge, especially on skills that may be overlooked by teachers whose attention may be demanded by children who are struggling to master the basics.

Finally, teachers recognized My Math Academy as a valuable resource that keeps children motivated and engaged in learning math. They were impressed with My Math Academy's ability to personalize the learning experience and meet children at their current level of understanding. They also appreciated how children could easily navigate through the app independently and direct their own learning, enabling

teachers to engage in other activities with students who may need small group or one-on-one instruction.

The results of this study replicate those of an efficacy study conducted in 2017 with pre-K and kindergarten classrooms that used My Math Academy over a period of 12–14 weeks (Thai et al., 2021). Similar to the 2017 study, the current study targeted Title I schools with predominantly Latinx students and large proportions of English Language Learners in an effort to address the opportunity and achievement gaps that often leave children from low socioeconomic (SES) families at a disadvantage. The evidence from this study supporting the effectiveness of My Math Academy in improving students' early math knowledge and increasing their confidence and interest in learning math suggests that My Math Academy is a promising intervention that can benefit children from low-SES families, helping to reduce the gap in their math knowledge relative to their middle-class peers.

Limitations and Future Research

A limitation of this study is that treatment teachers used other interventions despite instructions to refrain from using them for the duration of the study. The array of digital interventions used in both treatment and control classrooms indicates that business-as-usual instruction in this information age is the implementation of a range of digital programs rather than the absence of them. The impact of My Math Academy might have been diluted between treatment and control groups because some teachers in the control condition reported using other digital math program(s) with features similar to the intervention. While the research team cannot prevent the control teachers from using other programs, the product development team continues to refine My Math Academy to differentiate it even more from its competitors, and the small effect observed in this study can be interpreted as the incremental impact of My Math Academy in addition to the effects that may have been observed due to the implementation of several other digital math programs.

Another limitation is that the assessment used to measure students' math skills is not a standardized test. When selecting the assessment, a review of existing standardized assessments of early math skills revealed a lack of adequate alignment between the assessments and My Math Academy, especially given that most standardized assessments are designed to measure progress that occurs over a period of a school year, while My Math Academy is a supplemental program designed to be used in conjunction with a core math curriculum. This led the research team to construct an assessment using items from the Certica assessment item bank. This self-designed assessment may be limited for capturing what the students learned through My Math Academy,

and a more comprehensive standardized test should be considered or developed for future studies in order to provide stronger evidence of the extent to which My Math Academy has a meaningful and practical impact on improving students' math skills.

Additionally, a large proportion of the study sample was Latinx, and it is likely that a significant proportion of these students did not speak English as a first language. Some of the schools involved in the study did not agree to share data about individual student's home language or English language proficiency; therefore, the analyses could not account for the effect of student's language skills on the outcome. Moreover, teachers of only three treatment and nine control students opted to have the students take the Spanish version of the assessment at pretest, and no student took the Spanish version of the assessment at post-test. Several of the assessment items required students to read short sentences, and it is possible that the post-test scores are partly a reflection of students' English language proficiency, in addition to their math knowledge. Efforts will be made in future studies to minimize the potential effects of students' language skills on their math assessment performances.

A strike in one of the school districts delayed the start of the study, leaving only about 12 weeks (less than 1/2 of a school year) for the implementation of My Math Academy. Like any new program, teachers needed time to learn how to effectively use My Math Academy. Given the strong positive correlation between the post-assessment score and the number of skills mastered, it is likely that a longer intervention period would have generated greater growth in skills among the treatment students. A longer intervention period would also allow more time for teachers to learn to effectively integrate My Math Academy into their instruction. Longitudinal studies across multiple grades in the future could help us determine the long-term effects of using the program.

Future studies should also examine the relationships between My Math Academy use and other constructs that affect development, such as children's social interactions, persistence in learning, and preference for more tangible play objects versus digital / abstract games. It is also possible that factors such as children's personalities, degree of prior experience with digital games, and language development affected the extent to which children engaged with My Math Academy and the ease with which they could learn using the program. Additionally, future investigations of My Math Academy will focus on how specific student-facing games (including interactions in the app and different levels of supports in the games) and educator-facing resources (developed after the study) supplement core math curricula / general math instruction in classrooms.

Conclusion

This study provides evidence that My Math Academy, a supplemental mastery-based personalized math learning program designed with research-based, data-informed learning engineering process, can significantly improve young children's early math skills while keeping them interested in learning math. Such evidence is important, especially given the increased interest in [game-based learning](#) resources and the need for educators and parents to evaluate the utility of tools that children can use independently.

Findings from this study, such as the teachers' recognition of My Math Academy as a resource that can successfully adapt to individual student's learning needs, will inform ongoing refinements to My Math Academy, such as the development of educator tools that would enable teachers to further personalize instruction for individual students outside of the app. Additionally, given the important role that parents play in their children's education, future enhancements of My Math Academy will include resources for parents as well, which will equip them with information and tips on activities they can do with their children to help these young learners further expand and apply their math knowledge in the real world. Future efficacy studies of My Math Academy will investigate the impact of these new resources and reveal insights on how the program can be successfully implemented in various learning contexts.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10643-022-01332-3>.

Author Contributions HJB serves as the Senior Director of Efficacy Research and Evaluation at Age of Learning, Inc. She oversees the design and implementation of studies to maximize the effectiveness of digital programs aimed at improving children's academic and social-emotional skills. Areas of her research interests include use of technology to personalize instruction, language acquisition, and education of immigrant origin youth. LL, a Project Director / Senior Research Associate at WestEd, leads research work on the areas of developmental psychology, early math and science intervention, inclusion of children with disabilities in the regular education classroom, and family engagement. Her recent work involves using interactive games to design and evaluate interventions for students living in poverty and at risk for academic difficulties. KF, Director of Research and Evaluation in WestEd's Science, Technology, Engineering, and Mathematics (STEM) program, leads multiple large-scale efficacy studies that are federally funded by the Institute of Education Sciences and Office of Innovation and Improvement, as well as large contracts with educational companies, such as Scholastic, Newsela, and Age of Learning, Inc. Her research interest focuses on testing interventions for younger students at risk for less than optimal school outcomes.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are

included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arnott, L. (2016). An ecological exploration of young children's digital play: Framing children's social experiences with technologies in early childhood. *Early Years: An International Journal of Research and Development*, 36(3), 1–18. <https://doi.org/10.1080/09575146.2016.1181049>
- National Association for the Education of Young Children, & Fred Rogers Center for Early Learning and Children's Media at Saint Vincent College. (2012). Technology and interactive media as tools in early childhood programs serving children birth through age 8. Washington, DC. Retrieved from https://www.naeyc.org/sites/default/files/globally-shared/downloads/PDFs/resources/position-statements/ps_technology.pdf
- Bingham, A. J., Pane, J. F., Steiner, E. D., & Hamilton, L. S. (2018). Ahead of the curve: Implementation challenges in personalized learning school models. *Educational Policy*, 32(3), 454–489.
- Bloom, B. (1968). Learning for mastery. In J. H. Block (Ed.), *Mastery learning: Theory and practice* (pp. 47–63). Holt, Rinehart, & Winston.
- Brody, J. (2015). Screen Addiction Is Taking a Toll on Children -The New York Times, https://well.blogs.nytimes.com/2015/07/06/screen-addiction-is-taking-a-toll-on-children/?_r=0
- Bronfenbrenner, U. (1995). Developmental ecology through space and time: A future perspective. In P. Moen, G. H. Elder, & K. Luscher (Eds.), *Examining lives in context: Perspectives on the ecology of human development* (pp. 619–647). American Psychological Association.
- Campuzano, L., Dynarski, M., Agodini, R., & Rall, K. (2009). Effectiveness of reading and mathematics software products: Findings from two student cohorts. NCEE 2009–4041. National Center for Education Evaluation and Regional Assistance.
- Chu, F. W., vanMarle, K., & Geary, D. C. (2016). Predicting children's reading and mathematics achievement from early quantitative knowledge and domain-general cognitive abilities. *Frontiers in Psychology*, 7, 1–14. <https://doi.org/10.3389/fpsyg.2016.00775>
- Clarke-Midura, J., Code, J., Dede, C., Mayrath, M., & Zap, N. (2012). Thinking outside the bubble: Virtual performance assessments for measuring complex learning. In M. Mayrath, J. Clarke-Midura, D. Robinson, & G. Schraw (Eds.), *Technology based assessment for 21st century skills: Theoretical and practical implications from modern research* (pp. 125–147). New York, NY: Information Age Publishing.
- Clements, D. H. & Wright, T. S. (2022). Teaching content in early childhood education. In NAEYC (2022). Developmentally appropriate practice in early childhood programs serving children from birth through age 8. 4th edn, (pp. 69–72). Washington, DC: NAEYC.
- Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning*, 6(2), 81–89. https://doi.org/10.1207/s15327833mtl0602_1
- Clements, D. H., & Sarama, J. (2014). Learning trajectories: Foundations for effective, research-based education. In A. P. Maloney, J. Confrey, & K. H. Nguyen (Eds.), *Learning over time: Learning trajectories in mathematics education* (pp. 1–30). IAP Information Age Publishing.
- Denton, K., & West, J. (2002). Children's reading and mathematics achievement in kindergarten and first grade (No. (NCES-2002–125)). Washington, D.C.: National Center for Education Statistics.
- Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5–8.
- DiCerbo, K.E. & Bertling, M., Stephenson, S., Jia, Y., Mislevy, R., Bauer, M. & Jackson, G. T. (2015). An Application of Exploratory Data Analysis in the Development of Game-Based Assessments. *Serious Games Analytics*. 319–342.
- Dietze, B., & Kashin, D., (2011). *Playing and Learning in Early Childhood*. Pearson Education Canada.
- Dixon, F. A., Yssel, N., McConnell, J. M., & Hardin, T. (2014). Differentiated instruction, professional development, and teacher efficacy. *Journal for the Education of the Gifted*, 37(2), 111–127. <https://doi.org/10.1177/0162353214529042>
- Dohring, D., Hendry, D., Gunderia, S., Hughes, D., Owen, V. E., Jacobs, D. E., Betts, A., & Salak, W. (2019). U.S. Patent No. 20190236967 A1. Washington, DC: U.S. Patent and Trademark Office.
- Early Childhood STEM Working Group (January, 2017). Early STEM matters: Providing high-quality STEM experiences for all young learners. Chicago, ILL: author. Retrieved at https://urldefense.com/v3/ecstem.uchicago.edu!!F8LM3WSaiTgsDg!2DF0X64sZ9X6ZanNq4mOM2_7xsCq3YXimycX156ZfcV6k6jigOicK3IS6noxI_woS
- Eberle, S. G. (2014). The elements of play: Toward a philosophy and definition of play. *Journal of Play*, 6(2), 214–233.
- Edwards, S. (2016). New concepts of play and the problem of technology, digital media and popular-culture integration with play-based learning in early childhood education. *Technology, Pedagogy and Education*, 25(4), 513–532.
- Engel, M., Claessens, A., & Finch, M. A. (2013). Teaching students what they already know? The (Mis) Alignment between mathematics instructional content and student knowledge in kindergarten. *Educational Evaluation and Policy Analysis*, 35(2), 157–178.
- Engel, M., Claessens, A., Watts, T., & Farkas, G. (2016). Mathematics content coverage and student learning in kindergarten. *Educational Researcher*, 45(5), 293–300.
- Entwisle, D. R., & Alexander, K. L. (1989). Early schooling as a “critical period” phenomenon. In K. Namboodiri & R. G. Corwin (Eds.), *Research in the Sociology of Education and Socialization*, 8 (pp. 27–55). JAI Press.
- Entwisle, D. R., & Alexander, K. L. (1990). Beginning school competence: Minority and majority comparisons. *Child Development*, 61, 454–471.
- Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2005). First grade and educational attainment by age 22: A new story. *American Journal of Sociology*, 110(5), 1458–1502.
- Fisch, S. (2000). A capacity model of children's comprehension of educational content on television. *Media Psychology*, 2(1), 63–91.
- Fisher, K., Hirsh-Pasek, K., & Golinkoff, R. M. (2012). Fostering mathematical thinking through playful learning. *Contemporary debates on child development and education*, 81–92.
- Flynn, R. M., & Richert, R. A. (2015). Parents support preschoolers' use of a novel interactive device. *Infant and Child Development*, 24, 624–642. <https://doi.org/10.1002/icd.1911>
- Frank, K. A., Maroulis, S. J., Duong, M. Q., & Kelcey, B. M. (2013). What would it take to change an inference? Using Rubin's causal model to interpret the robustness of causal inferences. *Educational Evaluation and Policy Analysis*, 35(4), 437–460.
- Friedman, S., Masterson, M., & Wright, B. L. (2022). Context matters: Reframing teaching in early childhood education. In NAEYC (2022). Developmentally appropriate practice in early childhood programs serving children from birth through age 8. 4th ed, (pp. 48–51). Washington, DC: NAEYC.

- Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' Functional Numeracy Is Predicted by Their School Entry Number System Knowledge. *PLoS ONE*, 8(1), e54651. <https://doi.org/10.1371/journal.pone.0054651>
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Palgrave Macmillan.
- Gestwicki, C. (2017). *Developmentally appropriate practice: Curriculum and development in early education* (6th ed.). Wadsworth Cengage Learning.
- Goddard, Y., Goddard, R., & Kim, M. (2015). School instructional climate and student achievement: An examination of group norms for differentiated instruction. *American Journal of Education*, 122(1), 111–131.
- Gross, B., & DeArmond, M. (2018). Personalized Learning at a Crossroads: Early Lessons from the Next Generation Systems Initiative and the Regional Funds for Breakthrough Schools Initiative. *Center on Reinventing Public Education*.
- Hatzigianni, M. (2018). Transforming early childhood experiences with digital technologies. *Global Studies of Childhood*, 8(2), 173–183.
- Hecht, S. A., Burgess, S. R., Torgesen, J. K., Wagner, R., & Rashotte, C. (2000). Explaining social class differences in growth of reading skills from beginning kindergarten through fourth-grade: The role of phonological awareness, rate of access, and print knowledge. *Reading and Writing*, 12(1–2), 99–128. <https://doi.org/10.1023/A:1008033824385>
- Higgins, P. S., Xiao, Z., & Katsipataki, M. (2012). *The Impact of Digital Technology on Learning: A Summary for the Education Endowment Foundation* (p. 52). Education Endowment Foundation.
- Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in “educational” apps: Lessons from the science of learning. *Psychological Science in the Public Interest*, 16(1), 3–34.
- Hoffman, B., & Nadelson, L. (2010). Motivational engagement and video gaming: A mixed methods study. *Educational Technology Research and Development*, 58(3), 245–270.
- Hopkins, L., Brookes, F., & Green, J. (2013). Books, bytes and brains: The implications of new knowledge for children's early literacy learning. *Australian Journal of Early Childhood*, 38(1), 23–28. <https://doi.org/10.1177/183693911303800105>
- Hubber, P. J., Outhwaite, L. A., Chigeda, A., McGrath, S., Hodgen, J., & Pitchford, N. J. (2016). Should touch screen tablets be used to improve educational outcomes in primary school children in developing countries? *Frontiers in Psychology*, 7, 839. <https://doi.org/10.3389/fpsyg.2016.00839>
- Ke, F., Shute, V., Clark, K. M., & Erlebacher, G. (2019). *Interdisciplinary design of game-based learning platforms. A Phenomenological Examination of the Integrative Design of Game, Learning, and Assessment*. Springer.
- Kosko, K., & Ferdig, R. (2016). Effects of a tablet-based mathematics application for pre-school children. *Journal of Computers in Mathematics and Science Teaching*, 35(1), 61–79.
- Krebs, R. J. (2009). Proximal processes as the primary engines of development. *International Journal of Sport Psychology*, 40, 219–227.
- Kucirkova, N. (2017). *Digital personalization in early childhood: Impact on childhood*. Bloomsbury Academic.
- Laurel, B. (2003). *Design research: Methods and perspectives*. MIT Press.
- Levine, M.H. (April, 2018). What does the research say about tech and kids' learning? Part 1 & 2. New York: NY: The Joan Ganz Cooney Center at Sesame Street Workshop. [https://urldefense.com/v3/joanganzcooneycenter.org/blog/F8LM3WSaiTgsDg!2DF0X64sZ9X6ZanNq4mOM2_7xsCq3YXimyxCI56ZfcV6k6jigOICk3IS6lOL191r\\$](https://urldefense.com/v3/joanganzcooneycenter.org/blog/F8LM3WSaiTgsDg!2DF0X64sZ9X6ZanNq4mOM2_7xsCq3YXimyxCI56ZfcV6k6jigOICk3IS6lOL191r$)
- McClure, E. R., Guernsey, L., Clements, D. H., Bales, S. N., Nichols, J., Kendall-Taylor, N., & Levine, M. H. (2017). STEM starts early: Grounding science, technology, engineering, and math education in early childhood. The Joan Ganz Cooney Center at Sesame Workshop. [https://urldefense.com/v3/joanganzcooneycenter.org/publications/F8LM3WSaiTgsDg!2DF0X64sZ9X6ZanNq4mOM2_7xsCq3YXimyxCI56ZfcV6k6jigOICk3IS6vXsvj3z\\$](https://urldefense.com/v3/joanganzcooneycenter.org/publications/F8LM3WSaiTgsDg!2DF0X64sZ9X6ZanNq4mOM2_7xsCq3YXimyxCI56ZfcV6k6jigOICk3IS6vXsvj3z$)
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods source-book* (3rd ed.). SAGE.
- Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). *A Brief Introduction to Evidence-centered Design* (p.37). Princeton, NJ: Educational Testing Service. Retrieved from <https://www.ets.org/Media/Research/pdf/RR-03-16.pdf>
- Moore, H. L., & Adair, J. K. (2015). “I’m just playing iPad”: Comparing prekindergarteners’ and preservice teachers’ social interactions while using tablets for learning. *Journal of Early Childhood Teacher Education*, 36, 362–378. <https://doi.org/10.1080/10901027.2015.1104763>
- Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five-year growth trajectories of kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities*, 42(4), 306–321.
- National Center for Education Statistics (NCES) (2019). *NAEP Data Explorer*. The Nations Report Card. <https://www.nationsreportcard.gov/ndecore/landing>
- National Association for the Education of Young Children (2022). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8*. 4th ed. Washington, DC: NAEYC.
- Nores, M., & Barnett, S. (2014). *Access to high-quality early care and education: Readiness and opportunity gaps in America*. New Brunswick, NJ: Center for Enhancing Early Learning Outcomes and National Institute for Early Education Research.
- Outhwaite, L. A., Gulliford, A., & Pitchford, N. J. (2017). Closing the gap: Efficacy of a tablet intervention to support the development of early mathematical skills in UK primary school children. *Computers & Education*, 108, 43–58. <https://doi.org/10.1016/j.compedu.2017.01.011>
- Outhwaite, L. A., Faulder, M., Gulliford, A., & Pitchford, N. J. (2019). Raising early achievement in math with interactive apps: A randomized control trial. *Journal of Educational Psychology*, 111(2), 284–298. <https://doi.org/10.1037/edu0000286>
- Paciga, K. A., & Donohue, C. (2017). Technology and interactive media for young children: A whole child approach connecting the vision of Fred Rogers with research and practice. Fred Rogers Center for Early Learning and Children's Media at Saint Vincent College. [https://urldefense.com/v3/fredrogerscenter.org/frctecreportF8LM3WSaiTgsDg!2DF0X64sZ9X6ZanNq4mOM2_7xsCq3YXimyxCI56ZfcV6k6jigOICk3IS6oPFPwuk\\$](https://urldefense.com/v3/fredrogerscenter.org/frctecreportF8LM3WSaiTgsDg!2DF0X64sZ9X6ZanNq4mOM2_7xsCq3YXimyxCI56ZfcV6k6jigOICk3IS6oPFPwuk$)
- Pane, J., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2014). Effectiveness of cognitive tutor algebra I at scale. *Educational Evaluation and Policy Analysis*, 36(2), 127–144.
- Pane, J., Steiner, E. D., Baird, M. D., & Hamilton, L. S. (2015). Continued Progress: Promising Evidence on Personalized Learning. RAND Corporation.
- Piaget, J. (1962). *Play, dreams, and imitation in childhood*. *Morton library* (Vol. 24, pp. 316–339). Norton Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology*, 104 (3), 700–712. <https://doi.org/10.1037/a0027268>
- Plass, J. L., & Pawar, S. (2020). Toward a taxonomy of adaptivity for learning. *Journal of Research on Technology in Education*, 52(3), 275–300.

- Plass, J. L., Homer, B., & Kinzer, C. (2015). Foundations of Game-Based Learning. *Educational Psychologist*, 50(4), 258–283. <https://doi.org/10.1080/00461520.2015.1122533>
- Rathbun, A., & West, J. (2004). *From kindergarten through third grade: Children's beginning school experience*. National Center for Education Statistics. Retrieved from <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2--4--7#>
- Reardon, S. F. (2013). The widening income achievement gap. *Educational Leadership*, 70(8), 10–16.
- Rupp, A.A., Gushta, M., Mislevy, R.J., & Shaffer, D.W. (2010). Evidence-centered Design of Epistemic Games: Measurement Principles for Complex Learning Environments. *Journal of Technology, Learning, and Assessment*, 8 (4). Retrieved from <http://www.jtla.org>.
- Rutherford, T., Farkas, G., Duncan, G., Burchinal, M., Kibrick, M., Graham, J., Richland, L., Tran, N. A., Schneider, S. H., Duran, L., & Martinez, M. E. (2014). A randomized trial of an elementary school mathematics software intervention: spatial-temporal (ST) math. *Journal of Research on Educational Effectiveness*, 4, 358–383. <https://doi.org/10.1080/19345747.2013.856978>
- Sarama, J., & Clements, D. H. (2004). Building Blocks for early childhood mathematics. *Early Childhood Research Quarterly*, 19(1), 181–189. <https://doi.org/10.1016/j.ecresq.2004.01.014>
- Sarama, J., & Clements, D. H. (2013). Lessons learned in the implementation of the TRIAD scale-up model: Teaching early mathematics with trajectories and technologies. In T. Halle, A. Metz, & I. Martinez-Beck (Eds.), *Applying implementation science in early childhood programs and systems* (pp. 173–191). Brookes.
- Schenke, K., Rutherford, T., & Farkas, G. (2014). Alignment of game design features and state mathematics standards: Do results reflect intentions? *Computers & Education*, 76, 215–224.
- Scott, F. L. (April 2021). Digital Technology and Play in Early Childhood.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503–524). Information Age Publishers.
- Shute, V. J., & Kim, Y. J. (2014). Formative and stealth assessment. In *Handbook of research on educational communications and technology* (pp. 311–321). Springer, New York, NY.
- Simon, M. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, 26(2), 114–145.
- Slavin, R. E., Lake, C., & Groff, C. (2017). Effective programs in middle and high school mathematics: a best-evidence synthesis. *Review of Educational Research*, 79(2), 839–911. <https://doi.org/10.3102/0034654308330968>
- Slutsky, R., & DeShetler, L. M. (2016). How technology is transforming the ways in which children play. *Early Child Development and Care*, 187(7), 1138–1146.
- Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly*, 19(1), 99–120.
- Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology*, 105(4), 970–987.
- Sullivan, P., Zevenbergen, R., & Mousley, J. (2003). The contexts of mathematics tasks and the context of the classroom: Are we including all students? *Mathematics Education Research Journal*, 15(2), 107–121. <https://doi.org/10.1007/BF03217373>
- Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Thai, K. P., Bang, H. J., & Li, L. (2021). Accelerating early math learning with research-based personalized learning games: A cluster randomized controlled trial. *Journal of Research on Educational Effectiveness*. <https://doi.org/10.1080/19345747.2021.1969710>
- Tudge, J. R. H., Merçon-Vargas, E. A., Liang, Y., & Payir, A. (2017). The importance of Urie Bronfenbrenner's bioecological theory for early childhood education. In L. E. Cohen & S. Waite-Stupiansky (Eds.), *Theories of early childhood education: Developmental, behaviorist, and critical* (pp. 67–79). New York, NY: Routledge.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher*, 43(7), 352–360. <https://doi.org/10.3102/0013189X14553660>
- Watts, T. W., Duncan, G. J., Clements, D. H., & Sarama, J. (2018). What is the long-run impact of learning mathematics during preschool? *Child Development*, 89(2), 539–555. <https://doi.org/10.1111/cdev.12713>
- White, I. R., & Thompson, S. G. (2005). Adjusting for partially missing baseline measurements in randomized trials. *Statistics in Medicine*, 24(7), 993–1007. <https://doi.org/10.1002/sim.1981>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.