Effectiveness of the Cognitive Tutor on Mathematics Achievement of High School Students

Theresa Basadre
B.S.E., Philippine Normal University, 1993
M.A., Friends University, 2016

Submitted to the Graduate Department and Faculty of the School of Education of Baker University in partial fulfillment of the requirements for the degree of Doctor of Education in Educational Leadership

____
James Robins, Ed.D.
Major Advisor

____
Susan Rogers, Ph.D.

____
Harold B. Frye, Ed.D.

Date Defended: November 19, 2020

Abstract

Investigated in this study was the impact the implementation of the Cognitive Tutor (CT) curriculum on tenth-grade student achievement in mathematics, as measured by the Kansas Assessment Program (KAP) test. Also, the impact of CT on the mathematics achievement of students with disabilities, English language learner students, and free-and-reduced students as measured by the KAP assessment was examined. A quasi-experimental, quantitative research design was used to examine the difference of the mathematics test scores of tenth-grade students enrolled in District X before and after the implementation of the CT as measured by the Mathematics KAP test. Performance level scores of tenth-grade students on the Mathematics KAP test in the spring of 2017 before the implementation of CT were compared to the tenth-grade students' Mathematics KAP performance level scores in the spring of 2018, which occurred after the implementation of CT. Demographic variables included in the study to select student subgroups for further comparison were disability status, English language learner status, and free and reduced lunch status. The results of the current study indicated no significant difference between the Mathematics KAP scores of all tenth-grade students, English language learner students, and students with free- and reduced-lunch who did not receive CT instruction and the Mathematics KAP scores of all tenth-grade students, English language learner students, and students in free- and reduced-lunch who received CT instruction. The results also revealed that the Mathematics KAP scores of students with disabilities who did not receive CT instruction were significantly higher than the Mathematics KAP scores of students with disabilities who received CT instruction. The effect size of the traditional curriculum was small.

Dedication

This dissertation is dedicated to my children, Lucas and Isabel Mamaoag, who have been my inspiration to be the best I could be.

Acknowledgments

I wish to thank the following people for all their support in making this study possible:

Dr. James Robins, I am incredibly grateful for all your support throughout this academic journey. Your guidance, feedback, and encouragement made this study possible.

Dr. Li Chen-Bouck, thank you for all your constructive feedback on the statistical aspect of my research study.

To Dr. Susan Rogers, thank you for your extensive and valuable comments. Your attention to detail was advantageous.

Dr. Harold Frye, thank you for serving on my committee. Your dedication to education is remarkable.

Dr. Lisa Lutz, thank you for helping me with the student achievement data for my study.

To all my professors in my doctoral courses, thank you for inspiring, educating, and challenging me to be a good administrator.

Table of Contents

List of Tables

Chapter 1

Introduction

Improving student mathematics achievement is a continuing concern among policymakers, educational leaders, district administrators, teachers, parents, and the public (Pane, Griffin, McCaffrey, & Karam, 2014). The low academic achievement performance of students in the United States (U.S.) has been compared to students around the world using the Trends in International Mathematics and Science Study (TIMSS), and the possible effect that these low levels of achievement might have on college and career readiness has prompted educational reforms regarding the mathematics curricula. The 2015 TIMSS results showed the average scores of U.S. eighth-graders in mathematics were higher than the average scores of students in 24 international education systems but lower than the average scores of students in 8 other education systems (Provasnik, Malley, Stephens, Landeros, Perkins, & Tang, 2016; TIMSS, 2015). In the U.S., 15-year old students' average score on the 2015 Program for International Student Assessment (PISA) was lower than the average score in more than half of the other international education systems (PISA, 2015). The U.S. 15-year-old students' average scores were also lower than the Organization for Economic Cooperation and Development (OECD) average (PISA, 2015).

The low mathematics assessment performance of students has prompted several education reforms across the U.S. designed to improve student achievement. In 2007, the National Governors Association Center (NGA) and the Council of Chief State School Officers (CCSSO) initiated the development of the Common Core State Standards (CCSS) to prepare students for college, career, and life. The CCSS includes research-based standards for mathematics and English/language arts, which provide rigorous content and application of knowledge consistent across all states to ensure that students acquire deeper conceptual understanding and higher-order thinking skills regardless of where they live. The CCSS are comparable to the top-performing countries in the world, so U.S. students are "well prepared to collaborate and compete with their peers in the United States and abroad" (CCSS, 2018, p. 1).

"As of 2018, 41 states, the District of Columbia, four territories, and the Department of Defense Education Activity (DoDEA) have adopted the Common Core and are implementing those standards according to their own timelines" (Common Core Standards Timeline, 2018, para. 9). Kansas is one of the states that adopted the CCSS. In August 2017, the Kansas State Board of Education adopted the 2017 Mathematics Standards. The school districts within Kansas started the initial implementation of the 2017 Mathematics Standards during the 2018-2019 school year. The full implementation of the 2017 Mathematics standards was in the 2019-2020 school year. The adoption of the 2017 Mathematics Standards called for a restructuring of the curriculum and the implementation of new textbooks and instructional resources to prepare students for post-secondary education and the workforce.

Background

The National Assessment of Educational Progress (NAEP) results showed a low performance of both eighth and twelfth-grade students across the U.S. at or above the proficient level score on the mathematics portion of the assessment. NAEP grades 8 and 12 mathematics results are reported as average scores on a 0 to 500 and 0 to 300 scale and as percentages of students performing at or above three achievement levels. The three different levels include basic, proficient, and advanced. Scores of 299 for eighth grade and 276 for twelfth-grade students are considered to be at or above the proficient level on NAEP. The national average score and the percentage of students at or above the NAEP proficient score of eighth grade and twelfth-grade students on the mathematics NAEP during each assessment year are shown in Table 1 and Table 2.

Compared to the first assessment year in 1990, the average NAEP math score of eighth-grade students across the U.S. increased until 2013. The average score then decreases during 2015, increases in 2017, and decreases again during the 2019 assessment. Although the average NAEP score of eighth-grade students in the year 2019 increased by 19 points compared to the first assessment score in 1990, the average score is still below the cut-off score for students to perform at or above the NAEP Proficient level on the mathematics assessment. Also, the percentage of eighth-grade students who performed at or above the NAEP proficient level in mathematics is 34%, which is about one-third of the students who were assessed.

Table 1

Description of Eighth-Grade Students Average Score and Percentage at or Above NAEP Proficient Level

Year	Average Score	Percentage at or Above NAEP Proficient
1990*	263	15
1992*	268	21
1996*	272	24
1996	270	23
2000	273	26
2003	278	29
2005	279	30
2007	281	32
2009	283	34
2011	284	35
2013	285	35
2015	282	33
2017	283	34
2019	282	34

Note. Adapted from "*The Nation's Report Card: NAEP Mathematics and Reading Assessment Highlighted Results at Grades Four and Eighth for the Nations, States, and Districts,"* National Assessment of Educational Progress (2019). Retrieved from https://www.nationsreportcard.gov/highlights/mathematics/2019/

*Indicates that accommodation was not permitted

The average NAEP math score of twelfth-grade students across the U.S. during the testing year 2005 increased from 150 to 153 in the year 2009. The average score did not change from 2009 to 2013 but decreased from 153 to 152 between 2013 and 2015. The average NAEP math score of twelfth-grade students across the U.S. during the testing year 2015 was not statistically different from the average score in 2005, 2009, and 2013 (NAEP, 2015).

Table 2

Description of Twelfth-Grade Students Average Score and Percentage at or Above NAEP Proficient

Year	Average Score	Percentage at or above NAEP Proficient
2005	150	23
2009	153	26
2013	153	25
2015	152	25

Note. Adapted from "*The Nation's Report Card: NAEP Report Card: 2015 NAEP Mathematics & Reading at Grade 12,"* National Assessment of Educational Progress (2015). Retrieved from https://www.nationsreportcard.gov/reading_math_g12_2015/#mathematics/scores

*Indicates that accommodation was not permitted

On the 2017 Kansas Mathematics Assessment Test, 23.07% of tenth-grade students demonstrated proficiency to understand and use the mathematics skills and knowledge needed for college and career readiness, while in 2018 and 2019 that percentage slightly increased to 23.72% and 24.5% (Kansas Department of Education [KSDE], 2019). With the low performance of students on the mathematics state assessment, school districts in Kansas are continually searching for textbooks and instructional resources that will enhance their curriculum to help students acquire the necessary knowledge and skills to improve their mathematics achievement. Considering the low mathematics achievement of eighth and tenth-grade students on the international, national, and state levels, mathematics curriculum might have played a critical role in the achievement of these students.

Carnegie Learning Algebra 1 is one of the curricula adopted by about 2,000 schools across the U.S to improve student achievement in mathematics (Carnegie Mellon University, 2015). The Carnegie Learning Algebra 1 and Geometry blended curriculum is also referred to as the Cognitive Tutor (CT) curriculum. For this study, CT will be used to indicate Carnegie Learning's Algebra 1 and Geometry blended curriculum. CT utilizes a blended approach that integrates group instruction employing a consumable textbook with independent learning using an adaptive software called MATHia (Carnegie Learning, 2018, para.1). MATHia is a computerized tutoring program that uses a cognitive model to provide immediate feedback and useful hints to students while they are working on any given mathematics problem.

This study intended to investigate the impact of CT curriculum on student achievement of tenth-grade students and each student subgroups, including students with disabilities, English language learner students, and students with free and reduced lunch status. The current study was conducted in District X, a large, urban district in Kansas. During the 2020-2021 school year, District X had an enrollment of 47,236 (District X Snapshot, 2020). Of this population, 76.5% were classified as economically disadvantaged, and 15.5% received special education services. The student subgroups were reported at Hispanic 36.1%, Caucasian 30.8%, African-American 20%, Multi-racial 7.6%, Asian 4.5%, Native America, .8%, and Pacific Islander .2% (USD 259 District Snapshot, 2020).

Statement of the Problem

Students' overall KAP assessment performance level scores in District X showed a three-year low performance in mathematics from the 2016-2017 school year to the 2018-2019 school year, requiring the district to take action to increase student performances (KSDE, 2019). In the 2017 Kansas Math Assessment, only 12.01 % of tenth-grade students in district X performed at or exceeding grade level/on track for college and career readiness, while in 2019, that percentage slightly increase to 13.49% (KSDE, 2019).

To improve student's low mathematics performance District X, a large urban public school district in South Central Kansas, adopted CT Curriculum for middle school and high school students beginning in the school year 2017-2018 up to the present. Multiple studies regarding CT's impact on student achievement have been conducted (Barton, 2016; Bibi, 2010; Cabalo, Ma, & Jaciw, 2007; Cabalo, Jaciw & Vu, 2007; Morgan, & Ritter, 2002). While these studies examined the effectiveness of mathematics curriculum on student achievement, the results were mixed. For example, Morgan and Ritter (2002) found that students taught using the CT curriculum performed better than students who were taught with a traditional curriculum as measured by the Educational Testing Service (ETS) Algebra end-of-course assessment and students' course grades. Also, Sarkis (2014) found that ninth-grade students who completed Algebra 1 CT were significantly more likely to achieve higher overall scores than those students who completed conventional Algebra 1 classes. However, Cabalo, et al. (2007) found that CT has no significant effect on student achievement as measured by the overall score in the Northwest Evaluation Association (NWEA) end-of-course Algebra 1 test.

Sarkis (2004) suggested that exceptional student education and limited English proficiency students who were instructed with CT performed significantly better on the mathematics Florida Comprehensive Assessment Test than their counterparts who were instructed with a traditional curriculum. The results of the study revealed that students' final grades across two groups regarding exceptional student education and limited English proficiency students were not different from the whole group results (Sarkis, 2004). However, limited empirical research has been conducted on the impact of CT on other student subgroups, for example, students with disabilities, English language learner students, and free- and-reduced students.

Also, there is very little research on the relationship between the number of hours spent by students using CT and their mathematics achievement. For example, Fancsali et al. (2018) provided evidence to suggest that the use of MATHia software is associated with students' mathematics test scores. The study used data from over 23,000 ninth-grade students. The results indicated that the number of problems, skills, and topics mastered and the amount of time taken, the rate at which content was mastered, the number of errors and hints made in the software was able to predict outcomes on the students' state mathematics test scores (Carnegie Learning, 2018). Examining CT's impact on different student subgroups and the relationship between the number of hours spent by students using CT and their mathematics achievement may provide valuable information to educational leaders.

Purpose of the Study

The primary purpose of this study was to investigate the effectiveness of the CT curriculum on mathematics achievement of tenth-grade students in District X as measured by the KAP assessment. Another purpose of this study was to determine the impact of CT on the mathematics achievement of students with disabilities and English language learner students as measured by the KAP assessment. Lastly, the study examined the impact of CT on students with free- and reduced-lunch status as measured by the KAP assessment.

Significance of the Study

The results of this study could contribute to the body of research regarding the effectiveness of CT to high school students' performance. The effectiveness of CT related to student achievement has been examined in recent years; however, this study could provide more valuable information to evaluate the impact of the CT specifically for students with disabilities, English language learners, and economically disadvantaged students. The results of this study could also provide evidence of the impact the curriculum would most likely produce in a typical school setting with similar demographics. Examining the effectiveness of the CT curriculum may provide evidence to aid curriculum specialists, directors, and administrators in selecting resource materials for the students in their respective school districts.

Delimitations

According to Lunenburg and Irby (2008), “Delimitations are self-imposed boundaries set by the researcher on the purpose and scope of the study” (p. 134). The following delimitations were placed upon this study.

1. Data were collected from one large urban school district in the Midwest.
2. The data used for this study were gathered during the 2016-2017 school year before the students received CT instruction and during the 2017-2018 school

year after the students received CT instruction. The study did not use the data during the 2018-2019 school year because the KAP test was based on different mathematics standards.

3. The mathematics achievement scores utilized for this study were from the Mathematics KAP results.
4. Participants of this study were limited to tenth-grade students who took the Mathematics KAP.

Assumptions

Lunenburg and Irby (2008) defined assumptions as "postulates, premises, and propositions that are accepted as operational for purposes of research" (p. 135). This study included the following assumptions:

1. It was assumed that the participants were sufficiently instructed using the CT curriculum.
2. It was assumed that the participants who logged in to use MATHia were working on the math problems with fidelity.
3. It was assumed that the participants performed their best during the assessment.
4. It was assumed that all teachers were adequately trained on how to use CT.

Research Questions

The research questions guiding the current study were as follows:

RQ1. To what extent is there a difference in the Mathematics KAP scores between tenth-grade students who did not receive CT instruction and tenth-grade students who received CT instruction?

RQ2. To what extent is there a difference in the Mathematics KAP scores between tenth-grade students with disabilities who did not receive CT instruction and tenth-grade students with disabilities who received CT instruction?

RQ3. To what extent is there a difference in the Mathematics KAP scores between tenth-grade English language learners who did not receive CT instruction and tenth-grade English language learners who received CT instruction?

RQ4. To what extent is there a difference in the Mathematics KAP scores between tenth-grade students with free- and reduced lunch who did not receive CT instruction and tenth-grade students with free- and reduced-lunch status who received CT instruction?

Definition of Terms

The following terms are defined to ensure the reader understands their use throughout the study.

Cognitive Tutor (CT). According to Carnegie Learning (2014), CT is a secondary mathematics curriculum developed by Carnegie Learning that focuses on how students think about and learn mathematics. Teachers facilitate student learning as students acquire and apply new information and discuss their work. The curriculum can be implemented using a textbook, adaptive software, or a combination of textbook and software activities (Carnegie Learning, 2014).

English Language Learner (ELL). According to the National Council of Teachers of English (2018), an ELL student is an active learner of the English language who may benefit from various language support programs. This term is used mainly in the U.S. to describe K–12 students.

Free- and Reduced-Lunch Students (FRL). A student from a household with an income at or below 130% of the poverty income threshold is eligible for free lunch. A student from a household with an income between 130% and up to 185% of the poverty threshold is eligible for reduced-price lunch (US Department of Agriculture, 2020).

Kansas Assessment Program (KAP). A program of the KSDE that fulfills a mandate from the Kansas Legislature to provide general education assessments, alternate assessments, career and technical education assessments, and an English language proficiency assessment. The English language arts, mathematics, and science assessments are part of the federal elementary and secondary education legislation. KAP tests and tools are designed to support educators and policymakers in evaluating student learning and meeting the requirements for federal and state accountability (KAP, 2020).

MATHia. Carnegie Learning (2014) indicated that MATHia is an Intelligent Tutoring System that is a part of Carnegie Learning's high school blended mathematics curriculum.

National Assessment of Education Progress (NAEP). A congressionally mandated project administered by the National Center for Education Statistics (NCES) within the U.S. Department of Education and the Institute of Education Sciences (IES). It is the only assessment that measures what U.S. students know and can do in various subjects across the nation, states, and in some urban districts (NAEP, 2018).

Performance Levels. Scale scores on KAP assessments are used to place students' assessment scores into one of four possible categories: Level 1, Level 2, Level 3, and Level 4. Level 1, a student is below grade level; Level 2, a student is at grade level, but not on track for college or career readiness; Level 3, a student is at grade level and on track for college or career readiness; and Level 4, a student exceeds grade-level expectations and is on track for college or the workplace (KAP, 2015)

Program for International Student Assessment (PISA). The Organization for Economic Co-operation and Development (2015) stated that PISA measures 15-year-olds' ability to use their reading, mathematics, and science knowledge and skills to meet real-life challenges.

Students with disabilities (SWD). A child is considered to be SWD with mental retardation, hearing impairments (including deafness), speech or language impairments, visual impairments (including blindness), serious emotional disturbance, orthopedic impairments, autism, traumatic brain injury, other health impairments, or specific learning disabilities; and who, by reason thereof, needs special education and related services (National Center for Education Research, 2020)

Trends in International Mathematics and Science Study (TIMSS). The International Association for Evaluation of Educational Achievement administers an international assessment conducted in the United States by the NCES that monitors trends in student achievement in mathematics, science, and reading. TIMSS provides reliable and timely data on the U. S. students' mathematics and science achievement compared to that of students in other countries (U.S. Department of Education, IES, NCES, 2020).

Organization of the Study

This research study is divided into five chapters. Chapter 1 includes the background of the study, statement of the problem, the purpose of the study, the significance of the study, delimitations, assumptions, research questions, and definition of terms. Chapter 2 includes a review of related literature, which includes the impact of CT on student achievement, the impact of the curriculum on different student groups, technology as a pedagogical tool, and intelligent tutoring system. Chapter 3 contains the methodology, which includes the research design, selection of participants, measurement, data collection procedures, data analysis, and hypothesis testing, and limitations of the study. Presented in chapter 4 are the results of the study, which includes descriptive statistics and the results of the hypothesis testing. Provided in chapter 5 are the summary of the entire study, finding related to the literature, and conclusions.

Chapter 2

Review of the Literature

The purpose of this study was to investigate the effectiveness of the CT curriculum on the mathematics achievement of tenth-grade students. This study also examined the impact of CT on the mathematics achievement of tenth-grade students with disabilities, English language learner students, and free- and reduced-lunch students, as measured by the KAP assessment. This chapter includes a review of the historical overview of CT development, the impact of CT on student achievement, intelligent tutoring, and the impact of technology in the classroom.

Historical Overview of Cognitive Tutor

In 1982, Anderson completed a theory of learning and cognition called Adaptive Control of Thought-Rational (ACT-R), which has been the theoretical foundation for the current cognitive tutors (Anderson, 1983; Koedinger & Corbett, 2006). ACT-R elaborates the ACT theory about how human cognition works (Anderson, 1976, 1983). The ACT-R model embodies three types of memory structures: declarative, procedural, and working memory. According to the ACT-R theory, all knowledge begins as declarative information, including factual information that a person knows and can report. The procedural knowledge is learned by making inferences from the already existing factual knowledge, and the working memory is responsible for temporarily holding information available for processing (Anderson & Schunn, 2000). ACT-R model assumes that with continuous practice, both declarative and procedural knowledge are intensified so that student performance grows faster and more reliable (Corbett & Anderson, 1995). The representation of knowledge fundamental in this kind of model is called cognitive modeling, and the approach of using a cognitive model in a tutoring system is called a Cognitive Tutor (Anderson, Boyle, Corbett, & Lewis, 1990).

Anderson, Corbett, Koedinger, Pelletier (1985) from Carnegie Mellon University developed curriculum and software for teaching and learning high school mathematics in 1983. These software programs were geometry proof tutor and list processing (LISP) tutor, a programming language that was commonly used for artificial intelligence (Anderson et al., 1995). These tutors embodied several key ideas about how computer-based instruction should be realized (Anderson, Corbett et al, 1985). In 1984, Geometry and LISP tutors were completed to validate the ACT theory of cognition in a computer tutor. Students who worked with the LISP Tutor completed problems in one-third of the time required by students working in a programming environment without the tutor and scored 25% higher on succeeding tests (Corbett & Anderson, 1991).

In 1985, more than 10,000 educational software tutors existed, but very few had the components of the intelligent computer-aided instruction (ICAI). Commonly known as an Intelligent Tutoring System, ICAI simulates an understanding of the domain they teach and responds specifically to each student's problem-solving strategies. Anderson, Boyle, and Reiser (1985) built a computer tutoring system that combined cognitive psychology, computer technology, and artificial intelligence.

The Geometry Proof Tutor was piloted in Peabody High School in Pittsburgh from the fall of 1985 until 1987. The results of the study indicated that the tutor was about one-half as effective as a human tutor, but two to three times more effective than traditional computer-aided instruction. Koedinger and Anderson (1993) tested the effectiveness of a cognitive-based geometry proof tutor called A New Geometry Learning Environment (ANGLE). The results of the study indicated no significant difference between the post-test scores of students in the ANGLE classes and students with non-ANGLE classes. An evaluation of algebra tutors was performed during the 1987-1988 school year in Peabody High School. The results indicated no significant difference between the experimental and control classes. Anderson, Conrad, and Corbett (1989) developed programming languages, including LISP, Pascal, and Prolog in a self-paced programming course at Carnegie Mellon University. These languages were used to control the behavior of a computer (Anderson et al., 1989).

In the 1990s, Pittsburgh Advanced Cognitive Tutor (PACT) Center created Cognitive Tutor Algebra. This intelligent tutoring system for teaching algebra was believed to be the most successful ITS application for use in K-12 classrooms (Carnegie Learning, 2007; Koedinger, Anderson, Hadley, & Mark (1997). In 1991, the Pittsburgh Urban Mathematics Project (PUMP) was formed through collaboration among the teachers in the Pittsburgh Public Schools, ACT Research Group of Carnegie Mellon University's Psychology Department, and researchers from Carnegie Mellon's Human-Computer Interaction Institute. PUMP created an algebra 1 curriculum and an intelligent tutoring system called Practical Algebra Tutor (PAT), which was considered as the second ITS. PAT was aligned with the National Council of Teachers Mathematics' (NCTM) 1990 recommendation emphasizing problem-solving, reasoning, and multiple representations on algebra and geometry lessons (Koedinger et al., 1997). PAT was piloted in Pittsburgh Public High Schools at Langley High School in 1992, then at Brashear and Carrick High Schools during the 1993-1994 school year. The results of the study indicated that students in PAT classes outperformed students in comparison classes, with a statistically moderate achievement gain of 15% to 25% on basic skills and 50% to 100% improvement on a test that emphasized real-world problem solving and multiple mathematical representations (Corbett, Koedinger, & Hadley, 2002). In 1995, Algebra 1 Cognitive Tutor, an adaptation of the PAT, was piloted at two colleges. The results indicated that Cognitive Tutor students performed more than 50% better than students in regular classes on a performance-based assessment (Koedinger & Sueker, 1996). In 1998, Carnegie Melon University founded Carnegie Learning to accumulate market-driven research, development, and dissemination of curricula, textbooks, and mathematics software for middle and high school students. By 1998-1999, the Cognitive Tutor was implemented in 75 schools and increased to 1,400 schools by 2003.

In 1999, Carnegie Melon University received an exemplary award from the U.S. Department of Education and was awarded a grant from the National Science Foundation to study human tutoring and create a more effective computer-based tutor. Researchers from Carnegie Melon Curriculum began developing Cognitive Tutor courses for middle school mathematics, which was then used in 150 private and public schools in 14 different states (Corbett et al.,2002).

During the 2000-2001 school year, the Cognitive Tutor math six course was implemented at two Pittsburgh schools, and the results indicated that students in Cognitive Tutor classes outperformed students in traditional classes in two assessments in both schools (Corbet et al., 2002). In 2004, Cognitive Tutor Algebra 1 was used in approximately 2000 schools in the U.S by half a million students for about 20 million hours (Koedinger & Corbett, 2006). In 2005, Carnegie Learning piloted Bridge to Algebra, an intervention product for middle school students, then released it to the public the following year. As of August 2008, CT curricula have been used by more than 500,000 students in approximately 2,600 urban, suburban, and rural school districts across the United States (U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse, 2009).

In 2011, Carnegie Learning launched MATHia, an intelligent math software based on cognitive science that imitates a human tutor and provide a personalized learning experience to every student. MATHia adapts to students' problem-solving strategies and provides real-time feedback to help students understand why they get a math concept wrong and develop skills to solve problems (Carnegie Learning, 2011). In June 2020, Carnegie Learning released MATHiaFlex, a new software experience powered by MATHia that provided teachers real-time data on what the students were working on, areas students were struggling, and sent alerts to teachers identifying which students need additional support. MATHiaFlex provides the flexibility to put every student on a personalized, continuous learning path that teaches and adapts to their unique needs to focus on specific skills that they need to master (Carnegie Learning, 2020).

Cognitive Tutor's Impact on Student Achievement

CT is a mathematics curriculum developed by Carnegie Learning based on cognitive science research and the use of intelligent tutoring systems to guide students in mathematics problem-solving. CT can be implemented using print material corresponding to a consumable textbook with skills practice workbook, an adaptive software, or a combination of a textbook and software (U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse, 2016). In a blended implementation, students spend 60% of their time in the classroom doing group activities, problem solving, and class presentations of their work and 40% in the computer lab working with CT software. The name CT is also referred to as a type of intelligent tutoring produced by Carnegie Learning for middle and high school mathematics.

The CT courses' software component was based on Anderson's ACT-R model and had been proven to be successful. Students worked through practice exercises in one-third of the time while performing as well or better on tests (Anderson, Corbett et al., 1985; Anderson & Reiser,1985). However, some students continued to struggle despite general success, which led the researchers from Carnegie Melon University to incorporate knowledge tracing into the computer tutors to implement mastery learning. The frameworks of Bayesian Knowledge Tracing (BKT) and CT were known to be a particular type of Intelligent Tutoring System (ITS) that is widely used in school settings (Corbett & Anderson, 1995). Using the probabilistic framework of BKT, CT continually assesses student knowledge of fine-grained skills or knowledge components that are a part of CT's underlying cognitive model (Corbett & Anderson 1995). Using BKT's assessment of student skill mastery, CT provides each student with activities that emphasize the skills they need to learn. In using the CT software, students are presented with instruction in mathematics problem-solving while working at their own pace. As part of the cognitive model component, students' knowledge and skills are assessed. At the same time, they respond to mathematical problem questions, scaffolding is provided for immediate feedback on errors, then the instruction is adjusted to adapt to each student's solution strategies.

This section includes studies on CT that reported a statistically positive significant difference between the CT group and the comparison groups on outcomes in the algebra domain. This section also included studies that indicated no statistically significant difference between the CT group and the comparison groups on outcomes in the algebra domain. Lastly, this section presented studies that revealed a statistically significant negative difference between the CT and comparison groups on outcomes in the geometry domain.

Cognitive Tutor studies with positive significant effect. Shneyderman (2001) conducted a quasi-experimental study in six senior high schools in Miami–Dade County Public Schools to examine the impact of Cognitive Tutor Algebra 1, as measured by the Florida Comprehensive Assessment Test–Norm-Referenced Test 2001 (FCAT-NRT) and the ETS End-of-Course Algebra I test. Students in the intervention group were taught using CT Algebra I for the entire school year, while students in the comparison group received Algebra I instruction using a curriculum other than Cognitive Tutor Algebra 1. The results of the study indicated that students in the CT Algebra 1 group demonstrated significantly higher results on the ETS Algebra I Test than did students in the comparison group. However, the overall mathematics and algebra performances measured by the 2001 FCAT-NRT were not different for students in the two groups (Shneyderman, 2001).

Morgan and Ritter (2002) conducted a study to determine the effectiveness of the CT Algebra I program for the Algebra 1 instruction of ninth-grade students in five junior high schools in Moore Independent School District, in Moore, Oklahoma. The students were randomly assigned to either the CT Algebra I course or a traditional Algebra I course using McDougal-Littell's Heath Algebra I (Morgan & Ritter, 2002). The researchers assessed student achievement and attitude using the Educational Testing Service (ETS) Algebra I end-of-course exam, course grades, and a survey of attitude towards mathematics. Morgan and Ritter (2002) restricted their analysis to the teachers who taught both curricula to control for teacher effect. The results of the study indicated that students taught using the CT curriculum performed better than students who were taught with a traditional curriculum on the Educational Testing Service (ETS) Algebra end-of-course assessment and their course grades. According to the What Works Clearinghouse (WWC) criteria, the effect size was considered a large effect. WWC considered an effect size of at least 0.25 to be a large positive effect. WWC was established by the Institute of Education Sciences (IES) within the U.S. Department of Education, which reviews existing research on different education programs, products, practices, and policies to provide educators with the information they need to make evidence-based decisions (U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse, 2002). The results also indicated that students using the CT were more confident and more likely to think that mathematics is useful outside the classroom than students in the traditional math curriculum (Morgan & Ritter, 2002).

Likewise, Wolfson, Koedinger, Ritter, and McGuire (2008) conducted a quasi-experimental design study across three high schools in the Pittsburgh Public School District involving 26 Algebra 1 classrooms. Students in the intervention classrooms utilized an early version of Cognitive Tutor and used textbook materials four times per week, and the cognitive tutor program once per week, as part of the standard curriculum. Students in the comparison group were taught using their schools' traditional Algebra I curricula. The results indicated a positive and statistically significant difference between the Cognitive Tutor Algebra I and comparison groups on the Multiple Representation Test and the Iowa Algebra Aptitude Test. Wolfson et al. (2008) indicated that Cognitive Tutor students' performance significantly increased by 26 percentile points on the Multiple Representation Test and by 12 percentile points on the Iowa Aptitude Test.

Bibi (2010) investigated the impact of the CT Geometry curriculum on student academic achievement, the different ways of implementing the curriculum, and the teachers' perceptions of the curriculum. The participants of the study were 549 students and 12 Geometry teachers in eight schools in Iowa. Bibi (2010) utilized both quantitative and qualitative methods. The researcher administered a pre- and post-test designed by the Mississippi Bend Area Education Agency to measure the impact of CT Geometry on student achievement. Teachers were interviewed to determine the level of implementation and to examine their views about the CT Geometry software and the textbook. The results of the study indicated that the group of students who used the CT Geometry textbook and its companion software had higher gains on the Mississippi Bend

Education Agency test compared to students who used CT Geometry with a different textbook. The results of the interview revealed that teachers favor the CT software compared to the CT textbook. Teachers stated that the CT software paces instruction according to each student's level. They also remarked that CT software assisted in implementing active learning, whereas the textbook lacks coherence based on the poor sequence of the topics. They stated that the textbook made simple concepts more complicated and lengthier (Bibi, 2010).

The U.S. Department of Education awarded a $6 million grant to RAND Corporation, an independent institution that helps improve policy and decision-making through research and analysis, to conduct a study on the effectiveness of the Carnegie Learning Algebra 1 blended curriculum from 2007 to 2009 (Carnegie Learning, 2017). The two-year study included over 18,000 students in 74 middle schools and 73 high schools from the school districts in Alabama; Connecticut; New Jersey; Texas; suburban districts near Detroit, Michigan; and rural districts in Kentucky and Louisiana. Schools that were randomly assigned to the treatment group implemented the Cognitive Tutor Algebra I curriculum, and those assigned to the control group used their existing algebra I curriculum. Schools were randomly assigned to the experimental or control group. The results of the study indicated no significant effect in the first year of implementation of Carnegie Learning blended approach on an algebra proficiency exam. However, a statistically significant improvement of approximately 8% was demonstrated for high school students in the second year of implementation, but not for middle school students (Pane et al., 2014).

Cognitive Tutor studies with no significant effect. However, several other experiments that tested the efficacy of Cognitive Tutor found no significant effect as measured by the overall score on the Northwest Evaluation Association (NWEA) end-of-course Algebra 1 test. Cabalo, Jaciw, and Vu (2007) tested the impact of Cognitive Tutor Algebra 1 curriculum on mathematics achievement of middle school students in Hawaii's Maui School District. Cabalo, Ma, and Jaciw (2007) also studied the effectiveness of Cognitive Tutor Bridge to Algebra curriculum. The researchers in both studies utilized a pre-and post-test experimental design. Teachers implemented the Cognitive Tutor to randomly selected students in intervention classes, while others used the traditional curriculum in comparison classrooms. The result of each of these studies did not show a significant, positive impact of the Cognitive Tutor on student achievement. However, Cognitive Tutor did increase student engagement in math (Cabalo, Ma, Jaciw, & 2007).

In addition, Campuzano, Dynarski, Agodini, and Rall (2009) investigated the impact of Cognitive Tutor Algebra 1 for eighth and ninth-grade students in nine schools and four districts in urban and urban fringe areas. The researchers utilized an experimental design and administered pre and post-tests to 142 intervention students and 128 comparison students. Teachers were randomly assigned to Cognitive Tutor for the treatment group and a traditional algebra 1 curriculum for the control group. However, students were not randomly assigned to teachers. The results of the study indicated that using Cognitive Tutor Algebra 1 has no significant effect on the mathematics scores of students as measured by the Educational Testing Service End-of-Course Algebra test (Campuzano et al., 2009).

Miami-Dade County Public Schools implemented the iPrep math program during the school year 2013-2014 in 49 middle schools, with approximately 240 students in each school. The iPrep program employed the Carnegie Learning Middle School Math Solutions, a curriculum that includes a consumable textbook and an adapted software application called MATHia. Carnegie Learning Math Solutions is also known as the Cognitive Tutor curriculum. Fontana, Beckerman, Levitt, and Levitt (2014) reported that proficiency rates for iPrep Math students were higher compared to the non-iPrep math students on the Florida Comprehensive Assessment Test (FCAT) math measures in both 2013 and 2014; however, proficiency rates increased more from 2013 to 2014 for the non-iPrep Math students. On the Algebra End-of-Course exam, there was no significant difference in the test performance of seventh-grade iPrep Math students and non-iPrep Math students. However, the scores for the eighth-grade iPrep Math students were significantly lower than those for the non-iPrep Math students. On the course academic content grades, iPrep Math students received lower grades and higher effort grades than non-iPrep Math students. The results also indicated that students' progress with the Carnegie MATHia software was positively connected to better performance on both the FCAT Math and Algebra EOC tests (Fontana et al., 2014).

Pane et al. (2014) examined the effectiveness of Cognitive Algebra 1 Tutor in mathematics student achievement. The study also investigated the effects of Cognitive Tutor Algebra 1 on the attitudes and confidence of students. The study was conducted in 51 school districts in seven states, including 73 high schools and 74 middle schools, where each school participated for two years. The researchers used a pair-matched, randomized cluster design to assign schools to the treatment and control groups. Students in the treatment group used the Cognitive Tutor Algebra 1 curriculum, while the control group used the traditional math curriculum. The researchers administered an Algebra Readiness Exam as a pretest three weeks after the start of the study, and the Algebra Proficiency Exam at the end of the course as a post-test. The results indicated no significant effect in the first year of implementation. However, the results revealed a significant positive effect for high schools, but not for middle schools in the second year of implementation, compared to similar schools that employed a variety of existing textbook-based algebra curricula. The median student performance increased by approximately eight percentile points. The results indicated no effects of Cognitive Tutor Algebra 1 on student attitudes and confidence in any of the groups (Pane et al., 2014).

Cognitive Tutor study with significant negative effect. Pane, McCaffrey, Slaughter, Steele, and Ikemoto (2010) conducted a three-year study in eight high schools on the impact of CT geometry on student achievement and their attitudes toward mathematics and technology. The study revealed that students who were instructed with CT geometry scored statistically lower than their counterparts in the comparison group on the Baltimore County Public School District geometry assessment (Pane et al., 2010). The WWC characterizes this study finding as a statistically significant negative effect (U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse, 2016).

Technology's Impact on Student Achievement

Middleton and Murray (1999) examined the relationship between technology implementation levels in the classroom and student achievement in reading and mathematics. A survey was used to determine the level of technology implementation of 107 teachers. Standardized test scores in mathematics and reading of 2,574 fourth- and fifth-grade students were analyzed to compare to the teachers' level of technology implementation. The results indicated a significant difference in math and reading scores among the fifth-grade students, but no significant difference was found among the fourth-grade students (Middleton & Murray, 1999).

Du, Havard, Sansing, and Yu (2004) analyzed data from the Education Longitudinal Study of 2002 to examine the impact of computer use at school and home on low-income, minority students and their peers. The participants of the study were tenth-graders from 752 public, Catholic, and other private schools. The data collected for this study included an achievement test for mathematics and reading, a facilities checklist, and five separate questionnaires for students, parents, school administrators, and library media center completed by 15,362 tenth-graders, 13,488 parents, 7,135 teachers, 743 principals, and 718 librarians. The results of the study indicated that computer use at home was far more significant than computer use at school in relation to high academic performance; however, this effect was absent for minority and low socioeconomic status students. The results of the study also indicated that socioeconomic status impacted student access to computers and was a stronger indicator than race regarding the use of computers and students' achievement (Du et al., 2004).

Antonijevec (2007) investigated the effect of using computers and calculators on student achievement as measured by Trends in International Mathematics and Science

Study 2003 (TIMSS) assessment. The study included eighth-graders from four countries, including the United States, the Netherlands, Bulgaria, and Serbia, and explored the impact of modern technology equipment on students' overall achievement. Antonijevec suggested that using computers in teaching does not significantly improve students' achievement in mathematics but did show some significant influence on students' achievement in science. The students' international average achievement across all four countries in mathematics and science were 457 and 468.48, respectively. The results also indicated that using calculators in mathematics teaching improved overall students' achievement (Antonijevec, 2007).

Wenglinsky (2006) conducted a series of studies to examine the effects of technology on student achievement as measured by the test score and questionnaire data from the National Assessment of Educational Progress (NAEP). Wenglinsky analyzed data from the 1996 NAEP in mathematics and science, the 1998 NAEP in reading for students in fourth and eighth grades, and the 2001 NAEP U.S. history assessment for twelfth-graders. The results from the NAEP assessments in mathematics, science, and reading for fourth- and eighth-graders suggested that the quality of computer work was more important than the quantity. The study results indicated that using computers to help students work through complex problems to develop higher-order thinking skills produced higher benefits than using computers to drill students on a set of routine tasks. The results from the U.S. history study reported an effect size of 16 for the amount of time students had used computers for schoolwork outside the school and an effect size of 36 for students' socioeconomic status. An effect size of 15 and above is considered substantial, which indicates that students who make frequent use of computers for academic tasks outside school had higher achievement in history, and students' socioeconomic status was the factor that was strongly related to their history achievement. Moreover, Wenglinsky (2006) indicated that high schools would receive the most significant improvement if technology was used to deepen students thinking and enhance their work products through technology-driven processes. While some of these studies were not about mathematics, the impact of technology on student achievement, as measured by a standardized test, was explored.

Lei (2007) examined the relationship between technology use and student outcomes by comparing the association between the amount of time technology was used and student outcomes with the association between the quality of technology use and student outcomes. The amount of time technology used was measured by the time spent on computers every day, while the quality of technology was examined by looking at how the technology was used. The five general technology uses investigated in this study included subject-specific technology uses, social-communication technology uses, construction technology uses, and exploration/entertainment technology uses. The participants of the study were seventh- and eighth-graders and teachers in a Northwestern middle school in the U.S. with a total enrollment of 237. The results of the study revealed that the amount of time technology was used had no significant relationship with student outcomes. When the quality of technology used was examined, the results suggested that none of the five types of technology used had statistically significant effects on student outcomes as measured by their grade point average (Lei, 2007).

Pellerito (2011) examined the effectiveness of a computer-based instruction system on student achievement, attitudes toward mathematics, school climate, attendance, and discipline referrals for at-risk students in an alternative high school. The participants of the study included 30 at-risk high school students using computer-based instruction and 40 students using textbook-based instruction to teach Algebra 1 concepts from a suburban school district south of Kansas City, Missouri. The results of the study indicated a significant increase in grades and state assessment scores for at-risk students in the computer-based instruction group compared to the students in the traditional group. The mean difference between students' grades and state assessment scores from year one to year two was 2.54 and 0.54, respectively. Students in the computer-based instruction group also indicated more positive attitudes toward mathematics than did the students in the traditional class. The mean difference from year one to year two was 0.41 (Pellerito, 2011).

Cheung and Slavin (2013) examined research on the effects of educational technology applications on mathematics achievement in K-12 classrooms. Included in the analysis were 45 elementary studies and 29 secondary studies with a total sample size of 56,886 K-12 students. The results of the analysis suggested that educational technology applications generally produced a positive, though modest, effect when compared to traditional methods. The results revealed supplemental computer assisted instruction to have the largest effect, while computer-management learning and comprehensive programs, had a much smaller effect on student achievement (Cheung & Slavin, 2013).

While some researchers have indicated a positive impact of technology on student achievement, Reynolds (2013) found the use of technology in school to be negatively associated with student achievement. Reynolds (2013) examined the results of the 2009 NAEP 8th grade mathematics assessment to determine if factors including student technology use inside the school, student technology outside school, test motivation, academic self-concept, and home environment influence student's mathematics achievement in school. The participants of the study included 167,300 eighth-graders who participated in the 2009 NAEP mathematics assessment. The results of the study revealed that student technology use in school is negatively associated with student mathematics achievement for all 8th-grade students and race/ethnicity groups. Eighth-grade students who hardly or never use technology in school are likely to score 42 units lower on the NAEP mathematics assessment than the mean score. In contrast, student technology use outside school was positively associated with student mathematics achievement for all 8th-grade students and race/ethnic groups. Students who use technology outside school are more likely to score nine units higher than the mean mathematics score. The study results also indicated that student math self-concept is positively associated with student mathematics achievement for all eighth-grade students and race/ethnicity groups. Students with a high self-concept are likely to score 33 units above the mean. However, motivation is negatively associated with predicting student achievement. The results indicated that eighth-grade students who did not take the assessment seriously or did not exert effort on the test were likely to score 17 units lower than the mean score. Home environment was not a significant variable (Reynolds, 2013).

Arbain and Shukor (2015) conducted a study on the effects of Geogebra on student achievement. GeoGebra is a mathematics software for all levels of education that brings together geometry, algebra, spreadsheets, graphing, statistics and calculus in one package (Geogebra, 2020). The researchers utilized a quasi-experimental design and included 62 high school students in Malaysia. The results of the study indicated that the difference in the mean scores for the pre-achievement test and the post-performance test for the two groups shows that the experimental group performed better than the control group. The findings indicated that using GeoGebra software positively impacted students' achievement in Mathematics (Arbain & Shukor, 2015).

Cline (2017) explored differences in third- through sixth-grade student achievement, as measured by the change in Measures of Academic Progress mathematics assessment scores, between students who participated in a one-to-one iPad initiative and students who participated in a one-to-many initiative and whether those differences were affected by student gender, race, or socioeconomic status. The participants of the study were third through sixth graders from a suburban school district in Northeast Kansas who participated in one-to-one and one-to-many iPad initiatives. The results of the study suggested no significant difference in mathematics gain between students who participated in the one-to-many initiative and students who did not participate in the one-to-one initiative. The data analysis showed no interaction effects for student gender, race, or student SES, but there were four main effects. Fifth grade White students had a significantly higher mean gain than did Black and Hispanic students. Fifth grade students on full pay lunch status had a significantly higher mean gain than did free/reduced lunch status students. In sixth grade, females had a significantly higher mean gain than did males. Sixth grade students on full pay lunch status had a significantly higher mean gain than free/reduced lunch status students. When comparing students who participated in the initiative for one year versus two years, the results of the analysis indicated that there were significant differences in the mean gain for fourth and sixth grade students. The results of the data analysis also showed that student race and student SES affected the mathematics gain of sixth-grade students (Cline, 2017).

Similarly, Howard (2017) investigated the impact of a one-to-one iPad initiative program on eleventh-graders' student achievement as measured by the Michigan Merit Exam in mathematics, science, and social studies. Participants in the study were 11th graders from one rural school district in Michigan. The results of the study indicated that the iPad's implementation had significantly increased eleventh-graders' test scores in the mathematics, science, and social studies test over the years 2013-2016. Mathematics test scores increased by 219.30 points, Science test scores increased by 215.22 points, and Social Studies test scores increased by 215.11 points after the implementation of the iPad program (Howard, 2017).

Hillmayr, Ziernwald, Reinhold, Hofer, and Reiss (2020) conducted a comprehensive meta-analysis of 92 studies including 14,910 students on how the use of technology can enhance learning in secondary school mathematics and science for grades 5–13 students. The purpose of the studies was to compare the learning outcomes of students in the experimental group using digital tools to those of a control group taught without the use of digital tools. The results of the study indicated that digital tool use had a medium significant effect on student learning outcomes. Thus, secondary school students who were instructed using digital tools in science or mathematics classes had significantly greater learning outcomes than students who were taught without the use of digital tools (Hillmayr et al., 2020).

Seitan, Ajlouni, and Al-Shra (2020) investigated the impact of integrating flipped learning and information and communication technology on secondary school students' academic achievement and attitudes. The study used a quasi-experimental design and included 40 high school students from Mashrek International School in Amman, Jordan.

An academic achievement test was utilized, and an attitude scale was implemented for the experimental group who was instructed using a flipped learning model and different information and communication technology. The findings revealed that integrating the flipped learning approach and information communication technology enhanced the academic achievement of high school students and created a positive attitude towards peers, teachers, the environment, and motivation. The results indicated that the flipped learning model had a large effect on the academic achievement of the secondary students' computer discipline (Seitan et al., 2020).

Zulnaidi, Oktavika, and Hidayat (2020) conducted a similar study on the effects of using GeoGebra's software on student achievement. A total of 53 students ages 16 and 17 years from Sekolah Menengah Perempuan Jalan Ipoh, Kuala Lumpur, participated in the study. The study results indicated a significant difference between the mean of the students' scores in favor of the students who used GeoGebra. The findings indicated that computer-assisted instruction as a supplement to traditional classroom instruction was more effective than traditional instruction alone (Zulnaidi et al., 2020).

Intelligent Tutoring System

Carbonell (1970) introduced Intelligent Computer Assisted Instruction (ICAI) when designing a computer program called SCHOLAR. SCHOLAR was considered the first Intelligent Tutoring System (ITS) because of its tutorial capabilities to not only contain knowledge about the domain but also knowledge about the student and how to teach that student (Steele & Steele, 1999). Sleeman and Brown (1982) envisioned an advanced Computer Assisted Instruction (CAI) system they called an intelligent tutoring system (ITS). Their system's emphasis was learning by doing and representing a learner's knowledge. Sleeman and Brown (1982) were credited for creating the term "intelligent tutoring system," and they were the first to use the term "student model," to describe an abstract representation of the learner within the computer program. ITS can be understood as a type of Computer-Based Tutoring (CBT) called cognitive tutor.

Anderson, Corbett et al. (1985) reviewed the results of some computer-based tutoring, including geometry tutors and algebra tutors. These tutors were developed based on the cognitive model and the use of intelligent tutoring systems. During the 1986-1987 school year, high school student performance in classes with geometry tutors was compared to classes without tutors but with the same teacher using regression analysis. The results indicated that students who had one-to-one access to geometry tutor gained 14 points on a geometry proving theorem paper and pencil test, which is more than one standard deviation in the population. However, students who were paired with another student using the computer, creating two-on-one access to geometry tutor, gained four points on the paper and pencil test on proof skills, which was not significant compared to the control group. An evaluation of an algebra tutor was also performed in the 1987-1988 school year. The results indicated no statistical difference between the experimental groups who had access to the algebra tutor and the control groups who had no access to the algebra tutor (Anderson, Corbett et al., 1985).

Koedinger and Anderson (1993) developed and investigated the effectiveness of A New Geometry Learning Environment (ANGLE), an ITS for proving geometric theorems in high school mathematics class. The study included four experimental classes that used ANGLE and four control classes that did not use ANGLE at Pittsburgh public schools. Classes were taught by three teachers who were assigned to teach at least one experimental class and one control class. One of the teachers included in the study participated in the ANGLE development project. Students were instructed using ANGLE in the spring of 1992 for about 20-25 class periods, each lasting about 44 minutes. The results of the study indicated no significant difference between the post-test scores of students in the ANGLE classes and students with no ANGLE classes. However, the students who were taught by the teacher who participated in the ANGLE development project scored much higher on average than students in the other classes. From an informal observation, the teacher who was involved in the ANGLE project was more involved in the content related interaction with the students because he already had knowledge of the problem solving and tutoring approach of ANGLE, compared to the other teachers who had limited experience in using the computer tutor (Koedinger & Anderson 1993).

Corbett, Koedinger, and Anderson (1997) reported that ITS yielded approximately one standard deviation benefit compared to traditional practice. ITS adaptively responds and identifies a student's problem-solving strategy through their responses and compares correct and incorrect responses (Graesser, Conley, & Olney, 2012). The comparison process is called tracking knowledge or model tracing. Corbett (2001) found a large effect size when students were exposed to a model tracing system compared to no model tracing. Koedinger and Corbett (2006) estimated that approximately half a million students had used CT for a total of about 20 million hours. Students who were instructed with CT Algebra 1 have scored twice as high on end-of-course open-ended problem-solving tests and 15% higher on objective tests as students enrolled in a traditional Algebra course (Koedinger & Corbett, 2006).

Koedinger et al. (1997) evaluated the effect of the Pittsburgh Urban Mathematics Project (PUMP) curriculum and an intelligent tutor called Practical Algebra Tutor (PAT) in comparison to the traditional curriculum without PAT. The study included three Pittsburgh high schools with similar demographics and student aptitudes. Students in the experimental class consisted of 20 algebra classes and involved 470 students and 10 teachers, while students in the comparison group consisted of five algebra classes that included 120 students and three teachers. The study utilized two standardized tests, including the Iowa Algebra Aptitude test and a subset of the Math SAT appropriate for 9th graders. The students were also administered multiple mathematical representation and real-world problem-solving tests to assess objectives reflecting the National Council of Teachers of Mathematics (NCTM) recommendations and the PUMP curriculum. The study results indicated that students in the PUMP and PAT classes outperformed students in comparison classes by about 1.0 standard deviation on the NCTM-oriented test. The PUMP+PAT students' scores were about 100% better or double those of the comparison classes, and they scored about 15% on the standardized tests (Koedinger et al., 1997).

The first-generation of computer tutors were known as computer-assisted instruction tutors (CAI tutors), and the second-generation tutors were usually called intelligent tutoring systems (ITS) (VanLehn, 2011). According to VanLehn (2011), CAI tutors were generally believed to increase examination scores by 0.3 standard deviations, or from the 50th to the 62nd percentile, while the ITS were considered to be more effective, raising test performance by approximately one standard deviation, or from the 50th to the 84th percentile. Advances in ITS have progressed to being used in the educational setting. CT is thought to be the most widely used of all the ITS in school settings and is referred to as a particular type of ITS developed by Carnegie Learning for middle and high school mathematics. Graesser, Conley, and Olney (2012) defined ITS as "computerized learning environments that incorporate computational models in the cognitive sciences, learning sciences, computational linguistics, artificial intelligence, mathematics, and other fields that develop intelligent systems that are well specified computationally" (p.2). ITS was based on artificial intelligence concepts and cognitive theory that guides learners through every step of a problem solution by providing hints and feedback as needed from expert-knowledge databases (VanLehn, 2011).

VanLehn (2011) conducted a meta-analysis of 54 studies comparing the learning outcomes for ITS and non-ITS groups. The analysis results indicated an improvement in the student test scores with an average effect size of 0.58 standard deviation representing a medium effect size. VanLehn (2011) categorized the ITS used in the 54 comparisons as either step-based tutoring or sub-step-based tutoring. Step-based tutoring provides hints and explanations on steps that students typically take when solving problems. In contrast, sub step-based tutoring is an approach that provides scaffolding and feedback at a more detailed level. The effect size of step-based tutoring was large, raising test scores by 0.76 standard deviations, whereas sub step-based tutoring had a small effect size raising test scores by only 0.40 standard deviation (VanLehn, 2011).

Researchers have not reached a consensus on the effect of ITS on student learning. Steenbergen-Hu and Cooper (2013) conducted a meta-analysis of 26 studies involving ITS on K-12 mathematics learning. The results of the analysis indicated that ITS has no overall effect on learning in K-12 students. Test scores of students who were instructed with ITS and control group students differed overall by 0.05 standard deviations, which was insignificant. The results of Kulik and Flecher (2016) were in contrast to Steenbergen-Hu and Cooper's (2013) meta-analysis, which found a small effect size in 18 studies of the effectiveness of ITS in elementary and high school mathematics. The two studies defined ITS differently. Steenbergen-Hu and Cooper (2013) defined ITS as "self-paced, learner-led, highly adaptive, and interactive learning environments operated through computers" (p. 983), which led them to include some computer systems that do not adhere to the standard features of ITS. Whereas, Fletcher (1985) identified three key features of ITS as follows: (a) represent a relevant knowledge domain which contains the foundations, concepts, and rules that experts understand and use in solving problems in the domain; (b) represent the student's state of knowledge, or model, of the subject matter; and (c) represent an expert tutor, which chooses tutoring strategies and actions to apply in specific situations for specific students.

Kulik and Flecher (2016) conducted a meta-analytic review of the effectiveness of the Intelligent Tutoring System. The 50 evaluations in the meta-analysis included 39 studies conducted in the United States, and 11 were conducted outside the U.S. The results of the study indicated that students who received Intelligent Tutoring outperformed students from traditional classes in 92% of the studies that were analyzed. The results revealed a medium effect size of Intelligent Tutoring System to studies conducted in the U.S., while the studies conducted outside the U.S. revealed a large effect. What Works Clearinghouse (2002) considered an effect size of at least 0.25 standard deviation to be a substantially important positive effect. Cohen (1988) cited that an effect of 0.20 standard deviations is small, 0.50 standard deviations are medium size, and 0.8 standard deviations are large. By these standards, the average effect size for intelligent tutoring was moderate to large.

Usage of Cognitive Tutor

As students use CT, the program records timestamps at the beginning and end of each activity. Analyzing the relationship between the number of hours the students spent using the CT software and learning outcomes provides valuable information to district leaders in deciding the type of curriculum to adopt. This section focuses on studies of the adaptive software component of CT and its impact on student achievement.

Pane, McCaffrey, Slaughter, Steele, and Ikemoto (2010) conducted an experiment to evaluate the efficacy of Cognitive Tutor Geometry in eight high schools in the Baltimore County Public School District over three years. Cognitive Tutor Geometry instruction was provided to 329 students in the intervention group, and 340 students in the comparison group received the standard geometry curriculum. The study was conducted over three school years. The results of the study indicated that students who received CT instruction performed significantly lower than students who received the traditional geometry curriculum. The results indicated statistically significant negative effects on student achievement, as measured by the Baltimore County Public School District Geometry Assessment. The study results also revealed that the average classroom usage of the tutoring software was not related to student achievement. The classroom software usage was based both on software logs and observers' ratings of computer lab usage. However, the number of skills a student mastered using the software and that student's post-test scores revealed a positive and significant relationship (Pane et al., 2010).

Sales and Pane (2015) explored the causal mechanism of the effectiveness of CT Algebra 1 from a large-scale randomized control trial revealing that students who completed more than 27 CT sections significantly outscored equivalent students using a standard curriculum (Pane et al., 2014). According to Pane et al. (2014), approximately 72% of students completing more than 27 sections passed the end of year Algebra 1 proficiency exam. Sales and Pane (2015) utilized data from the high school sample that experienced a substantial CT Algebra I effect to explore the relationship between student usage and causal effects. The researchers could not establish that students who spent more time using CT Algebra 1 benefited from its usage. However, students who completed a larger number of sections had a much larger effect than those who completed fewer sections. The effect size was between 0.13 and 0.47, which was considered a large effect (Sales & Pane, 2015).

Fancsali, Ritter, Yudelson, Sandbothe, and Berman (2016) investigated the effectiveness of teachers' implementations of ITS and how implementations affect student efficiency in ITS usage and long-term learning outcomes. The study included CT Algebra usage data from 2,025 students across 18 high schools and three middle schools in a large Southeastern school district in the U. S. The school district policy mandated that the high school students who failed their mathematics year-end standardized tests in the previous academic year use CT in their high school Algebra course. The results indicated no relationship between the time the students used the software and the percentage of students passing the Algebra 1 exam. The results also revealed a strong relationship between the number of sections mastered per hour and students' scores on the end-of-year exam.

Fancsali et al. (2018) provided evidence that the use of MATHia, an Intelligent Tutoring software, was associated with students' test scores as measured by the Florida Standards Assessment (FSA) and Florida Comprehensive Assessment Test (FCAT). The researchers used data from 23,374 sixth to eighth-grade students who used MATHia over three academic years in Miami-Dade County Public Schools. The results indicated that statistical models based on process variables including the number of problems, skills, and topics mastered and the amount of time taken, the rate at which content was mastered, the number of errors and hints made in the software were able to predict outcomes on the students' state test scores (Fancsali et al., 2018). Schools that use the MATHia software consistently and with high proficiency also showed higher proficiency levels on the state assessments (Carnegie Learning, 2018).

Trecek-Schaffer (2019) conducted a study on the impact of MATHia on student achievement in mathematics. The study participants included students who were placed in Tier 2 and 3 sections. Students in the Tier 2 section scored between the 17th and 31st percentiles on the Measures of Academic Progress (MAP) assessment, whereas students in the Tier 3 category scored below 16^{th} percentile on the MAP assessment. The results of the study related to software usage revealed a significant difference in Tier 2 and Tier 3 high school students' growth from fall to winter and winter to spring, as measured by the change in the NWEA MAP Mathematics assessment score, among students with low, moderate, and high utilization of the MATHia software (Trecek-Schaffer, 2019).

Effects of Demographic Variables on Cognitive Tutor Software

What Works Clearinghouse (WWC) conducted a systematic review of 11 studies about CT Algebra 1 and an additional study on CT Geometry. “CT Algebra 1 was found to have mixed effects on algebra and no discernible effects on general mathematics achievement for secondary students while CT Geometry was found to have potentially negative effects on geometry for secondary students" (U.S. Department of Education, Institute of Education Sciences, WWC Cognitive Tutor, 2016, p. 2). While the impact of Cognitive Tutor was examined in several studies, most of the studies focused on the general student population, with limited data on the impact of the Cognitive Tutor curriculum on the mathematics achievement of students in different subgroups, including students with disabilities, English language learners, and students with free- and reduced-lunch status.

Steele and Steele (1999) identified an ITS called DISCOVER, which could help students with learning difficulties in solving word problems. The system used direct teaching and was made up of several independent computer programs. These programs captured students' knowledge and abilities, then provided different strategies, hints, and coaching techniques, giving students individualized instruction on the succeeding sequential steps for solving word problems. DISCOVER was “beneficial for students with learning problems because it incorporates many of the advantages of direct teaching, including verbal rehearsal, clear sequencing, repetition, modeling, structure, and review” (Steele & Steele, 1999, p. 4).

Sarkis (2004) examined the effectiveness of the CT curriculum compared to the conventional curriculum as measured by the Florida Comprehensive Assessment Test and students' final grades and teachers' and administrators' opinions about the programs. The study participants were 4,649 ninth-grade students in ten high schools in the Miami-Dade County Public Schools who completed Algebra I. Separate analyses were conducted for

770 exceptional education students and 976 students with limited English proficiency. The results indicated that ninth-grade students who completed Algebra I Cognitive Tutor scored significantly better than students who completed conventional Algebra I classes, as indicated by the 4.4 score gain on the Florida Comprehensive Assessment Test (FCAT). Exceptional education students, including those with learning and behavioral difficulties and limited English proficiency who were instructed with CT, performed significantly better than their counterparts who were instructed with a traditional curriculum. The exceptional education students in the Cognitive Tutor group gained 44 points higher on the FCAT mathematics scale score than their comparison group. Similarly, limited English proficiency students gained 16 points higher on the FCAT mathematics portion compared to the comparison group. There was no significant difference in the students' final grades in the Cognitive Tutor group than the comparison group. Outcomes for final grades across two groups regarding exceptional student education and limited English proficiency students were similar to the whole group results (Sarkis, 2004).

Ritter, Joshi, Fancsali, and Nixon (2013) conducted a study to predict the standardized test scores of students based on usage of Cognitive Tutor, as measured by the outcomes on the Virginia Standards of Learning (VSOL) and NWEA's MAP Assessment. The study participants included 3,224 sixth-to eighth-grade students from a school district in Eastern Virginia who used CT software during the 2011-2012 school year. The data included demographic variables such as gender, age, race, socioeconomic status, English language status, and student disability status. The results of the study indicated that the number of sections completed in the CT software was the strongest predictor in the VSOL assessment. The results also revealed that students of low socioeconomic status performed significantly worse on the VSOL assessment. Students' age indicated a small significant effect, where older students tended to underperform their younger classmates. The other demographic variables did not significantly predict student scores (Ritter et al., 2013).

Huang, Craig, Xie, Graesser, & Hu (2016) investigated how the intelligent tutoring system, ALEKS, can reduce the student mathematics achievement gap among race, gender, and socioeconomic status. ALEKS instructs the students on the topics they are ready to learn, tracks their knowledge, and adaptively corrects mistakes and misconceptions. The study participants were 533 sixth-grade students from five middle schools in West Tennessee who attended an after-school program. The program ran for three consecutive years, but only students from years two and three were included in the study because the first year was a pilot year. The results revealed that students' scores in the math standardized state test were higher in the ALEKS class, but the difference was not significant. The results also indicated no significant interaction in the ALEKS-lead class among race, gender, and socioeconomic status (Huang et al., 2016).

Trecek-Schaffer (2019) conducted a study on the impact of MATHia on student achievement in mathematics, as measured by the NWEA Measures of Academic Progress (MAP) Mathematics assessment. The participants of the study were students in Tier 2 and 3 math intervention classes. The Tier 2 section included students who scored between the 17th and 31st percentile on the MAP assessment. In contrast, Tier 3 sections include students who scored below the 16th percentile on the MAP assessment. One of the study's major findings in relation to student demographics indicated no significant difference in academic growth between students by gender, socioeconomic status, or learning disability status (Trecek-Schaffer, 2019).

Summary

Chapter 2 included literature about the historical overview of CT, Intelligent Tutoring System, and the impact of technology on student achievement. Current studies on the impact of CT on student achievement of students with disabilities, English language learner students, and students in free- and reduced-lunch status, and the amount of time spent using CT was also discussed. Chapter 3 includes the research design, measurement, data collection procedures, data analysis and hypothesis testing, and the limitations of the study.

Chapter 3

Methods

The primary purpose of this study was to examine the impact of the Cognitive Tutor (CT) curriculum on students' Mathematics KAP test scores after a year of implementation. Chapter 3 contains several sections that explain the methodology contained in this study. The sections include the research design, selection of participants, measurement, data collection procedures, data analysis, and hypotheses testing, and the review of study limitations.

Research Design

A quasi-experimental quantitative research design was used to examine the difference of the mathematics test scores of tenth-grade students at District X before and after the implementation of the CT as measured by the Mathematics KAP test. The independent variable was the curriculum type (CT or traditional) that tenth-grade students received. The dependent variable was the mathematics KAP performance level scores of tenth-grade students. The tenth-grade student's performance level scores on the Mathematics KAP test in the spring of 2017 before the implementation of CT were compared to the tenth-grade students' Mathematics KAP performance level scores in the spring of 2018, which occurred after the implementation of CT. Demographic variables included in the study to select student subgroups for further comparison were: disability status, English language learner status, and free- and reduced-lunch status.

Selection of Participants

The population for this study included all tenth-grade students in eight comprehensive high schools in District X. Lunenburg and Irby (2008) defined purposive sampling as “selecting a sample based on the researcher’s experience or knowledge of the group to be sampled” (p. 175). In the current study, purposive sampling was selected based on the researcher’s knowledge that high schools in District X implemented the CT curriculum, and students participated in the KAP test. Individual student data were included in this study if the following criteria were met:

1. The student was a 10th grader enrolled in District X during the 2016-2017 and 2017-2018 school years;
2. The student received traditional curriculum instruction during the 2016-2017 school year and the student received CT instruction during the 2017-2018 school year;
3. Students completed the Mathematics KAP test in the spring of 2017 or 2018 of their tenth-grade year.

The sample for this study included 3,040 tenth-grade students who received traditional mathematics instruction during the 2016-2017 school year and 3,061 tenth-grade students who received mathematics instruction using the CT during the 2017-2018 school year.

Measurement

This section includes a discussion of how the independent variables, curriculum type, and the number of hours students MATHia were measured. A discussion on how the dependent variable, mathematics achievement, was measured is also presented in this section. Lastly, this section also explains how each student subgroup, including students with disabilities, English language learner students, and students with free and reduced lunch status, was measured.

Curriculum type. Students' curriculum type was measured by the school year for which they received mathematics instruction during their tenth-grade. Students who were in their tenth-grade on or before the school year 2016-2017 were instructed with a traditional mathematics curriculum, so they were categorized as the traditional curriculum group. Students who were in the tenth-grade on or after the 2017-2018 school year were instructed using the CT curriculum, so they were categorized as the CT curriculum group.

MATHia usage. Each time a students log in to MATHia, each student's data is constantly recorded. The number of hours students utilized MATHia was measured by the total number of hours collected from the software at the end of the student's tenth-grade. The MATHia software automatically records the number of hours the students used the program whenever they are logged in to the program.

Mathematics achievement. Mathematics achievement was measured by Mathematics KAP testing. The Mathematics KAP is a standards-based test that measures specific skills defined for each grade by the state of Kansas. The 2017 and 2018 Mathematics KAP tests measure specific claims related to Kansas College and Career Ready Standards (KCCR) (KSDE, 2018). The Mathematics KAP test is comprised of four subtests: the areas of concepts and procedures, problem-solving, communication and reasoning, and modeling and data analysis. The KAP test is an adaptive assessment that adjusts the question difficulty at the midpoint based on individual student responses (University of Kansas Achievement & Assessment Institute, KAP Technical Manual, 2017). The test was administered for two sessions. All students take math Session 1. After completing Session 1, the computer calculates a student's score and chooses the appropriate level of Session 2 testing for the student. The 2017 and 2018 KAP assessment is a multiple-choice and machine-scored and compares student responses to the correct keys stored with the items and assigns the predetermined scores accordingly. The total score is derived using the number-correct method in which each item's scores are added together to obtain the raw score (University of Kansas Achievement & Assessment Institute, 2017).

Each of the raw scores was transformed into a scale score to facilitate the interpretation of test scores. A scaled score is a mathematical conversion of the total number of points a student earned on an assessment into a score along a predefined scale, which allows for consistent reporting of assessment results across years for any specific grade and subject area (KAP, 2017). Scale scores on all KAP assessments range from 220 to 380. The cut score for proficiency is 300 for all KAP assessments. The Mathematics KAP test result was reported with individual student scores and the students' performance level. KAP (2017) uses four performance levels to assess students' mathematics proficiency, from level 1, indicating a below grade level performance, to level 4, indicating an exceeding grade-level performance.

Lunenburg and Irby (2008) defined validity as "the degree to which an instrument measures what it purports to measure" (p. 181). The Achievement & Assessment Institute (AAI) at the University of Kansas cited validity and reliability evidence of the Mathematics KAP assessment development and reported that item development followed well-established procedures. "Evidence of content validity for the KAP assessment comes from the alignment between KAP items and the KCCRS and the congruence between the test and test blueprint" (University of Kansas Achievement & Assessment

Institute, 2017, p. 31). To evaluate the validity of the KAP assessment, AAI examined the alignment between the KAP test blueprint with the knowledge and skills identified in the KCCRS (University of Kansas Achievement & Assessment Institute, 2017). AAI also conducted a content review using a panel of content experts to evaluate each of the items, ensure the alignment of all items to KCCR, consider grade-level appropriateness, depth of knowledge, distractor reasonableness, answer correctness, graphics, grammar, punctuation, and appropriate language demand (University of Kansas Achievement & Assessment Institute, 2017). The KAP test also undergoes an external bias, fairness, and sensitivity review for issues related to gender, diversity, and other factors (University of Kansas Achievement & Assessment Institute, 2017).

Test reliability is "the degree to which an instrument consistently measures whatever it is measuring (Lunenburg & Irby, 2008, p. 182). The Mathematics KAP test had been tested for reliability, and the results showed that for each of the tenth-grade student subgroups, including race, Hispanic, students with disabilities, and English language learners, the internal reliability for Mathematics KAP is 0.93. The reliability coefficient for the Mathematics KAP tests at high school grade is 0.93, which reflects a high level of internal consistency within each of the tests (University of Kansas Achievement & Assessment Institute, 2017, p. 46).

Disability status. Students' disability status was measured using an assessment by the school district or from a medical provider. District X provided a free comprehensive evaluation to any students suspected of having disabilities and may need school intervention. During enrollment, all parents were required to answer a questionnaire about student information, including their child's disability status, and provide supporting documentation if their child has been previously diagnosed with disabilities. Parents made a verbal and written request to the school or district's special education office for their child to be evaluated for suspected disabilities. Teachers and other staff also referred students they observed have disabilities that impact their learning to the Child Study Team (CST). Working with the parents, the CST evaluated the identified student in all areas related to students' potential disability and decided if the student qualified for a disability as defined by Individuals with Disabilities Education Act (IDEA). Students who qualified as a child with disability were recorded in the district student information system, and they were identified as a special education student in the information system. Students' disability status was measured and recorded in the same manner from the years 2016-2017 to 2017-2018.

English language learner status. Students' English language learner (ELL) status was measured by an English language assessment, which is commonly known as placement or screener test. Schools asked families to complete a home language survey to identify students who are potential English learners and will require an assessment of their English language proficiency to determine whether they are eligible for Speakers of Other Languages (ESOL) services (U.S. Department of Education, Office of Language Acquisition, 2016; KSDE ELL, 2020). In addition, during enrollment, parents and guardians of all new kindergarten and incoming students in grades K-12 were asked information related to the child's primary language at home, the language most spoken by the student, and the language that the student first acquired. Students and families who are new to the school district and have a native language other than English were directed to District X Center for Multilingual Education Services to assess students' English language skills and confirm the student's language status. District X administered the Kansas English Language Proficiency Assessment-Placement (KELPA-P) to all students whose first language or home language is not English. KELPA-P evaluates a student's language proficiency in all four language domains, including speaking, listening, reading, and writing, and classified them into one of four proficiency level categories on each of the four domains: beginning, intermediate, advanced, and fluent (KSDE, 2020). High school students were considered fluent in each domain if they scored 86% correct in listening, 90% correct in reading, 88% correct in writing (multiple-choice), 100% correct in writing (constructed response), and 100% correct in speaking. Any students who achieve a score that is lower than the fluency cut off score in listening, reading, writing, and speaking was identified as an ELL. District X used KELPA-P to annually assess students' English language proficiency from 2016 to 2018, and ELL students' statuses were recorded in the same way to the district's student information system (KSDE, 2018).

Free- and reduced-lunch status. A student's free- and reduced-lunch status was measured by students' family income eligibility. During enrollment, all parents answer questions about the amount and source of income received by each family member living in the student household and the number of family members living in the same household. The school then compares the household size and total income to the Federal Income Eligibility guidelines, which determine who is eligible for free and reduced-price meals. Children from families with incomes at or below 130% of the poverty level are eligible for free meals, while those with incomes between 130 and 185% of the poverty level are eligible for reduced-price meals (Food Research and Action Center, 2020).

Moreover, schools utilized data from the state to certify categorically eligible students or coordinate with the school districts homeless and migrant education liaisons to obtain documentation to certify children for free school meals. Children in households participating in the Temporary Assistance for Needy Families, Supplemental Nutrition Assistance Program, and the Food Distribution Program on Indian Reservations, as well as homeless, migrant, foster youth, or runaway youth, and Head Start participants are categorically eligible for free school meals and can be certified without submitting a school meal application (Food Research and Action Center, 2020). Any students who met the income and categorical eligibility were added to the free-and reduced program, and their economic status was recorded in the district's student information system. A student’s free and reduced lunch status was measured using the same Federal Income Eligibility guidelines and recorded similarly in the school district data system from the 2016-2017 to the 2017-2018 school year.

Data Collection Procedures

Data collection began with an informal discussion about the purpose of the current research study with the executive director of the Research and Assessment of District X. A letter of request was sent to the Executive Director and Research Council Committee of District X on August 31, 2020 to request permission to collect and study archived Mathematics KAP data of tenth-grade students during the 2016-2017 and 2017-2018 school years from district data files. The Research Council Committee granted permission for the current study and for collecting Mathematics KAP data on September 9, 2020, with the agreement to adhere to X district guidelines to protect personal student information (see Appendix C). Upon approval to conduct a research study from District

X, an application was submitted to the Baker University Institutional Review Board (IRB) on September 21, 2020 seeking permission to conduct the current study (see Appendix A). The Baker IRB granted permission for the current study on September 22, 2020 (see Appendix B).

After permission to conduct a study and receive data had been granted, the researcher collected and downloaded the data from District X computer into an Excel document on September 22, 2020. Identifying student information (i.e., students' names) was excluded from the Excel document to protect students' anonymity. Data provided to the researcher included only student information essential to the study, including mathematics KAP scores of all tenth-grade students, special education status, English language learner status, free- and reduced-lunch status, and the curriculum type.

Data Analysis and Hypothesis Testing

Data were collected and entered into the IBM Statistical Package for the Social Sciences (SPSS) version 25 for analysis. The research questions guiding the current study were as follows;

RQ1. To what extent is there a difference in the Mathematics KAP scores between tenth-grade students who did not receive CT instruction and tenth-grade students who received CT instruction?

H1. There is a significant difference in the Mathematics KAP scores between tenth-grade students who did not receive CT instruction and tenth-grade students who received CT instruction?

An independent-samples *t* test was conducted to test H1. The mean of the math scores before CT instruction and the mean of the math scores after CT instruction were compared. An independent-samples *t* test was chosen for the hypothesis testing since it examines the mean difference between two mutually exclusive independent groups, and both means of two groups are continuous variables. The level of significance was set at .05. When appropriate, an effect size is reported.

RQ2. To what extent is there a difference in the Mathematics KAP scores between tenth-grade students with disabilities who did not receive CT instruction and tenth-grade students with disabilities who received CT instruction?

H2. There is a significant difference in the Mathematics KAP scores between tenth-grade students with disabilities who did not receive CT instruction and tenth-grade students with disabilities who received CT instruction?

An independent-samples *t* test was conducted to test H2. The mean of the math scores of students with disabilities before CT instruction and the mean of the math scores of students with disabilities after CT instruction were compared. An independent-samples *t* test was chosen for the hypothesis testing since it examines the mean difference between two mutually exclusive independent groups, and both means of two groups are continuous variables. The level of significance was set at .05. When appropriate, an effect size is reported.

RQ3. To what extent is there a difference in the Mathematics KAP scores between tenth-grade English language learners who did not receive CT instruction and tenth-grade English language learners who received CT instruction?

H3. There is a significant difference in the Mathematics KAP scores between tenth-grade English language learners who did not receive CT instruction and tenth-grade English language learners who received CT instruction?

An independent-samples *t* test was conducted to test H3. The mean of the math scores of English language learners before CT instruction and the mean of the math scores of English language learners after CT instruction were compared. An independent-samples *t* test was chosen for the hypothesis testing since it examines the mean difference between two mutually exclusive independent groups, and both means of two groups are continuous variables. The level of significance was set at .05. When appropriate, an effect size is reported.

RQ4. To what extent is there a difference in the Mathematics KAP scores between tenth-grade students with free- and reduced-lunch status who did not receive CT instruction and tenth-grade students with free- and reduced-lunch status who received CT instruction?

H4. There is a significant difference in the Mathematics KAP scores between tenth-grade students with free- and reduced-lunch status who did not receive CT instruction and tenth-grade students with free- and reduced-lunch status who received CT instruction?

An independent-samples *t* test was conducted to test H4. The mean of the math scores of students with free- and reduced-lunch status before CT instruction and the mean of the math scores of students with free- and reduced-lunch status after CT instruction were compared. An independent-samples *t* test was chosen for the hypothesis testing since it examines the mean difference between two mutually exclusive independent groups, and both means of two groups are continuous variables. The level of significance was set at .05. When appropriate, an effect size is reported.

Limitations

“Limitations are factors that may have an effect on the interpretation of the findings or on the generalizability of the results” (Lunenburg & Irby, 2008, p. 133). Other factors may influence the outcome because extraneous forms of variance from the environment are less controlled in quasi-experimental research. In the current study, the following factors were perceived limitations;

1. Teachers may have varied teaching strategies due to professional experience and academic qualifications and may have varied formative and summative assessments to evaluate students’ performance of specific concepts and provided different types of feedback.
2. All the sample is from one school district, and they are only tenth-grade students. So, the results of the study should not be generalized to other grades or districts with different settings.
3. This is a comparison study, so the results could not establish a causal relationship of the impact of the intervention.

Summary

This study was conducted to examine the difference in the Mathematics KAP scores between tenth-grade students and three student subgroups who did not receive CT instruction and tenth-grade students and three student subgroups who received CT instruction. The study also investigated the relationship between the number of hours tenth-grade students participated with CT’s Mathia and their overall Mathematics KAP scores. This chapter included the research design, selection of participants, instrumentation, data collection procedures, data analysis and hypothesis testing, and limitations of the study. The results of the current study will be presented in chapter four.

Chapter 4

Results

The primary purpose of this study was to investigate the extent of the difference in the mathematics KAP scores between tenth-grade students in District X who did not receive CT instruction and tenth-grade students who received CT instruction. The second purpose of this study was to determine to what extent there was a difference in the mathematics achievement of tenth-grade students with disabilities, English language learner students, and free- and-reduced students as measured by the KAP assessment after the implementation of the CT curriculum. This chapter contains the descriptive statistics and the results of the hypothesis testing for the five stated research questions.

Descriptive Statistics

The mathematics KAP test scores of students before the implementation of CT and after the implementation of CT were used to gather student math achievement information. The mean and standard deviation of the mathematics KAP scores for all tenth-grade students, students with disabilities, English language learner students, and students in free- and reduced-lunch who were instructed with the traditional curriculum and the CT curriculum are found in Table 3.

Table 3

Descriptive Statistics of Tenth-Grade Students Mathematics KAP by Subgroups

	N	*M*	*SD*
All Tenth-Grade Students			
Traditional	3,040	238.10	55.23
Cognitive Tutor	3,061	236.51	52.74
Students with Disabilities			
Traditional	394	235.55	32.52
Cognitive Tutor	402	230.94	30.01
English Language Learner Students			
Traditional	773	244.93	44.68
Cognitive Tutor	832	242.21	43.30
Students with Free-and Reduced Lunch			
Traditional	2,323	242.47	48.37
Cognitive Tutor	2,214	240.84	44.85

Note: M = mean and SD = standard deviation.

Hypothesis Testing

The results of the hypothesis testing that addressed the five research questions utilized in this study are discussed in this section.

RQ1. To what extent is there a difference in the Mathematics KAP scores between tenth-grade students who did not receive CT instruction and tenth-grade students who received CT instruction?

H1. There is a significant difference in the Mathematics KAP scores between tenth-grade students who did not receive CT instruction and tenth-grade students who received CT instruction?

The results of the independent samples *t* test indicated no statistically significant difference between the two means, $t(6081.86) = 1.15$, $p = .251$. The sample mean for tenth-grade students who did not receive CT instruction ($M = 238.10$, $SD = 55.23$, $n = 3040$) was not different from the sample mean for tenth-grade students who received CT instruction ($M = 236.51$, $SD = 52.74$, $n = 3061$). The research hypothesis was not supported. There is no difference between the mathematics KAP scores between tenth-grade students who did not receive CT instruction and tenth-grade students who received CT instruction.

RQ2. To what extent is there a difference in the Mathematics KAP scores between tenth-grade students with disabilities who did not receive CT instruction and tenth-grade Students with disabilities who received CT instruction?

H2. There is a significant difference in the Mathematics KAP scores between tenth-grade students with disabilities who did not receive CT instruction and tenth-grade students with disabilities who received CT instruction?

The results of the independent samples *t* test indicated a statistically significant difference between the two means, $t(786.09) = 2.08$, $p = .038$, $d = .15$. The sample mean for tenth-grade students with disabilities who did not receive CT instruction ($M = 235.55$, $SD = 32.52$, $n = 394$) was higher than the sample mean for tenth-grade students with disabilities who received CT instruction ($M = 230.94$, $SD = 30.01$, $n = 402$). The research hypothesis was supported. The Mathematics KAP scores of tenth-grade students with disabilities who did not receive CT instruction was significantly higher than the Mathematics KAP scores of tenth-grade students with disabilities who received CT instruction. The effect size indicated a small effect.

RQ3. To what extent is there a difference in the Mathematics KAP scores between tenth-grade English language learners who did not receive CT instruction and tenth-grade English language learners who received CT instruction?

H3. There is a significant difference in the Mathematics KAP scores between tenth-grade English language learners who did not receive CT instruction and tenth-grade English language learners who received CT instruction?

The results of the independent samples *t* test indicated no statistically significant difference between the two means, $t(1603) = 1.24$, $p = .215$. The sample mean for tenth-grade English language learner students who did not receive CT instruction ($M = 244.93$, $SD = 44.68$, $n = 773$) was not different from the sample mean for tenth-grade English language learner students who received CT instruction ($M = 242.21$, $SD = 43.30$, $n = 832$). The research hypothesis was not supported. The Mathematics KAP scores of tenth-grade English language learner students who did not receive CT instruction was not different from the Mathematics KAP scores of tenth-grade English language learner students who received CT instruction.

RQ4. To what extent is there a difference in the Mathematics KAP scores between tenth-grade students with free- and reduced-lunch status who did not receive CT instruction and tenth-grade students with free- and reduced-lunch status who received CT instruction?

H4. There is a significant difference in the Mathematics KAP scores between tenth-grade students with free- and reduced-lunch status who did not receive CT instruction and tenth-grade students with free- and reduced-lunch status who received CT instruction?

The results of the independent samples *t* test indicated no statistically significant difference between the two means, $t(4531.56) = 1.17$, $p = .241$. The sample mean for tenth-grade students with free- and reduced-lunch status who did not receive CT instruction ($M = 242.47$, $SD = 48.37$, $n = 2323$) was not different from the sample mean for tenth-grade students with free- and reduced-lunch status who received CT instruction ($M = 240.84$, $SD = 44.85$, $n = 2214$). The research hypothesis was not supported. The Mathematics KAP scores of tenth-grade students in free- and reduced-lunch who did not receive CT instruction was not different from the Mathematics KAP scores of tenth-grade students in free-and reduced lunch who received CT instruction.

Summary

Chapter 4 included a presentation of the descriptive statistics and the data analysis and the hypothesis testing for the research questions related to the implementation of the CT curriculum in District X. The results of the current study indicated that the Mathematics KAP scores of all tenth-grade students, English language learner students, and students in free-and reduced lunch who did not receive CT instruction was not different from the Mathematics KAP scores of all tenth-grade students, English language learner students, and students in free-and reduced lunch who received CT instruction. The results also revealed that the Mathematics KAP scores of students with disabilities who did not receive CT instruction was significantly higher than the Mathematics KAP scores of students with disabilities who received CT instruction. The effect size was small. Chapter 5 presents a summary of the current study, finding related to the literature, and the conclusions.

Chapter 5

Interpretation and Recommendations

This research was focused on the impact of the CT curriculum on tenth-grade student achievement and student subgroups, including students with disabilities, English language learner students, and students with free- and reduced-lunch status. This chapter begins with a study summary, which comprises an overview of the problem, purpose statement and research questions, a review of the methodology, major findings, and findings related to the literature. Conclusions from the study's findings are discussed in relation to the CT curriculum's impact on student achievement. Lastly, implications for action and recommendations for further research are presented and discussed

Study Summary

This section presents a summary of the research conducted for this study. The summary begins with an overview of the problem, followed by the purpose statement and the research questions related to the CT curriculum. A review of the methodology is presented, and major findings are discussed.

Overview of the problem. Students' overall KAP performance level scores in District X showed a three-year low performance in mathematics from the 2016-2017 school year to the 2018-2019 school year, requiring the district to take action to increase student performance (KSDE, 2019). To improve student's low mathematics performance District X, a large urban public school district in South Central Kansas, adopted CT Curriculum for middle school and high school students beginning the school year 2018-2019 up to the present. The current study examined the effect of the CT curriculum on student achievement of high school students as measured by the mathematics KAP test.

Purpose statement and research questions. The primary purpose of this study was to investigate the effectiveness of the CT curriculum on mathematics achievement of tenth-grade students in District X as measured by the KAP assessment. The other purpose of this study was to determine CT's impact on the mathematics achievement of students with disabilities, English language learner students, and free- and-reduced students as measured by the KAP assessment. Four research questions were developed to determine whether CT curriculum use affected the change in students' academic achievement as measured by their mathematics KAP test scores.

Review of the methodology. A quasi-experimental quantitative research design was used to examine the difference of the mathematics test scores of tenth-grade students at District X before and after the implementation of the CT as measured by the Mathematics KAP test. The independent variable was the curriculum type (CT or traditional) that tenth-grade students received. The dependent variable was the mathematics KAP performance level scores of tenth-grade students. The tenth-grade student's performance level scores on the Mathematics KAP test in the spring of 2017 before the implementation of CT were compared to the tenth-grade students' Mathematics KAP performance level scores in the spring of 2018, which occurred after the implementation of CT. Demographic variables included in the study to select student subgroups for further comparison were: disability status, English language learner status, and free and reduced lunch status.

Major findings. The results of the current study yielded mixed results. The results of the current study indicated that the Mathematics KAP scores of all tenth-grade students, English language learner students, and students in free-and reduced lunch who did not receive CT instruction were not different from the Mathematics KAP scores of all tenth-grade students, English language learner students, and students in free-and reduced lunch who received CT instruction. The study results revealed that the Mathematics KAP scores of students with disabilities who did not receive CT instruction were significantly higher than the Mathematics KAP scores of students with disabilities who received CT instruction.

Findings Related to the Literature

The findings from this study related to the literature on the impact of the CT curriculum on tenth-grade students' achievement are presented in this section. The differences in student achievement based on student learning disability status, socioeconomic status, and English language learning status are also included in this section. The literature related to each student subgroups is limited as there were only a few research studies conducted on the impact of the CT curriculum on each student subgroup.

The current study results supported the research conducted by Cabalo, Jaciw, and Vu (2007) and Cabalo, Ma, and Jaciw (2007). These studies revealed that the Cognitive Tutor had no significant effect on student achievement. Campuzano et al. (2009) investigated the impact of CT on eighth and ninth-grade students. The results of the current study were similar to those of Campuzano et al. The students' scores on the ETS End-of-Course Algebra test were not different from those who utilized CT Algebra 1. Several researchers conducted research similar to the current study. The results of Shneyderman’s (2001) study indicated that while students who were instructed with CT Algebra 1 demonstrated higher results on the ETS Algebra 1 test, their overall mathematics and algebra performances in 2001 FCAT-NRT were not different from the students who were not instructed with CT Algebra 1. Fontana et al. (2014) reported no significant difference in the Algebra End-of-Course exam between the performance of 7th-grade iPrep Math students who employed CT and non-iPrep Math students but reported higher proficiency rates for iPrep Math students on the FCAT math exam. Pane et al. (2014) examined the effectiveness of CT Algebra 1 on student achievement. The results of the study indicated no significant effect in high schools and middle schools during the first year of implementation but revealed a significant positive effect on high school students in the second year of implementation (Pane et al., 2014).

Some studies included in the literature that were not supported by the current study were conducted by Morgan and Ritter (2002), Wolfson et al. (2008), and Bibi (2010). Morgan and Ritter 2002) examined the effectiveness of the CT Algebra 1 curriculum on ninth-grade students. Data collected from the ETS Algebra end-of-course assessment revealed that students who were taught using the CT curriculum performed better than students taught with a traditional curriculum. Bibi (2010) investigated the impact of Geometry CT on student achievement. The results indicated that the students who used Geometry CT have a higher average score than students who did not. The results of Wolfson et al. (2008) study contrasted with the current study results. The study indicated a positive and statistically significant difference between CT Algebra 1 and comparison groups on the multiple representation test and the Iowa Algebra Aptitude Test.

Academic achievement of students with disabilities was examined. The current study results revealed that students with disabilities who did not receive CT instruction had higher scores in the Mathematics KAP test compared to the students who received CT instruction. The finding contrasted with Sarkis (2004), who indicated that exceptional education students, including those with learning and behavioral difficulties, performed significantly better in the FCAT than the conventional group. Trecek-Schaffer (2019) also found no difference for students with disabilities and without disabilities in Tier 2 and 3 students math intervention program growth from fall to winter and winter to spring. The result of the current study was also in contrast to Steele and Steele (1999), who indicated that an Intelligent Tutoring System called DISCOVER was beneficial for students with learning difficulties.

English language learner status was examined as it relates to academic achievement. The current study results revealed no significant difference between the mathematics KAP scores of English language learner students who received CT instruction and those who did not. The finding was in contrast to Sarkis (2001), who reported that students with limited English proficiency who were instructed with CT performed better in the FCAT test than the conventional group.

The academic achievement of students with free- and reduced-lunch status was also examined. The result of the current study supported Trecek-Schaffer (2019), who revealed no significant difference in students' academic growth in either group. The results of the current study also supported Huang et al. (2016), who indicated no significant interaction in an intelligent tutoring system called ALEKS, as it relates to students' socioeconomic status.

Conclusions

This section includes a conclusion formed by the researcher based on the research findings on the impact of the use of Cognitive Tutor (CT) curriculum on the academic achievement of tenth-grade students and the effects of learning disability status, English language proficiency status, and socioeconomic status to academic achievement. This section includes implications for action and recommendations for future research. Lastly, this section contains the researcher’s concluding remarks.

Implications for action. The findings of this study have far-reaching implications for District X administrators who decide the type of curriculum to adopt for its students. District administrators find evidence of the curriculum's effectiveness to all tenth-grade students and each student subgroups, including students with disabilities, English language learner students, and students with free and reduced status, very useful. For curriculum specialists, this study offers insight into the extent that the CT curriculum impacts student achievement. It also gives administrators a good idea of the type of curriculum that may influence student subgroups. In particular, the findings suggest no significant difference between students' academic achievement before and after the first year of implementation of the CT curriculum. The results addressing research questions one, three, and four revealed these findings. In all three research questions, the CT curriculum's use did not yield a significant difference between the mathematics KAP test scores of all tenth-grade students, English language learner students, and students with free and reduced status. The study also be useful to special education leaders interested in the impact of the CT curriculum on students with disabilities. The study results indicated that students with disabilities performed better using the traditional curriculum than the CT curriculum. Overall, the data indicated that the CT curriculum did not impact student achievement during the first year of implementation. The information in this research can be used by administrators to build future research in evaluating the impact of the CT curriculum during the subsequent years of implementation of the CT curriculum. The current study results can also serve as baseline data to evaluate student achievement growth throughout using the CT curriculum. District administrators should consider the CT curriculum's impact on student achievement before adopting the CT curriculum for long-term use.

Recommendations for future research. The purpose of the current study was to explore the effect that the CT curriculum had upon students' Mathematics KAP test scores after the first year of implementation. The present study examined the Mathematics KAP scores for tenth-grade students, and each student subgroup during the 2017-2018 school year. While the results of the current study provided some information to District X, further investigation should be conducted on the impact of CT on academic achievement. The present study could serve as the pilot study upon which to build future research. The following are recommendations for future research.

1. The results of the study indicated no significant difference between the Mathematics KAP test scores of all tenth-grade students, English language learner students, and students with free and reduced lunch status. The findings, although not significant, have some limitations. One limitation is that the study only utilized the data during the first year of CT curriculum implementation, and the use of MATHia was not mandatory. Future research could be done using a longitudinal study following a similar student

population and using the same methodology during the CT curriculum's continued use.

2. The results of the study revealed that students with disabilities performed significantly better using the traditional mathematics curriculum. Future research on this variable should be conducted using data beyond the initial year of implementation of the CT curriculum.
3. A study could be conducted to determine the impact of CT curriculum on student achievement during the school year when the district has implemented a one-to-one laptop initiative, where all the students are provided a laptop to use in school or out of school using a similar student population and methodology.
4. A study could be conducted in which district teachers are required to adopt Cognitive Tutor's recommendation in a blended implementation to spend three periods per week using the textbook for classroom activities and two periods per week in the computer lab using the MATHia software.
5. A study could be conducted in which district teachers are required to implement the curriculum with fidelity and not utilized non-CT supplementary resources.
6. The current study could be conducted using a mixed-method research design. Teachers' and students' perceptions on the use of the CT curriculum could be collected and analyzed.

Concluding remarks. The findings of this study expanded the work of previous researchers in the area of CT curriculum and student achievement. The current study determined the effects of the CT curriculum on student achievement of all tenth-grade students and learning disability status, English language learning status, and socioeconomic status. The type of curriculum instruction the students receive has a potential influence on increasing academic achievement. Future research on the CT curriculum's effect should be conducted to examine the possible positive effects on student achievement.

References

Antonijevec, R. (2007). *Usage of computers and calculators and students' achievement: Results from TIMSS 2003*. Retrieved from ERIC database. (ED497737)

Anderson, J. (1976). *Language, memory and thought.* Hillsdale, NJ: Erlbaum Associates.

Anderson, J. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, J. R., Boyle, C. F., Corbett, A. T., & Lewis, M. W. (1990). Cognitive modeling and intelligent tutoring. *Artificial Intelligence*, *42*(1) 7-49. doi:10.1016/0004-3702(90)90093-F

Anderson, J., Boyle, C., & Reiser, B. (1985). Intelligent Tutoring Systems. *Science, 228*(4698), 456-462. Retrieved from http://www.jstor.org/stable/1694721

Anderson, J. R., Conrad, F. G., & Corbett, A. T. (1989). Skill acquisition and the LISP Tutor. *Cognitive Science*, *13*, 467-506. Retrieved from http://people.cs.vt.edu/shaffer/cs6604/Papers/Anderson-skill.pdf

Anderson, J. R., & Corbett, A.T. (1985). *Knowledge tracing: Modeling the acquisition of procedural knowledge*. Retrieved from http://act-r.psy.cmu.edu/wordpress/wp-content/uploads/2012/12/893CorbettAnderson1995.pdf

Anderson, J. R., Corbett, A.T., Koedinger, K. R., & Pelletier, R. (1985). Cognitive tutors: Lessons learned: *Journal of Learning Sciences,4*(2), 167-207. Retrieved from https://www.learnlab.org/opportunities/summer/readings/AndersonEtAlJrnlLearningSci95.pdf

Anderson, J. R., & Reiser, B. J. (1985). The Lisp Tutor. *Byte*, *10*(4), 159-175. Retrieved from http://act-r.psy.cmu.edu/wordpress/wp-content/uploads/2012/12/113TheLISPTutor.pdf

Anderson, J. R., & Schunn, C. D. (2000). Implications of the ACT-R learning theory: No magic bullets. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 5). Mahwah, NJ: Erlbaum. Retrieved from http://www.lrdc.pitt.edu/Schunn/research/papers/NoMagicBullets.pdf

Arbain, N., & Shukor, N. A. (2015). The effects of Geogebra on student achievement. *Procedia - Social and Behavioral Sciences*, *172*, 208-214. doi:10.1016/j.sbspro.2015.01.356

Barton, J. M. (2016). *Mathematics intervention utilizing Carnegie Learning's Cognitive Tutor and Compass Learning's Odyssey Math* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 10010883)

Bibi, T. (2010). *Analysis of Cognitive Tutor Geometry curriculum* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 3418899)

Cabalo, J. V., Jaciw, A., & Vu, M. (2007). *Comparative effectiveness of Carnegie Learning Cognitive Tutor Algebra I curriculum: A report of a randomized experiment in Maui School District*. (Research Rep. No. EEI_EdCT-05-FR-Y1-O.2). Palo Alto, CA: Empirical Education Inc. Retrieved from https://www.empiricaleducation.com/pdfs/CTfr.pdf

Cabalo, J. V., Ma, B., & Jaciw, A. (2007). *Comparative effectiveness of Carnegie Learning's Cognitive Tutor Bridge to Algebra curriculum: A report of randomized experiment in the Maui school district* (Research Report No. EEI_EdCT2-06-FR-Y2-O.3). Palo Alto, CA: Empirical Education Inc. Retrieved from https://www.empiricaleducation.com/pdfs/CT2fr.pdf

Campuzano, L., Dynarski, M., Agodini, R., & Rall, K. (2009). *Effectiveness of reading and mathematics software products: Findings from two student cohorts* (NCEE 2009-4041). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from ERIC database. (ED504657)

Carbonell, J. R. (1970) AI in CAI: an artificial intelligence approach to computer-assisted instruction. *IEEE Transactions on Man-Machine Systems, 11*(4) 190–202. doi:10.1109/TMMS.1970.299942

Carnegie Learning. (2011). *MATHia*. Retrieved from https://www.carnegielearning.com/products/software-platform/mathia-learning-software/

Carnegie Learning. (2014). *High school math solution*. Retrieved from https://www.carnegielearning.com/products/blended-solution/high-school-math/

Carnegie Learning. (2017). *Effectiveness research: The RAND study*. Retrieved from http://cdn.carnegielearning.com/assets/page-images/RAND-Study-Flyer.pdf

Carnegie Learning. (2018a). *Advanced math learning software for grades 6-12*. Retrieved from https://www.carnegielearning.com/products/software-platform/mathia-learning-software/

Carnegie Learning. (2018b). *Carnegie Learning: An ESSA evidence-based approach*. Retrieved from https://cdn.carnegielearning.com/assets/page-images/2018-CL-ESSA.pdf

Carnegie Learning Math Solutions. (2018c). *Grades 6-8 summary of alignment and usability*. Retrieved from https://www.edreports.org/math/middle-school-math-solution/index.html

Carnegie Learning. (2020a). *Customer stories*. Retrieved from https://www.carnegielearning.com/why/customer-stories/

Carnegie Learning. (2020b). *MATHiaFlex*. Retrieved from https://discover.carnegielearning.com/MATHiaFlex.html

Carnegie Mellon University. (2015). *Cognitive Tutor authoring tools: Timeline of Cognitive Tutor history*. Retrieved from http://ctat.pact.cs.cmu.edu/index.php?id=timeline

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-as classrooms: A meta-analysis. *Educational Research Review, 9*, 88-113. doi:10.1016/j.edurev.2013.01.0

Cline, E. K. (2017). *The impact of one-to-one iPad technology on student achievement as measured by MAP mathematics assessment* (Doctoral dissertation). Retrieved from http://www.bakeru.edu/images/pdf/SOE/EdD_Theses/ClineEmily.pdf

Cohen, J. (1988). *Statistical power analysis for the social sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Corbett, A. T. (2001). Cognitive computer tutors: Solving the two-sigma problem. *User Modeling: Proceedings of the Eighth International Conference, UM 2001,* 137-147. Retrieved from https://link.springer.com/chapter/10.1007%2F3-540-44566-8_14

Corbett, A. T., Koedinger, K. R., & Anderson, J. R. (1997). Intelligent tutoring systems. In M. Helander, T. K. Landauer, & P. Prabhu (Eds.), *Handbook of human-computer interaction* (2nd ed., pp. 849-874). Amsterdam, Netherlands: Elsevier Science.

Corbett, A. T., Koedinger, K. R. & Hadley, W. S. (2002*).* Cognitive Tutors: From the research classroom to all classrooms. In P. Goodman (Ed.) *Technology enhanced learning: Opportunities for change*. Mahway, NJ: Lawrence Erlbaum Associates.

Creswell, J. W. (2014). *Research Design (4th ed.)*. Thousand Oaks, CA: Sage Publication

Du, J., Havard, B., Sansing, W., & Yu, C. (2004). *The Impact of Technology Use on Low-Income and Minority Students' Academic Achievements: Educational Longitudinal Study of 2002.* Retrieved from ERIC database. (ED485086)

Education Commission of the States. (2014). *What methods are used to identify English language learners?*. Retrieved from http://ecs.force.com/mbdata/mbquestNB2?rep=ELL1403

Fancsali, S. E., Ritter, S. Yudelson, M., Sandbothe, M., & Berman, S. (2016). *Implementation factors and outcomes for intelligent tutoring systems: A case study of time and efficiency with Cognitive Tutor Algebra*. Florida Artificial Intelligence Research Society Conference**,** North America, Retrieved from https://www.aaai.org/ocs/index.php/FLAIRS/FLAIRS16/paper/view/12965/12614

Fancsali, S. E., Zheng, G., Tan, Y., Ritter, S., Berman, S. R., Galyardt, A. (2018). *Using embedded formative assessment to predict state summative test scores*. Proceedings of the International Conference on Learning Analytics and Knowledge (pp. 161-170): ACM, New York. doi:10.1145/3170358.3170392

Fletcher, J. D. (1985). Intelligent instructional systems in training. In S. A. Andriole (Ed.), *Applications in artificial intelligence* (pp. 427-451). Princeton, NJ: Petrocelli Books. Retrieved from https://www.researchgate.net/publication/291714377_INTELLIGENT_INSTRUCTIONAL_SYSTEMS_IN_TRAINING#fullTextFileContent

Fontana, L., Beckerman, A., Levitt, J., & Levitt, M. (2014). *iPrep Math*. *August 2014 report*. Retrieved from http://pdfs.dadeschools.net/portal/August2014Report.pdf

Food Research and Action Center. (2020). *School meal eligibility and reimbursements*. Retrieved from https://frac.org/school-meal-eligibility-reimbursements#:~:text=The%20child's%20school%20then%20compares,are%20eligible%20for%20free%20meals.

Graesser, A. C., Conley, M. W., & Olney, A. M. (2012). Intelligent tutoring systems. In S. Graham & K. Harris (Eds.), *APA educational psychology handbook: Volume 3. Applications to learning and teaching* (pp. 451–473). Washington, DC: American Psychological Association.

Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computer & Education, 153,* 1-25. doi:10.1016/j.compedu.2020.103897

Huang, X., Craig, S. D., Xie, J., Graesser, A., & Hu, X. (2016) Intelligent tutoring systems work as a math gap reducer in 6th grade after-school program. *Learning and Individual Differences*. Retrieved from https://www.researchgate.net/profile/Xudong_Huang8/publication/294289790_Intelligent_tutoring_systems_work_as_a_math_gap_reducer_in_6th_grade_after-school_program/links/5a9848ce45851535bce0be76/Intelligent-tutoring-systems-work-as-a-math-gap-reducer-in-6th-grade-after-school-program.pdf

University of Kansas Achievement & Assessment Institute (2017). *Kansas assessment program technical manual 2016.* Retrieved from https://www.ksassessments.org/sites/default/files/documents/KAP_Technical_Manual_2016.pdf

University of Kansas Achievement & Assessment Institute (2017). *Kansas Assessment Program Technical Manual 2017*. Retrieved from https://ksassessments.org/sites/default/files/documents/KAP_Technical_Manual_2017.pdf

Center for Educational Testing & Evaluation (2018). *Cut scores for KAP summative assessments*. Retrieved from https://ksassessments.org/sites/default/files/ documents/KAP_summative_cut_scores.pdf

Kansas Assessment Program. (2020). *General education assessment*. Retrieved from https://ksassessments.org/sites/default/files/documents/scorereports/Scale_Scores_Performance_Level_Descriptors_Cut_Scores.pdf

Kansas Can. (2017). *Kansas report card*: 2016-2017. Retrieved from https://ksreportcard.ksde.org/assessment_results.aspx?org_no=D0259&rptType=2

Kansas Department of Education. (2018). *Career, standards and assessment services*. Retrieved from https://www.ksde.org/Portals/0/CSAS/Content%20Area%20(A-E)/English_Language_Proficiency/Assessments/KSDE%20Approved%20English%20Language%20Proficiency%20Assessments%20rev082918.pdf?ver=2018-08-29-152129-963

Kansas Department of Education. (2020). *English language proficiency assessment placement (screener: Kelpa-P scoring guidelines)*. Retrieved from https://www.ksde.org/Agency/Division-of-Learning-Services/Career-Standards-and-Assessment-Services/Content-Area-A-E/English-Learners/Assessment/KELPA-P-Screener

Koedinger, K. R., & Anderson, J. R. (1993). *Effective use of intelligent software in high school math classroom*. Retrieved from http://pact.cs.cmu.edu/koedinger/pubs/Koedinger%20&%20Anderson%2093.pdf

Koedinger, K. R., Anderson, J. R. Hadley, W. H., & Mark, M. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, *8*(1), 30-43. Retrieved from http://act-r.psy.cmu.edu/wordpress/wp-content/uploads/2012/12/232jaied97.pdf_2.pdf

Koedinger, K. R., & Corbett, A. (2006). *Cognitive tutors: Technology bringing learning science to the classroom*. Retrieved from https://www.researchgate.net/publication/255519934_Cognitive_Tutors_Technology_Bringing_Learning_Science_to_the_Classroom#fullTextFileContent

Koedinger, K. R., & Sueker, E. L. F. (1996). *Monitored design of an effective learning environment for algebraic problem solving*. Retrieved from http://pact.cs.cmu.edu/pubs/Koedinger,%20Sueker%20.pdf

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

Lafferty, I. O. (2006). Carnegie learning bridge to algebra. *Technology & Learning, 27*(1), 15. Retrieved from https://search-proquest-com.www2.lib.ku.edu/docview/212113600?accountid=14556

Lei, J. (2007). Quantity versus quality: A new approach to examine the relationship between technology use and student outcomes. *British Journal of Educational Technology*, *41*(3), 455-472. doi:10.1111/j.1467-8535.2009.00961.x

Lewis, R. E. (2010). *The effectiveness of computer-assisted instruction on student math achievement* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 3441943)

McEwen, T. O. (2008*). A study of differentiated mathematics curricula in Kansas and their effect on student achievement as demonstrated by the Kansas state mathematics assessments* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 3297624)

McFarland, J., Hussar, B., Wang, X., Zhang, J., Wang, K., Rathbun, A., …& Bullock Mann, F. (2018). *The condition of education 2018* (NCES 2018-144). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved from https://nces.ed.gov/pubs2018/2018144.pdf

Morgan, P., & Ritter, S. (2002). *An experimental study of the effects of Cognitive Tutor Algebra I on student knowledge and attitude*. Retrieved from https://carnegielearning.com/static/web_docs/morgan_ritter_2002.pdf

National Assessment of Educational Progress. (2015). *NAEP report card: 2015 NAEP mathematics & reading at grade 12*. Retrieved from https://www.nationsreportcard.gov/reading_math_g12_2015/#mathematics/scores

National Assessment of Educational Progress. (2017). *The nation's report card: NAEP mathematics and reading assessment highlighted results at grades four and eighth for the nations, states, and districts*. Retrieved from https://www.nationsreportcard.gov/reading_math_2017_highlights/

National Assessment of Educational Progress. (2019). *The nation's report card: NAEP mathematics and reading assessment highlighted results at grades four and eighth for the nations, states, and districts*. Retrieved from https://www.nationsreportcard.gov/highlights/mathematics/2019/

Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2014). Effectiveness of Cognitive Tutor Algebra I at scale. *Educational Evaluation and Policy Analysis, 36*(2), 127–144. doi:103102/0162373713507480

Pane, J. F., McCaffrey, D. F., Slaughter, M. E., Steele, J. L., & Ikemoto, G. S. (2010). An experiment to evaluate the efficacy of Cognitive Tutor Geometry. *Journal of Research on Educational Effectiveness, 3*(3), 254–281. doi:10.1080/19345741003681189

Pellerito, F. L. (2011). *The effects of technology enriched mathematics instruction on at-risk secondary school students* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 3458228)

Program for International Student Assessment (2015). *Results in Focus*. Retrieved from https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf

Preston, A. (2016). *Effects of Singapore model method with explicit instruction on math problem solving skills of students at risk for or identified with learning disabilities*, ProQuest Dissertations and Theses database. (UMI No. 10139928)

Provasnik, S., Malley, L., Stephens, M., Landeros, K., Perkins, R., and Tang, J. H. (2016). *Highlights from TIMSS and TIMSS Advanced 2015: Mathematics and science achievement of U.S. students in grades 4 and 8 and in advanced courses at the end of high school in an international context* (NCES 2017-002). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved from https://nces.ed.gov/pubs2017/2017002.pdf

Reynolds, A. (2013). *Does a student's use of technology outside school affect mathematic achievement in school?* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3569398)

Ritter, S., Joshi, A., Fancsali, S. E., & Nixon, T. (2013). Predicting standardized test scores from Cognitive Tutor interactions. Proceedings of the 6th International Conference on Educational Data Mining. *International Educational Data Mining Society*. Retrieved from https://www.educationaldatamining.org/EDM2013/papers/rn_paper_25.pdf

Sales, A., & Pane, J. F. (2015). Exploring causal mechanisms in a randomized effectiveness trial of the Cognitive Tutor. *Proceedings of the 8th International Conference on Educational Data Mining. International Educational Data Mining Society*. Retrieved from https://pdfs.semanticscholar.org/b25b/627b1ad5d940a94497d81a83f79704a9953f.pdf

Sarkis, H. (2004). *Cognitive Tutor algebra 1 program evaluation: Miami-Dade Public Schools*. Retrieved from http://relgroup.net/articles/cognitive_evaluation_final.pdf

Shneyderman, A. (2001). *Evaluation of the Cognitive Tutor Algebra I program*. Miami, FL: Miami–Dade County Public Schools, Office of Evaluation and Research. Retrieved from http://oer.dadeschools.net/algebra.pdf

Sleeman, D., & Brown, J. S. (1982). *Intelligent tutoring systems*. New York: Academic Press.

Steele, M. M., & Steele, J. W. (1999). Discover: An intelligent tutoring system for teaching students with learning difficulties to solve word problems. *Journal of Computers in Mathematics and Science Teaching, 18*(4), 351-359. Retrieved from https://www.learntechlib.org/primary/p/9000/

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. doi:10.1037/a0032447

Trecek-Schaffer, C. L. (2019*). The impact of intervention software on student achievement in mathematics* (Doctoral dissertation). Retrieved from http://www.bakeru.edu/images/pdf/SOE/EdD_Theses/Schaffer_Ceresa.pdf

Trends in International Mathematics and Science Study. (2015). *International Association for the Evaluation of Educational Achievement*. Retrieved from https://nces.ed.gov/timss/

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2009). *Cognitive Tutor Algebra 1*. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_cogtutor_072809.pdf

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2016). *Secondary mathematics intervention report: Cognitive Tutor®*. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_cognitivetutor_062116.pdf

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2020). *Trends in International Mathematics and Science Study*. Retrieved from https://nces.ed.gov/timss/

U.S. Department of Education, Office of Special Education. (2020). *Individuals With Disabilities Education Act, Section 1401*. Retrieved from https://sites.ed.gov/idea/statute-chapter-33/subchapter-i/1401

U.S. Department of Education, Office of English Language Acquisition. (2016). *English learner tool kit* (Rev. ed.). Washington, DC. Retrieved from https://www2.ed.gov/about/offices/list/oela/english-learner-toolkit/eltoolkit.pdf

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

Wenglinsky, H. (2006). Technology and achievement: The bottom line. *Educational Leadership, 63*(4), 29-32. Retrieved from https://imoberg.com/files/Technology_and_Achievement_--_The_Bottom_Line_Wenglinsky_H._.pdf

Wolfson, M., Koedinger, K., Ritter, S., & McGuire, C. (2008). *Cognitive Tutor Algebra I: Evaluation of results (1993-94)*. Pittsburgh, PA: Carnegie Learning, Inc. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_cogtutor_072809.pdf

Wright, P., & Wright, P (2019). *The child find mandate: What does it mean to you?.* Wrightslaw: Special Education Law. Retrieved from https://www.wrightslaw.com/info/child.find.mandate.htm#:~:text=The%20Individuals%20with%20Disabilities%20Education%20Act%20includes%20the%20Child%20Find%20mandate.&text=The%20IDEA%20requires%20all%20States,and%20which%20children%20are%20not.

Zulnaidi, H., Oktavika, E., & Hidayat, R. (2020). Effect of use of GeoGebra on achievement of high school mathematics students. *Education and Information Technologies*, *25*(1), 51–72. doi:10.1007/s10639-019-09899-y

Appendices

Appendix A: District Approval for Research

[redacted]

September 9, 2020

To: Theresa Basarde
[redacted]

Re: Dissertation Proposal

Dear Theresa Basarde,

This letter is in response to your recent request regarding your research titled: *Effectiveness of Cognitive Tutor (Carnegie Learning's Blended Math Curriculum) on mathematics achievement of high school students* in the [redacted] Schools. The Research Council has conditionally approved your request due to the following:

- In 2017-18, the district had just adopted Carnegie Math, the use of Mathia was not mandatory due to the lack of devices available in schools. The Research Council strongly encourages the scope of the study to expand beyond the initial year and possibly compare the initial year to the current year that includes remote learning. This would allow a robust study that could include implementation research.
- Another consideration is to focus on grade 8 rather than grade 10 where there is greater use and different attitudes toward the state assessment.
- 2017-18 data in a shareable format has been requested from the vendor, however a timeframe for receiving the data is unknown.
- Lastly, there are existing studies in this area that could lend insight. Some studies are provided in the email with this response letter.

If you elect to modify what was originally presented in your proposal, the Research Council will review your proposal upon re-submission.

On behalf of the [redacted] Research Council,

[redacted]

Appendix B: Baker University Approval to Conduct Research

BAKER UNIVERSITY
Own Confidence

Baker University Institutional Review Board

September 22nd, 2020

Dear Theresa Basadre and James Robins,

The Baker University IRB has reviewed your project application and approved this project under Exempt Status Review. As described, the project complies with all the requirements and policies established by the University for protection of human subjects in research. Unless renewed, approval lapses one year after approval date.

Please be aware of the following:

1. Any significant change in the research protocol as described should be reviewed by this Committee prior to altering the project.
2. Notify the IRB about any new investigators not named in original application.
3. When signed consent documents are required, the primary investigator must retain the signed consent documents of the research activity.
4. If this is a funded project, keep a copy of this approval letter with your proposal/grant file.
5. If the results of the research are used to prepare papers for publication or oral presentation at professional conferences, manuscripts or abstracts are requested for IRB as part of the project record.
6. If this project is not completed within a year, you must renew IRB approval.

If you have any questions, please contact me at npoell@bakeru.edu or 785.594.4582.

Sincerely,

Nathan Poell, MLS
Chair, Baker University IRB

Baker University IRB Committee
Sara Crump, PhD
Nick Harris
Christa Manton, PhD
Susan Rogers, PhD