10-2022

Assessing The Academic Impact Of Two Adaptive Learning Technology Math Programs Using Hattie's Visible Learning Theory

Charles C. Swan

ASSESSING THE ACADEMIC IMPACT OF TWO ADAPTIVE LEARNING TECHNOLOGY MATH PROGRAMS USING HATTIE'S VISIBLE LEARNING THEORY

By

Charles C. Swan

B.A. History/Honors, Eastern College, 1999
M. Ed. Multicultural Education, Eastern College 2001
M. Ed. School Leadership, University of Phoenix, 2006

A DISSERTATION

Presented to the Affiliated Faculty of
The College of Graduate and Professional Studies
at the University of New England

Submitted in Partial Fulfillment of Requirements
For the Degree of Doctor of Education

It was presented on
09/19/2022
and reviewed by:

Debra Welkley, Ed.D., Lead Advisor
University of New England

Jennifer Scott, Ed.D., Secondary Advisor
University of New England

Brenda Gammon, Ph.D., Affiliate Committee Member
University of Maine

UNIVERSITY OF NEW ENGLAND

College of Graduate and Professional Studies

Doctor of Education Program Final Dissertation Approval Form

This Dissertation was reviewed and approved by:

Lead Advisor Signature: Debra L. Welkley

Lead Advisor (print name): Debra L. Welkley, Ed.D.

Secondary Advisor Signature: Jennifer Scott

Secondary Advisor (print name): Jennifer Scott, Ed.D.

Date: September 19, 2022

Charles C. Swan
September 2022
Educational Leadership

ASSESSING THE ACADEMIC IMPACT OF TWO ADAPTIVE LEARNING TECHNOLOGY MATH PROGRAMS USING HATTIE'S VISIBLE LEARNING THEORY

ABSTRACT

This quantitative study examined the academic impact of two adaptive learning technology math programs (Espark and iXL) alongside the social validity of teachers in a rural western Maine school district who used the program in their classrooms. Using Hattie's Visible Learning Theory (2008) as the theoretical framework, the study tested two different hypotheses. The first hypothesis stated that both Espark and iXL would have an effect size of .40 or greater when the pre and post tests were compared. The study used the Northwest Evaluation Association (NWEA) assessment for its pre and post assessment because it is a nationally normed assessment that was already in use in this school district. Data for the study was taken from 24 grade five students, 29 grade seven students and 17 grade eight students for a total of 70 student data points. Student data was only taken from students who took both the pre and the post assessment, were enrolled in one of the two participating schools, and used either Espark (grade 5) or iXL (grades 7 & 8) for forty-five minutes or more a week. The study confirmed the first hypothesis that both adaptive learning programs met the .40 Hattie yearly growth threshold, with Espark having an effect size of .439 and iXL an effect size of .532. These findings should be accepted with caution given several factors including: low participant numbers, discrepancy

between the effect size calculations and the NWEA expected growth measurements, and the varying design use of the programs in the three different grade levels.

The second hypothesis for this quantitative study stated that there would be a positive mean score of greater than 3 on a study-specific social validity survey given to five participants who teach math in grades 5, 7 and 8 in the site school district. The study asked four math teachers to provide their perception of the adaptive program that they were using in their classroom and rate the program in terms of student motivation, impact on student math standards, impact on the NWEA, ease of implementation, and a number of other factors in the ten-question survey.

The Espark responders had an overall score of 3.32 (average of all 10 questions given by all three responders) and the iXL responders (one seventh/eight grade and one grade 6 math teacher) had an overall mean score of 4.20. These mean scores confirmed the initial hypothesis. However, caution is noted when accepting these results for the following reasons: there were only five participants for this part of the study, the survey was created specifically for the study and therefore has not been tested in other settings, and there was a methodology change that cause the study to have to use seventh and eighth grade responders instead of the grade 6 math teacher only which was part of the original design.

Keywords: Adaptive learning technology programs, social validity, NWEA, Visible Learning Theory.

DEDICATION

This dissertation is the end product of hours of research, writing, studying, revising, and learning. Through this journey there were many individuals who inspired and encouraged me, but there was one who through her life and legacy as an educator and as my mom was not able to see me through to the final steps of this tremendous accomplishment. Mom, Elizabeth Swan Stearns, thank you for inspiring me to be not just an educator, but one completely dedicated to the pursuit of learning as a passion, not just as a career. Thank you for always asking how the journey was going and providing the right level of encouragement and but-kicking to keep me moving in the right direction. I wish you were here to say “well done”, but I know those words are on your lips as you look down on me from above. I dedicate both this dissertation and my entire educational journey to you for inspiring me as a teacher, a mentor, and most of all as a mom.

ACKNOWLEDGEMENTS

Whenever you venture on a journey such as this there are a number who journey with you. I would like to thank the following individuals for their commitment to my success.

Dr. Debra Welkley -You were more than a lead advisor in this process. You provided me with direction, purpose, and the spirit of perseverance.

Dr. Brenda Gammon – I very much appreciate you being a part of my dissertation team and providing regular encouragement. I am grateful to have followed in your footsteps in pursuit of this goal.

Dr. Jennifer Scott – I appreciate your role as an advisor through this journey. Your feedback and participation in my presentations were both supportive and informative.

Dr. Jessica Swan – Thank you for not letting me give up on this goal! Also, thank you for the hours helping me read and revise while also constantly reminding me that the end would come.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

In1983, a report titled *A Nation at Risk,* was given to Secretary Bell, the acting Secretary of Education under President Ronald Reagan, and like previous educational reports it highlighted a number of concerns with the declining quality of education that students were receiving in America's public schools (United States National Commission on Excellence in Education, 1983). The report's notation of the fact that America was losing ground internationally is a theme that has surfaced a number of times since. A 2011 report from Harvard's Educational Policy and Governance echoed the global concern brought forth almost three decades earlier by the Nation at Risk Report. In this Harvard publication, it highlighted that the United States had a 32% proficiency rate in math and just 31% in reading, making it far below a number of other developed countries (Peterson, et. al., 2011). In an even more recent report from 2017 it was again highlighted that the United States was underperforming globally in the area of math, reading and science (Desilver, 2017). Despite the fact that the Nation at Risk report is now nearly forty years old, the United States does not seem to have gained much ground in terms of student academic achievement in math, reading, and science.

There was, however, one component of the 1983 Nation at Risk report that seemed to have left a lasting impression on the American public-school landscape. This report was the first time that there was a call for the need for an investment in technology and technology programming (Bell et al., 1983). With this technology investment, however, came a significant financial barrier that many districts are still wrestling with today, which forces schools to weigh the cost of implementing technology programs with the benefit that they ultimately provide (Coulson, 2006). Despite this, schools have taken on the challenge of budgeting for, purchasing,

and implementing technology aimed at making learning more efficient and differentiated. In fact, spending on technology in schools reached almost two billion dollars in 2014 (Koba, 2015) highlighting the fact that US public schools have taken to heart the technology warnings and suggestions brought forth in the 1983 Nation at Risk report (United States National Commission on Excellence in Education, 1983).

Almost four decades ago, national educational leaders recognized the impact that technology was already having on both the American and world societies and stressed the need for our schools to begin to shift gears to better prepare our students for the technological world they were moving into. These national education leaders at the heart of the Nation at Risk Report (United States National Commission on Excellence in Education, 1983), however, most likely could not have predicted the significance that technology investment would have on student achievement in the wake of the Covid-19 pandemic. This global health battle shut down most K-12 public schools across the United States, causing a major disruption to the traditional educational program available to most students (Office of Civil Rights, 2021). The academic impact of this global pandemic is still being assessed, but early indicators from research suggest that schools will be scrambling to recover from the student academic loss for years to come (Kuhfeld et al., 2020). Despite this significant disruptive wave, schools received billions of dollars in Elementary and Secondary Emergency Relief (ESSER) funds from the federal government (Gordon & Reber, 2020). This funding provided schools with the means to not only cope with initial effects of COVID on education in the United States, but also provided much needed funds for technology hardware and programs to support remote learning efforts that persisted through the 2020-2021 academic school year.

At this intersection of student achievement and the advancement of technology the nucleus of this proposed quantitative study is found. U.S. public schools were already academically struggling prior to the onset of the pandemic in 2019, with 40% of fourth and eighth grade public school students meeting proficiency in expectations in math and only 34% in reading (The Nation's Report Card, n.d.). In addition, U.S. schools were failing to properly implement technology systems that have been deemed significant to impacting student achievement (Stepman, 2018). A 2018 publication titled Report Card on American Education, highlighted the fact that only two states, Florida and Utah, received a grade of an A for digital learning, while nearly 50% or 24 states received a grade below a C (Stepman, 2018). The worldwide pandemic in 2020 did, however, increase the significance of online program evaluation, for it derailed more traditional learning methods in schools across the United States (Office of Civil Rights, 2021). This occurrence forced schools to adopt online digital programs that could continue to support student learning when direct instruction was not readily available.

It is critical for schools to continue to implement policies, programs and practices that have a high impact on student achievement (Hattie, 2018) if they wish to prepare students to be both successful post high school and to compete on the international stage. This research study looked at the impact that two adaptive learning technology math programs had on the academic achievement of students in grades five and six. The study took place in a rural western Maine school district, which has assessment scores in math that mirror the struggling scores found across the United States. In addition to assessing the impact of two different adaptive learning technology math programs, this quantitative study sought to analyze how the social validity (e.g., the satisfaction and acceptability of the online math program used) of teachers who are implementing the math programs align with the results that were gleaned from it. This important

layer allowed this study to explore further connections to how teacher perceptions of how these technology programs aligned with or impacted the results the programs have on student achievement.

Definition of Key Terms

There are a number of terms that are used in connection with this study, and it is essential to have a solid understanding in order to both glean the purpose of the study and its impact on discussions around student achievement and the impact of online digital math programs.

Espark: An adaptive learning technology program that differentiates instruction in both reading and math for students in grades K-5. The program aligns with several other outside programs and assessments like the NWEA. It allows the student to learn at their own level and at their own pace, while providing teachers the option of assigning individual and whole group topics in both math and reading (esparklearning.com/faqs).

ESSER Funds: ESSER, or Elementary and Secondary School Emergency Relief Fund, which allotted 13.2 billion dollars to public schools under the Coronavirus Aid Relief and Economic Security Act (Office of Elementary and Secondary Education, n.d.).

iXL: An adaptive learning technology system that provides personalized online math instruction for grades Pre-K to grade 12. The program includes a comprehensive K-12 curriculum and is both aligned with NWEA and has its own internal placement assessments. iXL can be used individually to support struggling learners or enrich the learning experience of students who need to be challenged (iXL.com, n.d.).

Northwest Evaluation Association (NWEA): NWEA began as a nationally normed assessment that examines the growth of students in math and reading over the course of a school year or multiple years (NWEA.org, n.d.).

Adaptive Learning Technology: A synonymous term with online learning, e-learning, used to refer to digital programs that provide academic instruction for students (Sangra, Vlachopoulos, & Cabrera, 2012). For the sake of this proposed study, Adaptive Learning Technology is defined as "technologies that dynamically adjust to the level or type of course content based on an individual's abilities or skill attainment, in ways that accelerate a learner's performance with both automated and instructor interventions" (Capuano & Caballe, 2020, p. 96).

Social Validity: Defined as, "the satisfaction and acceptability of the interventions and procedures affecting behavior change, based on the opinions of the individuals who receive services and implement them" (Basir & Presberg, 2019, para. 7).

Statement of the Problem

Review of literature highlighted that student achievement in the area of mathematics is a significant problem in the United States (The Nation's Report Card, 2021; Desilver, 2017; Boaler & Zoido, 2016). This problem pre-dates the onset of the Covid-19 pandemic and is one that finds the United States continuing to underperform against other developed countries across the globe. In 2019, only 41% of United States grade four students were considered proficient in mathematics on the country's national assessment (Nation's Report Card, 2019). At the site of this study, in Regional School Unit (RSU)-A, the percentage of students meeting grade level expectations on state assessments was even lower than the national average. On the 2020 Maine state assessment, only 37 % of students in grades 3-5 in RSU-A were considered proficient in mathematics (Maine Department of Education, n.d.). With more than half of the nation's public-school students and nearly 65% in RSU-A failing to meet mathematical proficiency, these reports point to a significant problem and an urgent need to address this problem.

Since the 1983, A Nation at Risk Education Report, technology has been pushed as a means to support the ailing performance of U.S. public school students (Culp, Honey & Mandinach, 2003). Technology has consistently been viewed as a means of improving student engagement in the classroom and therefore positively impacting student achievement (Reese, 2021). More recently, however, technology has become a means of diversifying instruction and support for students (Stanford, Crowe & Flice, 2010). A number of studies addressed the effective use of technology in increasing individual student achievement (Downey, 2008; Fokides, 2018; Mavridis et. al., 2017; O'Rourke et al., 2017; Outhwaite et. al., 2019; Smith, 2020; Yeh et al., 2019). Therefore, this quantitative study assessed the academic impact of two adaptive learning technology math programs on students' mathematical achievement.

This study contributes to the conversation regarding the effectiveness of adaptive learning technology math programs in supporting the math achievement of students in grades five and six. There has been a plethora of studies which looked at individual digital math programs and their impact on student achievement (Donnelly, 2021; Lyons, 2020; SRI Education, 2014). However, Hollands and Pan (2018), was the only research study identified that reported findings on the academic impact of using Espark and iXl among students in a large urban middle school. Yet, there was not any apparent research done regarding how these programs impact student success in a rural community or an examination of the social validity of the technology programs relative to student achievement. This study examined the impact of the same two math programs in a rural district and examined the social validity of the teachers who implemented the programs in their classrooms.

Statement of the Purpose of the Study

The purpose of this study was to examine the academic impact of two adaptive learning technology programs on fifth and six graders while comparing their impact with the social validity of the teachers who are implementing the programs at their school. During the data collection stage of the study, the target population changed to fifth, seventh, and eight graders. Adaptive learning technology has provided a number of benefits in the classroom including differentiated pacing, less student stress, increased student engagement, and increased academic performance (Kurt, 2021). With only one known study that examined the impact of two adaptive learning technology programs in the same setting (Hollands & Pan, 2018) and no known studies that also examined the social validity of teachers using this technology, this study sought to further the knowledge in both of these critical educational areas. The data from this study was used to continue to address America's ill performance in the area of mathematics, while also continuing the conversation about the impact of adaptive learning technology on both students and teachers alike.

Research Questions

This quantitative study sought to address the following research questions and hypotheses:

RQ1: What is the student achievement impact (in terms of NWEA growth scores) of two adaptive learning technology programs used in a rural district in Western Maine?

H1: Both Espark and iXL will have a .40 or higher effect size on student achievement in the area of mathematics as recognized by their growth between a pre and post NWEA assessment data. The .40 effect size is considered to be a

year or more's worth of academic growth in Hattie's Visible Learning Theory (Hattie, 2008).

RQ2: Does the social validity of the teachers who are using Espark and iXL in their classrooms align with the identified student achievement results?

H2: There will be a positive (mean score higher than 3) connection between the social validity of teacher's implementing the adaptive learning technology programs and the student achievement impact on the NWEA.

John Hattie's Visible Learning Theory (2018) sought to both empower educators to implement practices that have a high impact on student achievement (those that have an effect size of .40 or greater) and push educators to be reflective of their practices and the impact that they have on student growth (Hattie, 2008). The former of these two studies guided the first research question of this study, for technology implementation in the classroom is considered a high impact strategy from Hattie's meta-analysis (Hattie, 2018). It was essential, however, to test the effectiveness of technology implementation, and in the case of this study the two adaptive learning technology programs (Espark and iXL), allowed educators to reflect upon their perception of program effectiveness. It is here that Hattie's theory guided the second research question regarding teacher social validity.

This study looked at Northwest Evaluation Association (NWEA) using Espark during the 2021/22 school year. Espark data came from students in grade five at B Elementary School and iXL data was initially anticipated to come from grade six students at C Middle School. It was expected that there would be a minimum of twenty students meeting the time requirements in each of the programs and there would be a balanced mix of students who were at grade level, above grade level, and below grade level on the winter NWEA assessment. The size of the

student population did not allow for student data to be broken up and analyzed between students who were on grade level during the and students who were not. Finally, an anonymous social validity survey was administered to the six participating teachers, to seek their perceptions of the efficacy of the program used by their students.

Conceptual Framework

A conceptual framework rests upon three distinct legs: personal interest, topical research, and a theoretical framework (Ravitch & Riggan, 2017). The personal interest in this study began a decade ago when the leadership team at the researcher's school began to ask the question "what do effective schools do well?" This inquiry first led our leadership team to connect with information published from the Department of Education in the state of Washington, which released a publication providing schools with a list of nine characteristics of effective schools (Shannon & Bylsma, 2007). A review of this report soon bled into the rising work of Professor John Hattie (2018). Hattie also sought to answer the question of what works in education in terms of the academic impact that practices, programs, and procedures have on students (Hattie, 2018). Hattie's large meta-analysis of over 80,000 studies related to impacts on student achievement resulted in a cumulative list of programs and practices that have high, low, and negative impacts on student achievement (Fisher, Frey & Hattie, 2016). This effect size list along with Hattie's comprehensive meta-analysis study provided this researcher's school with a new framework from which to begin to inform decisions in regards to practice, methodology, and programs.

These two studies melded together when the researcher's school (one of the sites of this proposed study) was selected to participate in a two-year literacy/technology program with the Maine Department of Education called MoMEntum. The MoMEntum pilot provided schools

with literacy and technology-based training with the intention of implementing literacy practices that would most impact student achievement. This pilot also introduced the researcher's school to the adaptive learning technology program called Espark (one of the programs that was examined in this study), which was first adopted in the area of literacy in grades Kindergarten to grade 3. The technology coaching received from MoMEntum and the introduction to adaptive learning technology, helped stimulate personal interest associated with this research study.

Topical research revealed four critical components to this study. First, the severity of the decline of math achievement in the United States as evident on the Nation's Report Card, published by the National Assessment of Education Progress (NAEP), which continues to highlight that the majority of America's K-12 students are performing below expected proficiency levels in mathematics (NAEP, n.d.). With American students struggling to meet expectations it continues to force the question as to whether or not American students are able to compete globally both now in terms of academics and in the future in terms of employment and societal advancements (Peterson, et al., 2011). Studies in the area of student achievement highlighted the significant impact that the Covid-19 pandemic has had in disrupting learning (Institute of Educational Sciences, 2021; West & Lake, 2021).

Second, the topical research highlighted a number of studies that focused on the impact of single digital or adaptive learning technology programs on student achievement (Donnelly, 2021; Hinds, 2017; Shectman, et al., 2019;). These studies stressed the fact that technology can have a positive impact on student achievement which was supported by a recent mixed method case study that examined the impact of technology implementation on student achievement (Smith, 2020). Therefore, a positive correlation between the implementation of technology and

the growth of student achievement was well established in the literature, which is an important research foundation for this study.

Third, the topical research surfaced only one relevant study in the past decade, which assessed the impact of two different adaptive learning technology programs head-to-head at the same study site (Hollands & Pan, 2018). This study, conducted by Hollands and Pan (2018), examined the impact of Espark and iXL on student achievement among third through six graders. For this reason, the study by Hollands and Pan provides a springboard for this study, which is seeking to further the discussion initiated by this original study.

Fourth, social validity is not a new concept in academic studies. Recently there have been studies conducted that looked at teacher perceptions about both math and technology implementation (Hayak & Avidov-Ungar, 2020; Walker, 2019). One example of the former looked at teacher perception during the implementation of a Eureka Math program (Walker, 2019). While an example of the latter looked at the impact of seniority on the use and acceptance of technology in the classroom (Hayak & Avidov-Ungar, 2020). A review of the literature did not provide any depth to the understanding of social validity's impact on the implementation of the two adaptive learning technology programs being used in this study.

The major theoretical foundation for this study rested on the Visible Learning Theory (Hattie, 2018), which was selected because it provides the opportunity to compare the academic productivity of two similarly functioning online math programs. The Visible Learning Theory is relatively new and promotes the idea that the more teachers are given the power and knowledge to assess their own practices the more impactful those practices will be on student learning (Tehart, 2011). This theory suggested that everything that school systems do during a school year has a varying degree of positive or negative impact on student achievement. When educators test

the effectiveness or impact of their actions or programs it better informs the distinction of high impact learning strategies and programs (Tehart, 2011).

Hattie's Visible Learning Theory (2018) was woven through a number of components of this study. First, it served as the philosophical cornerstone of the study, which is based on the premise that if RSU-A understands the academic impact of the two adaptive learning technology programs it has adopted it will be better prepared to defend or move on from them in the future. Hattie suggested that one of the critical components of his research was that schools need to both adopt effective methods and remove educational practices that do not have an effect size of .40 or greater (Hattie, 2018). Second, Visible Learning Theory (Hattie, 2018) promoted the idea of using effect size calculations, or Cohen's d, to calculate the impact that programs, practices, and interventions have on student achievement. This study used effect size calculations as the primary statistical measure to compare the academic impact of both Espark and iXL. Finally, this study also examined the social validity of teachers who are using Espark and iXL in their classroom. Part of Hattie's Visible Learning Theory promotes the concept of teachers being reflective about their practices. Exploring the social validity of teachers can help their reflective process while also allowing the study to discuss potential impacts of the teacher's perception on the program's academic impact.

Rationale/Significance

This study is significant in several realms (student achievement realm, technology adoption realm, post-covid education realm etc.). First, the study is significant in a broader perspective, as the review of the literature only brought to the surface a single study which assessed the effectiveness of two competing adaptive learning technology programs in a single district. The Hollands and Pan (2018) study was the only study that surfaced in the literature

review that examined the academic impact of two different adaptive learning technology programs at the same site. Hollands and Pan (2018) tested the impact of the same two adaptive learning programs that were examined in this study, Espark and iXL. This study used Hollands and Pan's research as a platform, however, instead of examining achievement and the cost of the program this study examined student achievement and teacher social validity. This study can help to further the discussion initiated by Hollands and Pan (2018), as well as provide exploration regarding the comparison of the impact of two adaptive learning technology programs at one site.

Second, this study is also relevant because in the wake of Covid-19, the implementation of technology and tech-based programs increased (Bushweller, 2020; De, Pandey & Pal, 2020). The upward shift in technology use has been exacerbated by the recent remote learning opportunities that have emerged amidst the COVID-19 pandemic (Institute of Educational Sciences, 2021). Covid-19 brought to the U.S. public schools an increase in technology competence and use among teachers, while also driving forward the question of whether the newly adopted technology will continue when schools are confronted with more challenging financial times (Bushweller, 2020). Therefore, with the adoption of more technology programming in schools and the reality of a looming economic downturn it is critical for studies such as this, to be conducted in order to assess both their impact and the teacher's perceptions of the impact of the technology that is implemented.

Third, this study is both significant and relevant because it can help inform school leaders in RSU-A regarding the academic impact of two adaptive learning technology programs that they have implemented. RSU-A, like a number of school district's, will be making some tough programming decisions when the wave of Elementary and Secondary School Emergency Relief

Fund (ESSER) money starts to fade. It will be critical for them to be informed about how the programs are impacting student learning and teacher perceptions regarding those programs.

Finally, the study is significant because it assessed the ability and feasibility of the Visible Learning Theory (Hattie, 2018) at the school level, where it can be used to inform decision making around programs and practices. The Visible Learning Theory is a relatively new theory and its statistical backbone of using effective sizes is promoted as being not only beneficial in determining program effectiveness, but also easy to implement (Hattie, Masters & Birch, 2016). In addition to using the Visible Learning Theory as both the theoretical foundation of this study, Hattie's use of effect size calculations to determine the impact of practices, procedures and programs (Hattie, 2008), was adopted in the methodology of the study. This also tested the functionality of this theory as it applies to the impact of adaptive learning technology on student achievement. These four rationales provide a critical argument for why this study is important at both the local and the national level. If schools are going to be able to function and provide quality programming, there needs to be an effective and efficient means of assessing the impact of technology programs, such as adaptive learning technology, on student achievement. This study can aid in that endeavor and contribute to further discussion of the application of the Visible Learning Theory, teacher social validity, and the effective use of technology in the classroom.

Summary

A critical problem in schools in the United States is students underperforming in the area of mathematics (The Nation's Report Card, n.d.). This problem has persisted for nearly forty years (United States National Commission on Excellence in Education, 1983) and schools have tried to solve it in a variety of ways. The use of technology to impact student achievement was

suggested in the Nation at Risk education report in 1983, and since that time schools have spent billions of dollars implementing both technology hardware and software programs (Koba, 2015). With technological advancements like adaptive learning technology programs, which diversify instruction or intervention based on the individual needs of the student, technology continues to have a significant impact on student achievement (Capuano & Caballe, 2020). With this advancement in technology the need to assess its impact on student achievement is critical (Coulson, 2006).

This study intended to examine the academic achievement impacts of two adaptive learning technology programs, Espark and iXL, on fifth and six grade students in a rural Western Maine school district. During the data collection phase of the study the target population pivoted to fifth, seventh, and eight grade students. Using Hattie's Visible Learning Theory (2018) as the theoretical framework, the study sought to address two critical questions. The first research question asked what is the student achievement impact (in terms of NWEA growth scores) of two adaptive learning technology programs used in a rural district in Western Maine? This question was hypothesized by theorizing that both Espark and iXL will have a .40 or higher effect size on student achievement in the area of mathematics as recognized by their growth between a pre and post NWEA assessment data. The .40 effect size is considered to be a year or more's worth of academic growth in Hattie's Visible Learning Theory (Hattie, 2008). The second research question asked if the social validity of the teachers who are using Espark and iXL in their classrooms align with the identified student achievement results? It was hypothesized that there will be a positive (mean score higher than 3) connection between the social validity of teacher's implementing the adaptive learning technology programs and the student achievement impact on the NWEA. This second question allowed for exploration

regarding how teacher perceptions of a technology-based program align with the student achievement results gained from the program.

This quantitative study is both relevant and significant on a number of levels. It is relevant because in the wake of the Covid-19 pandemic there has been an increased focus on less traditional and more technology-based means of instruction. This study can contribute to schools continuing to weigh the value of the technology they are currently adopting. The study is also significant, because only one study within the past decade was found that looked at more than one adaptive learning technology in place at the same site (Hollands & Pan, 2018). Comparing two different adaptive learning technology programs head-to-head can better inform the local district as well as elevate discussions around both programs and their imprint on the students using the adaptive learning programs and the teachers implementing them. Additionally, no known studies have looked at social validity alongside the student achievement of an adaptive technological program. By examining the social validity relative to Espark and iXL, this study can open the door for discussions related to how teacher perception of adaptive learning technology programs connects to student performance. Finally, this study may directly impact the program decision making in RSU-A, the host site of the proposed study, as it finds itself wrestling with low student proficiency scores in mathematics and questioning how to assess the technology-based programs that it has adopted.

CHAPTER 2

LITERATURE REVIEW

Nearly a decade ago a book titled *Inevitable: Mass Customized Learning in the Age of Empowerment* (Schwahn & McGarvey, 2012), created a controversial wave in educational circles across the United States. The book was co-authored by Charles Schwahn and Beatrice McGarvey, two educational researchers, who visioned a newly redefined educational system in the United States, one that was less reliant on the current archaic system designed to sort rather than educate students (Schwahn & McGarvey, 2012). It was not necessarily the concept of mass customized or individualized learning that was at the heart of the wave of backlash towards mass customized learning (MCL), but more a general fear of the loss of educational connection with teachers and an over-reliance on technology (Herald, 2021). The literature around the concept of individualized learning or MCL can be traced back to the 1970's where it was discussed as a general means of providing individual students with both ownership of their educational experience and tailoring it to their individual needs (Baker & Goldberg, 1970; Heathers, 1977). These early concepts of individualized learning blossomed into Schwahn and McGarvey's (2012) promotion of MCL.

The work of Schwahn and McGarvey (2012) helped to resurface the call made in the *1983 A Nation at Risk*, education report, which called for schools to invest in technology and technology programming (Bell et al., 1983). Technology in schools has increased opportunities for communication, knowledge integration, and opportunities for more individualized instruction, but unfortunately schools are still not equipped to use technology in deeper and more meaningful ways (Spector et. al., 2014). Adaptive learning technology has become a means of shifting education from what Schwahn and McGarvey (2012) would call the factory system to

one that moves away from a one-size-fits-all to more of an individualistic approach to learning (Yang, et. al., 2019).

This literature review explores research on student achievement, adaptive learning technology, and teacher's perception, or social validity, of the two adaptive learning technology programs that are currently in use in a small rural district in western Maine. Student achievement was examined in relationship to the subject of mathematics, where it was highlighted that currently only approximately 40% students in the United States are proficient in math in grades 3-8 (The Nation's Report Card, n.d,). This national data coincided with the local data in the site district (RSU-A), where less than 40% of students were considered proficient in the area of mathematics (Maine Department of Education, n.d.). The reality of low performance in the area of mathematics led to research around technology's impact on student achievement, specifically in the area of mathematics. With the call for the implementation of technology and increased technology use in schools by the United States government as far back as 1983 (United States, 1983), there has been almost four decades of research surrounding this topic. The explosion of adaptive learning technology has it set to be a 5.4-billion-dollar industry by the year 2024 (GlobeNewswire, 2021), highlighting what McGarvey and Schwan (2012) coined an "Inevitable" force on the educational landscape. The focus of this literature review is on adaptive learning technology programs that specifically address the impact of this technology on mathematics achievement. Third, social validity, or "the satisfaction and acceptability of the interventions and procedures affecting behavior change, based on the opinions of the individuals who receive services and implement them" (Basir & Presberg, 2019, para. 2), was a critical component of this study and therefore this literature review. Social validity for this study allowed for not only further connections into the emerging literature around teacher perceptions of

technology implementation, but because it seeks reflection on practice (or intervention) it is also very much aligned with the Visible Learning Theory (Hattie, Fisher, & Frey, 2016). At the heart of the concept of social validity is the idea of reflection on implementation (Basir & Presberg, 2019), which is also a critical component highlighted in Hattie's effective size meta-analysis and a cornerstone of his visible learning principles (Hattie, Fisher, & Frey, 2016). These three concepts, student achievement, adaptive learning technology, and teacher social validity) are explored further in this literature review.

Conceptual Framework/Theoretical Framework

A conceptual framework rests upon three distinct legs: personal interest, topical research, and a theoretical framework (Ravitch & Riggan, 2017). Each of these three legs are critical in understanding the background of any study for they provide the lenses from which the study was viewed. Each of these three concepts and their connection to this study are examined and explained further in this section.

Personal Interest

The personal interest in this study begins with the desire to see a local school district make decisions that support the academic needs of its students in the area of mathematics. In addition, the results from this study can allow the district to make informed decisions about programs that they both seek to invest in and/or cut. In January of 2020, the district that was the focus of this study, finally obtained local taxpayer approval for their 2019/20 budget six months and five referendum votes after their initial proposal in early June of 2019. The passing of the 2019/2020 budget was met with both relief and concern as the district continued to weigh the balance between the economic strain on the local communities and the need to provide a rich educational experience that allowed for both student growth and student opportunities post-high

school. This intersection between programming and cost is often a tightrope act that pits small rural communities against their schools, which are trying to improve both their educational product as well as the overall prosperity of their students and their families.

Wedged in the middle of this battle between cost input and educational output is the ever-advancing world of technology. Technology has transformed the landscape of America's educational system over the past thirty decades as it has provided schools with improved informational systems, better access to the internet and digital content, increased social networking and gaming, and increased tools for student engagement (Sheninger, 2019). Technology has also greatly impacted the way in which educators teach (Costley, 2014), which has been vividly on display during the 2019-2020 pandemic. Technology has allowed schools across the globe to adapt to a global health crisis while still providing millions of students with the opportunity to access quality education and social connections.

Technology, however, has significantly contributed to the budgetary strain that many rural schools face when trying to pass their budgets (Coulson, 2006). Schools have had to weigh the specific cost of implementing technology with the value of the outcomes it produces. Unfortunately, schools have not been very successful in analyzing this input versus output data (Hollands & Pan, 2018). As a result, millions of dollars have been spent on technology hardware, software, and online digital programs and for many school districts the investment has done more to exacerbate their financial strain than it has solved their lagging educational performances. One of the silver linings of the 2019-2021 pandemic for schools has been the millions of dollars the federal government has poured into schools to provide assistance in educating students in non-traditional ways (Lieberman & Ujifusa, 2021). Some schools have used a large portion of this money to provide teachers and students with access to technology to

facilitate learning opportunities (Liberman & Ujifusa, 2021). This new investment has proven in many cases to be critical, but the long-term impact will most likely bring schools, particularly rural poverty-stricken schools, back to the table of economic choice. On the one hand they can choose to continue to invest in technology and consequently make dramatic cuts to other portions of their budget. Or on the other hand, they can continue to invest in technology while increasing the economic strain that this investment places on their local communities.

The argument at the nucleus of this research study, however, contended that there may be a more viable third option, one that rests in a better analysis and evaluation of technology based academic programs. Evaluating the impact of programs and technology, however, is often labor intensive and therefore schools are not successful in moving this critical piece forward (Cavanagh, 2015). This research study, therefore, not only highlighted the personal interest that the researcher has in technology investment and student achievement, but it also aimed to support him and others in making rational decisions on what to spend precious taxpayer dollars on in a rural American school district.

Theoretical Framework

The Visible Learning Theory was born out of a meta-analysis of over fifty thousand empirical studies by Professor John Hattie, and like all learning theories it seeks to explain what causes students to learn (Terhart, 2011). Hattie's meta-analysis theorizes that when teachers reflect on their practice and essentially take on the role of the student, and when the student reflects on their learning and essentially takes on the role of the teacher, visible learning take places (Hattie, 2009). Hattie's theory pushes educators to examine their practices in reference to a set of one hundred and thirty-eight variables which have been determined to have varying degrees of impact on student learning (Terhart, 2011). The main premise is that the higher

impactful strategies an educator adopts the more academically successful their students will be. Hattie's Visible Learning Theory provides a foundation for examining choice in program adoption in the current study. Therefore, Visible Learning Theory was selected for the theoretical framework of this study for three main reasons:

1. It pushes the current study to effectively evaluate the impact that the two online math programs have on student learning. This study is about evaluating online programs, but the overall genuine purpose in doing so is to increase student achievement. Employing strategies or programs that have the greatest impact on student achievement is at the heart of Hattie's theory.
2. The theory provides a concrete means of assessing the impact that variables have on student learning. Hattie's theory rests on the premise that simple effect size calculations can be used to determine the impact of programs on student achievement. Effect sizes are not only simple to calculate, but also easier to understand by the layperson.
3. Visible Learning Theory is grounded in a meta-analysis that examined thousands of previous studies related to the concept of what works in education. This makes the theory well-grounded and easily applicable to several studies which seek to increase student achievement by properly assessing and implementing programs and pedagogy that have the highest impact on it (Hattie, 2009).

These three factors help to illuminate why the Visible Learning Theory was selected as the theoretical foundation of this study and how the work of Hattie (2018) helps to support the methodology selected for it.

Strengths and Weaknesses of the Theoretical Frameworks

As previously highlighted, there are several key strengths associated with the Visible Learning Theory in relationship to this study. First the theory is extremely adaptable, especially when analyzing factors promoting growth in student achievement (Hattie, 2009). A second strength is that it is grounded in decades of research that span tens of thousands of studies related to the idea of student achievement (Hattie, 2018). Therefore, the literature reviewed demonstrates a connection to a long line of studies that have sought to better understand what is effective in improving student academic achievement (Terhart, 2011). Another strength is that the theory is easily applied at the research and the implementation level (Hattie, 2009). The theory rests on the idea that the more educators understand the impact of their practices the better equipped they will be to develop and use strategies that are proven to create substantial growth.

Thomas Romer (2018) highlights five areas of concern with Hattie's Visible Learning Theory, the first of these five is the most critical in terms of a weakness pertaining to this study. Romer (2018) suggests that the theory is not one that should be aligned with education practice as it is more a theory of evaluation. This key criticism for Romer (2018) is based on the suggestion that Hattie's work seems to exclude volumes of knowledge and research on sound educational practices and theories. (Romer, 2018).

Literature Review

A literature review is a systematic approach to researching, highlighting, and discussing relevant information to a study and its research questions (Bloomberg & Volpe, 2008). Even more specific, "a key objective of a literature review is to provide a clear and balanced picture of current leading concepts, theories and data, relevant to your topic or subject of study" (Bloomberg & Volpe, 2008, p. 46). In this section information on student achievement, adaptive

learning technology and teacher social validity will all be discussed with relevant literature materials found in these areas.

Student Achievement

The overall goal of this research study was to examine two factors that potentially impact student achievement; adaptive learning technology and teacher social validity. In so doing, it is hoped that this study will add to the on-going conversation about how to schools can make decisions to positively impact their student learning rates. This task, however, is daunting for many institutions to say the least, because as Hattie (2018) highlighted there are hundreds of factors that both positively and negatively contribute to this realm of student achievement. Student engagement has clearly been established as an important factor in the ability of students to grow academically (Lei et al., 2018; Orosco, 2016). Student engagement, however, is a tremendously broad topic that encompasses everything from student emotional well-being to student boredom. The myriad of factors that influence achievement are found at the intersection of students' thoughts, feelings, and behaviors (Abid & Akhtar, 2020). Student engagement can be seen through the lens of boredom, which is the product of disengagement and therefore should be a critical factor that educators seek to address when pushing for academic achievement (Tze et al., 2016). One of the critical roles of the modern educator is to find strategies and programs that increase student motivation so that they can consequently increase the likelihood that a student will be more successful academically (Johnson, 2017). Since the arrival of technology into the classroom nearly four decades ago, schools around the world have relied upon it as a means to make the process of learning easier and more efficient, and to increase student motivation (Costley, 2014).

Technology and Student Engagement

The role of technology in addressing academic disengagement is also quite expansive and voluminous (Bond et al., 2020; D'Angelo, 2018; Morris, 2014; Power, 2017). There are, however, two critical components that need to be addressed in terms of examining the impact of digital programs on students' achievement. The first is the impact of devices on student learning and the second is the impact of online programs. These two factors are obviously inseparable, as one cannot deliver a technology-based program without a device. Examining them separately, however, lends credit to the idea that both components impact student engagement, and therefore achievement. The use of technological devices, in the form of personal computers, really began to expand in the 1980's and grew significantly from the 1990's into the new millennium (Lumen, n.d.). During this time the percentage of schools with computers jumped from 18 to 98 percent (Cuban, 1992), however, they were still limited to computer labs and isolated classrooms.

Twenty years ago, this initial trend in the use of computer devices was dramatically changed when state governors, like Maine's Angus King, led the way in supporting one-to-one devices (Maine.gov, n.d.). In the state of Maine, this initiative became known as the MLTI or Maine Learning Technology Initiative (Muir, Knezek, & Christensen, 2004). According to the Maine Department of Education (DOE), over the past two decades the program has changed, seeking to provide Maine students with unique opportunities to learn through technology-rich experiences (Maine Department of Education, n.d.). This shift in the use of technology in the classroom led to the current discussion regarding one-to-one devices and their current impact on student achievement. There are a number of studies which highlighted the positive impact that devices have on student learning (Harris et al., 2016; Hilton, 2018; Larking & Jorgensen, 2016). These same studies also highlighted that the most significant impact of one-to-one devices is

their impact on student motivation or engagement. Other studies, however, have suggested that technology and devices do not play a critical role in student success (Kul et al., 2018; Nelson et al., 2016). These studies, however, are significantly outweighed by those that take the opposing opinion. Studies that minimize the impact of technology tend to highlight other factors, such as teacher knowledge, (Kul et al., 2018), school climate (Maxwell et al., 2017), finances (Learning Policy Institute, 2018), and even classroom management (Herman, et al., 2020) that also have significant impacts on student achievement.

Origins of Adaptive Learning Programs

The exact origins of adaptive learning technology can be somewhat hard to trace. Some would contend that the use of television in providing educational programming, such as Sesame Street which started in 1969 (History.com, n.d.), was a foundation for our current educationally based digital programs (Jordan & Romer, 2014). Others suggested that the rise in the use of computers in the 1980's really set the stage for the development of digital programming in schools (Christensen, 2019). Research around differentiated instruction not only helped to transform American educational pedagogy, but how technology began to be used to support this change (Stern, 2015). Changes in curriculum demands through the Common Core State Standards, also created a surge in the use of academically based digital programs as the U.S. made technology an integrated part of learning expectations (Costley, 2014). A combination of the need for programs to diversify material and instruction to meet the individual needs of students, with the advances in digital technology, set the stage for an explosion of this type of learning technology through the start of the new millennium to our current use (Oremus, 2015).

Adaptive Learning Programs Today

At the onset of the use of one-to-one devices, the primary function of the devices was to make learning more efficient and to help connect students to more diversified information through the web. Over time, however, as more and more programs and education applications developed, the key to engagement through one-to-one technology came through the use of online programs and games (Ganimian et al., 2020). The research was very supportive of the generally accepted idea that digital programs do positively impact student learning in general and more specifically in the area of mathematics (Cozad & Riccomini, 2016; Rourke et al., 2017; Watson-Huggins & Trotman, 2019).

In addition to the general positive effect of online programs on student learning, other studies have illuminated the reality that these programs can address the need for differentiation in the classroom as they can support learning for struggling students and high achieving students alike (Cozad & Riccomini, 2016; Jimenez & Besaw, 2020; Stetted, 2018). There have been several studies on specific math programs and their impact on student learning. Some programs such as LEGO, ESPARK, iXL and ViLLE proved to have moderate to significant impacts on student math gains (Altakhayneh, 2020; Hollands & Pan, 2018; Kurvinen et al., 2020). A study conducted in the Midwest showed that the whole scale adoption of programs like ORTIGO, a digital online math program, did not provide expected gains (Corcoran, 2018). As one would expect this highlights the idea that programs vary significantly in their ability to impact student achievement. In a study conducted in Great Britain of three and four-year-old learners (Outhwaite et al., 2019), it was found that math applications varied greatly in the depth of their content and the structure of their learning. This variance in online applications led the

researchers to conclude that not all programs are equal in their ability to support the development of mathematical concepts in young students (Outhwaite et al., 2019).

More interestingly and pertinent to the conversation of student achievement are studies that have highlighted the depth and longevity of learning that some programs help students to achieve.-In addressing the former, depth of knowledge, a recent study indicated that online programs addressing math related skills were much more likely to access Bloom's higher-level thinking than were programs that targeted science and language art skills (Crompton et al., 2019). The ability to access higher-level learning through online math programs also landed credit to the latter idea, that adaptive learning technology math programs can have a longer lasting impact on students' development in the area of mathematics (VanderArk & Schneider, 2012).- Some of the research (Berrett & Carter, 2018; Larkin & Jorgensen, 2016), in this area suggested that one of the greatest impacts of digital programs, both in general and in the area of math, was on student's attitudes and desire to continue to persist in their learning. The Office of Educational Technology (2017) outlined several reasons why the conversations in schools shifted from whether technology can improve student achievement to how and what types of technologies are best suited to do so. In a publication entitled *Reimagining the Role of Technology in Education: 2017 Educational Plan Update* (2017), this office suggested that personalized learning through online learning programs and games increases student engagement and therefore significantly impacts student achievement.

The research on various digital programs was far more extensive than the research that arose in this literature review around assessing digital programs. The general trend in education has been to rely on websites such as What Works Clearinghouse (WWC) or the Evidence for ESSA, to vet programs that schools are looking to adopt (Gordon, 2018). The downside of

exclusively using these pre-assessed lists is twofold. First it eliminates the opportunity to assess the impact directly on a school's population and second the lists are not all encompassing, which sometimes eliminates potentially impactful programs from being considered and adopted (Gordon, 2018).

The review of the literature identified a number of studies focused on the Continuous Improvement Model (Best & Dunlap, 2014; Blanton & Harmon, 2005; Shakman et. al., 2017; Tichnor-Wagner et. al., 2018), but this practice was more generalized and usually used to identify problems and then plan a course for the resolution of the problem. There have been a number of studies that have examined the academic impact of individual programs in school settings. Programs such as iXL (Donnelly, 2021), Think Through Math (Hinds, 2017), and Reasoning Minds (Shectman, et al., 2019), are a few examples of studies which have examined the impact of these individual programs on student achievement outcomes. There was only one study which looked specifically at comparing the impact of two different online math programs and that was conducted by Hollands and Pan (2018). Hollands and Pan (2018). examined the impact of Espark and iXL on elementary and middle school students in a relatively large urban district. In a review of the study (Hollands & Pan 2018) there was a major methodological flaw in that they used two different assessments to analyze the impact of the two different programs. This study, however, is important because it provided a starting point for measuring the impact of two programs that are used simultaneously in a single district, which was the intent of this research study.

Teacher Social Validity

A critical component of this study was to examine teacher perceptions of the adaptive learning technology's impact on student achievement and its alignment with the student

achievement data that will be analyzed. There are numerous definitions and synonyms for the concept of social validity. One that seems to have encapsulated the current use and understanding states that, social validity is “the satisfaction and acceptability of the interventions and procedures affecting behavior change, based on the opinions of the individuals who receive services and implement them” (Basir & Presberg, 2019, para. 2). Social validity in the context of this study aligned with this definition, as it focused on how the perceptions of teachers using adaptive learning technology may impact the results of the programs on student achievement.

Examining teachers' social validity for school-based interventions was certainly not a new concept. A 2016 study (Vancel, et. al., 2016) examined the social validity of elementary, middle and high school teachers in districts implementing school-wide Positive Behavior Intervention Supports (PBIS), and how implementation of PBIS as an intervention aligned with teacher’s perception of it. Another recent social validity study (Dore et al., 2021) examined teacher and care-givers perception of virtual Pre-K as a pandemic intervention and its ability to prepare students for kindergarten. Social validity has also been used to assess the impact of technology interventions in a school setting, in one such study, teacher and student social validity was used to assess the perceptive effectiveness of using iTouch flashcard program with special education students (Jameson et al., 2012). When you merge the use of social validity in these three studies it illuminates the fact that this system of assessing teacher perception is grounded in educational research and even more specifically in technology-based educational research.

Despite the use of social validity to assess intervention impacts in schools and even in the realm of technology, the review of the literature did not locate studies which included social validity in an assessment of adaptive learning technology. One study conducted in 2018 by Smith examined future teacher perceptions of adaptive learning technology in k-8 mathematics

education. In this study, 17 college students enrolled in education programs were interviewed about their perceptions of adaptive learning programs, and they positively responded that they felt this technology would be a positive factor in their classrooms (Smith, 2018). Outside of this study on teacher perceptions of adaptive learning technology, and the previously mentioned use of social validity in classroom settings, review of the literature did not reveal studies that assessed social validity in terms of the impact of adaptive learning technology on student achievement. This points to the relevance of this aspect of this quantitative study's design and again further weaves the Visible Learning Theory into both the design and implementation of the study.

Conclusion

There are three primary factors that intersected in this study: student achievement, the impact of adaptive learning technology, and the social validity of the teachers implementing the technology. The student achievement factor for this study rests in the fact that students in RSU-A, like students across the United States, are struggling in the area of mathematics (The Nation's Report Card, n.d.). This academic struggle has persisted for four decades when The Nation at Risk report sounded the alarm on America's struggling academics (United States National Commission on Excellence in Education, 1983). The adaptive learning technology aspect of this study involved the use of two different adaptive learning technology programs currently in use in the site districts, Espark and iXL. Adaptive learning technology has become an important factor in schools both prior to and even more so during the pandemic to support the diverse needs of students in classrooms (Lempinen, 2020).

The final component of this study was the social validity of the teachers who used the adaptive learning technology programs in their classroom. Social validity in this study was

important for two main reasons. First, as Runco (2011) highlighted in his realm of therapy, interventions are often validated but that validation is sometimes only one sided. Using social validity allowed a study to see the perceived impact of the intervention (in this case the two adaptive learning technology programs) from the perspective of the users (in this case the teachers). Secondly, the review of the literature did not find any current studies that included social validity assessments alongside an analysis of the effectiveness of adaptive learning programs. These three components helped to build the framework for this research study.

The Visible Learning Theory, which empowers teachers to reflect upon their methodology to adopt practices that are most impactful to student achievement (Hattie, 2009), was selected as the theoretical foundation for this study because of its direct link to student achievement. The study's primary research question sought to answer whether or not two adaptive learning technology programs implemented at the proposed organization are having a positive impact on the mathematics achievement. The personal interest in this topic is rooted in both the general concept of student achievement and more specifically how technology can be used to positively impact student achievement in the area of mathematics. This personal interest, combined with both the theoretical framework of the Visible Learning Theory and the literature review related to student achievement, technology's impact on student achievement, and teacher social validity formed the three legs of the study's conceptual framework.

Chapter 3 explains the research methodology used in this study. This includes demographic information about the site location for the study. In addition, Chapter 3 highlights the types of data that was collected and how that data was analyzed in the study.

CHAPTER 3

METHODOLOGY

For nearly four decades, since the 1983 publication of the Nation at Risk report on education, United States public schools have attempted to remedy an antiquated education system without much success (Rothwell, 2016). The 2019 National Assessment of Education Progress (NAEP) National report card showed that just 41% of grade four students, 34% of grade eight students, and only 24% of twelfth graders were proficient in the area of mathematics (The Nation's Report Card, n.d.). This performance not only highlights the on-going struggle of our public school system to prepare students for both higher education and even life, but it also illuminates the reality that we are not closing the educational gap with other developed countries (Desilver, 2017). One of the critical components highlighted in the 1983 Nation at Risk report was the need for US schools to implement technology to support future student aspirations and societal needs (United States National Commission on Excellence in Education, 1983). U.S. public schools have overall taken on this challenge, spending billions on technology implementation each year (Epstein, 2021). The need for technology in US schools was even further exacerbated by the onset of global COVID-19 pandemic, which forced many schools to move from more traditional forms of educating students to designs that were more technologically based (Zhao & Watterston, 2021). With federal Elementary and Secondary School Emergency Relief (ESSER) funds available, US schools' spending on technology and technology programming became even more significant (Herald, 2021).

Despite being proficient at spending money on technology, schools have largely failed, at the local level, to assess the academic achievement impact of technology programs, systems, and hardware that they have adopted (Newcomb, 2020). It is here at the intersection of technology

programming and student achievement that the nucleus of this research study is found. The core problem addressed in this study is that US students are failing to meet academic expectations in the area of math (The Nation's Report Card, n.d.). To address this problem many schools have adopted adaptive learning technology programs, which are designed to differentiate academic instruction to meet the individual needs of students using the programs (Natriello, 2017). Two such adaptive learning technology programs, Espark and iXL, were implemented in Regional School Unit-A and this study assessed both their impact on student Northwest Education Assessment (NWEA) assessment scores as well as the social validity of the teachers who are using the programs in their classrooms.

This quantitative study planned to access NWEA data collected during the 2021/22 school year to compare the student growth of fifth and sixth grade students using two different math adaptive learning technology programs (Espark and iXL). However, during data collection pivoted to data from fifth, seventh, and eight grade students. In addition, an anonymous survey was given to the classroom teachers to examine their social validity allowing for an examination of how teacher perceptions of the program align with the program's academic impact. The study sought to address the following critical research questions and hypotheses:

RQ1: What is the student achievement impact (in terms of NWEA growth scores) of two adaptive learning technology programs used in a rural district in Western Maine?

H1: Both Espark and iXL will have a .40 or higher effect size on student achievement in the area of mathematics as recognized by their growth between a pre and post NWEA assessment data. The .40 effect size is considered to be a year or more's worth of academic growth in Hattie's Visible Learning Theory (Hattie, 2008).

RQ2: Does the social validity of the teachers who are using Espark and iXL in their classrooms align with the identified student achievement results?

H2: There will be a positive (mean score higher than 3) connection between the social validity of teacher's implementing the adaptive learning technology programs and the student achievement impact on the NWEA.

These research questions and supporting hypotheses were the driving force for the methodology selected for this study. Hypothesis 1 theorized that both adaptive learning technology programs will have a high impact on student achievement as recognized by Hattie's meta-analysis research on this topic (Hattie, 2008). Having the hypothesis directly linked to Hattie's research not only helped to connect the study to a larger context around student achievement, but it also tied it directly to the Visible Learning Theory which was the theoretical foundation for the study. This chapter provides information on the following components of the study: site information and demographic/setting, participants/sampling/methods, instrumentation and data collection, data analysis, limitations/delimitations and ethical issues, and trustworthiness.

Site Information and Demographics/Setting

The site for this study was Regional School Unit-A, which is a rural school district of approximately 750 students snuggled in the valley of the Androscoggin River in Western Maine. The district serves four rural towns that have a combined population of approximately 5500 people (United States Census Bureau, n.d.). As part of Maine's school regionalization movement, RSU-A merged with a larger school district for six years and recently in 2018 withdrew from the larger Regional School Unit (RSU) through a referendum vote of the citizens in the four towns. The goal of the citizens in withdrawing was to regain both programming and

financial control of their local schools. The district is approximately 98% Caucasian and over 60% of students who attend receive free or reduced lunch. RSUA has a special education population of approximately 19%. RSU-A has three schools: B Elementary School is the newest school (11 years old) and the largest with 355 students in grades Pre-K to grade five. C Middle School has approximately 175 students in grades 6-8 and was built in the 1980s. D High School currently has a population of around 220 students in grades 9-12 and first opened in the 1950s.

RSU-A, like many school districts in Maine, has struggled to attain student proficiency in both reading and mathematics for more than a decade. Table 1 highlights the percentage of students in each grade who have met proficiency in mathematics on the Northwest Education Assessment (NWEA) over the past ten years. Table 1 displays the fact that despite certain pockets of relative success in the area of mathematics, most grade levels K-8 have struggled to see more than 50% of students meet expected proficiency.

Table 1

RSU A NWEA Math Proficiency 2010-2021

NWEA Math, SPRING								CCSS NWEA			
	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21
Grade K	-	-	-	-	-	-	-	84	72	NO DATA	54*
Grade 1	-	-	-	-	-	-	-	72	60	NO DATA	48*
Grade 2	-	30	21	-	15	24	-	51	73	NO DATA	55*
Grade 3	29	37	42	29	35	30	34	51	42	NO DATA	41
Grade 4	32	25	21	36	32	19	21	33	46	NO DATA	43
Grade 5	28	29	21	25	21	17	17	21	38	NO DATA	26
Grade 6	39	31	33	31	22	18	27	26	12	NO DATA	35
Grade 7	44	40	32	40	27	26	20	33	28	NO DATA	25
Grade 8	61	51	51	30	48	29	42	45	50	NO DATA	34

To support positive change in math achievement, RSU–A has taken a number of proactive steps to support both staff and students. First, and the primary focus of this study, was the adoption of two adaptive learning technology programs, Espark in grades K-5 and iXL in grades 6-8. These two adaptive learning programs have been supported financially by the district

to help teachers provide better differentiated instruction to their students. Second, the district has adopted two new math curriculums to use as for primary instructional purposes while using the adaptive learning programs as a supplemental support for students. The first curriculum program, called Number Worlds, was adopted by the special education department. Number Worlds is designed to assess and address student deficiencies, especially with students who have known learning difficulties (McGraw Hill, n.d.). Second, in 2021, the entire district adopted math programs made by the same publisher, McGraw Hill, to help better align the vocabulary and scope and sequence of math instruction in the district. It is relevant to note that both the elementary and middle school, the sites for this study, are in their first year of adoption and therefore it was expected that this would help to eliminate some of the additional factors that can influence student achievement. Finally in 2021, the curriculum coordinator for the district secured the support of a local university education professor, who teaches math courses for education college students, to work with the staff at the elementary school. The goal of this outside resource was to:

1. Help provide teachers with tools and confidence to both teach math and help their students understand and excel at it.
2. Help teachers use both their core curriculums and their adaptive learning technology programs to differentiate instruction and support student needs.
3. Provide general coaching and feedback for teachers in the elementary school in terms of their instructional practices in the area of mathematics

Also relevant to this study is the fact that during the 2017-2018 school year the targeted district accomplished one of its technology-based goals when it was able to place a computing device in each K-12 students' hands. During the 2020-2021 school year this was expanded to

include the 40 pre-K students that attend B Elementary School. While the district was achieving this technology goal, the citizens of the district were sending the school board and district administrators a clear message regarding the district's budget. Previously the district enjoyed widespread support for its programming and its annual budget often passed with little concern or controversy. In 2019-2020, however, it took five rounds of referendum voting for RSU-A to pass its annual budget. This lengthy process highlighted the communities' realization of one of their withdrawal goals, which was regaining control of their school's expenditures and communicating to school leaders that financial and programming changes were necessary.

This merger of the community calling for lower school budgets and the failure of the district to effectively evaluate programs was slightly interrupted by the onset of the COVID 19 pandemic during the 2019-2020 and 2020-2021 school years. The global pandemic helped to reinforce the need for technology-based programs and hardware and fortunately financial aid through federal relief funds assisted the district to continue on its current course while maintaining a palatable budget for the communities. In reality, however, the COVID relief funding only delayed the district's necessary consideration of the programming funding and its impact on student achievement. One of the results of the significant disruption to school and students' ability to access continuous instruction between March of 2020 and June of 2021, the district also experienced significant impacts to its student achievement data. These impacts could have potentially affected the outcomes and intended focus of this study as it sought to examine the student achievement impact of two adaptive learning technology programs (Espark Learning and iXL).

The site for this study was selected for two main reasons. The first is accessibility; the researcher was an administrator in the school district. This allowed easier access to both data and

participants. It is understood that this also created a certain level of bias, especially considering one of the adaptive learning technology programs is already in place at the elementary school and the other is currently being used at the middle school. The desire, however, to provide the district with accurate information regarding the effectiveness of programming would help to keep these potential biases in check. The second reason the site was selected is because RSU-A currently has both Espark Learning and iXL in place (Espark at the elementary level and iXL at the middle school level), which helped to eliminate the need for implementation training. RSU-A also uses a common assessment, the NWEA, for students in grades K-10. These factors allowed for a solid discussion about the impact of the programs on students' achievement in math and further the discussion initiated by the Hollands and Pan (2018) study, which examined these same two adaptive learning technology programs. This study diverged from the Hollands and Pan study in three areas. First, one assessment was used to determine academic impact instead of two different assessments. Second, it was done in a much smaller setting. Third, this study also looked at the social validity of the teachers implementing the adaptive learning programs, something not done in the previous Hollands and Pan (2018) study.

Participants and Sampling Methods

There were two sets of data collected and analyzed for this quantitative study. First, the Northwest Education Assessment (NWEA) math data based on the annually scheduled RSU-A NWEA assessment for students in grades five and six was to be collected and analyzed. No students actively participated in this study, only their assessment data was used to assess the academic impact of Espark and iXL, and the data was deidentified to further secure any students' identity. The structure of math instruction at the elementary and middle school levels are different, which impacted this study. At the elementary school there are three self-contained

grade five classrooms. This means that the teachers in grade five teach all core subject areas (math, reading, writing, social studies and science) and students do not move from teacher to teacher based on subject matter. Subjects in the sixth grade are taught by different teachers. This means that grade six students have math with one teacher and reading or science with another teacher. This was believed noteworthy to this study for two reasons. First, the instruction provided by teachers in grade six might be more specialized in that it is provided by someone who just provides instruction in one or two areas versus the five or six provided by teachers at the elementary level. Second, it leaves just one teacher available at the middle school level in grade six to participate in the social validity part of this study.

The second set of data was to come from an anonymous social validity survey given to the three grade five teachers and one grade six teacher who have implemented the programs into their classrooms. It was understood that the one sixth grade math teacher's responses to the survey would not be anonymous given that they were the only person responding. However, the survey data from all four teachers was maintained confidentially. To address potential bias, direct permission from the teacher was secured in advance of executing this research study. In addition, the researcher was not the school administrator at the middle school, which helped to eliminate any pressure for the one respondent at that school to answer questions in a certain way.

The ten-question social validity survey was adapted from the Intervention Rating Profile (IRP) (Carter & Wheeler, 2019), which used a six-point Likert Scale, and asks participants to rate the success of the intervention (in this case the use of the adaptive learning technology) in the school or classroom setting (Carter & Wheeler, 2019). The IRP provided new avenues for clinical assessments of the impact of interventions, once primarily used in school settings (Carter & Wheeler, 2019). The survey used for this study, aligned with the IRP, but uses a five-point

Likert Scale, and was modified from twenty questions to ten to better align with the focus on adaptive learning technology's impact on student achievement (see Appendix A). The ten questions that were asked in the survey was administered confidentially using the REDCap system. The survey data was used to ascertain the social validity of the teachers in connection to the student achievement data.

The setting for this quantitative study was to be grade five at B Elementary School and grade six at C Middle School in RSU-A. These grades were selected because B Elementary School used Espark as an adaptive learning technology math program for students in grade five and C Middle School used iXL as an adaptive learning technology math program for grade six students. The two sites for this study were also in the first year of implementing the same new comprehensive math program. This allowed for cleaner data in connection to the program as all teachers had the same curriculum (though at different grade levels). The teachers all having the same experience level with the curriculum helped to minimize some of, but not all, of the influence of the teachers themselves.

Instrumentation and Data Collection

This study employed quantitative measures for collecting, analyzing, and discussing the data that was produced. This approach allowed this study to be an extension of the conversation initiated by Hollands and Pan (2018), but also provided a precise means of discussing data associated with the impact on student achievement. The decision to use effect size or Cohen's d for data analysis was made for three main reasons. First, this statistical measure is relatively easy to understand (Coe, 2002) and this study sought to inform school personnel on how to assess program effectiveness. Second, given the size of the pool of participants Cohen's d is a strong

choice for analyzing and comparing the impact of both adaptive learning technology programs (Schafer & Schwarz, 2019).

There were three grade five self-contained classrooms at B Elementary School and three grade six math classes at C Middle School (though just one math teacher). The steps involved in calculating, collecting, and analyzing the data from this study were to start with students in grades five and six receiving a minimum of ten weeks of math support from one of the two adaptive learning technology math programs. This period of time coincided with one of the predetermined NWEA testing windows which occurs three times during the school year. Data from the fall testing was to be compared with data from the spring testing. Once the post NWEA session was completed the data would be organized into two main categories, Espark and iXL. Within these groups, if numbers allowed, two subgroups, those performing at or above the fiftieth percentile and those below it from the pre-test (winter NWEA), would be organized to provide discussion on how the adaptive learning technology programs impacted assessment scores of students who were below or at and above the fiftieth percentile (what NWEA uses as a national mean score) (nwea.org, n.d.). This comparison was to allow for a discussion around whether one of the two adaptive learning technology programs had a larger academic impact on one of these two subgroups.

Once data was sorted into categories, an effect size or Cohen's d calculation was used to determine the effect size impact of each program on student achievement. Effect size was calculated by comparing the pre and post-test scores using the effect size calculation or Cohen's d, where the pre-test mean of each group was subtracted from the post-test mean and divided by the standard deviation (Ferguson, 2009). Cohen's d for both the Espark group and the iXL group was charted. In addition, data connected to students in the two subgroups (those at or above

grade level and those below) was charted to support analysis and discussion on the impact both programs had on the whole and on the two subgroups. Again, students were not actively involved in this study and any assessment data collected was deidentified.

The second set of quantitative data that was collected in this study came from a created ten-question social validity survey that participating teachers took. This survey included on a five-point Likert scale, which asked teachers to rate how they believe the programs impact student achievement. The survey used was designed for this study based on the Intervention Rating Profile and was administered using a REDCap survey. The ordinal data from this survey allowed for inferences related to participants expectations of the adaptive learning technology program they are using in their math classroom. In addition, this data allowed for a general comparison between teacher social validity and the achievement data gathered from an analysis of the NWEA assessment.

Data Analysis

Professor Hattie's Visible Learning Theory (2018) was one of the cornerstones of this study, and therefore, effect size calculation was used in calculating and analyzing the data. Hattie's theory rests on the premise that strategies that have a .40 or greater effect size have more than a year's worth of impact on student achievement (Fisher, Frey & Hattie, 2016). Therefore, schools and individual teachers should gravitate towards such practices (Hattie, 2018). Using effect size (Coehn's d) as the primary statistical calculation of this study allowed for an effective analysis of the first research question and hypothesis in this study:

RQ1: What is the student achievement impact (in terms of NWEA growth scores) of two adaptive learning technology programs used in a rural district in Western Maine?

H1: Both Espark and iXL will have a .40 or higher effect size on student achievement in the area of mathematics as recognized by their growth between a pre and post NWEA assessment data. The .40 effect size is considered to be a year or more's worth of academic growth in Hattie's Visible Learning Theory (Hattie, 2008).

As previously mentioned, the effect size is a simple statistical measurement that is easily understood in a number of different settings (Coe, 2002). Using this calculation allowed for the results of this study to be discussed and understood with confidence within and outside of the district. Effect size compared the mean of the pretest with the mean of the post-test dividing it by the spread or standard deviation of the testing groups (Madsen, Sayre, & McKagan, 2016). Using the means as a measure of comparison provided for more confidence in keeping individual student data secure, as for analyzing group averages not individual scores. In addition, this allowed for these two adaptive learning technology math programs to be part of Hattie's ongoing conversation about the importance of selecting impactful tools in the classroom.

The second set of quantitative data in this study, related to the social validity of teachers using the two programs, was mean and median. The mean and median data points were used in the analysis of the social validity survey. The mean scores were compared to the Cohen's d of the student achievement data using general comparative analysis and were compared to the effect size data of that adaptive learning technology.

Limitation, Delimitations, Ethical Issues

Bloomberg and Volpe (2018) stated, "the limitations are external conditions that restrict or constrain the study's scope or may affect its outcome. Delimitations, on the other hand, are conditions or parameters that you, as the researcher, intentionally impose in order to limit the

scope of your study" (p. 13). With this definition in mind there were several limitations and delimitations to consider with this study. The most prominent limitation was the size of the data and participant pools. Given that the study was conducted in a small rural district, the active participants were initially restricted to just four classroom teachers (three at the elementary school and two at the middle school). In addition, the data collected was limited, because it came from just 25-30 students in each grade, a total of 50-60 data points. The limited pool of student data and teacher participants impacted the conclusive statements made regarding the impact of the two adaptive learning technology math programs on student achievement.

A second limitation to consider in this study was the fact that students were not actively involved in the study, just their NWEA data was used. This is a limitation as it did not allow for the study to control many of the variables that come into play when examining student achievement. For example, teacher experience with the programs and the impact of teacher training were factors that were not controllable. Another example is the fact that teachers did not have a choice in using one program over the other. Randomly selecting such teachers could help to eliminate negative or positive presumptions regarding the adaptive learning technology programs. A third potential limitation for the study is the fact that the student population was not culturally diversified. This limited discussions around the impact of the adaptive learning technology programs on different subsets of school or society's populations.

In terms of delimitations for this study, the first is that a rural school district was selected, which decreased the data available for both student achievement and teacher social validity. One of the goals of the study was to provide the local school district with data that they can use to make program decisions. This goal, however, greatly limited the scope and impact of the study. A second delimitation of the study is that one student achievement data point was selected

instead of using two or more. This eliminated the idea of triangulation of data but allowed for a more direct and clear analysis and discussion on the data that was gained.

There were three ethical considerations for this quantitative study. The first consideration dealt with the fact that the study used student data found at the elementary school where the researcher was the acting administrator. This created the potential for bias towards the two programs in question and it also created a situation where teachers might be less neutral in their roles than they would be if there were an outside researcher. To help limit this bias, teacher identity with both the NWEA and the social validity survey at the elementary school was kept confidential and was deidentified and organized by a third party. A second ethical consideration was one of the programs was currently being used in the elementary school in grades kindergarten through fifth grade and the other in the middle school in grades six through eight. The study was planned to be conducted in grades five and six, which means that teachers in those grades could have already had a positive or negative association with the two adaptive learning technology programs, Espark and iXL. A third potential ethical concern was the fact that student data was involved. To eliminate, or best minimize this concern, the study was conducted "non-live", which means that it relied on the already planned assessment data and no changes to the student's programming or curriculum occurred as a result of the study. In addition, student names were de-identified by a third party so that only raw, non-identifiable data was analyzed. The study looked at collective growth not individual growth, which also helped to minimize this ethical concern.

Trustworthiness

Trustworthiness, or the idea that a study is both significant and valuable is comprised of several components (Bloomberg & Volpe, 2018). In this section the concepts of credibility,

transferability, dependability, and confirmability are discussed in relation to this study. The credibility of the study was based on its simple design and use of confidential data. The transferability was built upon the idea that it is using a similar study by Hollands and Pan (2018) as a springboard while also connecting it to previous studies around student achievement. The dependability was grounded in the quantitative methodology and the confirmability was established through the anonymity of the data and the national normed assessment that was selected.

Credibility

The credibility of this study rested in its simple design. For the first data set, data from already scheduled NWEA assessments was used, thus eliminating potential credibility issues with participant expectations. To increase credibility and decrease biases in this study a third party de-identified the data before the researcher managed it for analysis. This provided assurance to the teachers who participated that their survey responses were not directly linked to them, and also that student data was kept confidential at all times. Additionally, a third party helped to minimize the potential bias impact that the researcher could have as a building administrator in one of the site schools. In addition to supporting the credibility of the study among the participating teachers, participation in the survey was voluntary and their names were not used in any discussion related to the study. There is only one math teacher at the middle school for students in grade six. This created a potential confidentiality concern, so to address this the researcher gained direct permission from that teacher fully disclosing that given that they are the only teacher in their pool deidentifying their responses to the social validity survey were not always possible. However, in all written depictions of the study and its findings their identity was not specifically referenced, and every effort will be made to maintain confidentiality.

Transferability

Transferability, or external validity, is the idea that a study connects to the larger population (Bloomberg & Volpe, 2018). For this study transferability was found in the fact that it was building off from a previous study while adding to the ongoing discussions around both student achievement and the impact of adaptive learning technology. This study piggybacked the 2018 Hollands and Pan study, which assessed the same two adaptive learning technology programs. This study sought to improve upon the Hollands and Pan study by using a simplified methodology and a common assessment to examine the impact on student achievement. This study's simplistic design and wide scope allowed for the data to be used to expand conversations around the use of adaptive learning technology programs, strategies to impact student achievement in math, and the use of Hattie's Visible Learning Theory (2008) to assess program effectiveness.

Dependability

A quantitative approach to this study was selected to increase the dependability of the results that were generated. Using the Hollands and Pan (2018) study as a launching pad, this study allowed for the hard data that was gleaned to be discussed alongside the hard data associated with the original study. Effect size calculations were selected both because they are simpler and better understood by the layperson. Further, effect size provides an accurate means of determining the impact of each program on student achievement. The dependability of the study also rested in the fact that the data collected is part of a larger assessment system that is given three times a year to students in the site district. The NWEA assessment is a nationally normalized assessment that is given to millions of students each year (NWEA, n.d.), therefore using the data from this assessment significantly increased the dependability of the study.

Confirmability

The confirmability for this study rests in the fact that there was no means by which the researcher could adjust the data collected. For the student assessment data, the NWEA is a nationally normed assessment that is given to students three times a year in the site district. There was no opportunity for the data to be tampered with by the researcher because it was collected, organized and deidentified by a third party. REDCap was used to collect the teacher survey data. This surveying program provided better confidentiality for the teachers participating in the survey (project-redcap.org).

Summary

The purpose of this quantitative study was to examine the impact of two adaptive learning technology math programs in terms of the student achievement outputs indicated on the NWEA assessment for a small rural district in Maine. At the same time the study also sought to examine how the student achievement results aligned with the social validity of the teachers who were using the programs in their classrooms. The study's simple design of using Cohen's d or effect size when comparing student achievement data, allowed for it to be accessible by the layperson in the field of education. This simpler methodology design consequently opens the study up to some potential biases. The most striking biases highlighted in this section was that the researcher is the school administrator at B elementary school. This created the potential for bias in terms of the researcher's desired outcomes for the study as well as how teachers taking the social validity survey responded. To address these ethical concerns, a third party was used to collect, organize and deidentify the data.

This study sought to extend the conversation initiated by Hollands and Pan (2018) when they looked at the same two adaptive learning technology programs (Espark and iXL) and

assessed their impact on student learning in a much larger and urban-based school district. Changes to methodology in this study versus the one conducted by Hollands and Pan were made because of a more limited participant pool and to provide a better connection to its visible learning theoretical framework. The changes in the methodology may have impacted the reliability, validity and transferability of this current study. The former two were impacted because the limited size of the participant pool diminished both the reliability and validity of the results provided. The former element, that being the transferability was improved in this study because of the simpler design. Again, it was the intention that this study would be highly transferable to not only the impact of online digital programs on student achievement, but potentially to other school-based program analyses that schools are seeking to examine.

Chapter Four of this study presents both the results and an analysis of the results. This includes the effect size calculations gleaned from an analysis of the spring 2022 NWEA. The data compares the academic impact of both Espark and iXL on NWEA assessment. In addition, the data is presented in terms of the two sub-groupings, those from participants that were at and above grade level and those that were below. The teacher social validity data is also presented in Chapter Four. In Chapter Five, the implications and importance of the study are discussed along with recommendations that emerged from an analysis of the data in Chapter Four.

CHAPTER 4

RESULTS

The 1983 publication of the Nation at Risk Federal Education Report highlighted significant concerns around the academic performance of America students (United States National Commission on Excellence in Education, 1983). Three decades later in 2011, a report from Harvard's Educational Policy and Governance echoed the concerns raised in the Nation at Risk report. In this Harvard publication, it highlighted that the United States had a 32% proficiency rate in math and just 31% in reading, making it far below several other developed countries (Peterson, et. al., 2011). In an even more recent report from 2017 it was again highlighted that the United States was underperforming globally in the area of math, reading, and science (Desilver, 2017). Even though the Nation at Risk report is now nearly forty years old, the United States does not seem to have gained much ground in terms of student academic achievement in math, reading, and science.

In addition to illuminating the well-established trend of America's failing education system, the Nation at Risk report provided several recommendations for remedying these woes (United States National Commission on Excellence in Education, 1983). One such recommendation was a call for an investment in technology, technology that would both enhance student academics and make the organization of education more effective and efficient. With this technology investment, however, came a significant financial barrier that many districts are still wrestling with today, which forces schools to weigh the cost of implementing technology programs with the benefit that they ultimately provide (Coulson, 2006). Despite this, schools have taken on the challenge of budgeting for, purchasing, and implementing technology aimed at making learning more efficient and differentiated. In fact, spending on technology in schools

reached almost two billion dollars in 2014 (Koba, 2015) highlighting the fact that US public schools have taken to heart the technology warnings and suggestions brought forth in the 1983 Nation at Risk report (United States National Commission on Excellence in Education, 1983).

Despite investments in technology, US Schools have not necessarily done an adequate job of assessing the impact that technology is having specifically on student achievement (Hollands & Pan, 2018). The rapid rise in available educational technology has spurred the hope of transformational change in the US public school system, with calls for ideologies such as Mass Customized Learning to replace what is considered an archaic educational system (Schwahn & McGarvey, 2012). At the heart of these calls for change was a realization that new technology, especially that found in adaptive learning technology programs can provide diversification to support students' individual needs in ways more traditional programs of the past could not (Schwahn & McGarvey, 2012).

Therefore, the purpose of this quantitative study was to examine the academic impact of two adaptive learning technology programs on fifth and sixth graders while comparing their impact with the social validity of the teachers implementing the programs at their schools in Western Maine. The study specifically looked at the academic impact of iXL and Espark, which are two adaptive learning technology math programs used in Regional School Unit – A. However, the target population changed to fifth, seventh, and eighth graders through a modification submitted for IRB approval. Adaptive learning technology programs have provided a number of educational benefits (Kurt, 2021), and this study examined the impact of these two programs using Northwest Evaluation Association (NWEA) scores as the pre and post measures.

In addition to examining the impact of two adaptive learning math programs, this research study sought to analyze how the social validity of teachers implementing the math programs aligns with the results that are gleaned from it. Social validity for the purpose of this study referred to the teacher's perception of the effectiveness of the program they were implementing. The study sought to examine if the teacher's perception of program effectiveness matched the academic outcomes of the students using the programs. The social validity was assessed through a ten-question social validity survey, or teacher perception, which allowed for comparisons between student achievement data and teacher perception data. This important layer allowed this study to explore further connections regarding the extent that teacher perceptions of how these technology programs align with or impact the results the programs have on student achievement.

This research study contributes to the conversation regarding the effectiveness of adaptive learning technology math programs in supporting the math achievement of students in grades five, seven, and eight. The following research questions and hypotheses guided the study:

RQ1: What is the student achievement impact (in terms of NWEA growth scores) of two adaptive learning technology programs used in a rural district in Western Maine?

H1: Both Espark and iXL will have a .40 or higher effect size on student achievement in the area of mathematics as recognized by their growth between a pre and post NWEA assessment data. The .40 effect size is considered to be a year or more's worth of academic growth in Hattie's Visible Learning Theory (2008).

RQ2: Does the social validity of the teachers who are using Espark and iXL in their classrooms align with the identified student achievement results?

H2: There will be a positive (mean score higher than 3) connection between the social validity of teacher's implementing the adaptive learning technology programs and the student achievement impact on the NWEA.

Data from the NWEA scores was prepared by a noninvolved individual, to allow for greater confidentiality of the source of both the student academic data and the data collected from the teacher social validity survey. This third-party individual gathered the NWEA data from both the elementary school and the middle school and removed all identifiable information. The data was then organized and sent to the primary investigator. Statistical measures (mean, standard deviation, and Cohen's d) were all used in the analysis of the data that was collected. Cohen's d or effect size calculations were performed on the student academic data. The teacher social validity data was analyzed using mean scores, which were calculated using the results of the three participants who used Espark and the two participants who use iXL. The remainder of this chapter focuses on the analysis method and the findings.

Changes to Target Population in Chapter 3

At the proposal stage of the study, both grade five and grade six student NWEA data from Regional School Unit-A (RSU-A) as well as survey data from grade five and grade six math teachers in these schools were identified as the targeted population in this research study. After University of New England IRB approval (see Appendix B) of the proposed methodology and target population groups, it was discovered that RSU-A grade six teachers did not use iXL in their classroom at a level that would provide any student data (as the study required students to use the adaptive learning technology programs for at least 45 minutes a week). The principal of C Middle School indicated, however, that students in both grade seven and grade eight met this threshold of use with iXL. An amendment to the study was submitted to UNE's IRB for approval

to collect data generated by seven and eighth grades as well as to survey the math teachers in these grades. Approval for this change was granted (see Appendix B).

This change impacted the study in a few areas. First, including grades seven and eight instead of grade six increased the student data from the expected twenty-five participants to forty-nine participants. Second, because the grade six teacher had already taken the teacher survey and the survey was anonymous, that data could not be removed and was part of the survey data. This also increased the Social Validity Survey data from four to five participants. A final impact of this change is that it created a one-to-two-year gap between the ages of students in the two comparison groups. The original design selected grades five and six because they were closest in age and academic development, with a move to grades seven and eight, this changed.

Analysis Method

The steps involved in calculating, collecting, and analyzing the data in this research study began with students in grades five, seven, and eight who received a minimum of ten weeks of math support from one of the two adaptive learning technology math programs (Espark and iXL). This period of time will coincide with one of the predetermined NWEA testing windows which occurs three times during the school year. Data from the fall testing was compared with data from the spring testing. Once the post NWEA session was completed the data was organized into two main categories, Espark and iXL.

The student academic data was collected from the fall and spring Northwest Evaluation Association (NWEA) assessment, which is completed three times a year by students in grades K-10 in RSU A (the winter data was not used in this study). Students completed the fall assessment, which took place the second week of October 2021 and served as pre-test data. The spring administration of the NWEA assessment, which took place in mid-May 2022, served as the post-

assessment data. The data was collected using the NWEA grade report, by a third party, and was organized into a data chart which contained the following information: students (deidentified), fall NWEA math scores, spring NWEA math scores, and achieved growth. This data was collected for all students who met the following three criteria:

1. They were enrolled in either B Elementary School in grade 5 or at C Middle School in grades 7 and 8.
2. The students took both the fall and the spring NWEA assessment.
3. The students used Espark or iXL for an average of 45 minutes a week between the two testing periods.

There were twenty-four (24) students in grade 5 that met the three qualifying criteria and their NWEA assessment data was used to calculate the effect size of Espark on student math achievement. There were twenty-nine (29) seventh graders and seventeen (17) eighth graders who met the three qualifying factors making a total of forty-six (46) eligible participants for the iXL math achievement data. Pre and post assessments scores were entered into a data chart with all names deidentified by a third-party individual. This individual was a district staff member who was not eligible to participate in the study and had no direct means of benefiting from or having negative consequences on or from the study.

Cohen's d, mean, and standard deviation were used to answer the two research questions associated with this study, which related to the academic impact of adaptive learning technology math programs and the social validity of teachers implementing the programs in their classrooms. Mean scores for the 24 participants in the Espark group and the 46 participants in the iXL group were calculated and recorded. In addition, the standard deviation for each of the groups was calculated and recorded in the data chart. Four data points (pre-assessment group

mean, post assessment group mean, pre-assessment standard deviation and post-assessment standard deviation) were then used to calculate Cohen's d or effect size. This calculation was done both by hand and using an online effect size calculator to ensure that the calculations were accurate. To arrive at the effect size calculation, where there was no control group used, the pre-mean score was subtracted from the post-mean score and divided by the mean of the two standard deviation scores (from the pre and post data sets). This allowed for mean, standard deviation and Cohen's d to all be used to analyze the quantitative student data collected during this study, as demonstrated in Table 2.

Table 2

Quantitative Data Related to Related to Design Question 1

Cohort	N	Pre-test Mean	Post-test Mean	Mean Growth	Pre-Test SD	Post-test SD	Mean SD	Cohen's d
Grade 5 (Espark)	24	207.83	212.41	4.58	9.66	11.15	10.41	.439
Grade 7 (iXL)	29	215.52	221.34	5.83	11.10	13.13	12.12	.479
Grade 8 (iXL)	17	218.53	223.24	4.76	8.07	8.05	8.06	.584

The second set of quantitative data that was collected in this study came from a ten-question Social Validity Survey (see Appendix A) that participating teachers took. This survey was based on a six-point Likert scale, which asked teachers to rate how they believe the

programs impacted student achievement. The survey was designed after the Intervention Rating Profile for this proposed study and was administered using a REDCap survey. The survey took approximately ten to fifteen minutes to complete by the five eligible participants (teachers in grades 5, 7 and 8). Since the participant grade level had to be changed from grade six to grades seven and eighth (because grade 6 students did not use iXL for the required amount of time), an additional participant was added (the grade 7 and 8 math teacher). The social validity survey was anonymous and therefore the data collected from the grade six teacher could not be eliminated. This teacher, however, used the iXL program in their classroom, therefore the data was considered relevant to the purpose of the study. The ordinal data from this survey allowed for inferences related to participants expectations of the adaptive learning technology program they are using in their math classroom. In addition, this data allowed for a general comparison between teacher social validity and the achievement data gathered from an analysis of the NWEA assessment.

The average number of teaching years for the three Espark grade five respondents was nineteen years and the mean teaching experience for the iXL respondents (grades seven and eight) was fourteen years. There were two female grade five participants and one male grade five participant from the Espark group and one male and one female participant from the iXL group. All participating teachers used the program in their classroom for at least two years. Once all data was collected anonymously in the RedCap data system, two data sets were created based on the Espark and iXL groups. Mean scores for each of the ten questions for each participant group (Espark and iXL) were calculated and entered into a table.

Presentation of Results and Findings

The two research questions and hypotheses were analyzed for this study using quantitative data collected from both student achievement scores and a teacher social validity survey. The first research question posed was what the student achievement impact (on students' NWEA assessment scores) of two adaptive learning technology math programs was. The researcher hypothesized that both Espark and iXL would have a .40 or greater effect size due to Hattie's conclusion after examination of over 1,800 meta-analyses, which involved over 300 million students worldwide (Corwin Visible Learning, n.d.). The statistical marking of .40 in the hypothesis is in line with Hattie's Visible Learning Theory (2008), which promoted that practices and programs that have a .40 effect size have a year or more worth of growth impact on student achievement.

The Espark group consisted of 24 grade five students that used Espark math for an average of 45 minutes or more a week from October 2021 to May 2022. The iXL group included twenty-nine seventh grade and seventeen grade 8 assessment data for a total of forty-six data points between the two grades who used iXL math for 45 minutes or more during this same time frame. Table 2 highlighted the statistical measures that were taken to address the first research question and hypothesis with the mean of both the pre and the post NWEA assessments for each of the two groups (Espark and iXL).

The mean of the Espark (Grade 5) pre-test was 207.83 and the pre-test mean of the seventh-grade cohort was 215.52 and the grade eight cohort was 218.53. The post mean for the grade 5 Espark cohort was 212.41 representing an average growth of 4.58 between the two assessments. The post mean score for the grade 7 cohort (iXL group 1) was 221.34, representing

a mean growth of 5.83. The mean post-assessment data for the grade eight cohort (iXL group 2) was 223.24, representing a mean growth of 4.76.

To determine the effect size of Cohen's d, the standard deviation for each group was also calculated and represented in Table 2. The mean standard deviation between the pre and post test data for grade 5 (Espark group) was 10.41. Grade 7 (iXL) had a mean standard deviation for the pre and post data of 12.12. Grade eight (iXL) had a mean standard deviation for the pre and post assessment data of 8.06. To arrive at the Cohen's d, the Pre-assessment mean data (the average of the student scores in the cohort for the fall assessment) was subtracted from the post assessment mean data (the average of the student scores from the spring assessment) and divided by the mean of the standard deviations for each grade cohort. This allowed for a comparison of the growth between the first assessment (pre-assessment) and the final assessment (the post-assessment). The results showed that the grade 5 cohort (Espark) had an effect size of .439, the grade 7 (iXL cohort 1) had an effect size of .479, and the grade 8 (iXL cohort 2) had an effect size of .584. Each of these three effect size scores confirm the hypothesis that both Espark and iXL would have a .40 or greater effect size on student achievement data.

In addition to the data that was directly relevant to this study, the NWEA Grade level report, which was run for each cohort of students provided additional data that is relevant to this and potential future studies. A critical piece of data provided in the NWEA report was the project growth score as well as an indication of whether the student met NWEA's projected growth mark. This projected growth score was determined by using the millions of student data points that NWEA collects each year. These data points provide a performance estimation from fall to spring in the form of a projected score based on the hundreds of thousands of students of like age who started in the fall with a similar score on the NWEA assessment (NWEA.org, n.d.).

The data collected from the NWEA report for this study, which is outlined in Table 3,

Table 3

NWEA Projected Growth vs. Cohen's d Comparison Chart

Cohort Projection	N.	Cohen's d.	N met NWEA Projection.	% Who met NWEA
Grade 5 (Espark)	24	.439	7	29
Grade 7 (iXL)	29	.479	16	55
Grade 8 (iXL)	17	.584	8	47

indicated that of the 24 grade five students who met the study requirements seven or 29% met their projected NWEA yearly growth mark. Of the 29 grade seven student NWEA data points, the report showed that 16 or 55% met their projected NWEA growth mark. In grade 8, where there were seventeen participants, 8, or 47%, met their NWEA projected growth. This means that the iXL group had a total of 24 students or a combined average of 51% who met their NWEA projected score, whereas only 29% of the students who were in the Espark group met their projected score.

The NWEA Grade Report also includes a breakdown of student performance among students who were categorized as Low, Low Average, Average, Above Average, and High. NWEA considers a student low if they score below the 21%, low average if they are between the 21-40%, average if they are between 41-60 percentile, high average if they are between the 61-80% and high if they are above the 80%. Because of the lower number of overall participants for this study, where some categories only had 3-5 total student scores, comparisons between these subgroups were not conducted. Future studies with much larger participant numbers, however,

would be able to use this data to draw conclusions about how the adaptive learning technology programs performed within each of these NWEA percentile categories.

The second set of quantitative data collected addressed research question number two, which asked if the social validity of teacher's implementing the adaptive learning technology programs aligned with the student achievement results. The hypothesis for this question stated that there would be a positive connection or a mean score of 3 or higher between the teacher responses on the Social Validity Survey and the student achievement results reported in the NWEA reports. There were three respondents in the Espark group (Grade 5) and two for the iXL (Grades seven and eight) group for a total of five participants. One of the two participating data points from the grades seven and eighth iXL group came from a grade sixth teacher who completed the survey before the study moved from using grade six data to grades seven and eight. This data is included in the study because the survey was done anonymously and there was no way to remove the data from what was collected. The overall mean score (mean of all scores on all ten questions) for the Espark group was 3.32 and the overall mean score for the iXL group was 4.20. The Espark group scores ranged from 2.0 to 4.3 for the mean scores on each question. The range for the iXL group was 3.5 to 5.0. Both groups have an overall mean score of above 3.0, confirming the hypothesis related to research question number two in this study. Table 4 highlights the overall mean score for each question for both the Espark and the iXL groups.

Table 4

Teacher Social Validity Survey Data

#	Question	Espark Mean	iXL Mean
1.	Students in your classroom were motivated to use the adaptive learning technology program.	2.67	3.5
2.	The adaptive learning technology program was engaging for your students.	2.0	3.5
3.	The adaptive learning technology program provided individualized instruction at appropriate levels for the students in your class.	4.0	4.5
4.	The adaptive learning technology program was easy to implement into the structure of your classroom.	3.67	5.0
5.	The adaptive learning technology program required minimum background technology understanding to implement.	4.0	3.0
6.	The students did not need a lot of technical assistance when using the adaptive learning technology program.	4.33	5.0
7.	The use of the adaptive learning technology program in your class allowed you to diversify your instruction.	4.0	4.5
8.	The adaptive learning technology supported student growth on the NWEAs.	2.67	5.0
9.	The adaptive learning technology program had a positive impact on students' perceptions of their ability to successfully do mathematics.	2.33	4.0
10.	The adaptive learning technology program had a direct impact on your student's ability to master math related Standards.	2.67	4.0
Overall Mean		3.32	4.20

Table 4 identifies that motivation (question 1) and engagement (question 2) were the two lower scored elements shared between the two groups on the survey. On the other hand, students not needing a lot of technical assistance (question 6) was an overall higher scoring element shared by both the Espark and the iXL group. Another area that teachers who used both Espark and iXL rated above their overall average was question 3, which asked if the programs provided instruction at the student's level. The design of adaptive learning technology programs is to

adapt to the level of the user of the program. Therefore, this question seemed to confirm, in the eyes of the teachers implementing the program, that both programs were meeting this expectation (Espark responders rated this a 4.0 whereas iXL responders rated it a 4.5).

The greatest discrepancy of scoring between the two groups of teachers was question 8, which asked respondents if they believed the adaptive learning technology program supported student growth on the NWEA. Espark participants had a mean of just 2.67 while the iXL participants scored this a 5 on question 8. In a similar response to question 10, which asked if the teachers felt that the adaptive learning technology program had a direct impact on their students' ability to master math related standards, the iXl group rated this a 4.0, where the Espark group rated this at just a 2.57. As Table 4 highlighted the Espark group only scored one question (question 5), related to the technical background knowledge required to implement the program, higher than the iXL group.

Summary

This study sought to examine the academic impact of two adaptive learning technology math programs on student achievement, while also assessing the social validity of the teachers that were implementing the programs in their classrooms. The data gathered from the study confirmed the hypotheses the researcher made relative to the two research questions for this study. The first hypothesis which stated that both Espark and iXL would have a .40 effect size or greater on student achievement was confirmed as the Espark group had an effect size of .439 and the iXL group had an effect size of .479. Data derived from the NWEA grade report showed that only 29% of grade 5 students, 55% of grade 7 students and 47% of grade 8 students met their projected yearly growth on the NWEA assessment.

The second research question's hypothesis was that both Espark and iXL would have an overall mean score of 3.0 or higher on the staff Social Validity Survey, which was confirmed through the data provided in this study. The Espark (grade 5) Social Validity Survey had an overall mean score of 3.32 and the iXL group (grades 7 and 8) had an overall mean score of 4.20. As previously highlighted, the teachers using iXL provided a higher rating on 9 out of 10 questions than did the Espark. In terms of student motivation, engagement, growth on the NWEA and overall growth on math standards, the two iXL teachers gave a higher rating than did the three teachers using Espark. The only question that where the Esaprk teachers provided a higher rating than the iXL teachers was on the needed technology background to implement the program in their classrooms.

In Chapter Five, the results presented in this chapter are interpreted and discussed further. The discussion in Chapter Five provides more insight as to how the data relates to this study, future studies, and the stakeholders involved in the study. Chapter Five includes the following subjects: Interpretations and importance of findings of the study, implications, recommendations for action, and recommendations for further study.

CHAPTER 5

CONCLUSION

The identified critical problem that this study sought to address was the fact that students in United States schools are underperforming in the area of mathematics (The Nation's Report Card, n.d.). This problem has persisted for nearly forty years (United States National Commission on Excellence in Education, 1983) and schools have tried to solve it in a variety of ways. The use of technology to impact student achievement was suggested in the Nation at Risk education report in 1983, and since that time schools have spent billions of dollars implementing both technology hardware and software programs (Koba, 2015). With technological advancements like Adaptive Learning Technology programs, which diversify instruction or intervention based on the individual needs of the student, technology continues to have a significant impact on student achievement (Capuano & Caballe, 2020). With this advancement in technology the need to assess its impact on student achievement is critical (Hollands & Pan, 2018).

The national education leaders at the heart of the Nation at Risk Report (United States National Commission on Excellence in Education, 1983), however, most likely could not have predicted the significance that technology investment would have on student achievement in the wake of the Covid-19 pandemic. This global health battle shut down most K-12 public schools across the United States, causing a major disruption to the traditional educational program available to most students (Office of Civil Rights, 2021). The academic impact of this global pandemic is still being assessed, but early indicators from research suggest that schools will be scrambling to recover from the student academic loss for years to come (Kuhfeld et al., 2020). Despite this significant disruptive wave, starting in 2019 and continuing through 2022 schools

received billions of dollars in Elementary and Secondary Emergency Relief (ESSER) funds from the federal government (Gordon & Reber, 2020). This funding provided schools with the means to not only cope with initial effects of COVID on education in the United States, but also provided much needed funds for technology hardware and programs to support remote learning efforts that persisted through the 2020-2021 academic school year.

At this intersection of student achievement and the advancement of technology, the nucleus of this proposed quantitative study is found. U.S. public schools were already academically struggling prior to the onset of the pandemic in 2019, with 40% of fourth and eighth grade public school students meeting proficiency in expectations in math and only 34% in reading (The Nation's Report Card, n.d.). In addition, U.S. schools were failing to properly implement technology systems that have been deemed significant to impacting student achievement (Stepman, 2018). A 2018 publication, titled Report Card on American Education, highlighted the fact that only two states, Florida and Utah, received a grade of an A for digital learning, while nearly 50%, or 24 states, received a grade below a C (Stepman, 2018). The worldwide pandemic in 2020 did, however, increase the significance of online program evaluation, for it derailed more traditional learning methods in schools across the United States (Office of Civil Rights, 2021). This occurrence forced schools to adopt online digital programs that could continue to support student learning when direct instruction was not readily available.

It is critical for schools to continue to implement policies, programs and practices that have a high impact on student achievement (Hattie, 2018) if they wish to prepare students to be both successful post high school and to compete on the international stage. This research study looked at the impact that two adaptive learning technology math programs had on the academic achievement of students in grades five, seven, and eight. The study took place in a rural western

Maine school district, which has assessment scores in math that mirror the struggling scores found across the United States. In addition to assessing the impact of two different adaptive learning math programs, this quantitative study sought to analyze how the social validity (e.g., the satisfaction and acceptability of the online math program used) of teachers who are implementing the math programs aligned with the results that are gleaned from it. This important layer allowed this study to explore connections to teacher perceptions of how these technology programs align with or impact the results the programs have on student achievement.

This study examined the academic achievement impacts of two adaptive learning technology programs, Espark and iXL, on fifth, seventh and eighth grade students in a rural Western Maine school district. Using Hattie's Visible Learning Theory (2018), as the theoretical framework, this quantitative study sought to address two critical research questions and hypotheses:

RQ1: What is the student achievement impact (in terms of NWEA growth scores) of two adaptive learning technology programs used in a rural district in Western Maine?

H1: Both Espark and iXL will have a .40 or higher effect size on student achievement in the area of mathematics as recognized by their growth between a pre and post NWEA assessment data. The .40 effect size is considered to be a year or more's worth of academic growth in Hattie's Visible Learning Theory (Hattie, 2008).

RQ2: Does the social validity of the teachers who are using Espark and iXL in their classrooms align with the identified student achievement results?

H2: There will be a positive (mean score higher than 3) connection between the social validity of teacher's implementing the adaptive learning technology programs and the student achievement impact on the NWEA.

With these two research questions and hypotheses established, the aim of this study was to both assess the impact of using adaptive learning technology on student achievement, while also assessing how teacher social validity might be a critical component in both the success of a program and student achievement. This second layer, which addressed research question two, allowed for exploration regarding how teacher perceptions of a technology-based program align with student achievement results gained from the program.

This study is both relevant and significant on a number of levels. It is relevant because in the wake of the Covid-19 pandemic there was an increased focus on less traditional means of instruction to technology-based means. The results from this study can contribute to schools continuing to weigh the value of the technology they are currently adopting. This study is also significant, because only one study within the past decade was found that looked at more than one adaptive learning technology in place at the same site (Hollands & Pan, 2018). Comparing two different adaptive learning technology programs head-to-head can better inform the local district as well as elevate discussions around both technology programs (Espark and iXL) and their imprint on the students using these adaptive learning programs and the teachers implementing them.

Previous studies examined teacher social validity for school-based interventions. A 2016 study (Vancel, et. al., 2016) examined the social validity of elementary, middle and high school teachers in districts implementing school-wide Positive Behavior Intervention Supports (PBIS), and how implementation of PBIS as an intervention aligned with teacher's perception of it.

Another recent social validity study (Dore et al., 2021) examined teacher and care-givers perception of virtual Pre-K as a pandemic intervention and its ability to prepare students for kindergarten. Social validity has also been used to assess the impact of technology interventions in a school setting, in one such study, teacher and student social validity was used to assess the perceptive effectiveness of using iTouch flashcard program with special education students (Jameson et al., 2012). Only one previous known study examined teacher social validity in the context of adaptive learning technology. That study completed in 2018 by Smith, examined future teacher perceptions of adaptive learning technology in k-8 mathematics education. The fact that only one known study examined teacher social validity alongside adaptive learning technology programs, makes this study more relevant. By examining the social validity relative to Espark and iXL, this study may open the door for discussions related to how teacher perception of adaptive learning technology programs connects to student performance. Finally, the results from this study may directly impact the program decision making in RSU A, the host site, as it finds itself wrestling with low student proficiency scores in mathematics and questioning how to assess the technology-based programs that it has adopted.

The theoretical framework that served as the foundation of the study rested in John Hattie's Visible Learning Theory (Hattie, 2018). This theory promotes the idea that if teachers are armed with the knowledge of the impact that their programs and practices have on student learning, they will have more success in helping students meet their academic goals (Hattie, 2008). Hattie (2008) used effect size or Cohen's d as the statistical backbone of the real-life application of his theory, and therefore Cohen's d served as one of the critical statistical measures addressing the first research question in this study. The hypothesis for research question 1 indicated that both Espark and iXL would have more than a year's worth of effect on

the students' academic growth. A year's worth of growth according to Hattie's meta-analysis of over 1800 studies was calculated to be .40 (wlearning.com, n.d.). The results of the study indicated that both Espark (used by grade 5 students) with a mean effect size of .439 and iXL (used by seventh and eighth graders) with a mean effect size .532, according to Hattie's .40 effect size research (Hattie, 2008), had more than a year's work of impact on student math achievement in RSU A. The quantitative data for this part of the study came from an analysis of the mean and standard deviation of 24 students in grade 5 and 46 students in grades 7 and 8 -who met the qualifying parameters for the study: they were enrolled in the district, they took both the fall NWEA assessment (pre-test) and the Spring NWEA assessment (post-test), and they used Espark for 45 minutes or more each week from October 2021 to May 2023.

The second set of quantitative data collected from this study came from a ten-question Social Validity Survey given to teachers who teach math in grades 5 through eight (the grade 6 math teacher took the survey before the study had to shift to grades seven and eight). This survey, which addressed research question two, looked at the connection between the teacher perception and program effectiveness. Visible Learning Theory (Hattie, 2008) promotes the concept of educators being reflective practitioners, able to examine their pedagogy and adjust to better serve the needs of their students (Hattie, 2018). Social validity is defined as, "the satisfaction and acceptability of the interventions and procedures affecting behavior change, based on the opinions of the individuals who receive services and implement them" (Basir & Presberg, 2019, para. 2). For the purpose of this study, social validity was a means of examining how teacher's perceptions might influence student outcomes. By examining the social validity of the teacher's using the adaptive learning program, this study allowed for reflection on the potential impact of teacher perceptions on program success and ultimately student achievement.

The average or mean of each question and the overall mean was used as the statistical measure for analysis for the second research question in this qualitative study.

Five participants responded to the ten-question social validity survey (see Appendix A) as part of this study. Three participants actively used Espark in their classrooms and two used iXL. The questions on the survey asked the respondents to rate (using a 5-point Likert Scale) items from how they perceived that the program engaged their students, to how it aided them in diversifying their instruction to how much technical skill was required to implement the program.

There was an overall positive perception of both programs by the five participants of the survey. Espark had an overall mean score (overall score of all questions) of 3.32 and iXL had an overall mean score of 4.20. The lowest questions for the Espark group were:

- The adaptive learning technology program was engaging for your students (mean score of 2.0);
- the adaptive learning technology program had a positive impact on students' perceptions of their ability to successfully do mathematics (mean score of 2.33);
- students in your classroom were motivated to use the adaptive learning technology program (mean score of 2.67); and
- the adaptive learning technology program had a direct impact on your students' ability to master math related standards (mean score of 2.67).

For the iXL group the lowest ranking questions were:

- The adaptive learning technology program required minimum background technology understanding to implement (mean score of 3.0);

- the adaptive learning technology program was engaging for your students (mean score of 3.5); and
- students in your classroom were motivated to use the adaptive learning technology program (3.5).

Both groups of respondents rated student motivation and engagement lower than the other items surveyed, which is contrary to both what is promoted by Espark and iXL (Espark.com, n.d., iXL.com, n.d.) and what has been previously suggested in research (Walkingon, 2013; Haven, 2014). The fact that both the Espark and the iXL group gave motivation and engagement a lower score, is contrary to what is promoted for these technology systems and what is previously shown as critical components for student growth. The iXL group respondents gave iXL a higher rating than the Espark respondents on all questions except the questions which asked about the amount of technology background needed to implement the program (the overall mean for the iXL group for all questions was 4.2, where the overall mean for the Espark group was 3.32). Regarding technology background and implementation, Espark had a rating of 4.0 and iXL had a rating of 3.0. Overall, however, the two participants who were using iXL in their classrooms rated that adaptive learning program higher than those who use Espark with an overall mean of 4.2 for the former group 3.32 for the latter. In terms of analyzing the social validity of the two programs, it was evident, from this data, that the teachers using iXL had more confidence in the program's ability to impact student growth. The lower ratings provided by the Espark responders to the social validity survey also seem to indicate that they had less confidence in the program's ability to effectively impact student achievement.

There were some limitations of the data that needed to be addressed when analyzing and interpreting the results from this research study. First, the scope of the study was small. While

the number of student data points was larger than initially anticipated with grades 5 and 6 as the target population, it was still small with a total of seventy student data points for the student achievement part of the study (grades 5, 7, and 8). Five teacher participants made up the sample for the social validity portion of the study. The seventy available student generated data points were in line with what was expected at the onset of the study, but this was partially due to having to switch from including grades five and six to grades five, seven, and eight. By including two grades (seven and eight) instead of just using one grade (six) for the iXL group, it increased the overall available student scores from 31 to 70. A second limitation of this study is the fact that students were not actively involved in the study, just their NWEA data was used. This is a limitation as it did not allow the study to control many of the variables that come into play when examining student achievement. A third limitation regarding this study is the fact that the student population is not culturally diversified. Though demographic data of the students was not collected, the district has less than 3% of its student population represented by minority students. This limited potential discussions around the impact of the adaptive learning technology programs on different subsets of school or society's populations.

Another limitation of the student data was relying on the NWEA assessment as the pre and post assessment for the study. Despite this being a nationally normed assessment, assessing the impact of one factor (such as the use of an adaptive learning technology program) is hard to tease out from the many other factors that can contribute to or limit student progress. Another limitation related to the NWEA data was the fact that the data gained from calculating the Cohen's d did not seem to match the data provided in the NWEA Grade Report. The effect size calculation (using the .40 mark) showed that both programs had a year's worth of impact on student achievement. The NWEA projected data, however, showed that just 29% of grade 5

students, 55% of grade 7 students, and 47% of grade 8 students met their yearly projected growth. There are several factors that could be influencing the discrepancy shown in this data. The first is the fact that the NWEA assessment assumes growth and therefore is not the same identical assessment from the fall to the spring. This potentially can influence the effect size calculations which are typically done using measures that are the same for the pre and the post test. Another factor that could impact this is the fact that the mathematical areas that were addressed in the adaptive learning technology programs were not the same concepts that were assessed on the NWEA assessment. Despite both programs having the ability to align with the NWEA assessment, if teachers manually assigned tasks in the program this could move students out of alignment with the NWEA. Another consideration is again the fact that the study was conducted with a rather small population which could skew the effect size calculations and therefore, create a discrepancy between them and the projected NWEA growth scores.

The scope of the second set of data in this study was also limited by the number of participants responding to the teacher social validity survey, having just five participants. This was a known limitation given the size of the district and the number of math teachers in each of the grades that were studied. This limitation, however, narrows the potential discussions and conclusions that could be generated from the social validity side of this study. Finally, another limitation of the teacher Social Validity Survey is that it was created specifically for this study and therefore not able to be directly linked to other data points collected using the same instrument. This again limits the scope of the data to the study at hand. Despite these limitations, however, the data collected for the qualitative study confirmed the two hypotheses and contributes to a continued discussion around both the effectiveness of adaptive learning technology programs and the impact of teacher perceptions on student achievement.

Interpretations and Importance of Findings

There are two critical components addressed in this quantitative study. The first is reflected in the first research question, which asked if the two adaptive learning technology math programs (Espark and iXL) had a significant (defined as .40) effect on student achievement. To assess the impact Cohen's d or effect size, calculations were done using pre (fall NWEA assessment) and post (spring NWEA assessment) scores to determine if both Espark and iXL did or did not meet Hattie's .40 growth mark which was promoted by Hattie (2008) as being a year or more of growth. The data from the study showed that when Cohen's d was calculated, both programs had higher than a .40 effect size (Espark was .439 and iXL was .532). Students who used Espark grew on average 4.58 points from the fall to spring administration of the NWEA assessment However, iXL students grew an average of 5.30. On the surface this seems to indicate that both adaptive learning technology programs had a significant impact on student achievement. There are, however, two points of caution one must use when accepting this data and making conclusions regarding the importance of it in terms of supporting the on-going use of adaptive learning technology math programs. The first is the fact that there are many factors which contribute to student achievement in the classroom, Hattie (2008) himself lists dozens of them, and they range from teacher experience to student homelife to curriculum use. This study was not able to limit all of the variables that contribute to student success in the classroom, therefore it is not possible to draw the conclusion that these two adaptive learning technology programs were the sole source of student progress. Given the fact that the entire k-8 math program recently moved to a new curriculum and both programs were actively used in the classroom to support diversification of instruction, student enrichment and student intervention, it is likely that both the adaptive learning program and these changes had a positive impact on

student achievement gains. This is supported by the teacher's social validity data, which showed an overall positive view of the programs by the teacher in terms of the impact that they have on student growth. On the Social Validity Survey, Espark teachers have a mean average of 2.67 on the question which asked if the program had a positive impact on students mastering math skills and a 2.57 on the question which asked if the program positively impacted students' growth on the NWEA. This was even higher among the iXL respondents, where they rated the former as 4.0 and the latter as a 5.0.

The second critical component was that the use of the NWEA assessment limited the interpretation of the data in two profound ways. First, it did not allow for an analysis of specific math skills that the programs focused on and the growth of these specific skills from the fall to the spring assessment. This could have been accomplished with a study-specific assessment, but that would have lost the national norm and relevance of the assessment used. A second factor is the fact that despite making growth from fall to spring the collective growth for both groups was below the NWEA projected growth for these students based on their national norms. The mean average growth for the grade 5 (Espark users) was 4.58 and NWEA projected the mean growth for this group of students to be 9.79. For students in grade seven and eight, the mean growth was 5.30 and the NWEA expected growth for this group of students was 6.02. One of the causes of this discrepancy is the fact that the NWEA norms changed from fall to spring as the assessment accounts for expected growth. When calculating Cohen's d or effect size the pre and post assessments should be identical so that true growth can be calculated and assessed.

The data gathered from the second part of this study, related to the second research question, seems to be more reliable for interpretation. Despite having a small sample size, all invited participants responded to the ten-question social validity survey. The results of the survey

showed that the program which seemed to have the higher overall impact on student achievement (iXL), also had the higher scores on all but one of ten survey questions answered by the respondents. Despite both groups (Espark and iXl) having a mean score (for the total of the ten questions) of over 3.0, the respondents who used Espark in their classrooms -ranked had a mean response rate below a 3.0 on five of the ten questions: students in your classroom were motivated to use the adaptive learning technology program; the adaptive learning technology program was engaging for your students; the adaptive learning technology supported student growth on the NWEAs; the adaptive learning technology program had a positive impact on students' perceptions of their ability to successfully do mathematics; and the adaptive learning technology program had a direct impact on your students ability to master math related standards. The respondents who used Espark in their classrooms had a mean score higher than the respondents that used iXL in their classrooms on one question: the adaptive learning technology program required minimum background technology understanding to implement; where Espark had a mean score of 4.0 and the iXL group had a mean score of 3.0. The fact that the iXL respondents rated their perception of the program higher on 9/10 of the social validity questions paired with the fact that the student data in the iXL group was higher than that of the Espark group lends evidence to the fact that social validity could be an important factor in the success of adaptive learning technology programs in classrooms. These findings support previous studies that suggested social validity plays an important role in the outcome of an intervention (Hayak & Avidov-Ungar, 2020; Walker, 2019).

An interesting finding of the Social Validity Survey was that both groups of respondents scored student engagement and student motivation lower than other topics covered on the survey. The three Espark respondents had a mean score of 2.67 when asked about student motivation to

use the program, and the iXL respondents had a mean score of 3.5 for this same question. Though the iXL score was above the 3.0 mark, this was one of the lower scores recorded by this group. When asked about how well the programs engaged their students, the Espark group of respondents had a mean score of 2.0 and the iXL group scored this a 3.5. Again, these mean scores represented some of the lower scores provided by the groups towards the programs they were implementing. This is important because one of the critical factors promoted in the use of Espark and iXL is that they are both engaging and motivating for students (El-Sabagh, 2021; Schuetz, Biancarosa & Goode, 2018). If schools are investing in adaptive learning technology programs such as Espark and iXL to support learning through student engagement and motivation, this is an important factor to continue to explore.

Implications

Prior to this study only one other study, conducted by Hollands and Pan (2018), compared two adaptive learning technology math programs in the same site. That study found that Espark had a greater impact on student achievement than iXL, but one of the key weaknesses of that study was the fact that two different assessments were used. This study improved upon the Hollands and Pan (2018) study by using a common assessment. However, this study had slightly different results, as iXL had a larger effect size on student achievement than Espark. With both groups having greater than a .40 effect size, this study also validates the overall research which suggested that technology can have a positive impact on student achievement (Downey, 2008; Fokides, 2018; Outhwaite et. al., 2019; O'Rourke et al.,2017; Smith, 2020; Yeh et al., 2019). This study adds to the literature on adaptive technologies used in elementary and secondary schools indicating that digital math programs (such as Espark and

iXL) have a positive impact on student achievement (Donnelly, 2021; Lyons, 2020; SRI Education, 2014).

This study has implications that for those seeking to address the core problem, the persistent struggle for American students to meet expectations in the area of mathematics (Boaler & Zoido, 2016; Desilver, 2017; The Nation's Report Card, 2021). With the increase in technology spending to support student learning over the past four decades, there continues to be a need for studies like this to examine the impact that these technology programs are having on student achievement. This is especially true in the wake of Covid-19, where technology systems, like adaptive learning technology, were implemented to provide support to students in non-traditional ways. This study helped to support the concept that adaptive learning technology can increase academic performance (Kurt, 2021), the purpose of the study.

Another significant implication for this study is the fact that it is the first known study to link adaptive learning technology to both social validity and student achievement. Previous studies, such as Walker (2019), looked at teacher perception of digital math programs. Hayak and Avidov-Ungar (2020) looked at teacher perception of the implementation of technology. This research study adds to both of those discussions while also providing an avenue of discussion around how teacher perceptions influence the impact of adaptive learning technology programs on student achievement. There have been studies that have previously explored the perception of educators when implementing technology (Hayak & Avidov-Ungar, 2020). Other studies such as Walker (2019) have examined the perception of teachers when implementing a new technology-based math program. This study, however, is the only one known to examine the social validity of teachers implementing adaptive learning technology programs in the classrooms.

Another significant implication of this study as it relates to the previous body of knowledge on this topic is the link it provides to the work around student achievement analyzed by Hattie (2008). Hattie's Visible Learning Theory (2008) was selected as the theoretical foundation for this study, because it addressed the issue of improving student achievement through reflecting on both a teacher's practice and the programs that are used to support the classroom. Hattie's (2008) meta-analysis of thousands of studies regarding what supports student growth, used effect size to determine the level of impact that programs and practices had on student achievement. This research study used Hattie's effect size lens as a means of assessing the impact of Espark and iXL on student achievement. It also used social validity to link teacher reflection on the impact of adaptive learning technology programs and how this compared to the impact the programs had on student achievement. Despite concerns around the use of NWEA as the pre and post assessment, this study confirmed that effect size is a valuable tool to compare the impact of adaptive learning technology programs on student achievement. It adds to the significant body of research on what supports student achievement and also the body of work attached to Hattie's (2008) meta-analysis which formed the foundation for the Visible Learning Theory.

On a practical level, the final implication of the findings in this research study is that it provides a blueprint for use by schools to assess the impact of programming on student achievement. One area of concern identified in the literature review for this study was the lack of time, resources and attention schools pay in terms of properly assessing the impact of the programs they adopt (Coulson, 2006). This research study can provide districts with the opportunity to use basic statistical measures (like mean, standard deviation and Cohen's d) to analyze and compare the effect that programs have on student growth. At the same time this

study highlights the importance of considering teachers' perspectives on programs when assessing the impact of them.

Recommendations for Actions

These recommendations are based on the methods used in the study, the data collected, and the analysis of the data. During the research process of this study there were a number of factors that arose that can provide the basis for future action. These recommendations are grounded in research and are made first to provide RSU-A with data regarding adaptive learning programs that they are currently using as well as a blueprint for assessing and evaluating these programs and others it might adopt in the future. In addition, the recommendations are made to provide future researchers in the area of student achievement, adaptive learning technology, and social validity with the foundations for potential studies related to these topics. Overall, the recommendations are meant to provide further action on the local school level as well as further research actions for those looking to further explore student achievement, adaptive learning technology, and social validity.

Orthod (2014) explained that sample sizes that are too small or too large can influence the outcomes and consequently the decisions that are made based on the study. The student data collected from this study came from a total of 70 students, which according to Orthod (2014) allows for more interpretation as it aligns well with the effect size and mean calculations that were used in the methods section of the study. This, however, is not true for the social validity data collected regarding teacher perceptions. With just five participants, this data is less reliable, and more caution needs to be used when using the data to draw conclusions. Therefore, one of the recommendations for future action would be to increase the participant pool for the social validity part of this study.

A second recommendation from the results of this study is for the RSU A school district to look closer at the data provided before making overarching program decisions. This recommendation is made because both programs proved to have a higher than .40 effect size, but one program (iXL) outperformed the other (Espark) in terms of both student achievement and the overall perception (based on the ratings from the Social Validity Survey) of teachers using the program. The data could suggest that iXL better meets the needs of students in this district and also teachers in the district have a more positive view of the adaptive learning technology program than those using Espark. One critical piece to consider here is the age difference between those using Espark (grade 5 students) and those using iXL (grade seven and eight students). Further research is suggested regarding whether each program is better tailored towards the target audiences selected for this study. Researching this further could allow RSU-A to make a better-informed decision regarding if either program or both best suits the needs of its students and the teachers using adaptive learning technology in the district.

A third recommendation from this study for RSU-A is to further analyze how the programs it purchases are implemented and used at the classroom level. This recommendation is made because it is important for the success of any program or practice to have both teacher and student buy-in (Lee & Min, 2017). If these programs were adopted and implemented without considering that element it could prevent the programs from providing the best results possible for student success. Specifically, an area to explore would be whether staff and students at either location become accustomed to a program over time or whether it was being used to its fullest potential.

The data from this study indicated that adaptive learning technology math programs provide students with opportunities to grow academically. At the same time the social validity

data highlighted that these programs also help diversify instruction for students and provide teachers with some freedom in terms of planning and support for individual students and their academic needs. With this in mind the three above recommendations for action should be considered by both RSU-A and anyone seeking to build upon this study in the future.

Recommendations for Further Study

Through the data collected in this study, which showed that adaptive learning technology math programs have a positive impact on student achievement and that teachers implementing the programs have a positive perception of their impact a few recommendations for further study surfaced. This researcher recommends future researchers consider using a common assessment that is consistent from the pre to the post test. The NWEA assessment was used for this study because it was a common assessment that was already in place for both groups (Espark and iXL). The NWEA assessment was also used because it is a nationally normed assessment that would allow for the results of this study to be analyzed alongside a larger assessment pool. The downside of using the NWEA is that it is an adaptive assessment, which means that students do not receive the same questions. NWEA adjusts according to how students are answering the assessment. Additionally, the NWEA assessment adjusts from the fall to the spring, meaning it expects student growth and therefore the assessment is not identical from the fall administration to the spring administration. The recommendation would be to use an assessment that is consistent for each student and is the same from the pre to the post administration of the assessment. This could involve creating an assessment and possibly limiting which math skills are assessed using the adaptive learning technology math programs.

A second recommendation for further study would be looking at more than just time invested in the adaptive learning technology program. This study used 45 minutes a week using

the program as one of its qualifying characteristics for the use of student data. Both Espark and iXL log time use, but time on the program does not always mean student engagement in the tasks assigned by the programs. iXL provides more data around both time on individual questions and both programs provide data around success rate on individual skills. This might provide more accurate data and data that is able to be directly linked to specific math skills that the programs support or do not support growth in.

Exploring the use of the adaptive learning technology programs in the two school settings is a third recommendation for future research. This study did not consider how the programs (Espark and iXL) were being used in the classroom setting, only if they were being used for a certain period of time. Further research could consider adopting and using the programs more consistent between the Espark and the iXL group to provide more consistent and reliable data. Additionally, future studies could specifically look at how the different approaches to the use of adaptive learning technology impacts student achievement. In this latter case, comparisons between programs that are closely monitored by teachers versus those which are not as closely monitored could be examined. Also, one could investigate whether programs that are used for intervention or enrichment purposes are as successful as those that are adaptive for general use and support in the classroom.

A fourth recommendation for further research would be to examine the conflict between the fact that Cohen's d seemed to indicate more than a year's worth of growth for students using both Espark and iXL, whereas NWEA's projected growth data did not seem to support this. This is again linked to the idea that potentially the NWEA assessment is not a great assessment to use when calculating Cohen's d. Potentially, however, it could highlight an inconsistency in Hattie's theory (2008).

A fifth recommendation for further study associated with this quantitative study is to include a qualitative survey or interview that follows up on the Social Validity Survey. This would be an important aspect as it would allow for teachers to provide more insight into why they felt the programs they were using were or were not meeting the academic needs of their students. This research study used teacher input data related to the social validity survey alone, which allowed for general conclusions to be made. To gain more specific information related to the impact of teacher perception, interviewing teachers could provide more specific data from which to draw further conclusions. This data would also provide more insight into how changes in program implementation better support student academic growth.

A final recommendation is to continue exploring the benefits and drawbacks of using Hattie's Visible Learning Theory (2008) as a theoretical foundation for research linked to student achievement, adaptive learning technology programs, and social validity. This was the first known study to tie all these factors together and it would be important and valuable to continue to explore the benefit of using the Visible Learning Theory to explore technology implementation, its impact on student achievement, and the perception that teachers have regarding these programs and their view of their impact on student learning. Using Hattie's work as a theoretical foundation in future studies could allow for further conclusions around how applicable Visible Learning Theory (2008) is to directly assess the impact of technology programs in school settings. Hattie's Visible Learning Theory (2008) is relatively new and therefore further studies are needed to verify that it is a legitimate theoretical foundation for studies like this which are linked to student achievement and teacher social validity.

Conclusion

The 1983 Nation at Risk report highlighted a concerning trend in American education that has persisted for the past forty years; that being that American students are not meeting the expected performances in the area of mathematics (United States. National Commission on Excellence in Education, 1983). This same report also recommended to then President Reagan, that public schools should invest in technology to both support student academic achievement and to make teaching more effective and efficient (United States. National Commission on Excellence in Education, 1983). This quantitative study targeted this concern and recommendation by examining the fact that students in the United States are continuing to not meet expected grade level performance in the area of mathematics (The Nation's Report Card, n.d.). Technology has for the past decade been a solution pushed to meet the ailing performance of US students (Culp, Honey & Mandinach, 2003; Reese, 2021). More specifically, this study was seeking to examine how adaptive learning technology programs impact student achievement.

Previous studies showed evidence that such programs had a positive effect on student assessment scores (Donnelly, 2021; Hollands & Pan, 2018; Lyons, 2020; SRI Education, 2014). The purpose of this quantitative study, therefore, was to examine the academic impact that two adaptive learning technology math programs have on student achievement. At the same time, this study sought to explore the influence that teacher social validity or perception of the programs had on student achievement.

To guide this study, the following two research questions and supporting hypotheses were created:

RQ1: What is the student achievement impact (in terms of NWEA growth scores) of two adaptive learning technology programs used in a rural district in Western Maine?

H1: Both Espark and iXL will have a .40 or higher effect size on student achievement in the area of mathematics as recognized by their growth between a pre and post NWEA assessment data. The .40 effect size is considered to be a year or more's worth of academic growth in Hattie's Visible Learning Theory (Hattie, 2008).

RQ2: Does the social validity of the teachers who are using Espark and iXL in their classrooms align with the identified student achievement results?

H2: There will be a positive (mean score higher than 3) connection between the social validity of teacher's implementing the adaptive learning technology programs and the student achievement impact on the NWEA.

Both research questions were grounded in the theoretical framework of the study, which rested upon the Visible Learning Theory work of John Hattie (2008). This theory promoted the idea that through teacher reflection and the use of basic statistical analysis, educators can adequately assess the programs and practices they use, which in turn provides them with greater opportunities to adopt pedagogy that will better support student learning (Hattie, 2008).

The results of this study indicated that both adaptive learning technology programs (Espark and iXL) had a greater than .40 effect size on student achievement (as measured by a pre and post NWEA assessment). The .40 mark was used as this is what Hattie promotes to be a year's worth of growth in relationship to student achievement (Hattie, 2008). Of the two

programs, the iXL adaptive learning technology program, which had data from forty-six eligible students, outperformed the Espark group, which had data from twenty-four eligible students (the data could only be gathered from students who used the program for 45 minutes or more a week, who took both the pre and the posttest and who was enrolled in one of the two schools). The iXL group had a mean effect size of .532, where the Espark group had an effect size of .439. The ten-question social validity survey also showed that the teachers who are implementing iXL in the classrooms for this study had a more positive perception of the impact that iXL was making on their students than did the teachers using Espark. The data showed that Espark had a total mean of the questions of 3.32, where iXL had a total mean of 4.20.

These results, however, come with caution for a number of reasons. First there was a very limited pool of available participants for both the student data (just seventy-two student data points) and the social validity data (just five data points). Using the NWEA assessment helped to connect the results to national norms, but this assessment is not the same for a pre and posttest and the assessment adapts to the person taking it, meaning students were not all taking the same assessment. To better assess the impact of adaptive learning technology programs, future studies should also consider more than just time on the program, but also should include standards or skills met, scores on specific skills and how long it took students to master skills using the programs. It is also recommended that future studies follow up on the social validity survey with interviews of teachers, to gain a better insight into why they answered questions the way they did. This could provide more data around how teacher perceptions influence or impact the success of programs such as adaptive learning technology in the classroom. The results of this study confirmed both hypotheses, but also leaves room for further study and discussion related to adaptive learning technology programs, student achievement, social validity, and the use of John

Hattie's Visible Learning Theory (2008) to support studies linked to improving student proficiency.

References

Abid, N., & Akhtar, M. (2020). Relationship between academic engagement and academic achievement: An empirical evidence of secondary school students. *Journal of Educational Research*, *23(*1), 48.

Altakhayneh, B. (2020). The impact of using the lEGO education program on mathematics achievement of different levels of elementary students. *European Journal of Educational Research*, *9*(2), 603-610.

Baker, G. & Goldberg, I. (1970). The Individualized Learning System. *The Association for Supervision and Curriculum Development.* https://files.ascd.org/staticfiles/ascd/pdf/journals/ed_lead/el_197005_baker.pdf

Basir, J. & Presberg, N. (2019). *The definitive guide to ABA therapy and how it can help your child*. https://www.attentivebehavior.com/aba-therapy-guide/

Bell, T. Et al. (1983). *A nation at risk: The imperative for educational reform*. (A Report to the Nation and the Secretary of Education United States Department of Education). The National Commission on Excellence in Education. https://edreform.com/wp-content/uploads/2013/02/A_Nation_At_Risk_1983.pdf

Berrett, A. N. & Carter, N. J. (2018). Imagine math facts improves multiplication fact fluency in third-grade students. *Journal of Behavioral Education*, *27*(2), 223-239.

Best, J. & Dunlap, A. (2014). *Continuous improvement in schools and districts: Policy considerations*. McRel International. https://files.eric.ed.gov/fulltext/ED557599.pdf

Blanton, R., & Harmon, H. (2005). Building capacity for continuous improvement of math and science education in rural schools. *The Rural Educator*, *26*(2).

Bloomberg, L. & Volpe, M. (2018). *Completing your qualitative dissertation: A road map from beginning to end*. (4th Edition). Sage Publications.

Boaler, J. & Zoido, P. (2016). Why math education in the U.S. doesn't add up. *Scientific American.* https://www.scientificamerican.com/article/why-math-education-in-the-u-s-doesn-t-add-up/

Bond, M., Buntins, K., Bedenlier, S. *et al.* (2020). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. *Int J Educ Technol High Education, 17*(2). https://doi.org/10.1186/s41239-019-0176-8

Bushweller, K. (2020). How Covid-19 is shaping tech use. What that means when schools reopen. *Education Week.* https://www.edweek.org/technology/how-covid-19-is-shaping-tech-use-what-that-means-when-schools-reopen/2020/06

Capuano, N. & Caballe, S. (2020). Adaptive learning technologies. *Ai Magazine, 41*(2).

Carter, S. & Wheeler, J. (2019). *The social validity manual: Subjective evaluation of interventions* (2nd edition). Academic Press.

Cavanagh, S. (2015). Districts struggle to judge ed-tech pilot projects. *Education Week, 35(*14),

Christensen, D. (2019). *The history of the emergence of technology in education*. Classcraft. https://www.classcraft.com/blog/the-history-of-the-emergence-of-technology-in-education/

Coe, R. (2002). *It's the effect size, stupid: What effect size is and why it is important*. School of Education, University of Durham. https://f.hubspotusercontent30.net/hubfs/5191137/attachments/ebe/ESguide.pdf

Coulson, E. K. (2006). The technology balance beam. *Journal of Cases in Educational Leadership*, *9*(2), 11–16.

Corcoran, R.P. (2018) Preparing teachers' to raise students' mathematics learning. *International Journal of Science and Math Education, 16*, 1169–1185.

Corwin Visible Learning. (n.d.). www.visiblelearning.com.

Costley, K. (2014). *The positive effects of technology on teaching and student learning*. https://files.eric.ed.gov/fulltext/ED554557.pdf

Cozad, L. E., & Riccomini, P. J. (2016). Effects of digital-based math fluency interventions on learners with math difficulties: A review of the literature. *Journal of Special Education Apprenticeship, 5*(2), 1-19.

Crompton, H., Burke, D. & Lin, Y.-C. (2019). Mobile learning and student cognition: A systematic review of PK-12 research using Bloom's Taxonomy. *British Journal of Education Technology, 50*, 684-701.

Cuban, L. (1992). Computers meet classrooms: Classrooms win. *Education Week, 12(10)*, 27 & 36.

Culp K., Honey, M., & Mandinach, E., (2003). *A retrospective on twenty years of education technology policy*. U.S. Department of Education, Office of Educational Technology. https://www2.ed.gov/rschstat/eval/tech/20years.pdf

D'Angelo, C. (2018). *The impact of technology: Student engagement and success*. Press Books. https://techandcurriculum.pressbooks.com/chapter/engagement-and-success/

De, R., Pandey, N., Pal, A. (2020). Impact of digital surge during Covid-19 pandemic: A viewpoint on research and practice. *International Journal of Information Management, 55*.

Desilver, D. (2017). U.S. students' academic achievement still lags that of their peers in many other countries. *PEW Research Center.* https://www.pewresearch.org/fact-tank/2017/02/15/u-s-students-internationally-math-science/

Donnelly, S. (2021). *A study of the effectiveness of iXL math online software on student achievement in an urban middle school*. (Dissertation). https://scholar.stjohns.edu/cgi/viewcontent.cgiarticle=1159&context=theses_dissertation

Dore, R., Justice, L., Mills, A., Narui, M., & Welch, K. (2021). Virtual kindergarten readiness programming for preschool-aged children: Feasibility, social validity, and preliminary impacts. *Early Education and Development, 32(6),* 903-922. DOI: 10.1080/10409289.2021.1919041

Downey, C. J., et al. (2008) *50 Ways to Close the Achievement Gap*, SAGE Publications.

El-Sabagh, H. (2021). Adaptive e-learning environment based on learning styles and its impact on development students' engagement. *International Journal of Educational Technology in Higher Education,18*(53).

Epstein, B. (2021). *We have no clue how much the U.S. spends on edtech. But it is at least 2x what we thought.* LinkedIn. https://www.linkedin.com/pulse/we-have-clue-how-much-us-spends-edtech-its-least-2x-what-bart-epstein

Esparklearning. (n.d.) *Frequently asked questions*. https://www.esparklearning.com/faqs

Ferguson, C., (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40*(5), 532-538

Fisher, D., Frey, N., & Hattie, J. (2016). *Visible Learning for literacy: Implementing practices that work best to accelerate student learning*. Corwin.

Fokides, E. (2018). Digital educational games and mathematics. Results of a case study in primary school settings. *Educ Inf Technol, 23,* 851–867.

Ganimian, A, Vegas, E. & Hess, F. (2020). *Realizing the promise: How can education technology improve learning for all?* Brookings. https://www.brookings.edu/essay/realizing-the-promise-how-can-education-technology-improve-learning-for-all/

Globenewswire. (2021). *Smart learning marked worth $56.5 billion by 2024-Proliferation of connected devices in the education sector drives growth*. https://www.globenewswire.com/news-release/2020/02/05/1980388/0/en/Smart-Learning-Market-Worth-56-5-Billion-by-2024-Proliferation-of-Connected-Devices-in-the-Education-Sector-Drives-Growth.html

Gordon, N., & Reber, S. (2020). Federal aid to school districts during the Covid-19 Recession. *National Tax Journal, 73*(3), 781+.

Gordon, N. (2018). *What works vs. what we can evaluate*. Brookings. https://www.brookings.edu/research/what-works-vs-what-we-can-evaluate/

Harris, J., Al-Bataineh, M., Al-Bataineh, A. (2016). One to one technology and its effect on student achievement and motivation. *Contemporary Educational Technology, 7(4)*, 368-381.

Hattie, J. (2018). *Hattie ranking: 252 influences and effect sizes related to student achievement*. VisableLearning.org. https://visible-learning.org/hattie-ranking-influences-effect-sizes-learning-achievement[1]

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hattie, J. (2008). *Visible learning*. London: Routledge.

Hattie, J., Masters, D., Birch, K. (2016). *Visible learning in action: International case studies of impact*. Routledge.

Haven, D. (2014). Measuring student engagement with learning technology. *New Schools Venture Fund.* https://nearpod.com/page/pdf/Measuring-Student-Engagement-with-Learning-Technology.pdf

Hayak, M. & Avidov-Ungar, O. (2020). The integration of digital game-based learning into the instruction: Teachers' perceptions at different career stages. *Tech Trends, 64*(4).

Heathers, G. (1977). *A working definition of individualized instruction.* The Association for Supervision and Curriculum Development. https://files.ascd.org/staticfiles/ascd/pdf/journals/ed_lead/el_197702_heathers.pdf

Herald, B. (2021). Schools are flush with stimulus money. Will they waste it on unproven technology money? *Education Week.* https://www.edweek.org/technology/schools-are-flush-with-stimulus-money-will-they-waste-it-on-unproven-technology/2021/04

Herman, K.C., Reinke, W.M., Dong, N. & Bradshaw, C. P. (2020). Can effective classroom behavior management increase student achievement in middle school? Findings from a group randomized trial. *Journal of Educational Psychology.*

Hilton, A. (2018). Engaging primary school students in mathematics: Can iPads make a difference? *International Journal of Science and Math Education, 16,*145-165.

Hinds, D. (2017). An evaluation of one district's Think Through math program and curriculum implementation. (Dissertation). https://digitalcommons.nl.edu/diss/263/

History.com. (n.d.).

Hollands, F. & Pan, Y. (2018). Evaluating digital math tools in the field. *Middle Grades Review. 4*(1).

iXl.com. (n.d.).

Institute of Educational Sciences. (2021). *Report on the condition of education*. Institute of Education Sciences. https://nces.ed.gov/pubs2021/2021144.pdf

Jameson, J. M., Thompson, V., Manuele, G., Smith, D., Egan, H. & Moore, T. (2012). Using an iTouch to teach core curriculum words and definitions: Efficacy and social validity. *Journal of Special Education Technology, 27*(3), 41–54. https://doi.org/10.1177/016264341202700304

Jimenez, B. A. & Besaw, J. (2020). Building early numeracy through virtual manipulatives for students with intellectual disability and autism. *Education and Training in Autism and Developmental Disabilities, 55(1)*, 28-44.

Johnson, D. (2017). *The role of teachers in motivating students to learn*. https://files.eric.ed.gov/fulltext/EJ1230415.pdf

Jordan, A. & Romer, D. (2014). *Media and the wellbeing of children and adolescents*. Oxford.

Koba, M. (2015). Education tech funding soars - but is it working in the classrooms? *Forbes*. https://fortune.com/2015/04/28/education-tech-funding-soars-but-is-it-working-in-the-classroom/

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., Liu, J. (2020). Projecting the potential impact of Covid-19 school closures on academic achievement. *Educational Research, 49*(8), 549-565.

Kul, U., Celik, S., Aksu, Z. (2018). The impact of educational material use on mathematics achievement: A meta-analysis. *International Journal of Instruction, 11(4)*, 303-324.

Kurt, S. (2021). Adaptive learning; What is it, what are the benefits, and how does it work? *Educational Technology*. https://educationaltechnology.net/adaptive-learning-what-is-it-what-are-its-benefits-and-how-does-it-work/

Kurvinen, E., Erkki, K., Laakso, M., & Salakoski, T. (2020). Long term effects on technology enhanced learning: The use of weekly digital lessons in mathematics. *Informatics in Education, 19(1)*, 51-75.

Larking, K., & Jorgensen, R. (2016). I hate maths: Why do we need to do maths? Using iPad video diaries to investigate attitudes and emotions towards mathematics in year 3 and year 6 students. *International Journal of Science and Math Education, 14*, 925 - 944.

Learning Policy Institute. (2018). *Research shows that when it comes to student achievement, money matters*. https://learningpolicyinstitute.org/press-release/research-shows-student-achievement-money-matters

Lee, S.W. & Min, S. (2017). Riding the implementation curve: Teacher buy-in and student academic growth under school comprehensive school reform programs. *The Elementary School Journal, 117*(3).

Lei, H., Cui, Y., & Zhou, W. (2018). Relationships between student engagement and academic achievement: A meta-analysis. *Social Behavior and Personality, 46(3)*, 517-528.

Lempinen, E. (2020). The pandemic could open the door to new technology-and dramatic innovation-in education. *Berkeley News*. https://news.berkeley.edu/2020/05/27/the-pandemic-could-open-a-door-to-new-technology-and-dramatic-innovation-in-education/

Liberman, M. & Ujifusa, A. (2021). Everything you need to know about schools and Covid relief funds. *EdWeek*. https://www.edweek.org/policy-politics/everything-you-need-to-know-about-schools-and-covid-relief-funds/2021/09

Lumen. (n.d.). https://courses.lumenlearning.com/informationliteracy/ chapter/personal-compute

Lyons, C. (2020). *Identifying opportunities and challenges of using Prodigy Math in a sixth grade math classroom*. (dissertation). https://sophia.stkate.edu/cgi/viewcontent.cgi?article=1372&context=maed

Madsen, A., Sayre, E. & McKagan, S. (2016). Effect size: What it is and when and how should I use it? *PhysPort*. https://www.physport.org/recommendations/Entry.cfm?ID=93385

Maine Department of Education. (n.d.). https://www.maine.gov/doe/Testing_Accountability/MECAS/results

Maine Department of Education. (n.d.). https://www.maine.gov/doe/learning/ltt/mlti

Maine Department of Education (n.d.). *MoMEntum K-3 literacy pilot overview*. https://www.maine.gov/ doe/learning/MoMEntum/overview

Maxwell, S., Reynolds, K., Lee, E., Subasic, E., & Bromhead, D. (2017). The impact of school climate and school identification on academic achievement: Multilevel modeling with student and teacher data. *Frontiers in Psychology*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02069/full

Mavridis, A., Katmada, A. & Tsiatsos, T. (2017). Impact of online flexible games on students' attitude towards mathematics. *Education Tech Research Dev.* 65, 1451–1470.

McGraw Hill. (n.d.). https://www.mheducation.com/prek-12/program/number-worlds-2015/MKTSP-TIG05M0.html

Morris, R. & Parker, L. C. (2014). Examining the connection between classroom technology and student engagement. *Journal of Teaching and Learning With Technology*, *3*(1), 1–15. https://doi.org/10.14434/jotlt.v3n1.4720

Muir, M., Knezek, G., & Christensen, R. (2004, November). The power of one to one: early findings from the Maine Learning Technology Initiative. *Learning & Leading with Technology*, *32*(3), 6+.

NAEP. (2019). Retrieved from: https://nces.ed.gov/nationsreportcard/?src=ft

Natriello, G. (2017). The adaptive learning landscape. *Teachers College Record*. https://www.tcrecord.org/Content.asp?ContentId=21830

Nelson, N., Fien, H., Doabler, C., & Clarke, B. (2016). Considerations for realizing the promos of educational gaming technology. *Teaching Exceptional Children, 48(6)*, 293-300.

Newcomb, T. (2020). *Educators spend $13 billion a year on ed tech for K-12 classrooms*. EdTech Genome Project is looking to help them make better choices. *LA School Report*. http://laschoolreport.com/educators-spend-13-billion-a-year-on-ed-tech-for-k-12-classrooms-edtech-genome-project-is-looking-to-help-them-make-better-choices/

NWEA. (n.d.) https://www.nwea.org/normative-data-rit-scores/

Office of Civil Rights. (2021) *Education in a Pandemic: The desperate impact of Covid 19 on America's Students*. https://www2.ed.gov/about/offices/list/ocr/docs/20210608-impacts-of-covid19.pdf

Office of Education Technology. (2017, January). Reimagining the role of technology in education: 2017 technology plan update. https://tech.ed.gov/netp/

Office of Elementary and Secondary Education. (n.d.). https://oese.ed.gov/offices/ educationstabilization-fund/ elementary- secondary-school-emergency-relief-fund/

Oremus, W. (2015). No more pencils, no more books: Artificially intelligent software is replacing the textbook-and reshaping American education. *Slate*. http://www.slate.com/articles/technology/technology/2015/10/adaptive_learning_software_is_replacing_textbooks_and_upending_american.html

Orosco, M. (2016). Measuring elementary student's mathematics motivation: A validity study. *International Journal of Science and Mathematics Education, 14*, 945-958.

O'Rourke, J., Main, S., & Hill, S. M. (2017). Commercially available digital game technology in the classroom: Improving automaticity in mental-maths in primary-aged students. *Australian Journal of Teacher Education, 42*(10), 50-70.

Orthod, J. (2014). How sample size influences research outcomes. *Dental Press Journal of Orthodontics, 19*(4).

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.

Power, R. (2017). *Technology and the curriculum*. PressBooks.

Peterson, P., Woessman, L., Hanushek, E, Lacastra-Anadon, C. (2011). *Globally challenged, are U.S. students ready to compete*? https://files.eric.ed.gov/fulltext/ED526954.pdf

Project-redcap.org. (n.d.).

Ravitch, S. & Riggan, M. (2017). *Reason and rigor: How conceptual frameworks guide research*. Sage.

Reese, A. (2021). *The effects of technology on student engagement and academic success*. (Dissertation). https://nwcommons.nwciowa.edu/cgi/viewcontent.cgi?article=1297&context=education_masters

Romer, T. (2018). A critique of John Hattie's theory of visible learning. *Education Philosophy and Theory, 51*(6), 587-598.

Rourke, J., Main, S., & Hill, S. (2017). Commercially available digital game technology in the classroom: Improving automaticity in mental-maths in primary-aged students. *Australian Journal of Teacher Education, 42(10),* 50-70.

Rothwell, J. (2016). The declining productivity of education. *Brookings.* https://www.brookings.edu/blog/social-mobility-memos/2016/12/23/the-declining-productivity-of-education/

Runco, M.A., (2011). Social validity. *Science Direct*. https://www.sciencedirect.com/topics/nursing-and-health-professions/social-validity

Sangra, A., Vlachopoulos, D. & Cabrera, N. (2012). Building an inclusive definition of E-Learning: An approach to the conceptual framework. *The International Review of Research in Open and Distant Learning, 13(2).*

Schafer, T. & Schwarz, M. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2019.00813

Schuetz, R., Biancarosa, J. & Goode, J. (2018). Is technology the answer? Investigating students'engagement in math. *Journal of Research on Technology in Education, 50*(4), 318-332.

Schwahn, C. & McGarvey, B. (2012). *Inevitable: Mass customized learning*. Chuck Schwahn and Bea McGarvey Publishing.

Sheninger, E. (2019). *Digital leadership: Changing paradigms for changing times*. (2nd ed.). Corwin Publishing.

Skakman, K, Wogan, D., Rodriguez, S., Boyce, J. & Shaver, D. (2017). Continuous improvement in education: A toolkit for school and districts. *National Center for Education Statistics*. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=REL2021014

Shannon, S. & Bylsma, P. (2007). *Nine characteristics of high performing schools*. https://files.eric.ed.gov/fulltext/ED499819.pdf

Shechtman, N., Roschelle, J., Feng, M. & Singleton, C. (2019). An efficacy study of a digital core curriculum for grade 5 math. *American Education Research Association, 5*(2).

Smith, S. T. (2020). *A mixed methods case study investigating the correlation between technology integration and student achievement and teacher perceptions of technology in an elementary school* (Order No. 28315757). Available from ProQuest One Academic. (2512360981).

Smith, S. (2020). A mixed methods case study investigating the correlation between technology integration and student achievement and teacher perceptions of technology in an elementary school (Dissertation). https://www.proquest.com/openview/6b0d06c4d59a029e66b86f69b7f38cc0/1?pq-origsite=gscholar&cbl=18750&diss=y

Spector, M., Merrill, D., Elen, J., & Bishop, M.J. (2014). *Handbook of research on educational communications and technology*. Springer.

SRI Education. (2014). *Research on the use of Khan Academy in schools*. https://www.sri.com/wp-content/uploads/pdf/2014-03-07_implementation_briefing.pdf

Stanford, P., Crowe, M.W., & Flice, H. (2010). Differentiating with technology. *Teaching Exceptional Children-Plus, 6.*

Stern, J. (2015). Enhancing learning through differentiated technology. *Edutopia*. https://www.edutopia.org/blog/enhanced-learning-through-differentiated-technology-julie-stern

Stepman, I. (2018). Report card on American education: 22nd edition. *American Legislative Exchange Council*. https://www.alec.org/publication/report-card-on-american-education-22nd-edition/

Stetted, M. (2018). The use of technology to assist school-aged students with high incidence special needs in reading. *Education Sciences, 8(61).*

The Nation's Report Card. (n.d.). www.nationsreportcard.gov

Terhart, E. (2011). Has John Hattie really found the holy grail of research on teaching? an extended review of "visible learning". *Journal of Curriculum Studies, 43*(3), 425-438.

Tichnor-Wagner, A., Allen, D., Socol, A. R., Cohen-Vogel, L., Rutledge, S., & Xing, Q. W. (2018). Studying implementation within a continuous-improvement process: What happens when we design with adaptations in mind? *Teachers College Record, 120*(5), 1–52.

Tze, V., Daniels, L., Klassen, R. (2016). Evaluating the relationship between boredom and academic outcomes: A meta-analysis. *Educational Psychology Review, 28, 119-144).*

United States Census Bureau. (n.d.). https://www.census.gov

United States National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform: A report to the Nation and the Secretary of Education, United States Department of Education. Washington, D.C.: National Commission on Excellence in Education: [Superintendent of Documents, U.S. Government Printing Office distributor]

Vancel, S., Missall, K., & Bruhn, A. (2016). Teacher ratings of the social validity of schoolwide positive behavior interventions and supports: A comparison of school groups. *Preventing school Failure: Alternative education for children and youth, 60*(4), 320-328**,** DOI: 10.1080/1045988X.2016.1157784

VanderArk, T., & Schneider, C. (2012). How digital learning contributes to deeper learning. *Getting Smart*. https://www.gettingsmart.com/wp-content/uploads/2012/12/Digital-Learning-Deeper-Learning-Full-White-Paper.pdf

Walker, L. (2019). *Elementary teacher perceptions of Eureka Math*. (Dissertation). https://www.researchgate.net/publication/340903992_The_Integration_of_Digital_GamBased_Learning_into_the_Instruction_Teachers'_Perceptions_at_Different_Career_Stag

Watson-Huggins, J., Trotman, S. (2019). Gamification and motivation to learn math using technology. *Quarterly Review of Distance Education, 20(4).*

West, M. & Lake, R. (2021). How much have students missed academically because of the pandemic: A review of the evidence to date. Center on Revitalizing Public Education. https://files.eric.ed.gov/fulltext/ED614130.pdf

Yang, S., Tian, H., Sun, L. & Yu, X. (2019). From one-size-fits-all teaching to adaptive learning: The crisis and solution of education in the era of AI. *Journal of Physics: Conference Series, 1237* (4).

Yeh, C. Y. C., Cheng, H. N. H., Chen, Z., Liao, C. C. Y. & Chan, T. (2019). Enhancing achievement and interest in mathematics learning through math-island. *Research and Practice in Technology Enhanced Learning, 14*, 1-19.

Zhao, Y. & Watterston, J. (2021). The changes we need: Education post COVID-19. *Journal of Educ Change*, 22, 3–12.

Appendix A

Teacher Social Validity Survey

Teacher Social Validity Survey		
For the questions below please use the following scoring guide: 1 = Completely Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Completely Agree		
#	Question	Scale
1.	Students in your classroom were motivated to use the adaptive learning technology program.	1-2-3-4-5
2.	The adaptive learning technology program was engaging for your students.	1-2-3-4-5
3.	The adaptive learning technology program provided individualized instruction at appropriate levels for the students in your class.	1-2-3-4-5
4.	The adaptive learning technology program was easy to implement into the structure of your classroom.	1-2-3-4-5
5.	The adaptive learning technology program required minimum background technology understanding to implement.	1-2-3-4-5
6.	The students did not need a lot of technical assistance when using the adaptive learning technology program.	1-2-3-4-5
7.	The use of the adaptive learning technology program in your class allowed you to diversify your instruction.	1-2-3-4-5
8.	The adaptive learning technology supported student growth on the NWEAs.	1-2-3-4-5
9.	The adaptive learning technology program had a positive impact on students' perceptions of their ability to successfully do mathematics.	1-2-3-4-5
10.	The adaptive learning technology program had a direct impact on your students' ability to master math related standards.	1-2-3-4-5

Appendix B

IRB Exemption Determination & Amendment Letter

UNE UNIVERSITY OF NEW ENGLAND
INNOVATION FOR A HEALTHIER PLANET

Institutional Review Board
Julie Longua Peterson, Chair

Biddeford Campus
11 Hills Beach Road
Biddeford, ME 04005
(207) 602-2244 T
(207) 602-5905 F

Portland Campus
716 Stevens Avenue
Portland, ME 04103

DATE OF LETTER:	April 25, 2022
PRINCIPAL INVESTIGATOR:	Charles Swan
FACULTY ADVISOR:	Debra Welkley, EdD
PROJECT NUMBER:	0322-18
PROJECT TITLE:	Assessing the Academic Impact of Two Adaptive Learning Technology Math Programs Using Hattie's Visible Learning Theory
SUBMISSION TYPE:	Exempt Project
SUBMISSION DATE:	3/23/2022
ACTION:	Determination of Exempt Status
DECISION DATE:	3/25/2022
REVIEW CATEGORY:	Exemption Category # 2 & 4

The UNE Institutional Review Board (IRB) for the Protection of Human Subjects has reviewed the materials submitted in connection with the above referenced project and has determined that the proposed work is exempt from IRB review and oversight as defined by 45 CFR 46.104.

Additional IRB review is not required for this project as submitted. However, if any changes to the design of the study are contemplated (e.g., revision to the protocol, data collection instruments, interview/survey questions, recruitment materials, participant information sheet, and/or other IRB-reviewed documents), the Principal Investigator must submit an amendment to the IRB to ensure the requested change(s) will not alter the exempt status of the project.

Please feel free to contact me at (207) 602-2244 or irb@une.edu with any questions.

Best Regards,

Bob Kennedy, MS
Director, Research Integrity

From: UNE Institutional Review Board <irb@une.edu>
Sent: Tuesday, May 31, 2022 9:42 AM
To: Charles Swan <cswan2@une.edu>; Debra Welkley <dwelkley@une.edu>
Cc: Deborah Jameson <djameson1@une.edu>
Subject: 0322-18-A - Swan - Exempt Amendment Determination Letter

Hi Charles & Debra,

RE: IRB # 0322-18-A; Assessing the Academic Impact of Two Adaptive Learning Technology Math Programs Using Hattie's Visible Learning Theory

The UNE Institutional Review Board (IRB) has reviewed the above-referenced amendment and has determined that this project continues to be exempt from IRB review and oversight. Please review the attached Exemption Determination Letter carefully and note any conditions you may be required to comply with.

Please let me know if there are any questions or concerns.

Best Regards,

Bob Kennedy, MS
Director of Research Integrity
University of New England
rkennedy1@une.edu
(207) 602-2244

www.une.edu

UNE UNIVERSITY OF NEW ENGLAND
INNOVATION FOR A HEALTHIER PLANET