

A CAUSAL-COMPARATIVE STUDY ON THE EFFICACY OF INTELLIGENT TUTORING  
SYSTEMS ON MIDDLE-GRADE MATH ACHIEVEMENT

by

Kevin Lamar Rholetter

Liberty University

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University

2020

A CAUSAL-COMPARATIVE STUDY ON THE EFFICACY OF INTELLIGENT TUTORING  
SYSTEMS ON MIDDLE-GRADE MATH ACHIEVEMENT

by Kevin Lamar Rholetter

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University, Lynchburg, VA

2020

APPROVED BY:

Meredith Park, Ed.D. Committee Chair

Gina Thomason, Ed.D. Committee Member

Scott Watson, PhD, Director of Quantitative Research

## ABSTRACT

This study is a quantitative examination of intelligent tutoring systems in two similar suburban middle schools (grades 6-8) in the Southeastern United States. More specifically, it is a causal-comparative study purposed with examining the efficacy of intelligent tutoring systems as they relate to math achievement for students at two similar middle schools in the Midlands of South Carolina. The independent variable, use of an intelligent tutoring system in math instruction, is defined as the supplementary use of two intelligent tutoring systems, Pearson's Math Digits and IXL, for math instruction. The dependent variable is math achievement as determined by the Measures of Academic Progress (MAP) SC 6+Math test. The student data examined is archived MAP SC 6+ Math scores from the 2017-2018 school year. A one-way ANCOVA was used to compare the mean achievement gain scores of both groups, students whose math instruction included intelligent tutoring systems and students whose math instruction did not include intelligent tutoring systems, to establish whether or not there was any statistically significant difference between the adjusted population means of the two independent groups. The results showed that the adjusted mean of posttest scores of students who did not receive math instruction that involved an intelligent tutoring system were significantly higher than those who did.

*Keywords:* intelligent tutoring systems, adaptive computer-assisted instruction, personalized learning, adaptive learning technologies

**Copyright Page**

### **Dedication**

I dedicate this work to my family. To my loving wife, Pamela, thank you for your ceaseless support and love. I could not ask for a better wife and best friend. To my two beautiful daughters, Madeline and Kennedy, thank you for being the wonderful people you are and for understanding and the long hours sometimes necessary to complete this process. To my parents, Jeff and Debbie, thank you for your high expectations, your love, your prayers, and the Christian environment in which I was raised.

### **Acknowledgments**

I am expressly grateful for the educators who have knowingly or not been a part of this journey. To Mr. Joe Bryant and Mr. Randy McManamay, thank you for showing me that men could be great teachers long before I knew that is what I wanted to be. To Dr. Wimberely, thank you for planting the seed that would become the focal point of this work. Finally, to Dr. Park and Dr. Thomason, I am grateful that God has blessed me with your guidance and prayers. I could not have asked for a better chair and committee member.

## Table of Contents

<u>ABSTRACT</u> .....	3
<u>Copyright Page (Optional)</u> .....	4
<u>Dedication (Optional)</u> .....	5
<u>Acknowledgments (Optional)</u> .....	6
<u>List of Tables</u> .....	9
<u>List of Figures</u> .....	10
<u>List of Abbreviations</u> .....	11
<u>CHAPTER ONE: INTRODUCTION</u> .....	12
<u>Overview</u> .....	12
<u>Background</u> .....	12
<u>Problem Statement</u> .....	13
<u>Purpose Statement</u> .....	14
<u>Significance of the Study</u> .....	15
<u>Research Question(s)</u> .....	16
<u>Definitions</u> .....	16
<u>CHAPTER TWO: LITERATURE REVIEW</u> .....	17
<u>Overview</u> .....	17
<u>Conceptual or Theoretical Framework</u> .....	17
<u>Related Literature</u> .....	25
<u>Summary</u> .....	46
<u>CHAPTER THREE: METHODS</u> .....	47
<u>Overview</u> .....	47

<u>Design</u> .....	47
<u>Research Question(s)</u> .....	48
<u>Hypothesis(es)</u> .....	49
<u>Participants and Setting</u> .....	49
<u>Instrumentation</u> .....	50
<u>Procedures</u> .....	53
<u>Data Analysis</u> .....	54
<u>CHAPTER FOUR: FINDINGS</u> .....	57
<u>Overview</u> .....	57
<u>Research Question(s)</u> .....	57
<u>Null Hypothesis(es)</u> .....	57
<u>Descriptive Statistics</u> .....	57
<u>Results</u> .....	59
<u>Hypothesis</u> .....	59
<u>CHAPTER FIVE: CONCLUSIONS</u> .....	66
<u>Overview</u> .....	66
<u>Discussion</u> .....	66
<u>Implications</u> .....	69
<u>Limitations</u> .....	71
<u>Recommendations for Future Research</u> .....	72
<u>REFERENCES</u> .....	74
<u>APPENDIX or APPENDICES</u> .....	93



### List of Tables

Table 1: Means Report.....	58
Table 2: Descriptive Statistics.....	59
Table 3: Estimates .....	59
Table 4: Test of Between-Subjects Effects.....	62
Table 5: Tests of Normality.....	62
Table 6: Levene's Test of Equality of Error Variances.....	64
Table 7: Test of Between-Subject Effects .....	65
Table 8: Pairwise Comparisons.....	65

**List of Figures**

Figure 1: Grouped Scatter of Posttest (Spring MAP) by Pretest (Fall MAP) by ITS Level.....61

Figure 2: Simple Scatterplot of Standardized Residual for Posttest by Predicted Value.....63

**List of Abbreviations**

Measures of Academic Progress (MAP)

Analysis of Covariance (ANCOVA)

Zone of Proximal Development (ZPD)

English for Speakers of Other Languages (ESOL)

Intelligent Tutoring System (s) (ITS)

Computer-Aided Instruction (CAI)

Northwest Evaluation Association (NWEA)

Computer Adaptive Test (CAT)

Rasch Unit (RIT)

Confirmatory Factor Analysis (CFA)

Non-Intelligent Tutoring System (NITS)

Program for International Student Assessment (PISA)

National Assessment of Educational Progress (NAEP)

## **CHAPTER ONE: INTRODUCTION**

### **Overview**

In recent decades, technology, including intelligent learning systems, has proliferated in K-12 education as a means of personalizing learning for individual students. The following chapter will provide an introduction to personalized learning with an emphasis on intelligent tutoring systems as a means of achieving it. The first section will provide some background on both personalized learning and intelligent tutoring systems, specifically their historical, social, and conceptual underpinnings. The next section will synthesis a problem found within current research on intelligent tutoring systems. Finally, the remaining sections will delineate the purpose of this study, its significance, the accompanying research question, and any pertinent definitions.

### **Background**

Personalized learning is a way to address the multitude of differences that exist in how people learn. Attributes of personalized learning often include student choice based on interest; minimal or relaxed sequencing of topics, concepts or skills (prerequisite dependencies excluded); pretests for diagnostic purposes; posttests; and immediate and customized feedback to promote reflection and the correction of misconceptions (Gudivada, 2017). Basham, Hall, Carter, and Stahl (2016) outlined a definition for personalized learning that not only includes tailoring instruction to each learner's strengths and needs but also one that permits learners some ownership and control of their learning as a means of gaining mastery. Bingham, Pane, Steiner, and Hamilton (2018) recently outlined four critical components of personalized learning - learner profiles that highlight students' strengths and weaknesses; tailored learning paths that adjust to learners' goals, progress, and motivations; flexible learning environments; and competency-based

progression. Paive, Ferreira, and Frade (2017) added that personalized learning is often self-paced and designed around teaching modules with very specific learning objectives and outcomes.

### **Problem Statement**

Personalized learning in education seeks to adapt instructional approaches and learning experiences to each student's interests, strengths, weaknesses, culture, and learning styles. Considering the overwhelming logistics of such an endeavor and the increasing diversification that exists in today's K-12 classrooms, many school leaders have turned to technology, specifically intelligent tutoring systems as a means of providing at least some degree of instructional personalization (Lee, Huh, Lin, & Reigeluth, 2018). Over the past five decades, computer-assisted learning has evolved, and with this evolution has come intelligent tutoring systems. These advanced, computer-assisted instructional tools boast a complicated array of customization on a variety of levels, complete with immediate feedback, often requiring little teacher intervention. Furthermore, they "model learners' psychological states to provide individualized instruction...for diverse subject areas (e.g., algebra, medicine, law, reading) to help learners acquire domain-specific, cognitive and metacognitive knowledge" (Ma et al., p. 901) and can be used as either a primary or supplementary means of instruction for a variety of learners.

Due in part to their potential to revolutionize modern education, intelligent tutoring systems have regularly been compared to non-intelligent tutoring system learning environments and evaluated for effectiveness. Several meta-analyses conducted in recent years have yielded mixed results, and a cursory search of databases yields evaluative research on a wide array of intelligent tutoring system tools, the intended students and domains of which run the gamut

(Kulik & Fletcher, 2016; Ma, Adesope, & Nesbit, 2014). Many of these studies, however, focus on very limited intervals of time, feature a relatively small sampling, and often do not include K-12 populations. As such, a gap in the literature exists. The problem is few to no studies have been conducted on a fairly large K-12 population over an extended period of time comparing achievement in a particular cognitive domain, e.g., math, between populations utilizing intelligent tutoring systems and populations not using intelligent tutoring systems.

### **Purpose Statement**

The purpose of this study is to examine the math achievement of students in two similar suburban middle schools (grades 6-8) in the Southeastern United States. Both schools have similar demographic and socioeconomic makeups, each serving approximately 1000 students. School A is 73% African American, 21% white, 4% Hispanic, and 2% other, and more than half of its students receive subsidized breakfast and lunch. School B is 87% African American, 6% Caucasian, 3% Hispanic, and 4% other, and sixty-five percent of its students are from low-income families and receive subsidized lunch. One school will have employed intelligent tutoring systems for math instruction and one will not. As such, the dependent variable in this study is math achievement, and the independent variable is the use of intelligent tutoring for math instruction. A quantitative approach to this study is appropriate due to the fact that scientific inquiry is being employed to examine the differences between the two groups (Rovai, Baker, & Ponton, 2013), middle school students whose math instruction included the use of an intelligent tutoring system and middle school students whose math instruction did not include the use of an intelligent tutoring system. Moreover, a causal-comparative design is particularly fitting because this study seeks to explore differences that already exist between the two groups

(Gall, Gall, & Borg, 2007), math achievement as measured by the Measures of Academic Progress (MAP) Math SC 6+.

### **Significance of the Study**

Personalization within educational contexts delineates an individualized approach to instruction that deviates from the one-size-fits-all approach to teaching that has dominated education for decades. A personalized approach to learning not only gives a level of control to learners, but in doing so, cultivates academic efficacy, awareness, reflection, and motivation (Chatti & Muslim, 2019). Additionally, personalized learning dictates that student learning experiences are customized to their specific abilities, goals, and interests (Childress & Benson, 2014), as well as style, content, background knowledge, pace, and even location (Hopkins, 2019). Personalization within classrooms is driven by more than simply aligning pedagogy to learner progression; rather, it is guided by a moral concern to promote life-long learning in addition to academic achievement (Hopkins, 2019).

Within traditional classrooms, teachers are unable to personalize learning for each student because customization within an interdependent system necessitates “a complete redesign of the entire product or service every time” (Christensen et al., 2011, p. 31). Because it is impractical for all students to have their own respective teachers, many in education are looking to technology for answers. The U.S. National Education Technology Plan recognized the importance of technology in providing students personalized learning experiences that included continuous evaluation of student learning, feedback, and record-keeping (U.S. Department of Education, 2010). As such, the significance of this study lies in its investigation and evaluation of instructional tools capable of individualizing and customizing instruction for diverse populations of students – an undertaking considered unlikely if not impossible in traditional

classrooms. More specifically, this study endeavors to examine the efficacy of intelligent tutoring systems concerning math achievement and contribute to the current literature related to adaptive learning pedagogies within middle school math classrooms.

### **Research Question(s)**

**RQ1:** Is there a difference in the math achievement of middle school students whose math instruction includes an intelligent tutoring system and middle school students whose math instruction does not include an intelligent tutoring system?

### **Definitions**

1. *Adaptive learning technologies* - Technologies that cater to learning styles, cognitive abilities, affective states and learning context (Kinshuk, 2015).
2. *Intelligent Tutoring Systems (ITS)* - Advanced computer-assisted instructional tools capable of customizing instruction on a variety of levels (Ma, Adesope, Nesbit, & Liu, 2014).
3. *Personalized Learning*- A way to address the multitude of differences that exist in how people learn. Attributes of personalized learning often include student choice based on interest; minimal or relaxed sequencing of topics, concepts or skills (prerequisite dependencies excluded); pretests for diagnostic purposes; posttests; and immediate and customized feedback to promote reflection and the correction of misconceptions (Gudivada, 2017).



## **CHAPTER TWO: LITERATURE REVIEW**

### **Overview**

Intelligent tutoring systems are designed to personalize instruction for individual students and therefore have the potential to impact education in a positive way. Though much research exists comparing these adaptive learning systems to traditional teaching environments, most reports have been limited to small sample sizes and cover only brief intervals of time. Moreover, many studies on these adaptive learning systems exclude K-12 populations and middle school populations in particular. The following chapter offers a synopsis of the conceptual underpinnings of personalized instruction, as well as related literature on intelligent tutoring systems as a vehicle for achieving some degree of learning personalization.

### **Conceptual or Theoretical Framework**

Chukwuedo and Uko-Aviomoh (2015) distinguished between conceptual and theoretical frameworks by explaining that conceptual frameworks are shaped when ideas are connected in studies to clarify the variables and findings in research. Moreover, Chukwuedo and Uko-Aviomoh (2015) asserted that the essential organization of the conceptual framework should include all pertinent variables related to the study, incorporate essential constructs, establish the problem and purpose of the research, and be connected to the study's conclusions and findings. Maxwell (2012) added that when considering a conceptual framework, it is essential for researchers to remember that it functions as a structural model for the proposed investigation, specifically what is happening within the study, the interaction among variables, and the phenomena at play. The overarching function of the conceptual framework Maxwell (2012) continued, is to inform the overall research design, evaluate and refine research goals, “develop

realistic and relevant research questions, select appropriate methods, and identify potential validity threats” to researchers’ conclusions (pp. 39-40).

This literature review will examine personalized learning within the framework of three theories of influence. It will also outline intelligent tutoring systems as a form of personalized learning. Specifically, this analysis will examine the causative mechanisms of personalized learning. These include the individualized learning and individual tutoring systems found in Bloom’s (1968) theory of mastery learning, the competency-based personalization of learning in Keller’s (1968) personalized system of instruction, and Vygotsky’s (1978) theories involving zone of proximal development and scaffolding necessary to achieve learning within this zone.

### **Personalized Learning**

No standard definition exists for personalized learning, but personalization within an educational context typically means that students have significant input in what they learn; pedagogy is specifically customized to students’ achievement levels; teaching and learning are student-paced, and instruction is heavily influenced by individual learner profiles and preferences (Hallman, 2019; Horn, 2017). In summary, personalized learning is a way to address the multitude of differences that exist in how people learn in order to optimize learning for all. Other characteristics of personalized learning often include student choice, an emphasis on out-of-school interest or non-cognitive factors (context personalization), adapting instruction based on learners’ prior knowledge, experiences, and competencies, (Bernacki & Walkington, 2018), minimal or relaxed sequencing of topics, and customized learning platforms delivering instruction how when and where students want it (Horn, 2017). Additionally, personalized learning relies heavily on pretests for diagnostic purposes, posttests to assess achievement, and

immediate and customized feedback to promote both reflection and the correction of misconceptions (Gudivada, 2017; Pardo, Jovanovic, Dawson, Gašević, & Mirriahi, 2019).

Personalized learning equates to instructional attention at the individual student level rather than the class level and involves student-centered pedagogy designed to help teachers differentiate instruction for their students (Bingham, 2017; Bingham 2019; Paz-Albo, 2017). Large classroom sizes and the heavy demands placed on teachers in most schools, however, make personalized learning an impractical if not impossible task. Much of the current literature on personalization within education is linked to educational technologies, the goal of which is to promote effective teaching through the use computer programs capable of providing digital curricula (Pepin, Choppin, Ruthven, & Sinclair, 2017; Godlen, 2017; Bingham, 2017; Bingham 2019), utilizing student data, (Bingham, 2017; Bingham 2019), and providing immediate, formative feedback (Bingham, 2017; Bingham 2019; Wongwatkit, Srisawasdi, Hwang, & Panjaburee, 2017). Additionally, advancements in technology allow for personalized learning that can gauge individual student progress, tailor personalized learning experiences and permit students to move at their own pace (Lee et al. 2018; Bingham, 2017; Kong, & Song, 2015). Hallman (2019) expounded on the connection between personalization in the classroom and technology explaining that personalized learning as pedagogy is “most often paired with 1:1 technology initiatives” and “connotes a shift in the teaching paradigm, one increasingly oriented toward students’ individual learning needs” (p. 301).

The growing emphasis on technology-driven, personalized learning has its roots in federal initiatives with policies related to technology use in U.S. classrooms going back several decades. *A Nation at Risk: 1983 Report of the Commission on Excellence in Education* emphasized computers, electronics, and other technologies, especially as they related to work

environments. *Transforming American Education: Learning Powered by Technology* (2010) outlined the importance of utilizing educational technologies to enhance student learning, as well as the significance of achieving personalized learning through technology in K-12 education (United States Office of Educational Technology, 2010). The U.S. Department of Education report, *Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics*, examined learning analytics as ways to determine students' learning patterns and predict academic outcomes (Bienkowski, Feng, & Means, 2014). Lastly, the U.S. National Education Technology Plan underscores the critical role that technology plays in personalized learning by "providing personalized instruction, continuously assessing students' learning, and tracking their mastery of skills and competencies" (Lee et al., 2018, p.1270). Patrick, Kennedy, and Powell (2013) surveyed educators who collectively defined personalized learning as authentic and meaningful learning experiences that take into account student's academic and personal needs, interests, and styles. Put differently, personalized learning is centered on individualization, and individualization has its theoretical roots in Bloom's (1968) learning for mastery or mastery learning (Lee et al., 2018).

### **Bloom's Mastery Learning**

The goal of mastery learning is fairly straightforward. All learners should master all of their educational objectives and curricula with as little variation in learning as possible (Zandvakili, Washington, Gordon, Wells, 2018; McGaghie, 2015). Dissimilarities in the *how* and *how long* of learning, however, do exist. Students differ in the amount of time it takes to reach mastery of specific objectives, standards, or topics. They also likely vary in how they attain individual mastery (McGaghie, 2015).

In his original discussions on Learning for Mastery, Bloom (1968) posited that most students, possibly more than 90% of all learners, can be successful learners if educators incorporate the appropriate pedagogy and materials requisite for reaching each individual student. Additionally, for mastery to be realized, students should be routinely assessed using formative tests and required to demonstrate a mastery level of 90% or better on these evaluations (Bloom, 1968). If students fall short of the established benchmark, before they move on to more advanced materials, remedial teaching and additional assessments should be employed until the student has met the predetermined criteria. Bloom (1968) further explained that in addition to regular formative assessment and periodic re-teaching, individual learning variables or differences must be considered including, aptitude, quality of instruction, ability to understand the task at hand, perseverance, and time allotted for learning for mastery learning to be realized.

In their discussion of mastery learning, Mitee and Obaitan (2015) supported Bloom's claims agreeing that almost every student can learn and learn well under optimal, appropriate conditions and that if teachers could see to these conditions, differences in achievement levels would almost disappear. McGaghie (2015) delineated several complementary features of mastery learning including diagnostic testing, clear learning objectives, sequenced units presented in increasing difficulty, relevant instruction focused on pertinent objectives, formative testing to assess a predetermined mastery of objectives, conditional progression through learning materials contingent on mastery of prerequisite knowledge, and continual practice until mastery is realized. Guskey (2007) highlighted two of these features as essential components for attaining mastery learning. He argued that frequent, specific feedback that is both diagnostic and prescriptive in nature should be utilized by teachers to reinforce learning expectations and define where students are regarding these expectations. Secondly, corrective measures should then be

employed to alleviate students' learning problems and position them on the appropriate learning path.

In one study, Bloom (1984) considered the findings of two doctoral students in education to expand his definition of mastery learning to include tutoring. This study compared student learning under three conditions of instruction, conventional instruction (control group), mastery learning, and a combination of tutoring and mastery learning (tutoring instruction followed by intermittent formative tests, feedback, and corrective procedures). Bloom (1984) discovered that the average student in the class that utilized mastery learning was above 84% of the students taught conventionally. He also found that the average student in the class that utilized both tutoring and mastery learning was approximately two standard deviations above the conventional group and approximately one standard deviation higher than the mastery-learning-only group. By expanding and thus revising his definition of mastery learning, Bloom (1984) had created a three-pronged model of individualized learning comprised of feedback, corrective measures, and tutoring.

Bloom (1984) alluded to the idiosyncrasies that exist among students beyond curricular and assessment related dissimilarities and advocated that teachers adapt what they do in the classroom to meet these varying needs. This process, called differentiation, proposes that teachers appropriately increase the variation in their teaching to maximize learning within their classrooms (Prast, Van de Weijer-Bergsma, Kroesbergen, & Van Luit, 2018). Within educational contexts, differentiation is established when teachers differentiate and individualize pedagogy, including content, process, products, and context to meet the individual needs of their students (Brevik, Gunnulfsen, & Renzulli, 2018; Anne & Haney, 2017; Tomlinson & Tomlinson, 2017). In addition, Bingham, Pane, Steiner, and Hamilton (2018) recently outlined four critical

components of differentiated, personalized learning: learner profiles that highlight students' strengths and weaknesses; tailored learning paths based on individual student's goals, needs, progress and motivations; flexible learning environments; and competency-based progression combined with ongoing assessment of students' progress toward delineated goals and objectives. These personalized learning mechanisms can find their predecessors in Bloom's assertions.

### **Keller's Personalized System of Instruction**

There is a natural correlation between a learner's strengths and weaknesses and assessment and feedback. Assessment and feedback are antecedents for determining and categorizing what a student's strengths and weaknesses are. Moreover, customized learning paths are akin to the differentiated, corrective measures Bloom advocated. The competency-based progression of intelligent tutoring systems, however, has its origins in Keller's (1968) personalized system of instruction.

In his article "Good-Bye Teacher," Keller (1968) delineated several features of individualized instruction similar to Bloom's, including the use of tutors, repeated assessment and feedback. Keller (1968) also added a "unit-perfection requirement," which allowed students to advance only after demonstrating mastery of all preceding material, and a self-paced feature that permitted students to grapple with instructional materials at a rate that matched their abilities (p. 83). In doing this, Keller's (1968) personalized system of instruction came into being and with it elements of personalized learning that are still present today, including mastery-based, self-paced instruction designed around study guides, small units of study, and teaching modules with very specific learning objectives and outcomes (Paiva, Ferreira, & Frade, 2017; Akera, 2017). Marzano (2017) argued that a data-driven, performance-based system, one built on the principle that every learner is unique and one capable of tailoring instruction to engage students

at their current aptitude levels via multiple learning paths, is necessary to ensure that teaching and learning are differentiated for learners with more prevalent needs, e.g., learners with disabilities, struggling learners, ESOL students, and gifted and talented learners. Vygotsky's theories of social constructivism have undoubtedly influenced this kind of learning personalization.

### **Vygotsky's Zone of Proximal Development**

Vygotsky (1978) promoted social constructivism in education. He argued that people learn within a community through the accumulation of knowledge through language and in their interactions with other people. He also described the gap between a students' potential intellect and their actual intellect or zone of proximal development (ZPD) as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (Vygotsky, 1978, p. 86). In other words, Vygotsky theorized that the zone of proximal development was the optimal area of instruction or the place where teachers or tutors should interact with their pupils to transition them from learning activities they can achieve independently to learning activities slightly above their current knowledge base (Vygotsky, 1978).

Concepts explored in Vygotsky's ZPD form the basis of his theories of human development and overtly emphasize that only material "within the very next developmental zone can be internalized via mediation from others, through social interactions" (Eun, 2019, p. 20). Instruction, therefore, has to concentrate on concepts and ideas that are "ready to develop" with support from the teacher or tutor, and these "developing functions, in turn, will be internalized and used by the learner independently after the support is withdrawn" (Eun, 2019, p. 20). To



achieve this transition, teachers are encouraged to scaffold their students' learning or actively direct their learning to a new and elevated position of understanding (Acedo & Hughes, 2014, p. 510). Even more challenging, teachers utilizing a Vygotskian approach must be able to identify when their students are on the verge of more advanced cognition and eventually be able to facilitate their transition over that border (Goggin, Rankin, Geerlings, & Taggart, 2016).

### **Related Literature**

Despite research supporting the educational effectiveness of Bloom's mastery learning and Keller's PSI, several factors have limited their use. Extended teaching time, curriculum pacing impediments, student self-discipline, differences in assignment completion rates, the time required for feedback and corrections, and difficulty modifying self-paced modules to an academic year have all been problematic (Pelkola, Rasila, & Sangwin, 2018; Paiva et al., 2017). Vygotsky's theories concerning the zone of proximal development have also been widely accepted as fundamental to teaching and learning (Guseva & Solomonovich, 2017). Nonetheless, the logistical complications of ensuring that every student is learning within an optimal zone are glaring (Mestad, & Kolstø, 2014). In recent decades, however, technologies have shown potential for alleviating some of these concerns.

### **Intelligent Tutoring Systems Defined**

Intelligent tutoring systems are computer learning systems purposed with helping students grasp knowledge or skills through the use of intelligent algorithms (Serrano, Vidal-Abarca, & Ferrer, 2018; Graesser et al., 2018; Mousavinasab et al., 2018; Wilson, & Scott, 2017). Intelligent tutoring systems monitor learners' psychological states, e.g., learning strategies, content knowledge, emotions, or motivations, and provide personalized, sequenced learning experiences with formative feedback (Serrano et al., 2018; Graesser et al., 2018; Huang, Craig,

Xie, Graesser, & Hu, 2016; Predrag, Jovan, Bojan, & Veis, 2016; Verdú et al., 2017). Intelligent tutoring systems are engineered to work with individual students to address their particular cognitive profiles and knowledge insufficiencies (Graesser et al., 2018; Sharada, Shashi, & Madhavi, 2015). Moreover, one of the most fundamental purposes of these systems is to assess learners' competencies within certain academic domains continually and to carefully select and propose activities to increase these proficiencies (Clément, Roy, Oudeyer, & Lopes, 2015).

Intelligent tutoring systems are adept at emulating human tutors to deliver personalized and adaptive, one-on-one instruction (Hooshyar, Ahmad, Yousefi, Yusop, & Horng, 2015; Malekzadeh, Mustafa, & Lahsasna, 2015; Wang, Han, Zhan, Xu, Liu, & Ren, 2015; Millis, Forsyth, Wallace, Graesser, & Timmins, 2017). Within these instructional paradigms, immediate cognitive analysis or student modeling (Najar, Mitrovic, & McLaren, 2016; Khodier, Elazhary, & Wanas, 2017; Basu, Biswas, & Kinnebrew, 2017) and adaptability (Esa, 2016; Najar et al. 2016; Basu et al., 2017) are emphasized. Learners are presented problems, and the intelligent tutoring system attempts to recognize whether or not students used the preferred strategy (as determined by the embedded curriculum) to solve these problems (Nye, Pavlik Jr, Windsor, Olney, Hajeer, & Hu, 2018)

Underlying student models examine questions and student answers to determine students' strengths and weaknesses. Moreover, they provide learners additional, targeted problems, along with suitable scaffolding (within their individual zones of proximal development) to achieve the desired learning objectives (Millis et al., 2017; Xin et al., 2017; Elazhary, & Khodeir, 2017; Ma, Adesope, Nesbit, & Liu, 2014). Specifically, this type of individualized instruction typically includes an explanation of the problem, solutions, and examples related to how to solve the problem, targeted feedback on learners' attempts at solving the problem, and recommended

learning sequences as determined by the learners' cognitive states (Wang et al., 2015; Ma et al., 2014; Millis et al., 2017; Wilson, & Scott, 2017).

### **History of Intelligent Tutoring Systems**

Intelligent tutoring systems have been examined through multiple lenses, e.g., technology, education, and psychology, for more than four decades. Wilson, C. & Scott, B. (2017). Intelligent tutoring systems were so named in the early 1980s, (Grasser et al., 2018), but the first computer tutoring systems were employed in classrooms in the 1950s (Kulik & Fletcher, 2016). These systems delivered instruction to students in brief frames or segments, asked questions, and offered opportune feedback (Kulik & Fletcher, 2016). In the 1970s and 1980s, first-generation computer tutors (CAIs) came into being as computer-assisted instruction began evolving and artificial intelligence and cognitive theory were employed to guide students through concepts and problems, step by step (VanLehn, 2011; Kulik & Fletcher, 2016; Mousavinasab et al., 2018). Although the modern terminology "intelligent tutoring system" had not yet come into play, in 1970 James Carbonell introduced SCHOLAR, now recognized as the first intelligent tutoring system. This system used natural language and limited discourse to facilitate instructional interchanges with students as they learned South America geography via a semantic web of facts and knowledge (Ma et al., 2014; Mousavinsab et al., 2018).

Over time, computer-assisted instruction increasingly used expert databases that offered hints and feedback to promote student learning and second-generation computer tutors were termed intelligent tutoring systems (Kulik & Fletcher, 2016). BIP, another predecessor of modern intelligent tutors, matched student's learning needs and capabilities to required domain-related tasks. In this early version of intelligent, adaptive software, like many that have been engineered since, the student model was a subcomponent of the domain model (Ma et al., 2014).

## **Architecture of Intelligent Tutoring Systems**

The conventional architecture of intelligent tutoring systems most often consists of three models: the expert module, the student model, and the tutoring model (Clément et al., 2015; Mousavinasab et al., 2018; Wang et al., 2015; El Mamoun, Erradi, & Mhouti, 2018; Ma et al., 2014). A fourth component, user interface, is also frequently included in the description of intelligent tutoring system components (Mousavinasab et al., 2018; El Mamoun et al., 2018; Ma et al., 2014). The expert model or domain model includes the subject knowledge the intelligent tutoring system intends to teach the student (Ma et al., 2014; El Mamoun et al., 2018; Mousavinasab et al., 2018) and is a technological depiction of a domain expert's subject knowledge and problem-solving capabilities (Sharada et al., 2015). Furthermore, the expert or domain model may be represented as a set of rational suggestions, production constraints, or natural language statements (Ma et al., 2014), and the domain knowledge contained within allows the intelligent tutoring system to compare student choices to those of an expert for evaluation purposes (Sharada et al., 2015).

The student model, also known as the learner or cognitive model, is constructed from students' learning styles, behaviors, and responses, as well as their individual domain knowledge proficiencies (Mousavinasab et al., 2018; Wang et al., 2015; Najar et al., 2016). This model exemplifies the cognitive, emotional, psychological, and affective condition of the student at the time of learning (El Mamoun et al., 2018). Moreover, the student model provides the basis for instructive methods chosen by the intelligent tutoring system (Rastegarmoghadam, & Ziarati, 2017; Rau, Michaelis, & Fay, 2015; Wang et al., 2015), and is essential for adapting and therefore personalizing instruction within the intelligent tutoring system to the specific needs of each student (Poitras et al., 2016; Rau et al., 2015; Rastegarmoghadam, & Ziarati, 2017). It

monitors the learner in an attempt to answer the question of what may be adapted and how the adaptations should look (Hafidi, 2015). Intelligent analysis in this model allows for adaptation within an intelligent tutoring system and consists of collecting information about the student, identifying domain gaps and misconceptions, examining student interactions in relation to the domain, choosing appropriate content and pedagogy, and delivering content (Poitras et al., 2016; Rastegarmoghadam, & Ziarati, 2017; Reddy, & Sasikumar, 2014). Analytics within the student model consider students' competencies within the domain as they progress through the chosen material and update student learning pathways accordingly (Poitras et al., 2016; Rau et al., 2015). Adaptive error feedback, e.g., hints, explanations, examples, practice problems, etc., are provided to learners to correct misconceptions related to relevant domain knowledge and problems are assigned based on individual learner progress (Rau et al., 2015; Ma et al., 2014; Sharada et al., 2015).

The tutor model or teacher or pedagogical model represents the aforementioned adaptive strategies and how they are chosen (Ma et al., 2014; Clément et al., 2015; Mousavinasab et al., 2018). This model identifies what deficits exist in learners' domain knowledge and determines what pedagogy is needed to address these deficiencies (Mousavinasab et al., 2018). The tutor model typically assigns tasks or additional practice just beyond students' current abilities or within their zone of proximal development (Ma et al., 2014; El Mamoun et al., 2018). Additional activities within this model include planning activities, providing explanations, determining when to intervene, and providing assistance to the learner (El Mamoun et al., 2018).

The interface model represents a visible and tangible means by which learners communicate with the intelligent tutoring system (Mousavinasab et al., 2018; Rastegarmoghadam, & Ziarati, 2017; El Mamoun et al., 2018). It is the environment or

graphical interface within which students interact with the system. The interface model typically consists of pointers, buttons, menus, icons, windows, or scroll bars (El Mamoun et al., 2018).

The interface model is also often domain-specific and regulates how learners navigate problem-solving strategies, how they seek and locate information, and how they respond to questions (Ma et al., 2014).

### **Types of Student Modeling**

Student modeling is the foundation of intelligent tutoring system design and is therefore worthy of consideration and analysis (Ma et al., 2014). One method of cognitive modeling used in intelligent tutoring systems is expectation and misconception tailoring (Ma et al., 2014; Nesbit, Adesope, Liu, & Ma, 2014; Cuirong, Weidong, & Hongtao, 2016). Expectation and misconception tailoring models student cognition by examining students' answers as they relate to learning goals with predicted misunderstandings in the domain (Ma et al., 2014; Nesbit et al., 2014; Cuirong et al., 2016). Another student modeling technique found in intelligent tutoring systems is model tracing. Model tracing aims to reduce the cognitive demands on students by assisting them through problematic areas, freeing them up to focus on areas that need to be practiced (Kessler, Stein, & Schunn, 2015). Ma et al. explained that in model tracing, the concept or skill being taught is modeled as a set of production rules that can be used to solve problems within the domain. Furthermore, they explained that the production rules mimic how a human would solve problems within the domain. As students make their choices, the model-tracing process employs definitive production rules specific to the particular domain. When students make errors, they are provided with feedback and offered a different way to approach the problem. Once model tracing has determined a student's particular use of the production

rule, Bayesian knowledge-tracing is used to estimate the likelihood that the information has been correctly taught (Ma et al., 2014).

Constraint-based modeling is a student knowledge modeling technique that represents domain knowledge as logical constraints by connecting or matching each constraint to probable solution difficulties (Ma et al., 2014; Khodeir, Wanas, & Elazhary, 2018). Each constraint has three parts, a relevance condition, a satisfaction condition, and a feedback condition (Ma et al., 2014; Khodeir et al., 2018). The relevance condition specifies when the constraint is appropriate. The satisfaction condition evaluates the student's solution, and the feedback condition advises the student of the error if the solution does not meet the satisfaction condition. As long as the student does not violate a constraint, he or she is headed toward a correct solution and no action is employed by the system (Ma et al., 2014; Khodeir et al., 2018). Conversely, a Bayesian network is a decision-making framework used to manage uncertainty based on probability theory, (Ma et al., 2014; Hooshyar et al., 2016; Hooshyar et al., 2015). In intelligent tutoring systems, Bayesian networks represent complex domain models containing numerous variables, and connections between variables are quantified into a network (Ma et al., 2014).

### **Artificial Intelligences in Intelligent Tutoring Systems**

Intelligent tutoring systems are computer programs that utilize artificial intelligence techniques for the purpose of simulating human tutors. They differ from educational technologies, such as Computer-Aided Instruction (CAI) because artificial intelligence allows intelligent tutoring systems to monitor both student learning and psychological characteristics (Alkhatlan, Kalita, 2018; Graesser et al., 2018). In doing so, intelligent tutoring systems are adept at adapting learning sequences, continually assessing learners, classifying them, and

updating student models as learners interface with the intelligent tutoring system (Graesser et al., 2018; Mousavinasab et al., 2018).

Mousavinasab et al. (2018) examined 53 intelligent tutoring system studies from 2007 to 2017. They found that condition-action rule-based reasoning techniques (33.86%), data-mining techniques (22.64%), and Bayesian-based techniques (20.75%) were artificial intelligence most frequently used. Other AI techniques included intelligent agents (15.09%), Fuzzy based techniques (13.20%), NLP techniques (11.32%), ANN-based techniques (9.43%), and case-based techniques (3.73%). Two of the more commonly used AI techniques used in recent intelligent tutoring systems are data mining and Bayesian knowledge tracing (Mousavinasab et al., 2018).

Data mining is a method of determining patterns in large data sets (Baker & Corbett, 2014). Techniques for this analytic include classification, association, clustering, and sequential pattern mining (Lin, Yeh, Hung, & Chang, 2013). The ever-increasing amounts of data available to educational practitioners and designers of educational technologies, especially from online learning environments, are currently being used to detect complex learning behaviors (Baker, & Corbett, 2014), model learning phenomena in online intelligent systems, gain insight into online learner's behaviors (Baker, 2014; Papamitsiou & Economides, 2014) and create adaptive, personalized learning environments (Lin et al., 2013). Bayesian knowledge tracing (BKT) is an algorithm used in many intelligent tutoring systems to model students' understanding of the subjects or domains represented in the intelligent tutoring system. It is a special case of a hidden Markov model. In BKT, skills are represented as known and unknown variables, and learning is specified as a transition between knowing and not knowing (Pelánek, 2017). The first BKT model delineated two learning states, learned and unlearned, which indicated whether or not a



student had mastered a specific knowledge component (Zhang & Yao, 2018). Later BKT models added a transitional state of learning to reflect when students were between unlearned and learned states (Zhang & Yao, 2018).

### **Intelligent Tutoring Systems and Personalization**

As previously stated, personalized learning looks to individualize instructional approaches and learning experiences to match students' individual interests, strengths, weaknesses, culture, and learning styles (Basham, Hall, Carter, & Stahl, 2016). Considering the diversification prevalent in today's K-12 classrooms and the subsequent logistics of meeting these diverse needs, some educational leaders and technology pacesetters have turned to technology, including intelligent tutoring systems as a means of providing personalization.

In recent years computer-assisted learning has evolved into what is currently known as intelligent tutoring systems. These advanced instructional tools often boast a complicated array of customization on a variety of levels, complete with immediate feedback and minimal teacher intervention. Furthermore, intelligent tutoring systems are able to consider students' psychological states and provide tailored instruction in every almost every discipline allowing learners to obtain cognitive knowledge, metacognitive knowledge, and domain-specific information (Ma et al., 2014). As tools for personalizing learning, intelligent tutoring systems include the causative mechanisms of this type of learning previously discussed, including the individual learning and tutoring systems rooted in Bloom's (1968) theory of mastery learning, Keller's (1968) competency-based learning personalization, and Vygotsky's (1978) scaffolding and zone of proximal development.

## **Feedback and Correctives in Intelligent Tutoring Systems**

Pardo et al. (2018) explained that students are by and large discontented with the feedback they receive from their teachers. Feedback is, nonetheless, acknowledged as one of the most crucial factors influencing academic achievement (Pardo et al., 2018). One obvious obstacle to providing personalized feedback is the amount of time it takes, but the solution may lie in the increased use of technologies and learning analytics to collect data and produce usable diagnostic information (Pardo et al., 2018; Rajendran, Iyer, & Murthy, 2019).

As their name suggests, intelligent tutoring systems are just that - intelligent tutors, and as such, one of their main functions is to assess and provide feedback. Learning assessment and feedback can be extremely time consuming and laborious. In traditional classrooms, adept teachers provide their students regular, constructive, and detailed written or verbal criticism and suggestions for improvement (Farrell & Rushby, 2015; Rajendran et al., 2019). Well-constructed and timely delivered feedback can positively affect students' attitudes, engagement, emotions, self-regulatory strategies, and learning outcomes (Muis, Ranellucci, Trevors, & Duffy, 2015; Hooshyar et al., 2016). The more differentiated and sophisticated the feedback, the more significant the effect will likely be (Farrell & Rushby, 2016). Unfortunately, this kind of feedback can be costly in terms of time and money, and human checking decreases the efficiency of feedback (Farrell & Rushby, 2016).

In recent years technology, including intelligent learning systems, has flourished in the area of assessment and feedback. Farrell & Rushby (2016) explained that technology has infiltrated the entire assessment process including, diagnostic assessment, monitoring student progress employing formative assessment, and even communicating essential learner data via summative assessment so that the necessary learning interventions can be employed. Within a

technology-based learning environment like an intelligent tutoring system, assessment and feedback are achieved through powerful learner analytics. Zilvinskis, Willis, and Borden (2017) defined learning analytics as a mechanism for utilizing real-time student data to forecast student achievement and provide learners personalized support. In a study on user-centered design and analytics in an adaptive learning system, Vesin, Mangaroska, and Giannakos (2018) posited that learning analytics can be used to track learners' development and accomplishments over time and provide feedback that enables both teachers and students to make informed decisions based on data. Within computer-assisted technologies like intelligent tutoring systems, well-informed, data-driven decisions can manifest in the form of corrective measures, applications like customized teaching plans, adaptive learning strategies, game-based learning strategies, teaching-learner gap analysis and customized teaching and learning strategies (Ahad, Tripathi, Agarwal, 2018).

### **Competency-Based Learning in Intelligent Tutoring Systems**

The primary objective of competency-based learning is to help learners attain recognized proficiency standards while taking into consideration the diversification that exists among students and employing criterion-referenced assessment rather than norm-referenced to promote learning (Hsu & Li, 2015). Furthermore, competency-based learning tools deliver individualized adaptive learning paths equipped with recurrent evaluation, feedback, and correction (Hsu & Li, 2015). The primary tenants of competency-based learning, as prescribed by Keller (1968), revolve around mastery of learning or unit perfection, as well as students being able to grapple with the material at their own pace. In technology-based learning environments, competency-based learning capable of achieving the functions as mentioned above are often domain-specific and contain features like curriculum sequencing, intelligent solution analysis, and problem-

solving support (Brusilovsky & Peylo, 2003). Curriculum sequencing involves suggesting to learners a planned or optimal sequence of learning activities within a certain domain (O'Neill, Donnelly, & Fitzmaurice, 2014). Intelligent analysis of solutions looks to discover not only the mistakes students are making but why they are making them in order to offer suggestions to help students over cognitive hurdles (Hafidi & Bensebaa, 2015).

Adaptability is an essential part of competency-based learning. Intelligent tutoring systems base their adaptability on didactic measures determined from the specific characteristics of each learner. This student profile or learner model is constructed from observing learners in their instructional environment and encompasses each learner's current state of knowledge and parameters concerning their personality, experience, and education (Rastegarmoghadam & Ziarati, 2017). Khodeir et al. (2018) further explained that a student model represents learner features, e.g., interests, goals learning styles, and knowledge, and it gathers information on the student's cognition within a certain domain. Students' problems solving behaviors, including correct and incorrect responses, number of attempts to achieve the correct answer are all attempts are analyzed to help build the student model (Khodeir et al., 2018).

An accurate student model is dependent on prediction accuracy and is essential for appropriate individualization of content and difficulty level. Common modeling techniques include logistic regression models, probabilistic models, and two of the most prevalent algorithms for estimating learner knowledge are Bayesian Knowledge Tracing and performance factor analysis (Kaser, Klingler, Schwing, Gross, 2017). The presentation of the actual learning content or domain knowledge, e.g., math, science, language arts, etc. is also based on students' learning characteristics or their learning model (Rastegarmoghadam & Ziarati, 2017).

Model-tracing tutors are technologies designed to mimic aspects of one-on-one tutoring by taking into account a student's cognitive load, what a student can handle without becoming overwhelmed, and dividing complicated tasks into smaller ones (Kessler et al., 2015). Kessler et al. (2015) explained that these technologies divide larger tasks into smaller, response-driven steps to both minimize cognitive load and assist learners in achieving automaticity with these individual components of the larger task. The technology can also solve conceptually unimportant parts of the larger task for the student. Put briefly, students' competencies are filtered through their respective cognitive loads, and their corresponding tutoring is adjusted accordingly.

### **Scaffolding and ZPD in Intelligent Tutoring Systems**

As previously noted, Vygotsky (1978) posited that learning improves when students are assisted by more knowledgeable or capable peers or teachers. Moreover, students learn best within their zone of proximal development. Vygotsky (1978) defined the zone of proximal development as the area between a student's independent problem solving and a context in which that learner is capable of problem-solving but only with assistance from or collaboration with a more capable adult or peer. Rus & Ștefănescu (2016) expounded on this idea adding that personalized learning requires assessing a student's learning and choosing larger appropriate instructional tasks or macro-adaptation and scaffolding these tasks for students using within-task or micro adaptation.

Wood, Bruner, and Ross (1976) coined the word "scaffolding" to describe the process in which a more knowledgeable or proficient person (within a particular area of student or knowledge) tutors a less competent individual. This exchange allows the less competent person to engage in tasks they would be unable to perform independently of the tutor. Considering

current educational technologies, namely intelligent tutoring systems, the definition of scaffolding must be extended to include contexts beyond interactions between individual humans (González-Calero, Arnau, Puig, & Arevalillo-Herráez, 2015; VanLehn, 2011). Scaffolding procedures must be dynamic and adaptable as they adjusted to fit each unique learning situation (Farias, Hastie, & Mesquita, 2018).

In recent decades educational technologies aimed at scaffolding student learning have increased significantly (González-Calero et al., 2015). Based on the learner or student model, constraint-based modeling is a technique used by many intelligent tutoring systems to model student knowledge and provide appropriate support (Khodeir et al., 2018). In constraint-based modeling, a learning domain is represented as a set of constraints, and students' knowledge is defined in terms of how well they satisfy or violate relevant constraints (Khodeir et al., 2018). Intelligent tutoring systems that employ constraint-based modeling act in place of human tutors often using scaffolding techniques to provide learners individualized support in the form of scaffolding questions (sequenced questions intended to helps students build understanding) and scaffolding feedback (incremental hints) (Khodeir et al., 2018). Other scaffolding techniques include explanations, prompts, hints, demonstrations, or reminders (Delen, Liew, & Wilson, 2014).

Instructional scaffolding specifically refers to the back-and-forth between learners and their tutors or teachers and the process that allows the learner to access what would otherwise be inaccessible (Delen et al., 2014). Vygotsky's zone of proximal development and instructional scaffolding are now frequently applied in technology-based instructional tools like intelligent tutoring systems (Yelland & Masters, 2007) in the form of visual cueing or hyperlinks and have been effective in scaffolding learning (Delen et al., 2014). Bartlet, Ghysels, Groot, Haelermans,

and van den Brink (2016) cautioned that within adaptive learning environments like intelligent tutoring systems, scaffolding must be present; otherwise differentiation will not guarantee learning.

Despite its established importance, scaffolding as an instructional tool must be employed judiciously. González-Calero et al. (2015) posited that “scaffolding is a delicate balancing act” because too little scaffolding could lead to student failure and a subsequent frustration and loss of motivation, while too much could hamper learners’ efforts in attaining their learning goals (p. 1191). Moreover, too much scaffolding can lead to learners manipulating the systems’ support strategies and solving problems without the requisite knowledge (González-Calero et al., 2015; Dale, & Scherrer, 2015).

### **Intelligent Tutoring System Advantages**

Intelligent tutoring systems offer several advantages over traditional, teacher-led or large-group classroom instruction, foremost of which is personalized, interactive, adaptive instruction (Ma et al., 2014; Sharada et al., 2015; Wilson & Scott, 2017; Huang et al., 2016; Verdú et al., 2017). This type of individualized instruction connotes sophisticated tutoring strategies like modeling or scaffolding (Serrano et al., 2018; Rastegarmoghadam & Ziarati, 2017; González-Calero, Arnau, Puig, & Arevalillo-Herráez, 2015), increased opportunities for feedback and practice (Serrano et al., 2018; Huang et al., 2016), well-organized pedagogy (Huang et al., 2016; Hooshyar et al., 2016), more teaching and learning opportunities than in a single-teacher classroom (Serrano et al., 2018) and learner-centered platforms that take into account students’ expectations, motivations, and learning habits (Rastegarmoghadam & Ziarati, 2017).

The adaptive features within various intelligent tutoring systems provide the foundation for their functionality and success. These features include immediacy of feedback and increased

learner control (Ma et al., 2014; VanLehn, 2011), providing students hints based on their specific needs (Rau et al., 2015; Hafidi, & Bensebaa, 2015), intelligent analysis of solutions to challenge and correct common misconceptions (Rau et al., 2015), and carefully chosen practice problems based on individual student's cognitive models (Rau et al., 2015). These designated practice problems are chosen because they fall within an individual learner's particular zone of proximal development. In other words, the student has the prerequisite knowledge to complete these problems but has not yet achieved mastery (Rau et al., 2015). Moreover, learning analytics in intelligent tutoring systems monitor learners' activities and growth over time (Mangaroska & Giannakos 2018) and collect information to assist in making data-driven decisions (Vesin et al., 2018).

By design, many intelligent tutoring systems take into consideration multiple learner idiosyncrasies. Some even consider students' multiple intelligences, i.e., verbal-linguistic, logical-mathematical, visual-spatial, bodily-kinesthetic, musical, interpersonal, intrapersonal, and naturalist (Hafidi, & Bensebaa, 2015; Rastegarmoghadam, & Ziarati, 2017; Verdú et al., 2017). Research has shown that intelligent tutoring systems can also level achievement across different demographic groups and between genders. They do so by reducing or eliminating culpable factors such as inconsistent teacher attention or biased teacher expectations (Huang et al., 2016) and heighten students' problem-solving abilities (Hooshyar, Binti Ahmad, Wang, Yousefi, Fathi, & Lim, 2018).

### **Intelligent Tutoring System Disadvantages**

Despite the many documented advantages to the use of intelligent tutoring systems, there are noted disadvantages. Due to student idiosyncrasies, it can be difficult even for artificial intelligence to accurately identify factors necessary to genuinely individualize learning, (Clément



et al., 2015), replicate human instruction (Hooshyar et al., 2016; Wang et al., 2015), and accurately address student misconceptions (Poitras et al., 2016). Implementing adaptive strategies in intelligent tutoring systems can also be both expensive and time-consuming (González-Calero et al., 2015; Graesser et al., 2018). Moreover, these strategies are typically domain-specific and not easily adapted to other subject areas (González-Calero et al., 2015; Nye et al., 2018).

Maintaining student engagement and motivation in adaptive computer environments can also be problematic when appropriate guidance and interaction are not readily available (Hooshyar et al., 2016; Millis et al., 2017). Also, many students misuse the program's hints to "game" the answers. In essence, students use the tutoring model's feedback to achieve the correct answer without truly mastering the concept the system is trying to teach (Millis et al., 2017). Finally, a significant disadvantage of adaptive software like intelligent tutoring systems can be their lack of flexibility and explorability, which can hamper metacognitive gains (Verdú et al., 2017; Wilson & Scott, 2017).

### **Efficacy of Intelligent Tutoring Systems**

Since their inception, computer-based learning platforms, including intelligent tutoring systems and other adaptive learning technologies, have been compared to non-intelligent tutoring system learning environments in terms of learner outcomes. Several studies have shown moderate to moderately strong positive effects between the use of intelligent tutoring systems and achievement gains (Ma et al., 2014; Steenbergen-Hu & Cooper, 2014; VanLehn, 2011). VanLehn (2011) evaluated outcomes from 54 intelligent tutoring systems and non-intelligent tutoring system groups and found that intelligent tutoring systems ( $d = 0.76$ ), were as effective as human tutors ( $d = 0.79$ ) at increasing student achievement across subject areas. VanLehn

(2011) found an increase in test scores of 0.58 standard deviations over traditional instruction and further delineated results into intelligent tutoring systems with step-based tutoring systems, which provide hints and explanations on typical problem-solving steps, and scaffolding and feedback systems. He discovered that step-based tutoring raised test scores by 0.76 standard deviations, while newer, more advanced scaffolding and feedback systems only raised test scores by only 0.40 standard deviations.

Ma, Ma et al. (2014) conducted a meta-analysis of 107 effects sizes involving 14,331 participants and found an average effect of 0.43 standard deviations. They compared intelligent tutoring system groups to three non-intelligent tutoring system control conditions: large-group instruction, non-intelligent tutoring system computer-based instruction, and textbook/workbook instruction. The effect sizes were 0.42, 0.57, and 0.35 respectively. Furthermore, they found that intelligent tutoring systems were associated with greater academic achievement at all levels of education when compared to traditional, teacher-led instruction, and in almost all subject domains. Gains were also significant regardless of whether or not the intelligent tutoring system modeled student misconceptions or provided feedback, and were also substantial irrespective of the degree or method of implementation, e.g., the primary form of instruction, supplemental to teacher instruction, homework aid, or an integral component of teacher-led instruction. No significant difference, however, was found between intelligent tutoring system environments and groups that benefited from individualized human tutoring ( $g = -0.11$ ) or small-group instruction ( $g = 0.05$ ).

Steenbergen-Hu and Cooper (2014) surveyed 39 studies involving the effectiveness of intelligent tutoring systems at the postsecondary level, spanning over 20 years (1990-2011) and surveyed 22 different types of intelligent tutoring systems. They found that intelligent tutoring

systems increased scores approximately 0.35 standard deviations; however, the exact results were contingent on the overall instruction the control group received. Intelligent tutoring systems increased scores by 0.86 standard deviations compared to the group receiving no instruction and 0.37 standard deviations higher than the traditional, teacher-led group.

Conversely, the intelligent tutoring system group scored 0.25 standard deviations lower than the control group received human tutoring.

In their meta-analysis of 50 controlled evaluations, Kulik and Fletcher (2016) found that in 46 of 50 controlled cases, students who received assistance from intelligent tutors achieved higher test scores than students who received only traditional instruction. Their studies revealed moderately strong positive effect sizes. The median effect size in this study was 0.66, comparable to an increase in test achievement from the 50<sup>th</sup> to the 75<sup>th</sup> percentile. Unlike similar meta-analyses, however, the study conducted by Kulik and Fletcher (2016) posited an effect size for intelligent tutoring system tutoring even higher than human tutoring. They also added that the moderate to strong effects were smaller in evaluations that assessed outcomes on standardized tests as opposed to locally developed assessments.

Graesser et al. (2018) posited a rational meta-meta estimate from all of these meta-analyses to be  $d = 0.60$ , analogous to human tutoring, between  $d = 0.42$  and  $d = .80$  contingent on the tutor's proficiency. Kulik and Fletcher (2016) found that students taught with intelligent tutoring systems outpaced those in traditional, teacher-led classes in 46 of 50 controlled evaluations. Their studies revealed moderately strong positive effect sizes and also saw intelligent tutoring system effects greater than those of human tutoring on locally developed tests, 0.73, but only 0.13 on standardized tests.

Xu, Wijekumar, Ramirez, Hu, BS, and Ireby (2019) conducted a meta-analysis of 19 studies examining the efficacy of intelligent tutoring systems on reading comprehension in first through tenth-grade classrooms. Their findings indicated that intelligent tutoring systems produced larger effect sizes than traditional, teacher-led classrooms or other educational applications, an effect size of 0.86, and an overall random effect of 0.60. However, as suggested in previous studies, they found that intelligent tutoring systems produced greater effect sizes on reading comprehension when compared to traditional teaching and much smaller effect sizes when compared to human tutoring.

Despite the research that suggests the use of intelligent tutoring systems are positively associated with increased academic gains, a few older reports suggest otherwise (Slavin, Lake, & Groff, 2009; Steenbergen-Hu & Cooper, 2013). The What Works Clearinghouse looked at 27 evaluations of intelligent tutoring systems used in Algebra I classrooms and found insignificant effect sizes. Slavin et al. examined intelligent tutoring system use in middle and high school mathematics, finding that intelligent tutoring systems increased student test scores by an average of only 0.12 standard deviations.

Steenbergen-Hu and Cooper (2013) analyzed 34 cases of intelligent tutoring system use in K–12 mathematics and discovered a difference between intelligent tutoring system groups and non-intelligent tutoring system groups of only 0.05 standard deviations. There also exists some discrepancy regarding whether or not stronger intensity and longer duration of intelligent tutoring system use is associated with larger effect sizes. Xu et al. (2019) found that stronger intelligent tutoring system use produced an effect size of 0.26 and weaker intelligent tutoring system use yielded an effect size of 0.78 when using mixed measures ( $p = 0.0097$ ) and when only standardized measures were analyzed there was no significant difference between the strong

and weak groups (0.23, and 0.26 respectively,  $p = 0.60$ ). These results differ from Cheung and Slavin (2013), who found stronger intensity and longer duration of an intelligent tutoring system to be positively associated with increased academic achievement.

### **Intelligent Tutoring Systems in Mathematics**

In the last decade, the use of intelligent tutoring systems in K-12 mathematics education has increased significantly (Graesser et al., 2018; El Mamoun et al., 2018; Huang et al., 2016). Domain-specific programs have been used to increase learning gains in algebra (Sabo, Atkinson, Barrus, Joseph, & Perez, 2013; Graesser et al., 2018; González-Calero et al., 2015) geometry (Sabo et al., 2013; Funkhouser, 2003; Graesser et al., 2018) statistics (Frith, Jaftha, and Prince, 2004), and elementary mathematics (Roschelle, Gaudino, & Darling, 2016). Additionally, due to their adaptive and interactive nature (Steenbergen-Hu & Cooper, 2013), intelligent tutoring systems in mathematics education are often employed as supplemental instructional tools set in remedial contexts (Bartelet et al., 2016).

Some math-oriented intelligent tutoring systems have even included a simulated student or SimStudent for real students to assist (Li, Matsuda, Cohen, & Koedinger, 2015; Mastuda et al., 2013). Within these systems, students assign problems to the SimStudent, which then attempts to decipher the problem using the curriculum embedded steps. If the SimStudent has “difficulty” solving the problem, it seeks help from the student who has assumed the role of the teacher (Bringula et al., 2016). Bringula et al. (2016) postulated that after 20 problems, the simulated student “could accurately predict students’ correct behavior on the mathematics problems more than 82% of the time” (p. 465).

### Summary

Personalized learning, the conceptual foundations of which can be traced to Bloom's (1968) theory of mastery learning, Keller's (1968) personalized system of instruction, and Vygotsky's (1978) sociocultural theory (specifically the tenets related to zone of proximal development (ZPD) and scaffolding), seeks to match instructional approaches and learning experiences to individual students. This is a logistical challenge; however, one that might be overcome through technology. Intelligent tutoring systems have been employed for nearly four decades by scholars in almost every academic field and are capable of providing personal training assistance to learners in virtually every area of academia, including the military, the corporate world, colleges, and university and K-12 education (Sharada et al., 2015).

Due in part to their potential to revolutionize modern education, adaptive learning computer environments like intelligent tutoring systems have regularly been compared to other learning environments and evaluated for effectiveness. Meta-analyses conducted in recent years have yielded findings suggesting that intelligent tutoring systems, especially as a supplement to existing pedagogies can increase student achievement. Moreover, a cursory database search yields evaluative research on a wide array of adaptive computer learning systems, including intelligent tutoring systems, the intended students, and domains of which run the gamut. Many of these studies, however, focus on very limited intervals of time, feature a relatively small sampling, and often do not include K-12 populations. Even fewer focus specifically on middle school populations over an extended period of time comparing achievement in a particular cognitive domain, e.g., math.

## **CHAPTER THREE: METHODS**

### **Overview**

The following chapter delineates the research design, question, hypothesis, participants, setting, instrumentation, procedures, and data analysis of a quantitative study. This study examines middle school math achievement in a school setting that uses an intelligent tutoring system and a school setting that does not. The significance of this study lies in its investigation and evaluation of instructional tools capable of individualizing and customizing math instruction for diverse populations of students. This level of individualization and customization is an undertaking considered unattainable in traditional classrooms.

### **Design**

This causal-comparative study endeavors to test the fundamental theories supporting personalized instruction and the use of intelligent tutoring systems as a means of achieving personalization. Specifically, this study will examine math achievement for students at two similar middle schools in the Midlands of South Carolina. The independent variable, use of an intelligent tutoring system in math instruction, will be defined as the supplementary use of two intelligent tutoring systems, Pearson's Math Digits and IXL (both adaptive computer software intended to deliver customized instruction and feedback to students without requiring teacher intervention) for math instruction. MAP Math is a computerized adaptive test for measuring math achievement administered multiple times each year in many schools across the US. An advantage of utilizing MAP scores is that this assessment uses an equal interval measurement scale that is stable and allows for comparing assessment scores among groups of students (NWEA, 2008).

A quantitative approach to this study is appropriate because scientific inquiry is being employed to examine the differences between two groups (Rovai, Baker, Ponton, 2013). A causal-comparative design is particularly fitting because this study seeks to explore differences that already exist between groups (Gall, Gall, & Borg, 2007), i.e., math achievement among middle school students who received math instruction that did not include the use of an intelligent tutoring system and middle school students who received math instruction that included the use of an intelligent tutoring system. Furthermore, a causal-comparative design is appropriate because the independent variable, intelligent tutoring system instruction, was manipulated before the research, and both the effect and possible causes have already occurred (Rovai et al., 2013). Other traits of this study consistent with causal-comparative studies include involving at least two groups, focusing on group differences rather than the relationships between variables, and omitting random assignment as the groups are based on their status (Rovai et al., 2013).

### **Research Question(s)**

As previously stated, this research will focus on differences in math achievement among middle school groups that experienced the use of intelligent tutoring systems in their math instruction and middle school groups that did not. The research question in the study is:

**RQ1:** Is there a difference in the math achievement of middle school students whose math instruction includes an intelligent tutoring system and middle school students whose math instruction does not include an intelligent tutoring system?



## **Hypothesis**

The null hypothesis for this study is:

**H<sub>0</sub>1:** There is no statistically significant difference in math achievement between middle school students whose math instruction includes an intelligent tutoring system and middle school students whose math instruction does not include an intelligent tutoring system, as shown by the Measures of Academic Progress (MAP) SC 6+Math test.

## **Participants and Setting**

To answer the research question and test the corresponding hypothesis, a causal-comparative study will be used. Moreover, a convenience sample will be employed to identify the 180 participants for this research, 90 from each middle school. Additionally, stratified sampling will be used to ensure that 30 students from each grade level (grades, 6, 7, and 8) are chosen. Gall, Gall, & Borg (2007) explained that to achieve population validity, researchers must randomly select the sample from the population they wish to generalize their results, and the sample must be large enough to reduce the possibility that it has different characteristics of the target population. A sampling size of 180 is more than the required minimum (N=166) for a medium effect size with a statistical power of 0.7 at the 0.05 alpha level, according to Gall et al., (2007).

The settings for this study include two middle schools, grades 6-8, situated in neighboring school districts in the South Carolina Midlands. Both schools are similarly sized, approximately 1000 students, and are a mix of rural and suburban populations with similar demographic and socioeconomic makeup. School A serves 930 students in grades 6-8 and employs 39 core teachers, 8 special education teachers, 1 ESOL teacher, 15 related arts teachers, 3 guidance counselors, and 4 administrators. The school is 73% African American, 21% white, 4%

Hispanic, and 2% other. Despite the school being located inside a fairly affluent neighborhood, many of School A's students are bused in from other neighborhoods, and more than half of its students receive subsidized breakfast and lunch. Furthermore, School A is 1:1 (provides each student with an electronic device to access the internet, online textbooks, curriculum-related software, etc.) and has been since 2013. School A has employed the use of various mathematics computer-assisted technologies since becoming a 1:1 school.

School B serves 1046 students in grades 6-8 and is staffed with 41 core teachers, 11 special education teachers, 1 ESOL teacher, 17 related arts teachers, 4 guidance counselors, and 4 administrators. Additionally, School B has 13 faculty members assigned to various instructional support and coaching roles. The school is 87% African American, 6% Caucasian, 3% Hispanic, and 4% other. Sixty-five percent of its students are from low-income families and receive subsidized lunch. Though School B is located within the city limits of the county seat, many of its students come from the large rural area surrounding the town in which it is situated. Finally, School B is not a 1:1 school and does not use computer-assisted technologies for math instruction.

### **Instrumentation**

The instrument that will be used to measure math achievement in this study is the Measures of Academic Progress Skills (MAP) Math SC 6+, a web-based, computerized adaptive test (CAT) published by the Northwest Evaluation Association (NWEA). MAP Math SC 6+ is a standards-based, mastery measure assessment system designed to provide information about students' mastery of foundational skills derived from the South Carolina mathematics standards (NWEA, 2011). Foundational skills for SC mathematics are organized into four strands - algebraic thinking and operation; real and complex number systems; geometry and measurement;

and data analysis, statistics, and probability. Moreover, the order in which students encounter these skills within the assessment is determined by the necessary knowledge to progress through the content within that strand, and the items themselves come from a large, flexible pool of questions rather than a rigid set of test items (NWEA, 2011).

To measure student progress, the MAP Math SC 6+ utilizes the RIT scale (Rasch Unit), which was developed by NWEA over 30 years ago, according to the Item Response Theory principles. The RIT scale uses individual item difficulty values to estimate student achievement, and test items are anchored to a vertically-aligned, equal-interval scale that covers all grades. This equal-interval scale ensures that the difference between scores is the same regardless of where the test-taker falls on the RIT scale, i.e., bottom, middle, or top. It also has the same meaning regardless of grade level, making MAP Math SC 6+ appropriate for measuring math achievement over time. MAP Math SC 6+ is not a fixed-form test but instead uses an adaptive algorithm in which test item selection is based on a momentary achievement or provisional ability estimate. That is to say, the range of ability is restricted, limiting items to those based on the student's provisional ability. The assessment includes 50 multiple-choice items with four or five options (NWEA, 2008).

RIT scores range from approximately 140-300. Although it is possible to score as high as 285 or more on the math test, 250 is a typical top score. Students typically start at the 140 to 190 level in the third grade and progress to the 240 to 300 level by high school. The expectation is that RIT scores will increase over time. Students at lower grade levels tend to show a greater increase in RIT scores during a school year than students in higher grade levels. At higher levels, questions become much more difficult, and the overall progress decreases (NWEA, 2011).

The adaptive nature of MAP tests requires reliability to be tested using alternative means. Test-retest procedures are problematic because dynamic item pool selection prohibits administering the same test. Parallel forms reliability is also impractical because the difficulty of items presented is based on students' responses to prior items, precluding identical content (NWEA, 2011). Testing the reliability of MAP tests, including MAP SC 6+, requires a combination of test-retest and parallel forms where the second test or retest is not the same test but is similar in structure and content but differing in item difficulty only (NWEA, 2011). This type of reliability testing is called stratified, randomly-parallel reliability (Green, Bock, Humphreys, Linn, & Reckase, 1984). NWEA reported reliabilities for the spring 2008-Fall 2008 MAP SC 6+ with different item pool structures in the high .80s (NWEA, 2011). NWEA (2011) also found similar reliability statistics (high .80s) for the fall 2008-spring 2009 test-retest correlations for the MAP SC 6+ with common pool structures.

Confirming internal consistency reliabilities for MAP is difficult as well because conventional approaches require all examinees to take a common test comprised of the same test items. Applying these techniques to adaptive tests is not only problematic but can also yield inaccurate results (NWEA, 2011). Samejima (1977, 1994) posited an equally valid alternative in the marginal reliability coefficient, which includes measurement error as a function of the test score. Calculating internal consistency in this way provides outcomes virtually identical to the coefficient alpha when both procedures are employed on identical fix-formed tests (NWEA, 2011). Marginal reliabilities for MAP Survey w/Goals 6+ were 0.965, 0.968, 0.970 for grades 6-8 respectively.

The validity of decisions made with MAP data assumes that they are capable of 1) determining if a student has a firm understanding of a skill and 2) identifying the foundational

skills within a grade-level content strand that a student needs to work on (Burns & Young, 2016). The claims were evaluated with an interpretation and use argument (IUA), based on Kane's (2013) framework. Third-party reviewers rated each item of MAP mathematics on a 4-point scale (4= item only aligns to identified skill). Ninety-seven percent received a rating of 4 (Burns & Young, 2016). Conducting factor analysis for Computer Adaptive Tests (CAT) like MAP is challenging because unlike fixed-form tests, with CAT tests, different participants respond to different items. In other words, there are no common forms, and an adaptive algorithm restricts covariance among items by limiting questions to the test-takers provisional ability. One way to circumvent the sparse data problem is to conduct CFA at the item cluster-level. Wang, McCall, Jio, & Harris (2013) conducted a CFA at the cluster level and came up with goodness-of-fit statistics for the South Carolina 6th grade MAP Math in the spring of 2011. All values of fit satisfied Hu and Bentler (1999) criteria and showed that each model fit data for content. Lastly NWEA (2011) expressed concurrent validity of MAP Math SC 6+ as Pearson product-moment correlations of concurrent performance on state accountability tests (6th grade  $r = 0.849$ ,  $n = 5974$ ; 7th grade  $r = 0.839$ ,  $n = 5920$ ; 8th grade  $r = 0.833$ ,  $n = 5570$ ); predictive validity as Pearson product-moment correlations of predicted performance on state accountability tests (6th grade  $r = 0.827$ ,  $n = 5740$ ; 7th grade  $r = 0.828$ ,  $n = 5748$ ; 8th grade  $r = 0.826$ ,  $n = 5396$ ); and criterion-related validity as Pearson product-moment correlations of criterion-related performance on state accountability tests (6th grade  $r = 0.676$ ,  $n = 5961$ ; 7th grade  $r = 0.660$ ,  $n = 5909$ ; 8th grade  $r = 0.690$ ,  $n = 5569$ ). All Pearson product-moment correlation coefficients suggested a strong positive linear association.

### **Procedures**

The researcher will first obtain written permission from both district and building-level

administrations for both School A and School B. Next, the researcher will seek permission for the study from Liberty University. After receiving consent from the Liberty University Institutional Review Board, student data will be acquired from each respective district office data manager and keyed into two separate Microsoft Excel worksheets – one for each school, along with corresponding fall and spring MAP SC 6+ scores. Students will next be separated into grades within each Microsoft Excel worksheet, and all student names will be removed to preserve anonymity. Next, an Excel formula will be used to randomly select 30 students from each grade level from each school. All data will be stored on a password-protected flash drive.

### **Data Analysis**

Participant data will be attained separated into two groups - middle school students whose math instruction included an intelligent tutoring system (ITS) and middle school students whose math instruction did not include an intelligent tutoring system (NITS). To ensure an equal number of participants from each grade level, the data will be further organized into six groups: 6th graders whose math instruction included an intelligent tutoring system (6ITS), 7th graders whose math instruction included an intelligent tutoring system (7ITS), 8th graders whose math instruction included an intelligent tutoring system (8ITS), 6th graders whose math instruction did not include an intelligent tutoring system (6NITS), 7th graders whose math instruction did not include an intelligent tutoring system (7NITS), and 8th graders whose math instruction did not include an intelligent tutoring system (8NITS). An Excel formula will be used to select 30 students from each grade level randomly, 180 students total ( $n=180$ ).

Utilizing a one-way ANCOVA will permit the researcher to compare the posttest scores of both groups, students whose math instruction included intelligent tutoring systems and students whose math instruction did not include intelligent tutoring systems, holding constant

pretest score differences and determine if there exists any statistically significant differences between the adjust means of these independent groups (Gall, Gall, & Borg, 2007). Post hoc testing will then be carried out to determine what those differences, if any, are.

A one-way ANCOVA has several assumptions related to study design that must be considered. A one-way ANCOVA assumes a continuous dependent variable (change in posttest scores), a continuous covariate (pretest scores), one categorical independent variable (type of math instruction) with two or more independent groups (NITS/ITS), and independence of observations (Laerd Statistics).

A one-way ANCOVA also has several assumptions related to data, including linearity, homogeneity of regression slopes, normality, homoscedasticity, homogeneity of variances, no significant outliers in the groups of the independent variable in terms of the dependent variable, and normality (Laerd Statistics).

The linearity assumption assumes that the covariate (pretest scores) is linearly related to the dependent variable (posttest scores) for each level of the independent variables (ITS/NITS). A visual inspection of a grouped scatterplot will be used to check for linearity. A test of homogeneity of regression slopes will be used to check if the linear relationships established in the linearity assumption have the same slope. A Shapiro-Wilk test for normality will used to assess that the dependent variable (posttest scores) are approximately normally distributed fore each group of the independent variable (ITS/NITS). The assumption of homoscedasticity states that there is homoscedasticity of error variance within each group and that the error of variances is equal between groups. Homoscedasticity will be assessed by a visual inspection of the standardized residuals plotted against the predicted values for each group. Homogeneity of variances requires that the variance of residuals should be equal for all groups of the independent

variable. This assumption will be assessed by Levene's test of homogeneity of variance.

Finally, testing for significant outliers is necessary for a one-way ANCOVA to ensure that there are no significant unusual points in the groups of the independent variable in terms of the dependent variable. Outliers in the data will be assessed to confirm there are no cases with standardized residuals greater than  $\pm 3$  standard deviations (Laerd Statistics).

A one-way ANCOVA will be conducted to determine if math achievement was different for middle school students whose math instruction included an intelligent tutoring system and middle school students whose math instruction did not include an intelligent tutoring system. An alpha level of 0.05 with a confidence level of 95% will be used for the one-way ANCOVA. For all rejected null hypotheses,  $\eta^2$  will be calculated to determine effect size.



## **CHAPTER FOUR: FINDINGS**

### **Overview**

This study is a causal-comparative examination of math achievement in two similar middle school settings, one that uses an intelligent tutoring system for math instruction and one that does not. The significance of this study can be found in its examination and assessment of instructional tools capable of personalizing math instruction for diverse populations of students. The following chapter outlines the study's findings and includes its research question, null hypothesis, and descriptive statistics. Moreover, this chapter will present the results of the statistical analysis comparing math achievement in the aforementioned middle school settings as measured by the Measures of Academic Progress (MAP) SC 6+Math test.

### **Research Question**

**RQ1:** Is there a difference in the math achievement of middle school students whose math instruction includes an intelligent tutoring system and middle school students whose math instruction does not include an intelligent tutoring system?

### **Null Hypothesis**

**H<sub>0</sub>1:** There is no statistically significant difference in math achievement (as shown by the Measures of Academic Progress SC 6+ Math test scores) between middle school students whose math instruction includes an intelligent tutoring system and middle school students whose math instruction does not include an intelligent tutoring system while controlling for pretest scores.

### **Descriptive Statistics**

The data provided from each school district's data manager was received in Excel format stripped of all identifiers except grade level. Schools were then recoded as NITS and ITS to

reflect the two groups examined in this study (students whose math instruction did not include an intelligent tutoring system and students whose math instruction included an intelligent tutoring system). Fall (pretest) and spring (posttest) Measures of Academic Progress SC 6+ Math test scores from 180 students, 90 from each school, 30 from each grade level within each school were randomly chosen from all scores provided using an Excel formula. Table 1 presents pretest and posttest unadjusted mean scores for both groups, as well as adjusted posttest scores for both levels of the intervention. Unadjusted pretest mean scores were higher in the ITS group (221.00) as compared the NITS group (217.70). Unadjusted posttest mean scores were also higher in the ITS group (222.777) as compared to the NITS group (221.70).

Table 1

*Means Report*

ITS level		Pretest (Fall MAP)	Posttest (Spring MAP)
NITS	Mean	217.43	221.70
Unadjusted	N	90	90
ITS	Mean	221.00	222.77
Unadjusted	N	90	90
NITS	Mean		223.491
Adjusted	N		90
ITS	Mean		220.976
Adjusted	N		90

Data obtained for the dependent variable posttest scores for participants whose math instruction both included intelligent tutoring software instruction (ITS), as well as those that did not include intelligent tutoring software instruction (NITS), were loaded into SPSS which generated the descriptive statistics found in Table 2. Table 2 presents the mean, standard deviation, and sample size for the dependent variable (posttest scores) for each group of the independent variable (ITS level). Group sizes were equal with 90 participants in each group ( $n=90$ ).

Table 2

*Descriptive Statistics*

Dependent Variable: Posttest (Spring MAP)

ITS level	Mean	Std. Deviation	N
NITS	221.70	16.924	90
ITS	222.77	18.131	90
Total	222.23	17.497	180

The Estimates table, Table 3, presents the adjusted mean, standard error, and 95% confidence interval of the adjusted mean for the dependent variable (posttest scores) for each group of the independent variable. The adjusted posttest mean for the NITS group (M=223.491) was higher than the adjusted posttest mean for the ITS group (M=220.976).

Table 3

*Estimates -Adjusted Means*

Dependent Variable: Posttest (Spring MAP)

ITS level	Adjusted Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
NITS	223.491 <sup>a</sup>	.740	222.030	224.952
ITS	220.976 <sup>a</sup>	.740	219.515	222.436

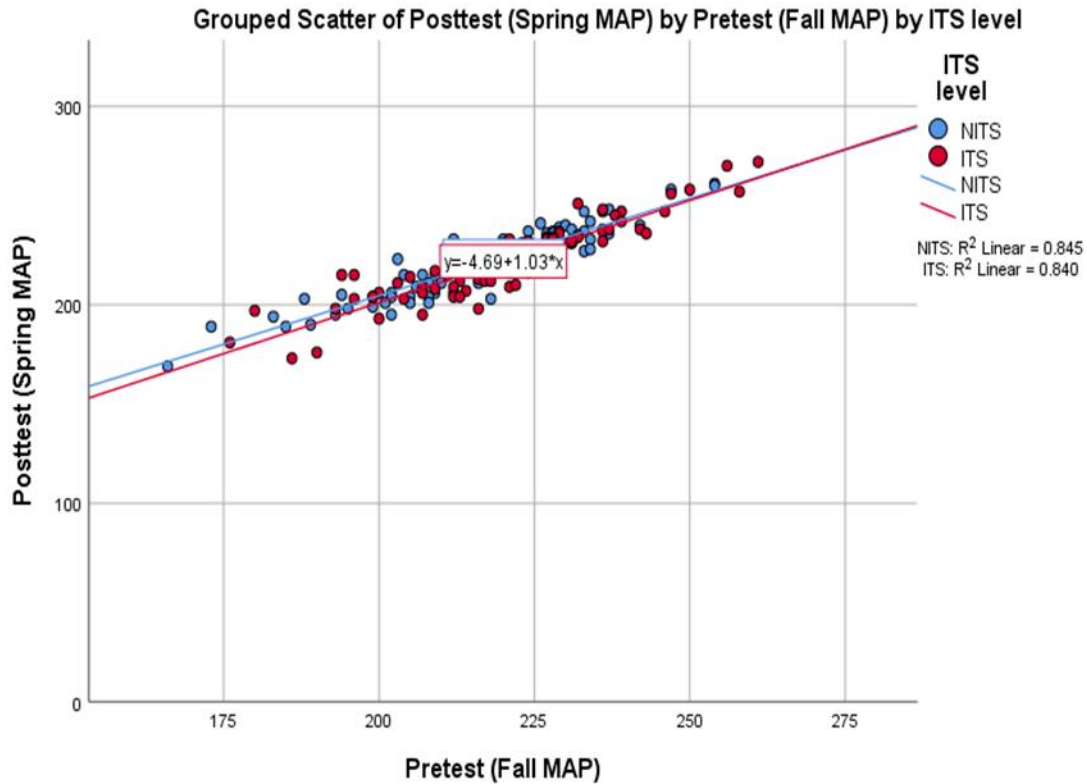
a. Covariates appearing in the model are evaluated at the following values: Pretest (Fall MAP) = 219.22.

**Results****Assumptions Tests**

A one-way ANCOVA was used to test the null hypothesis by comparing the posttest scores (Spring MAP SC 6+) of both groups (ITS/NITS), holding constant pretest (Fall MAP SC 6+) score differences and determining if there existed any statistically significant differences between the adjust means of these independent groups (Gall, Gall, & Borg, 2007). A one-way ANCOVA has four assumptions related to study design that must be considered. There must be one dependent variable measured at the continuous level (posttest scores); one independent variable (level of ITS intervention), which consists of two or more categorical, independent

groups (ITS/NITS); one covariate variable measured at the continuous level (pretest scores); and independence of observations (Laerd Statistics).

A one-way ANCOVA also has six assumptions related to how data fits the one-way ANCOVA model, including linearity, homogeneity of regression slopes, normality, homoscedasticity, homogeneity of variance, and no significant outliers in the groups of the independent variable in terms of the dependent variable (Laerd Statistics). The linearity assumption requires that the covariate (pretest scores) be linearly related to the dependent variable (posttest scores) for each level of the independent variable (NITS/ITS) (Laerd Statistics). There was a linear relationship between pretest scores and post-intervention posttest scores for each intervention group, as assessed by visual inspection of a scatterplot (see Figure 1).



*Figure 1.* Scatterplot showing linear relationship between pretest scores and posttest scores for each intervention group.

The homogeneity of regression slopes assumption necessitates that no interaction exists between the covariate (pretest scores) and the independent variable (ITS/NITS) (Laerd Statistics). Table 4 indicates that there was homogeneity of regression slopes as the interaction term was not statistically significant, ANCOVA model with and without interaction terms,  $F(1,176) = .593, p = .442$ .

Table 4

*Tests of Between-Subjects Effects (ITS\*pretest)*

Dependent Variable: Posttest (Spring MAP)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	46158.499 <sup>a</sup>	3	15386.166	313.288	.000
Intercept	4.167	1	4.167	.085	.771
ITS	43.634	1	43.634	.888	.347
pretest	46031.373	1	46031.373	937.275	.000
ITS * pretest	29.127	1	29.127	.593	.442
Error	8643.701	176	49.112		
Total	8944580.000	180			
Corrected Total	54802.200	179			

a. R Squared = .842 (Adjusted R Squared = .840)

The assumption of normality is also required for statistical significance testing utilizing a one-way ANCOVA. In other words, the dependent variable should be approximately normally distributed for each group of the independent variable (Laerd Statistics). Standardized residuals for the overall model were normally distributed, as assessed by Shapiro-Wilk's test ( $p > .05$ ) (see Table 5).

Table 5

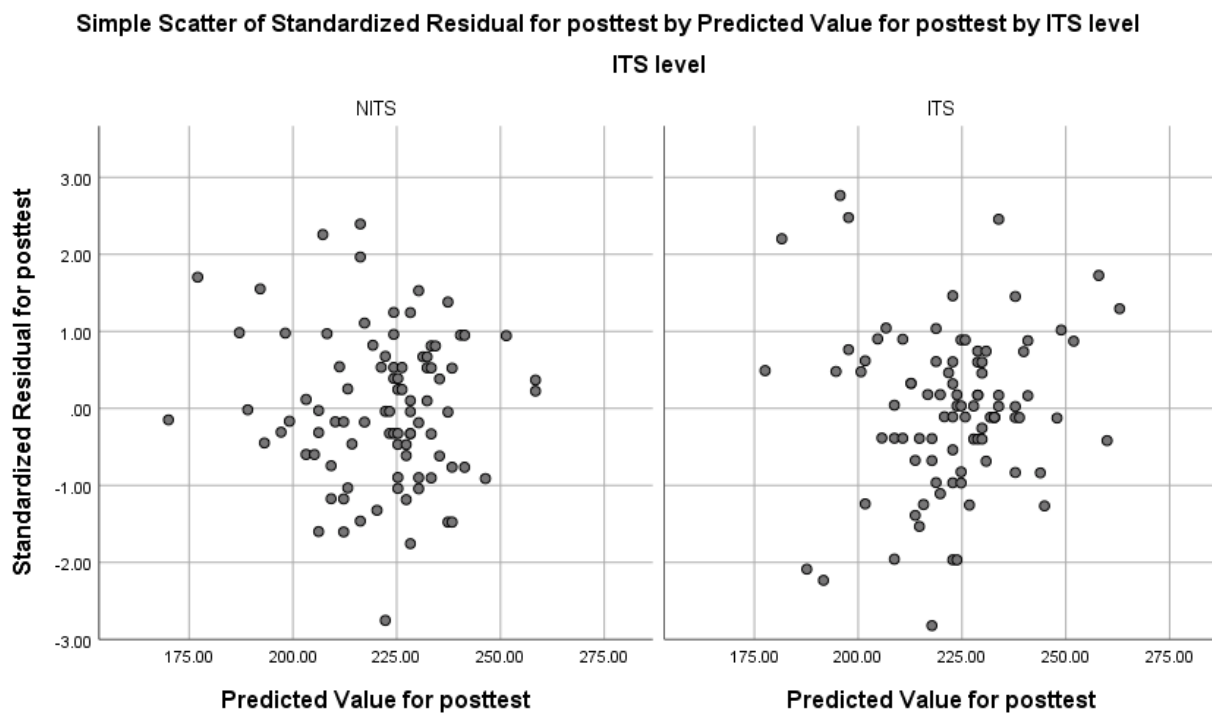
*Tests of Normality*

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual for posttest	.052	180	.200*	.992	180	.480

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Another assumption of a one-way ANCOVA, homoscedasticity, requires that there is homoscedasticity of error variance within each group and the error variances are equal between each group. The assumption of equal error variances can be assessed by a visual inspection of a plot of the standardized residuals against the predicted values for each ITS level (Laerd Statistics). There was homoscedasticity, as assessed by visual inspection of the standardized residuals plotted against the predicted values (see Figure 2).



*Figure 2.* Scatterplots of standardized residuals for posttest by predicted value for posttest by ITS level.

A one-way ANCOVAs also assumes that the variances of the residuals be equal for all groups of the independent variable (Laerd Statistics). As is shown in Table 6, there was homogeneity of variances, as assessed by Levene's test of homogeneity of variance ( $p = .828$ ).

Table 6

*Levene's Test of Equality of Error Variances<sup>a</sup>*

Dependent Variable: Posttest (Spring MAP)

F	df1	df2	Sig.
.047	1	178	.828

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + pretest + ITS

Finally, a one-way ANCOVA requires there should be no significant outliers among standardized residuals where the score is  $\pm 3$  standard deviations (Laerd Statistics). Potential outliers were assessed by consulting the standardized residuals within the data set. The largest standardized residual was 2.76, and the smallest was -2.82; therefore, there were no outliers in the data, as assessed by no cases with standardized residuals greater than  $\pm 3$  standard deviations.

**Hypothesis**

The null hypothesis states that there is no statistically significant difference in math achievement (as shown by the Measures of Academic Progress SC 6+ Math test scores) between middle school students whose math instruction includes an intelligent tutoring system and middle school students whose math instruction does not include an intelligent tutoring system while controlling for pretest scores. A one-way ANCOVA was run to determine the effect of the use of intelligent tutoring systems on math achievement after controlling for pretest scores. After adjustment for pre-test MAP SC 6+ scores, there was a statistically significant difference in posttest scores,  $F(1, 177) = 5.740$ ,  $p = .018$ , partial  $\eta^2 = .031$  (see Table 7).



Table 7

*Tests of Between-Subjects Effects*

Dependent Variable: Posttest (Spring MAP)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	46129.372 <sup>a</sup>	2	23064.686	470.717	.000	.842
Intercept	4.008	1	4.008	.082	.775	.000
pretest	46078.172	1	46078.172	940.390	.000	.842
ITS	281.243	1	281.243	5.740	.018	.031
Error	8672.828	177	48.999			
Total	8944580.000	180				
Corrected Total	54802.200	179				

a. R Squared = .842 (Adjusted R Squared = .840)

Posttest scores were statistically significantly greater in the NITS group as compared to the ITS group ( $M_{\text{diff}} = 2.516$ , 95% CI [0.443, 4.588],  $p = .018$ ) (see Table 8).

Table 8

*Pairwise Comparisons*

Dependent Variable: Posttest (Spring MAP)

					95% Confidence Interval for Difference <sup>b</sup>	
ITS level	ITS level	Mean Difference	Std. Error	Sig. <sup>b</sup>	Lower Bound	Upper Bound
NITS	ITS	2.516 <sup>*</sup>	1.050	.018	.443	4.588
ITS	NITS	-2.516 <sup>*</sup>	1.050	.018	-4.588	-.443

Based on estimated marginal means

\*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

## **CHAPTER FIVE: CONCLUSIONS**

### **Overview**

This study is an examination of middle school math achievement in two similar middle school settings, one that uses intelligent tutoring systems for math instruction and one that does not. Chapter Five includes the discussion, implications, limitations, and suggestions for future research for this study.

### **Discussion**

The purpose of this study was to examine the math achievement of students in two similar suburban middle schools (grades 6-8) in the Southeastern United States. More specifically, this study looked to find differences in the math achievement of middle school students whose math instruction included intelligent tutoring systems and middle school students whose math instruction did not include intelligent tutoring systems. Advocates of intelligent tutoring systems extol their capacity to personalize learning for students. The conceptual underpinnings of personalized learning can be found in Bloom's (1968) theory of mastery learning, the competency-based personalization of learning in Keller's (1968) personalized system of instruction, and Vygotsky's zone of proximal development (ZPD).

Bloom (1984) asserted that mastery learning necessitates optimal conditions in which feedback, corrective measures, and tutoring are readily available to students. Moreover, mastery learning requires instruction that is differentiated (Prast, Van de Weijer-Bergsma, Kroesbergen, & Van Luit, 2018). Complimentary features of mastery learning often include diagnostic assessment, clear-cut learning objectives, sequenced units presented in increasing difficulty, pertinent instruction, formative testing to assess a predetermined mastery of objectives, and conditional progression McGaghie (2015).

Similar to Bloom's mastery learning, Keller's (1968) personalized system of instruction also emphasized repeated assessment, feedback, and tutoring. Moreover, it included self-paced modules to ensure concept mastery before advancement (Paiva, Ferreira, & Frade, 2017). Keller's (1968) condition that students demonstrate mastery or unit-perfection before advancement evolved into the kind of modular, competency-based learning prevalent in many of today's intelligent tutoring systems (Paiva, Ferreira, & Frade, 2017).

Vygotsky (1978) theorized that the zone of proximal development (ZPD) was the peak instructional context. Said another way, ZPD is the optimal circumstances in which educators can interact with their students and move them from independent learning activities to activities just above their present understanding. Pedagogy that involves this type of scaffolding asks teachers to transition their students' learning by recognizing when they are on the brink of more advanced understanding (Goggin, Rankin, Geerlings, & Taggart, 2016).

Personalized learning is customized to students' abilities and involves learning that is student-paced (Hallman, 2019). Its delivery platforms are tailored to carry out instruction according to each student's needs (Horn, 2017), and it relies greatly on diagnostic pretests and posttests to assesses learning and provide timely, customized feedback (Pardo, Jovanovic, Dawson, Gašević, & Mirriahi, 2019). In recent decades personalization within the classroom has been pursued through 1:1 technology initiatives (Hallman, 2019) in which programs like intelligent tutoring systems are employed to deliver learning experiences that are customized to students' achievement levels, interests, and learning styles (Hopkins, 2019).

The research question in this study, "Is there a difference in the math achievement of middle school students whose math instruction includes an intelligent tutoring system and middle school students whose math instruction does not include an intelligent tutoring system?" is an

examination of intelligent tutoring systems as a vehicle for providing personalized learning experiences and whether or not these instructional tools boost math achievement.

Research comparing intelligent tutoring systems and other adaptive learning technologies to traditional learning environments in terms of achievement has been widely positive. Across multiple academic disciplines, intelligent tutoring systems have been linked to positive effects (Ma et al., 2014; Steenbergen-Hu & Cooper, 2014; VanLehn, 2011; Kulik & Fletcher, 2016) and even found comparable to human tutors (VanLehn, 2011). They have been linked to increased academic achievement at every level of education and in almost all subject domains. However, in some studies, differences were not significant when compared to students who received small-group or one-on-one instruction (Ma et al., 2014; Graesser et al., 2016). In other research, outcomes supporting the use of intelligent tutoring systems found them even more effective than human tutors (Kulik & Fletcher, 2016).

Still, other research suggests intelligent tutoring systems may not be significantly effective, especially in mathematics (Slavin, Lake, & Groff, 2009; Steenbergen-Hu & Cooper, 2013; U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse, 2009). Steenbergen-Hu and Cooper (2013) analyzed 34 cases of intelligent tutoring system use in K–12 mathematics and discovered a difference between intelligent tutoring system groups and non-intelligent tutoring system groups of only 0.05 standard deviations.

The one-way ANCOVA executed in this study found a statistically significant difference between ITS and non-ITS interventions on math achievement as measured by the posttest (Spring MAP Math SC 6+) while controlling for pretest scores (Fall MAP Math SC 6+),  $F(1, 177) = 5.740, p = .018, \text{partial } \eta^2 = .031$ . Additionally, posttest scores were statistically

significantly greater in the NITS group as compared to the ITS group ( $M_{\text{diff}} = 2.516$ , 95% CI [0.443, 4.588],  $p = .018$ ).

This research is at odds with broader studies that examined intelligent tutoring systems across multiple subject domains. VanLehn (2011) found intelligent tutoring systems comparable to human tutors at improving student achievement across subject areas, increasing test scores 0.58 standard deviations over traditional instruction. Ma, Ma et al. (2014) also found that intelligent tutoring systems were associated with greater academic achievement at all levels of education when compared to traditional, teacher-led instruction in almost all subject domains with an average effect of 0.43 standard deviations. Steenbergen-Hu and Cooper (2014) surveyed 22 different types of intelligent tutoring systems employed over a twenty-year span and found that intelligent tutoring systems increased scores approximately 0.35 standard deviations; whereas, Kulik and Fletcher (2016) found that students taught with intelligent tutoring systems outperformed those in traditional, teacher-led classes in 46 of 50 controlled evaluations.

Conversely, this research seems to agree with the findings that intelligent tutoring systems may not be effective in mathematics instruction. The What Works Clearinghouse looked at 27 evaluations of intelligent tutoring systems used in Algebra I classrooms and found insignificant effect sizes. Moreover, Slavin et al. (2009) examined intelligent tutoring system use in middle and high school mathematics, finding that intelligent tutoring systems increased student test scores by an average of only 0.12 standard deviations.

### **Implications**

The Program for International Student Assessment (PISA), a cross-national assessment administered once every three years, ranked the U.S. 38<sup>th</sup> out of 71 participating countries in mathematics (DeSilver, 2017). As such, it is not surprising that technology initiatives, including

intelligent tutoring systems, have flourished in K-12 classrooms in the U.S. as a way of addressing these disparities and improving math education as a whole. Findings from this research help close a gap in the literature on intelligent tutoring systems, especially as it relates to the efficacy of these instructional tools in middle school classrooms. Simply put, there are few studies that examine the effect of intelligent tutoring systems on math achievement, and there are fewer still that focus on middle school populations. As such, the implications of this added information are numerous.

The financial implications of technology use in the classroom necessitate a fair evaluation of these practices. According to the National Center for Education Statistics, total expenditures for public schools are currently over \$700 billion (approximately \$14,000 per/student) annually. Cumulative estimates on K-12 technology costs are sparse as these funds are often dispersed across multiple categories; however, one report analyzed \$2 billion in K-12 spending and found that one of the most prevalent underutilized resources in early, middle, and high school settings was educational technology (Glimpse, 2019).

Despite an emphasis on technology use to improve academic achievement, the National Assessment of Educational Progress (NAEP) pointed out that over 60% of students in the U.S. are deficient in math (National Center for Educational Statistics, 2016). Additionally, those students who have difficulties in math consistently lag behind their classmates throughout their school careers (Nelson, Parker, & Zaslofsky, 2016). Math deficiencies can have far-reaching ramifications for students, including diminished self-efficacy in this domain, which often leads to an avoidance of courses and careers that require math skills (Huang, Zhang, & Hudson, 2018).

This research is important for several reasons. As previously mentioned, this study examines a population that is often neglected in educational literature, middle school students.

Moreover, the sampling within this study was substantial when compared to related studies, included all pertinent grade levels, and was pulled from the entire student enrollments of both schools. Most importantly, this study serves as a reminder that pedagogical endeavors, most if not all of which are very costly in terms of both time and money, must be carefully scrutinized. The literature seems to suggest that for many academic domains, intelligent tutoring systems are positively linked with academic achievement. As the foundations of these systems are inherently grounded in providing personalized learning experiences, their success should come with little surprise. Nonetheless, delivering math instruction via an intelligent tutoring system may be problematic. The results of this research suggest that math achievement scores were actually higher in the group that did not receive the ITS treatment. Since these results do not suggest any increase in math achievement associated with the use of intelligent tutoring systems, at least in the population studied, it seems logical that more research is warranted.

### **Limitations**

As is the case with all research, this study has its limitations, which must be acknowledged in order to provide perspective for future studies. One such limitation stems from the populations being studied. Though the sampling utilized is relatively large when compared to similar research, it nonetheless involves rural and suburban populations with similar demographic and socioeconomic makeup. Both populations also share common proximity within a single state. As such, the extent to which these results are generalizable to dissimilar populations and settings must be considered. Moreover, there are also validity issues inherent to the use of convenience sampling which limits data collection to that which can be taken from available participants (Gall, Gall, & Borg, 2007). Certain limitations are also inherent in causal-comparative research in general. A causal-comparative design does not permit the researcher to

influence the independent variable to discern its effect on the dependent variable. As such, this type of design also does not permit strong conclusions about cause-and-effect (Gall, Gall, & Borg, 2007).

Another limitation stems from the independent variable considered in this study. For this research, the independent variable was specifically defined as the supplementary use of two intelligent tutoring systems, Pearson's Math Digits, and IXL, for math instruction. This research does not, however, delineate the degree to which either tutoring system was implemented, e.g., homework practice, remediation, in-class practice, etc. Nor does it consider the amount of time students worked with each program nor the disparities among teaching styles, competencies, and pedagogies that most likely existed within each context. It stands to reason, therefore, that any relationship between the independent variable and math achievement is likely influenced by other factors as well. Finally, because this study features only one school that utilizes two specific intelligent tutoring systems, its results may not be representative of educational settings that use other intelligent tutoring systems for math instruction.

### **Recommendations for Future Research**

Though this study of intelligent tutoring systems helps to fill a gap in the literature, the examination of adaptive learning software in middle school math instruction, there are still many questions that remain. Considering the positives associated with personalized instruction in K-12 environments and the inherent difficulties in achieving true customization within the classroom, intelligent tutoring systems make sense. Moreover, much of the literature surrounding this technology has been positive. It is, therefore, imperative to seek out opportunities to study this pedagogy and how it can be implemented successfully in all subject domains and within diverse educational settings.



One recommendation for future research would be to examine the use of intelligent tutoring systems for math instruction in larger populations of middle school students, e.g., studies that involve multiple school districts or even multiple states. It might also be beneficial to examine its use in populations with demographics and socioeconomics different than the ones included in this study. Another recommendation would be to take a more experimental approach to this type of research for a more rigorous assessment of the cause-and-effect relationship between intelligent tutoring systems and math achievement.

Much could also be learned if different types of intelligent tutoring software were included in a larger-scale study and if specific aspects of the populations were taken into consideration, i.e., gender, aptitude, and learning preferences. Specifying the instructional degree to which the intelligent tutoring system is used (i.e., primary, supplementary/complementary, homework, or remediation) or even delineating the specifics of when and how often it is used could provide critical information.

This study, like the studies before it, compared intelligent tutoring systems to traditional, teacher-led instruction. As intelligent tutoring systems and other computer-assisted instruction grow in use, additional research comparing these instructional tools to one another could be advantageous. In addition, research comparing the effect of similar intelligent tutoring systems on different domains might provide some insight into why they are more beneficial in some subject areas than others. Finally, as only the MAP Math SC 6+ was used to assess math achievement in this study, it would benefit the study of this topic to include other assessment tools to help gauge learning.

## REFERENCES

- A nation at risk: 1983 report of the commission on excellence in education. (1997).  
*Congressional Digest*, 76(11), 258.
- Acedo, C., & Hughes, C. (2014). Principles for learning and competences in the 21st-century curriculum. *Prospects*, 44(4), 503-525. doi: 10.1007/s11125-014-9330-1
- Ahad, M. A., Tripathi, G., & Agarwal, P. (2018). Learning analytics for IoE based educational model using deep learning techniques: *Architecture, challenges and applications*. *Smart Learning Environments*, 5(1), 1-16. doi: 10.1186/s40561-018-0057-y
- Akera, A. (2017). Bringing radical behaviorism to revolutionary Brazil and back: Fred Keller's personalized system of instruction and cold war engineering education. *Journal of the History of the Behavioral Sciences*, 53(4), 364-382. doi: 10.1002/jhbs.21871
- Alkhatlan, A., & Kalita, J. (2018). Intelligent tutoring systems: A comprehensive historical survey with recent developments. arXiv Preprint arXiv:1812.09628.
- Anne, L., & Haney, R. (2017). Differentiation (DI) in higher education (HE): Modeling what we teach with pre-service teachers. *Journal of Higher Education Theory and Practice*, 17(5), 28-38.
- Baker, R. S. (2014). Educational data mining: An advance for intelligent systems in education. *IEEE Intelligent Systems*, 29(3), 78-82. doi: 10.1109/MIS.2014.42
- Baker, R. S., & Corbett, A. T. (2014). Assessment of robust learning with educational data mining. *Research & Practice in Assessment*, 9(2), 38.
- Bartelet, D., Ghysels, J., Groot, W., Haelermans, C., & van den Brink, H.M. (2016). The differential effect of basic mathematics skills homework via a web-based intelligent tutoring system across achievement subgroups and mathematics domains: A randomized

- field experiment. *Journal of Educational Psychology*, 108(1), 1-20. doi: 10.1037/edu0000051
- Basham, J. D., Hall, T. E., Carter, R. A., & Stahl, W. M. (2016). An operationalized understanding of personalized learning. *Journal of Special Education Technology*, 31(3), 126-136. doi: 10.1177/0162643416660835
- Basu, S., Biswas, G., & Kinnebrew, J. S. (2017). Learner modeling for adaptive scaffolding in a computational thinking-based science learning environment. *User Modeling and User-Adapted Interaction*, 27(1), 5-53. doi: 10.1007/s11257-017-9187-0
- Bienkowski, M., Feng, M., & Means, B. (2014). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief
- Bernacki, M. L., & Walkington, C. (2018). The role of situational interest in personalized learning. *Journal of Educational Psychology*, 110(6), 864-881. doi: 10.1037/edu0000250
- Bingham, A. J. (2019). A look at personalized learning: Lessons learned. *Kappa Delta Pi Record*, 55(3), 124-129. doi:10.1080/00228958.2019.1622383
- Bingham, A. J., Pane, J. F., Steiner, E. D., & Hamilton, L. S. (2018). Ahead of the curve: Implementation challenges in personalized learning school models. *Educational Policy*, 32(3), 454-489. doi: 10.1177/0895904816637688
- Bloom, B. S. (1968). Learning for Mastery. Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reprints, Number 1.
- Bloom, B.S. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, 13(6), 4-16. Retrieved from <http://www.jstor.org/stable/1175554>

- Brevik, L. M., A. E. Gunnulfsen, and J. S. Renzulli. 2018. Student teachers' practice and experience with differentiated instruction for students with higher learning potential. *Teaching and Teacher Education* 71, 4–45. doi:10.1016/j.tate.2017.12.003.
- Bringula, R. P., Basa, R. S., Dela Cruz, C., & Rodrigo, M. M. T. (2016). Effects of prior knowledge in mathematics on learner-interface interactions in a learning-by-teaching intelligent tutoring system. *Journal of Educational Computing Research*, 54(4), 462-482. doi: 10.1177/0735633115622213
- Brusilovsky, P., Peylo, C. (2003). Adaptive and Intelligent Web-based Educational Systems. *International Journal of Artificial Intelligence in Education*, 13, 156-169. Retrieved from <https://www.pitt.edu/~peterb/papers/AIWBES.pdf>
- Burns, M. K., & Young, H. (2016). Test review: Measures of academic progress skills: Northwest Evaluation Association. *Journal of Psychoeducational Assessment*, 1(4). doi: 0.1177/0734282918783509
- Chatti, M. A., & Muslim, A. (2019). The PERLA framework: Blending personalization and Learning analytics. *The International Review of Research in Open and Distributed Learning*, 20(1) doi:10.19173/irrodl.v20i1.3936
- Cheung, A. C., & Slavin, R. E. (2013). Effects of educational technology applications on reading outcomes for struggling readers: A best-evidence synthesis. *Reading Research Quarterly*, 48(3), 277– 299.
- Childress, S., & Benson, S. (2014). Personalized learning for every student every day. *The Phi Delta Kappa*, 95(8), 33-38. doi: 10.1177/003172171409500808

- Christensen, C. M., Horn, M. B., & Johnson, C. W. (2011). *Disrupting class: How disruptive innovation will change the way the world learns* (2nd Ed.). New York, NY: McGraw-Hill.
- Chukwuedo, S.O., & Uko-Aviomoh, E.E. (2015). Building theoretical and conceptual frameworks for quantitative research report in education. *African Journal of Studies in Education*, 10(20), 83-101.
- Clément, B., Roy, D., Oudeyer, P., & Lopes, M. (2015). Multi-armed bandits for intelligent tutoring systems. *Journal of Educational Data Mining (JEDM)*, 7(2), 20-48.
- Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). Learning styles and pedagogy in post16 learning: A Systematic and critical review. Retrieved from:  
<http://www.leerbeleving.nl/wp-content/uploads/2011/09/learning-styles.pdf>
- Cuirong, Y., Weidong, C., & Hongtao, W. (2016). Sim-Agent and Student Interaction in IT'S: From Interdisciplinary Research and Design. *Modern Distance Education Research*, (6), 14.
- Dale, R., & Scherrer, J. (2015). Goldilocks discourse – math scaffolding that's just right. *The Phi Delta Kappan*, 97(2), 58-61.
- Delen, E., Liew, J., & Wilson, V. (2014). Effects of interactivity and instructional scaffolding on learning: Self-regulation in online video-based environments. *Computers & Education*, 78, 312-320. doi:10.1016/j.compedu.2014.06.018
- DeSilver, D. (2017). US students' academic achievement still lags that of their peers in many other countries. *Pew Research Center*, 15.
- El Mamoun, B., Erradi, M., & El Mhouti, A. (2018). Using an intelligent tutoring system to support learners' WMC in e-learning: Application in mathematics learning. *International*

- Journal of Emerging Technologies in Learning (iJET)*, 13(12), 142-156.  
doi:10.3991/ijet.v13i12.8938
- Elazhary, H. H., & Khodeir, N. (2017). A cognitive tutor of Arabic word root extraction using artificial word generation, scaffolding and self-explanation. *International Journal of Emerging Technologies in Learning (iJET)*, 12(5), 36. doi:10.3991/ijet.v12i05.6651
- Essa, A. (2016). A possible future for next-generation adaptive learning systems. *Smart Learning Environments*, 3(1), 1-24. doi: 10.1186/s40561-016-0038-y
- Eun, B. (2019). The zone of proximal development as an overarching concept: A framework for synthesizing Vygotsky's theories. *Educational Philosophy and Theory*, 51(1), 18-30.  
doi:10.1080/00131857.2017.1421941
- Farias, C., Hastie, P. A., & Mesquita, I. (2018). Scaffolding student-coaches' instructional leadership toward student-centered peer interactions: A yearlong action-research intervention in sport education. *European Physical Education Review*, 24(3), 269-291.  
doi: 10.1177/1356336X16687303
- Farrell, T., & Rushby, N. (2016). Assessment and learning technologies: An overview. *British Journal of Educational Technology*, 47(1), 106-120. doi:10.1111/bjet.12348
- Frith, V., Jaftha, J., Prince, R. (2004). Evaluating the effectiveness of interactive computer tutorials for an undergraduate mathematical literacy course. *British Journal of Educational Technology* 35(2): 159–171. doi:10.1111/j.0007-1013.2004.00378.x.
- Funkhouser, C. (2003). The effects of computer-augmented geometry instruction on student performance and attitudes. *Journal of Research on Technology in Education* 35(2): 163–175.

- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). New York, NY: Pearson/Allyn & Bacon.
- Gallardo-Echenique, E.E., Marques-Molias, L., Bullen, M., & Stijbos, J. (2015). Let's talk about digital learners in the digital era. *International Review of Research in Open and Distributed Learning* 16(3), 156-187. Retrieved from <http://files.eric.ed.gov/fulltext/EJ1067883.pdf>
- Glimpse K12 analysis of school spending shows that two-thirds of software license purchases go unused: Other top underutilized resources are professional development and printed materials. (2019, May 15). *NASDAQ OMX's News Release Distribution Channel* Retrieved from <http://ezproxy.liberty.edu/login?url=https://search-proquest-com.ezproxy.liberty.edu/docview/2225083178?accountid=12085>
- Goggin, T., Rankin, S., Geerlings, P., & Taggart, A. (2016). Catching them before they fall: A Vygotskian approach to transitioning students from high school to university. *Higher Education Research & Development*, 35(4), 698-711.  
doi:10.1080/07294360.2015.1137879
- Golden, N. A. (2017). Soft(a)ware in the English classroom. *English Journal*, 106(4), 86.
- González-Calero, J. A., Arnau, D., Puig, L., & Arevalillo-Herráez, M. (2015). Intensive scaffolding in an intelligent tutoring system for the learning of algebraic word problem solving. *British Journal of Educational Technology*, 46(6), 1189-1200.  
doi:10.1111/bjet.12183
- Graesser, A. C., Hu, X., Nye, B. D., VanLehn, K., Kumar, R., Heffernan, C., Baer, W. (2018). ElectronixTutor: An intelligent tutoring system with multiple learning resources for

- electronics. *International Journal of STEM Education*, 5(1), 1-21. doi: 10.1186/s40594-018-0110-y
- Green, B. F., Bock, R. D., Humphreys, L. G., Linn, R. L., Reckase, M. D. (1984). Technical guidelines for assessing computerized adaptive tests. *Journal of Educational Measurement*, 21, 347-360. doi:10.1111/j.1745-3984.1984.tb01039.x
- Grundmeyer, T. (2014). Adopting technology: Using student qualitative data and Gartner's Hype Cycle. *Journal of Education and Training Studies* 2(1), 207-216. doi:10.11114/jets.v2i1.228
- Gudivada, V. N. (2017). Cognitive analytics-driven personalized learning. *Educational Technology*, 57(1), 23-31. Retrieved from <https://www.jstor.org/stable/44430537>
- Guseva, L. G., & Solomonovich, M. (2017). Implementing the zone of proximal development: From the pedagogical experiment to the developmental education system of Leonid Zernov. *International Electronic Journal of Elementary Education*, 9(4), 775-786.
- Guskey, T. R. (2007). Closing achievement gaps: Revisiting Benjamin S. Bloom's "learning for mastery". *Journal of Advanced Academics*, 19(1), 8.
- Hafidi, M., & Bensebaa, T. (2015). Architecture for an adaptive and intelligent tutoring system that considers the learner's multiple intelligences. *International Journal of Distance Education Technologies (IJDET)*, 13(1), 1-21. doi:10.4018/ijdet.2015010101
- Hallman, H. L. (2019). Personalized learning through 1:1 technology initiatives: Implications for teachers and teaching in neoliberal times. *Teaching Education*, 30(3), 299-318. doi:10.1080/10476210.2018.1466874
- Hooshyar, D., Ahmad, R. B., Yousefi, M., Fathi, M., Abdollahi, A., Horng, S., & Lim, H. (2016). A solution-based intelligent tutoring system integrated with an online game-



- based formative assessment: Development and evaluation. *Educational Technology Research and Development*, 64(4), 787-808. doi: 10.1007/s11423-016-9433-x
- Hooshyar, D., Ahmad, R. B., Yousefi, M., Yusop, F. D., & Horng, S. (2015). A flowchart-based intelligent tutoring system for improving problem-solving skills of novice programmers. *Journal of Computer Assisted Learning*, 31(4), 345-361. doi:10.1111/jcal.12099
- Hopkins, D. (2019). Leadership of personalised learning. *Australian Educational Leader*, 41(1), 18-23.
- Horn, M. B. (2017). Now trending: Personalized learning. *Education Next*, 17(4).
- Hsu, W., & Li, C. (2015). A competency-based guided-learning algorithm applied on adaptively guiding e-learning. *Interactive Learning Environments*, 23(1), 106-125. doi:10.1080/10494820.2012.745432
- Hu, L. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling* 6, 55. doi: 10.1080/10705519909540118
- Huang, X., Craig, S. D., Xie, J., Graesser, A., & Hu, X. (2016). Intelligent tutoring systems work as a math gap reducer in 6<sup>th</sup>-grade after-school program. *Learning and Individual Differences*, 47, 258-265. doi:10.1016/j.lindif.2016.01.012
- Huang, X., Zhang, J., & Hudson, L. (2018). Impact of math self-efficacy, math anxiety, and growth mindset on math and science career interest for middle school students: the gender moderating effect. *European Journal of Psychology of Education* 34. 621-640. doi: 10.1007/s10212-018-0403-z
- Kane, M. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73. doi: 10.1111/jedm.12000

- Kaser, T., Klingler, S., Schwing, A. G., & Gross, M. (2017). Dynamic Bayesian networks for student modeling. *IEEE Transactions on Learning Technologies*, 10(4), 450-462. doi:10.1109/TLT.2017.2689017
- Keller, F.S. (1968). Good-bye teacher. *Journal of applied behavior analysis*, 1(1), 79-89. doi: 10.1901/jaba.1968.1-79
- Kessler, A. M., Stein, M. K., & Schunn, C. D. (2015). Cognitive demand of model tracing tutor tasks: Conceptualizing and predicting how deeply students engage. *Technology, Knowledge and Learning*, 20(3), 317-337. doi: 10.1007/s10758-015-9248-6
- Khodier, N., Elazhary, H., & Wanas, N. (2017). Rule-based cognitive modeling and model tracing in a math story problem tutor. *International Journal of Emerging Technologies in Learning (iJET)*, 12(4), 111. doi:10.3991/ijet.v12i04.6592
- Khodeir, N., Wanas, N., & Elazhary, H. (2018). Constraint-based student modeling in probability story problems with scaffolding techniques. *International Journal of Emerging Technologies in Learning (iJET)*, 13(1), 178-205. doi:10.3991/ijet.v13i01.7397
- Kinshuk, K. (2015). Adaptive learning software and platforms. In J. Spector (Ed.), *The SAGE encyclopedia of educational technology* (p. 8010). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781483346397.n12
- Kivunja, C. (2014). Theoretical perspectives of how digital natives learn. *Turkish Online Journal of Distance Education* 3 (1), 94-109. Retrieved from <http://eric.ed.gov/?q=Theoretical+perspectives+of+how+digital+natives+learn.++&id=EJ1067467>

- Kong, S. C., & Song, Y. (2015). An experience of personalized learning hub initiative embedding BYOD for reflective engagement in higher education. *Computers & Education*, 88, 227-240. doi:10.1016/j.compedu.2015.06.003
- Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, 86(1), 42-78. doi: 10.3102/0034654315581420
- Laerd Statistics (2018). One-way ANCOVA using SPSS Statistics. *Statistical tutorials and software guides*. Retrieved from <https://statistics.laerd.com/>
- Lee, D., Huh, Y., Lin, C., & Reigeluth, C. M. (2018). Technology functions for personalized learning in learner-centered schools. *Educational Technology Research and Development*, 66(5), 1269-1302. doi: 10.1007/s11423-018-9615-9
- Li, N., Matsuda, N., Cohen, W. W., & Koedinger, K. R. (2015). Integrating representation learning and skill learning in a human-like intelligent agent. *Artificial Intelligence*, 219, 67-91. doi:10.1016/j.artint.2014.11.002
- Lin, C. F., Yeh, Y., Hung, Y. H., & Chang, R. I. (2013). Data mining for providing a personalized learning path in creativity: An application of decision trees. *Computers & Education*, 68, 199-210. doi:10.1016/j.compedu.2013.05.009
- Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, 106(4), 901-918. doi: 10.1037/a0037123
- Malekzadeh, M., Mustafa, M. B., & Lahsasna, A. (2015). A review of emotion regulation in intelligent tutoring systems. *Journal of Educational Technology & Society*, 18(4), 435-445.

- Mangaroska, K., Giannakos, M.N, (2018). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies*. 11, 1–19.
- Marzano, R. J. (2017). *Beyond reform: Systemic shifts toward personalized learning*. Bloomington, Indiana: Marzano Research.
- Matsuda, N., Yarzebinski, E., Keiser, V., Raizada, R., Cohen, W. W., Stylianides, G. J., & Koedinger, K. R. (2013). Cognitive anatomy of tutor learning: Lessons learned with SimStudent. *Journal of Educational Psychology*, 105(4), 1152-1163. doi: 10.1037/a0031955
- Maxwell, J. (2012). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage Publications.
- Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). New York: Psychology Press.
- Mestad, I., & Kolstø, S. D. (2014). Using the concept of zone of proximal development to explore the challenges of and opportunities in designing discourse activities based on practical work. *Science Education*, 98(6), 1054 -1076. doi:10.1002/sce.21139
- Millis, K., Forsyth, C., Forsyth, C., Wallace, P., Wallace, P., . . . Timmins, G. (2017). The impact of game-like features on learning from an intelligent tutoring system. *Technology, Knowledge and Learning*, 22(1), 1-22. doi:10.1007/s10758-016-9289-5
- Mitee, T.L., & Obaitan, G.N. (2015). Effect of mastery learning on senior secondary school students' cognitive learning outcomes in quantitative chemistry. *Journal of Education and Practice*, 6(5), 34-28. Retrieved from [tps://files.eric.ed.gov/fulltext/EJ1083639.pdf](https://files.eric.ed.gov/fulltext/EJ1083639.pdf)

- Muis, K. R., Ranellucci, J., Trevors, G., & Duffy, M. C. (2015). The effects of technology-mediated immediate feedback on kindergarten students' attitudes, emotions, engagement and learning outcomes during literacy skills development. *Learning and Instruction*, 38, 1-13. doi:10.1016/j.learninstruc.2015.02.001
- McGaghie, W. C. (2015). When I say ... mastery learning. *Medical Education*, 49(6), 558-559. doi:10.1111/medu.12679
- Millis, K., Forsyth, C., Wallace, P., Graesser, A. C., & Timmins, G. (2017). The impact of game-like features on learning from an intelligent tutoring system. *Technology, Knowledge and Learning*, 22(1), 1-22. doi: 10.1007/s10758-016-9289-5
- Mitee, T.L., & Obaitan, G.N. (2015). Effect of mastery learning on senior secondary school students' cognitive learning outcomes in quantitative chemistry. *Journal of Education and Practice*, 6(5), 34-28. Retrieved from [tps://files.eric.ed.gov/fulltext/EJ1083639.pdf](https://files.eric.ed.gov/fulltext/EJ1083639.pdf)
- Mousavinasab, E., Zarifsanaiey, N., R. Niakan Kalhori, S., Rakhshan, M., Keikha, L., & Ghazi Saeedi, M. (2018). Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, 1-22. doi:10.1080/10494820.2018.1558257
- Najar, A. S., Mitrovic, A., & McLaren, B. M. (2016). Learning with intelligent tutors and worked examples: Selecting learning activities adaptively leads to better learning outcomes than a fixed curriculum. *User Modeling and User-Adapted Interaction*, 26(5), 459-491. doi: 10.1007/s11257-016-9181-y
- National Center for Education Statistics. (2016). *The condition of education 2016 (NCES 2016-144)*. National Center for Education Statistics, Institute of Education Sciences. Washington, DC: U.S. Department of Education.

- Nelson, P. M., Parker, D. C., & Zaslofsky, A. F. (2016). The relative value of growth in math fact skills across late elementary and middle school. *Assessment for Effective Intervention, 41*(3), 184-192. doi: 10.1177/1534508416634613
- Nesbit, J. C., Adesope, O. O., Liu, Q., & Ma, W. (2014). How effective are intelligent tutoring systems in computer science education? Paper presented at the 99-103.  
doi:10.1109/ICALT.2014.38
- Northwest Evaluation Association. (2008, September). *RIT scale norms: For use with Measures of Academic Progress*. Portland, OR: Author.
- Northwest Evaluation Association. (2011, January). *Technical manual for Measure of Academic Progress & Measure of Academic Progress for Primary Grades*. Portland, OR: Author.
- Nye, B. D., Pavlik Jr, P. I., Windsor, A., Olney, A. M., Hajeer, M., & Hu, X. (2018). SKOPE-IT (shareable knowledge objects as portable intelligent tutors): Overlaying natural language tutoring on an adaptive learning system for mathematics. *International Journal of STEM Education, 5*(1), 1-20. doi: 10.1186/s40594-018-0109-4
- O'Neill, G., Donnelly, R., & Fitzmaurice, M. (2014). Supporting program teams to develop sequencing in higher education curricula. *International Journal for Academic Development, 19*(4), 268-280. doi:10.1080/1360144X.2013.867266
- Paiva, R. C., Ferreira, M. S., & Frade, M. M. (2017). Intelligent tutorial system based on a personalized system of instruction to teach or remind mathematical concepts. *Journal of Computer Assisted Learning, 33*(4), 370-381. doi:10.1111/jcal.12186

- Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society*, 17(4), 49-64.
- Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalized feedback: Learning analytics to scale personalized feedback. *British Journal of Educational Technology*, 50(1), 128-138. doi:10.1111/bjet.12592
- Patrick, S., Kennedy, K., & Powell, A. (2013). *Mean what you say: Defining and integrating personalized, blended and competency education*. Vienna, VA: International Association for K-12 Online Learning.
- Paz-Albo, J. (2017). Is personalized learning the future of school? *Childhood Education*, 93(4), 295-299. doi:10.1080/00094056.2017.1343581
- Pekola, T., Rasila, A., & Sangwin, C. (2018). Investigating bloom's learning for mastery in mathematics with online assessment. *Informatics in Education*, 17(2), 363-380. doi:10.15388/infedu.2018.19
- Pelánek, R. (2017). Bayesian knowledge tracing, logistic models, and beyond: An overview of learner modeling techniques. *User Modeling and User-Adapted Interaction*, 27(3), 313-350. doi: 10.1007/s11257-017-9193-2
- Pepin, B. B., Choppin, J., Ruthven, K., & Sinclair, N. (2017). Digital curriculum resources in mathematics education: Foundations for change. *ZDM Mathematics Education*, 49(5), 645-661. doi: 10.1007/s11858-017-0879-z

- Poitras, E. G., Lajoie, S. P., Doleck, T., & Jarrell, A. (2016). Subgroup discovery with user interaction data: An empirically guided approach to improving intelligent tutoring systems. *Journal of Educational Technology & Society*, 19(2), 204-214.
- Prast, E. J., Van de Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. H. (2018). Differentiated instruction in primary mathematics: Effects of teacher professional development on student achievement. *Learning and Instruction*, 54, 22-34. doi: 10.1016/j.learninstruc.2018.01.009
- Predrag, D., Jovan, D., Bojan, C., & Veis, Š. (2016). A review of intelligent tutoring systems in e-learning. *Annals of the Oradea University: Fascicle Management and Technological Engineering*, XXV (XV)
- Rajendran, R., Iyer, S., & Murthy, S. (2019). Personalized effective feedback to address students' frustration in ITS. *IEEE Transactions on Learning Technologies*, 12(1), 87-97. doi:10.1109/TLT.2018.2807447
- Rastegarmoghadam, M., & Ziarati, K. (2017). Improved modeling of intelligent tutoring systems using ant colony optimization. *Education and Information Technologies*, 22(3), 1067-1087. doi: 10.1007/s10639-016-9472-2
- Rau, M. A., Michaelis, J. E., & Fay, N. (2015). Connection making between multiple graphical representations: A multi-method approach for domain-specific grounding of an intelligent tutoring system for chemistry. *Computers & Education*, 82, 460-485. doi:10.1016/j.compedu.2014.12.009
- Reddy, P. D., & Sasikumar, M. (2014). Student model for an intelligent language tutoring system. Paper presented at the 441-443. doi:10.1109/ICALT.2014.132



- Roschelle, J., Gaudino, S., Darling, S. (2016) Designing for consistent implementation of a 5<sup>th</sup>-grade digital math curriculum. *International Journal of Design for Learning*, 7(2), 41-56.
- Rovai, A.P., Baker, J.D., & Ponton, M.K. (2013). *A practitioner's guide to research methods and IBM SPSS analysis*. Chesapeake, VA: Watertree Press.
- Rus, V., & Ștefănescu, D. (2016). Non-intrusive assessment of learners' prior knowledge in dialogue-based intelligent tutoring systems. *Smart Learning Environments*, 3(1), 1-18.  
doi:<http://dx.doi.org.ezproxy.liberty.edu/10.1186/s40561-016-0025-3>
- Sabo, K. E., Atkinson, R. K., Barrus, A. L., Joseph, S. S., & Perez, R. S. (2013). Searching for the two sigma advantage: Evaluating algebra intelligent tutors. *Computers in Human Behavior*, 29(4), 1833-1840. doi:10.1016/j.chb.2013.03.001
- Samejima, F. (1977). A use of the information function in tailored testing. *Applied Psychological Measurement*, 1 (3), 233-247. doi: 10.1177/014662167700100209
- Samejima, F. (1994). Estimation of reliability coefficients using the test information function and its modifications. *Applied Psychological Measurement*, 18 (3), 229-244.  
doi.org/10.1177/014662169401800304
- Serrano, M. -, Vidal-Abarca, E., & Ferrer, A. (2018). Teaching self-regulation strategies via an intelligent tutoring system (TuiNLECweb): Effects for low-skilled comprehenders. *Journal of Computer Assisted Learning*, 34(5), 515-525. doi:10.1111/jcal.12256
- Sharada, N., Shashi, M., & Madhavi, D. (2015). A comprehensive study on intelligent tutoring systems. *International Journal of Advanced Studies in Computers, Science and Engineering*, 4(8), 1.

- Slavin, R. E., Lake, C., & Groff, C. (2009). Effective programs in middle and high school mathematics: A best-evidence synthesis. *Review of Educational Research*, 79, 839–911. doi: 10.3102/0034654308330968
- Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology*, 105(4), 970–987.
- Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology*, 106(2), 331–347. doi:10.1037/a0034752
- Tapscott, D. (2009). *Growing up digital: How the next generation is changing your world*. New York: McGraw-Hill.
- Tomlinson, C. A., & Tomlinson, C. A. (2017). *How to differentiate instruction in academically diverse classrooms* (3<sup>rd</sup> ed.). Alexandria, Virginia: ASCD.
- United States. Office of Educational Technology. (2010). *Transforming American education: Learning powered by technology: National education technology plan 2010*. (). Washington, D.C: U.S. Department of Education, Office of Educational Technology.
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46, 197–221. doi: 10.1080/00461520.2011.611369
- Verdú, E., Regueras, L. M., Gal, E., de Castro, J. P., Verdú, M. J., & Kohen-Vacs, D. (2017). Integration of an intelligent tutoring system in a course of computer network design. *Educational Technology Research and Development*, 65(3), 653–677. doi: 10.1007/s11423-016-9503-0

- Vesin, B., Mangaroska, K., & Giannakos, M. (2018). Learning in smart environments: User-centered design and analytics of an adaptive learning system. *Smart Learning Environments*, 5(1), 1-21. doi: 10.1186/s40561-018-0071-0
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* Cambridge, Mass.: Harvard University Press
- Wang, D., Han, H., Zhan, Z., Xu, J., Liu, Q., & Ren, G. (2015). A problem solving oriented intelligent tutoring system to improve students' acquisition of basic computer skills. *Computers & Education*, 81, 102-112. doi:10.1016/j.compedu.2014.10.003
- Wilson, C., & Scott, B. (2017). Adaptive systems in education: A review and conceptual unification. *The International Journal of Information and Learning Technology*, 34(1), 2-19. doi: 10.1108/IJILT-09-2016-0040
- Wongwatkit, C., Srisawasdi, N., Hwang, G., & Panjaburee, P. (2017). Influence of an integrated learning diagnosis and formative assessment-based personalized web learning approach on students learning performances and perceptions. *Interactive Learning Environments*, 25(7), 889-903. doi:10.1080/10494820.2016.1224255
- Wood, D., Jerome, S., Ross, B., & Ross, G. (1976). *The role of tutoring in problem-solving*. *Journal of Child Psychology and Psychiatry*, 17(2), 89-100. doi.org/10.1111/j.1469-7610.1976.tb00381.x
- Xin, Y. P., Tzur, R., Hord, C., Liu, J., Park, J. Y., & Si, L. (2017). An intelligent tutor-assisted mathematics intervention program for students with learning difficulties. *Learning Disability Quarterly*, 40(1), 4-16. doi: 10.1177/0731948716648740
- Xu, Z., Wijekumar, K. K., Ramirez, G., Hu, X., & Irey, R. (2019). The effectiveness of intelligent tutoring systems on K-12 students' reading comprehension: A meta-analysis:

- Meta-analysis of the effect of ITSs on reading comprehension. *British Journal of Educational Technology*, doi:10.1111/bjet.12758
- Yelland, N., & Masters, J. (2007). Rethinking scaffolding in the information age. *Computers & Education*, 48(3), 362-382. doi:10.1016/j.compedu.2005.01.010
- Zandvakil, E., Washington, E., Gordan, E., Wells, C. (2018). Master learning in the classroom: Concept maps, critical thinking, collective assessments (M3CA) using multiple-choice items (MCIs). *Journal of Education and Learning*, 7(6), 45-56. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1191203.pdf>
- Zhang, K., & Yao, Y. (2018). A three learning states Bayesian knowledge tracing model. *Knowledge-Based Systems*, 148, 189-201. doi:10.1016/j.knosys.2018.03.001
- Zilvinskis, J., Willis, J., & Borden, V. M. H. (2017). An overview of learning analytics. *New Directions for Higher Education*, 2017(179), 9-17. doi:10.1002/he.20239

## APPENDICES

### Appendix A

## DISSERTATION RESEARCH APPLICATION

**DIRECTIONS:** Complete this form by filling in the information requested. Attach the file to an email message and it to [REDACTED]. Please type "Research Application" in the subject line of your email.

---

#### SECTION 1: GENERAL INFORMATION

Kevin Rholetter

*(Applicant's first and last name)*

Meredith Park

*(Instructor's first and last name)*

krholetter [REDACTED]

*(Applicant's email address)*

*(Instructor's email address)*

Liberty University

*(College or University)*

---

#### SECTION 2: TIMEFRAME

What is the proposed start date? 12/1/2019

What is the proposed end date? 3/1/2020

---

#### SECTION 3: STATEMENT OF PURPOSE

State the purpose of the proposed research study. Limit your statement to one or two sentences that clearly identify the specific topic(s) and goal(s) of the study. (Example: This study will examine the effect of the ABC Reading program on the oral reading fluency of first-grade students from low-income homes.)

This study will examine the math achievement of students in two similar middle schools, one that used intelligent tutoring systems for math instruction during the 2017-2018 school year and one that did not.

---

#### SECTION 4: RESEARCH QUESTION(S) AND DEFINITION OF KEY TERMS

List the specific research question(s) to be investigated in this study. (Example: What is the effect of participation by students from high poverty home in the ABC Reading program on blending isolated phonemes to make words?)

Is there a difference in the math achievement of middle school students whose math instruction includes an intelligent tutoring system and middle school students whose math instruction does not include an intelligent tutoring system?

Is there a difference in the math achievement of 6th-grade students whose math instruction includes an intelligent tutoring system and 6th-grade students whose math instruction does not include an intelligent tutoring system?

Is there a difference in the math achievement of 7th-grade students whose math instruction includes an intelligent tutoring system and 7th-grade students whose math instruction does not include an intelligent tutoring system?

Is there a difference in the math achievement of 8th-grade students whose math instruction includes an intelligent tutoring system and 8th-grade students whose math instruction does not include an intelligent tutoring system?

**Provide definitions of terms that may be specific to your area of inquiry to ensure clarity and understanding.**

Intelligent tutoring systems are instructional programs that customize learning for students.

## **SECTION 5: DATA COLLECTION**

**Identify the data to be collected. If you will be using student performance data, you must specifically identify the data. (Not acceptable: test scores; Acceptable: SC READY Mathematics performance levels)**

2017-2018 Fall and Spring MAP SC 6+ Math scores

**Select the group(s) about or from which data will be collected. (Check all that apply)**

- ☒ Students ☐ Administrators  
☐ Teachers ☐ Parents / Families  
☐ Others (Specify):

**Select the group(s) about or from which data will be collected. (Check all that apply)**

- ☐ Early childhood (3 and 4-year-olds) ☐ High (Grades 9-12)  
☐ Elementary (Grades K-5) ☐ Adult Education  
☒ Middle (Grades 6-8)  
☐ Others (specify):

**How many participants are required?**

Minimum 300

Maximum 300

**Which [REDACTED] students, staff, or parents will be involved in this research? (Check all that apply)**

**If your research will be limited to magnet programs, list the programs below:**

**Explain the selection criteria for the participants.**

This study examines the math achievement of students who benefited from the use of intelligent tutoring systems for math during the 2017-2018 school year. Students at [REDACTED] School during the 2017-2018 school year utilized both Pearson's digits Math Program and Math IXL.

**Describe the data collection procedures. Include a timeline for each step as well as a description of any data collection activities and instruments.**

All data employed in this study is archival. In addition to grade level, the necessary data will consist of fall and spring MAP SC 6+ scores for all students who took this assessment in both the fall and spring.

**Describe the procedures and safeguards you will use to ensure the privacy and confidentiality of participants' data.**

All data will be stripped of all identifiers except grade level before being sent to the researcher. Data will then be kept on a password-protected flash drive.

**What are the potential risks to participants?**

N/A

**State the impact, if any, on instructional time.**

N/A

---

**SECTION 6: ATTACHMENTS**

**List all supporting documents, forms, surveys, etc. that you are submitting with this proposal.**

1. N/A

2.

3.

4.

5.

## Appendix B

# LIBERTY UNIVERSITY

## INSTITUTIONAL REVIEW BOARD

February 3, 2020

Kevin Lamar Rholetter

IRB Exemption 4136.020320: A Causal-Comparative Study on the Efficacy of Intelligent Tutoring Systems on Middle-Grade Math Achievement

Dear Kevin Lamar Rholetter,

The Liberty University Institutional Review Board has reviewed your application in accordance with the Office for Human Research Protections (OHRP) and Food and Drug Administration (FDA) regulations and finds your study to be exempt from further IRB review. This means you may begin your research with the data safeguarding methods mentioned in your approved application, and no further IRB oversight is required.

Your study falls under exemption category 46.101(b)(1), which identifies specific situations in which human participants research is exempt from the policy set forth in 45 CFR 46:101(b):

(1) Research, conducted in established or commonly accepted educational settings, that specifically involves normal educational practices that are not likely to adversely impact students' opportunity to learn required educational content or the assessment of educators who provide instruction. This includes most research on regular and special education instructional strategies, and research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.

Please note that this exemption only applies to your current research application, and any changes to your protocol must be reported to the Liberty IRB for verification of continued exemption status. You may report these changes by submitting a change in protocol form or a new application to the IRB and referencing the above IRB Exemption number.

If you have any questions about this exemption or need assistance in determining whether possible changes to your protocol would change your exemption status, please email us at [irb@liberty.edu](mailto:irb@liberty.edu).

Sincerely,

**LIBERTY**  
UNIVERSITY  
*Liberty University | Training Champions for Christ since 1971*



### Appendix C

Kevin,

██████████ research committee has approved your application to conduct research in our district. You must complete all research activities by June 30, 2020. You will need to request an extension from the research committee if you need to continue research activities beyond that date. Please remember the committee reserves the right to terminate the study at any time if circumstances change or the members feel it is in the best interest of our students, their families, or staff. Finally, you must submit a copy of all final reports, dissertations, or publications based on this research to me upon completion of your study.

Respectfully,

██████████

## Appendix D

Good afternoon Mr. Rholetter,

[REDACTED] has approved you to use our data as part of your research project. Please let me know if you have any questions.

[REDACTED]