IMPLEMENTATION AND EFFECTIVENESS OF SUCCESSMAKER MATHEMATICS FOR MIDDLE SCHOOL STUDENTS WITH DISABILITIES

A Dissertation

by

STEVEN KYLE MCKISSICK

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Robert M. Capraro
Co-Chair of Committee,	Mary Margaret Capraro
Committee Members,	Bugrahan Yalvac
	Oi-Man Kwok
Head of Department,	Lynn Burlbaw

May 2016

Major Subject: Curriculum and Instruction

ABSTRACT

Students with disabilities often display patterns of mathematics achievement significantly below grade level, and interventions are needed to address this gap. SuccessMaker Mathematics is an instructional learning system with a lengthy research base, though recent versions of the program and its effectiveness for students with disabilities have received little attention. Quantitative analyses using multivariate analysis of variance and multilevel modeling were employed. Qualitative investigation followed an instrumental case study approach.

Usage at or below vendor recommended levels was not found to impact the mathematics achievement gap for middle school students with disabilities. Longitudinal usage of the program over two years was not found to impact the mathematics achievement gap. Usage at levels greater than vendor recommendations was found to positively impact the mathematics achievement gap for middle school students with disabilities. Interviews conducted with teachers and campus administrators responsible for program implementation identified generally positive perceptions of the program as well as a variety of obstacles to implementation. Recommendations for improving program implementation are provided.

ACKNOWLEDGEMENTS

I would like to thank my committee chairs, Dr. and Dr. Capraro, as well as my committee members, Dr. Yalvac and Dr. Kwok, for their guidance and support throughout the course of this research. Drs. Capraro, thank you allowing me the opportunity to commute weekly for this program. More importantly, thank you for opening your home so your students come to know you for the wonderful and kind folks you truly are.

Thanks to my wife, Margaret, for her support and patience throughout this program. You have pushed me to achieve, and you have tolerated the demands it has placed on our family. Thank you as well, Zuleyka, for pushing a father to continue his education.

Thanks to my mother and father for their encouragement throughout my entire educational process. Thanks for the example you set as believers in the intrinsic value of education. Thank you for being my sounding board throughout this program.

I am appreciative of the leadership and support of Dr. Lynn Young, who graciously allowed me the flexibility at work to pursue this program. Without your support, I would have been unable to take advantage of this opportunity.

Thanks also go to my fellow students in Math Education. You have fostered and maintained a sense of community and friendship that continues beyond the classroom. I appreciate being included in that fellowship despite my age and distance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Statement of the Problem

Federal and state accountability systems, including No Child Left Behind and Race to the Top, have included passing standards in mathematics by which schools are expected to demonstrate adequate yearly progress. The accountability systems in Texas, such as the Texas Academic Performance Report (TAPR) and Performance-Based Monitoring and Assessment System (PBMAS), include standards for each significant demographic present at the school, including students with disabilities. These students have historically underperformed relative to their non-disabled peers, and it has become increasingly necessary for schools to provide needed supports for these students to meet national and state standards (Manning, 2004). Previous research suggests that students with disabilities may only make modest achievement gains during middle school (Graham, Bellert, Thomas, & Pegg, 2007), leading to an "ever-widening achievement gap with the passage of time" (Judge & Watson, 2011, p. 154). Pressures for students with disabilities to succeed is growing (Martindale, Pearson, Curda, & Pilcher, 2005), and the need to identify effective interventions for these students is great.

SuccessMaker Mathematics (SMM) provides one possible remedy for the disability achievement gap identified above. Foundations for the program trace back to the 1960s (Suppes & Morningstar, 1969), though significant research was done throughout the 1970s and 1980s to enable meta-analytic reviews to be conducted (Becker, 1992; Kulik, 1994). Few of the studies were subjected to peer review; most of the available research has consisted of dissertations and institutional reports as discussed in a recent review of the literature (McKissick, 2014). Advances in technology have allowed the program to evolve, changing the underlying theoretical foundations on which the program was built. Few recent studies have been conducted, leaving the effectiveness of recent program iterations unknown. Of nine studies conducted since 2001, none were published in a peer-reviewed journal. The technology-based format and use of behaviorist principles through programmed instruction should facilitate learning for students with disabilities by providing opportunities for developing automaticity and overlearning of concepts (Vockell & Mihail, 1993). However, the lack of recent peer-reviewed research leaves this claim with little support.

Purpose of the Study

The primary purpose of this study is to determine the effectiveness of SMM for improving the mathematics achievement of students with disabilities. A Central Texas school district purchased licenses allowing students with disabilities at all middle schools to use the program. The financial cost of these licenses, as well as accompanying upgrades in technology and infrastructure, were significant. The district also paid for frequent on-site vendor support during the first four years of program implementation. Recommended usage levels were identified by the on-site vendor representatives, somewhat consistent with vendor publications for recommended usage. Whereas vendor publications (Pearson, 2012) include a matrix based on initial achievement levels and anticipated growth, the on-site recommendation provided a condensed target for both daily and cumulative use. Vendor feedback consistently identified individual students who were making “large” achievement gains after extended program use. This research project will consider the program’s impact for all students who have received the intervention to determine what achievement gains have been effected. The first article compares groups of middle school students receiving varying program usage. In this article, a quantitative approach is taken regarding the question of SMM effectiveness. In the second article, a qualitative approach is taken as campus teachers and administrators discuss their perceptions regarding the program and resultant student success. A quantitative approach is resumed in the third article as multilevel modeling is used to identify the impact of multiple years of program use at or above recommended levels. At present, the What Works Clearinghouse has found insufficient research for review to determine the effectiveness of SMM. It is hoped that the research presented here provides a platform for determining the effectiveness of SMM for students with disabilities. The results herein will provide the school district with sufficient data to evaluate its current and ongoing use of the program.

Literature Review

For over thirty years, educators have utilized a class of technological interventions known as integrated learning systems (ILS). These systems share several characteristics that separate them from other instructional technologies: (a) target specific instructional objectives and connect these to specific lessons; (b) have the potential for integration into other curricula; (c) span multiple grade levels in one or more content areas; (d) utilize a networked system of computers; and (e) collect and maintain records of student performance (Bailey, 1992). A summarized description is provided by Kulik (2002), who noted that an ILS is a "software program that provides tutorial instruction at several grade levels and keeps extensive records of student progress on networked computer systems" (p. 1). Integration of instructional technology into the curriculum is encouraged by the National Council of Teachers of Mathematics (2000), though ILS use is frequently seen as supplementing existing curricula.

Recommendations for ILS use have not always been followed. The view of ILS as supplemental, stated above, has led to program usage at approximately 15-30% of recommended usage levels (van Dusen & Worthen, 1995). A matrix to evaluate implementation of SMM, a specific ILS, contrasted this "unacceptable use" with "ideal use" wherein the program is used "as a tool for regularly accomplishing classroom instructional objectives" (Mills & Ragan, 2000, p. 28). The potential for disparate usage of an ILS led Slavin (1987) to note that measurement of the effectiveness of an ILS should consider the amount of time actually spent using the program.

The history of SMM is important to trace as locating relevant work can be hindered by the program's multiple names and owners. The program is rooted in the work of Suppes and Zancotti at Stanford University in the late 1960s (Kulik, 1994; Wood, 2004). From their work grew the Computer Curriculum Corporation and, ultimately, the SMM program. This company was purchased by Simon & Schuster in 1990 and sold to NCS Learn, a division of the NCS Pearson family, in 1997 (Manning, 2004). Use of the product has identified it as Stanford-CCC, SuccessMaker, SuccessMaker Enterprise, or even by only a portion of the product such as Math Concepts and Skills (Manning, 2004). Though SuccessMaker has an available reading component, the present research will focus solely on the mathematics component of the program.

The experience of using SMM involves assessment followed by prescriptive use. Students using SMM undergo an “initial placement” to determine their entry achievement level. This process may take up to three hours (Pearson, 2012) or approximately 300 questions (Wood, 2004). Students are presented with questions of increasing or decreasing difficulty depending on the accuracy of their responses. The program uses a branching algorithm to work through various skill strands and difficulty levels (Svoboda, Jones, van Vulpen, & Harrington, 2012). Upon completion of initial placement, recommended usage of the program includes sessions of approximately 15-20 minutes in length (Pearson, 2013). Students may work on skills identified by the program at or near their measured ability level, or they may be assigned specific skills by the teacher. SuccessMaker Mathematics maintains a record of program usage statistics as well as student capabilities, allowing the teacher to generate current records of student use and progress. The program conducts a regular review of previously mastered skills to ensure continued understanding (Wood, 2004).

SuccessMaker Mathematics utilizes a multimedia environment to facilitate student learning and interaction. Students are provided with audio and visual materials as concepts are presented and explored. Students have access to virtual tools such as a highlighter and sticky notes that keep students active during their learning (Pearson, 2013), though no research was found that investigated the effectiveness of these tools. SMM provides immediate feedback for student responses, and a “cognitive coach who offers hints and insights" (p. 6) is provided when students answer incorrectly. Multimedia environments such as this have been found to improve student understanding of instruction (Cognition and Technology Group at Vanderbilt, 1990).

The Effectiveness of ILS for Students with Disabilities

Mathematics has been identified as a subject with which students with disabilities struggle (Fuchs, Fuchs, Hamlett, & Appleton, 2002). The addition of computer resources, including internet-based tools, have been found to be effective in raising scores on standardized tests for learning disabled students (Funkhouser, 2003; Lin, 2006). Recommendations for selection and implementation of technologies for use with students with disabilities have included targeted intervention, integration with existing curricula, and consistent monitoring of performance outcomes (King-Sears & Evmanova, 2007). Becker (1992) noted that positive outcomes may result from computer-based instruction, though these gains were likely to occur for only the highest- and lowest-performing students. Students in the middle of the achievement spectrum may not realize the same benefits as those at the extremes.

The use of repetitive practice and immediate feedback are behaviorist strategies common to many ILS programs. These strategies are based on principles of operant conditioning, and they have shown success for students with learning disabilities (Zafiropoulou & Karmba-Schina, 2005). There is reason to believe that integrating these strategies into computer-based interventions should also be effective for learning-disabled students (Burton, Morre, & Maglaiare, 2008). Recommendations have also been made for the use of step-by-step modeling of problem-solving, frequent progress monitoring, and the use of work sessions of increased frequency yet reduced intensity for students with disabilities (Cooley, 2007). Drill and practice methods are recommended as tools towards the development of automaticity of basic skills (Cummings & Elkins, 1999; Pellegrino & Goldman, 1987). "Those who lack automaticity at the basic skills level exhaust their cognitive resources trying to recall math facts and, therefore, have few resources left for solving problems" (Wendling & Mather, 2009, p. 173). SMM incorporates these behaviorist principles, though more recent versions of the program have incorporated concepts from other theoretical perspectives. The result is an eclectic conglomeration of strategies with a behaviorist foundation.

Previous Research Findings

Several meta-analyses have been conducted to determine the effectiveness of ILS programs and implementations. These meta-analytic studies have strived to identify consistent findings across studies and presented these synthesized findings in terms of an effect size. Various standards have been proposed for interpreting effect sizes. Cohen (1988) proposed that an effect size of d = .20 might signal a small effect, while Slavin and Fashola (1998) proposed that an effect size of d = .25 could be considered educationally meaningful. A review of five meta-analyses involving the use of computer-based instruction found Cohen's d effect sizes ranging from .25 to .48 (Lowe, 2002). A review by Kulik (2003) regarding the use of ILS for mathematics found Cohen's d effect sizes ranging from .14 to 1.05. He noted that the use of ILS had an "almost uniformly positive record of effectiveness in the 1970s, 1980s, and 1990s" (p.

x). Increases in student achievement, engagement, and motivation have also been identified when technology is integrated into the curriculum (Funkhouser, 2003; Lin, 2006). Student attitudes toward classes improve when technology is used, and students learn more in less time with computer-based instruction (Kulik, 1994).

A review of the findings for SMM as a specific ILS was not as encouraging (McKissick, 2014). Studies were identified from previous meta-analyses, and studies conducted after these meta-analyses were also reviewed. Most of the studies used by Becker (1992) and Kulik (1994) were at least 30 years old. Regional issues may have affected outcomes as previous studies have clustered in the United States' deep South, west coast, and north-central states. Methodological flaws were identified by Slavin and Lake (2008) that prevented inclusion in their meta-analytic review. Few true experiments have been conducted utilizing random assignment. Several lacked adequate control groups.

CHAPTER II

EFFECTIVENESS OF PEARSON'S SUCCESSMAKER MATHEMATICS FOR STUDENTS WITH DISABILITIES

Overview

SuccessMaker mathematics is an instructional learning system rooted in behaviorist instructional theory. Previous research efforts have left much to be desired and have produced inconsistent results. Recent research for this program appears to be tapering off, despite advances in technology signaling integration of concepts from other theoretical positions. A quasi-experimental review of data from a sample of students (*N* = 1186) from a central Texas school district over a five-year period was conducted. Multivariate analysis of variance identified that changes in state testing performance were not linked to program use. Changes in the rate of academic achievement were found to exist between usage groups. Students who met or exceeded usage recommendations (>20 hours of use) were found to have significantly greater rates of achievement (ES: *d* = 1.02). Recommendations for further studies and limitations of the current study are provided.

Introduction

Educators and researchers have spent more than thirty years investigating a class of technological interventions known as instructional learning systems (ILS). An ILS has been described as a "software program that provides tutorial instruction at several grade levels and keeps extensive records of student progress on networked computer systems" (Kulik, 2002, p. 1). Bailey (1992) expanded this description by identifying five key characteristics that separate an ILS from other instructional technology: (a) ability to target specific instructional objectives and connect these to specific lessons; (b) potential for integration into other curricula; (c) span multiple grade levels, possibly in multiple content areas; (d) the use of a networked computers; and (e) collection of student performance records. Though the National Council of Teachers of Mathematics (2000) has emphasized the inclusion of instructional technology in classrooms, the implementation and use of an ILS is more involved than the use of calculators or interactive smartboards. Various ILS technologies have been reviewed to include products developed by Wicat Systems and Jostens Learning Corporation, as well as programs such as Plato, Prescription Learning, and SuccessMaker (Becker, 1992).

Because ILS use is frequently treated as a supplemental curriculum, recommendations for ILS use have not always been followed. A number of ILS programs come with recommendations for minimum student usage (Gee, 2008; Manning, 2004). Failure to integrate the ILS with existing classroom curriculum has resulted in ILS usage of about 15-30% of program recommendations (van Dusen & Worthen, 1995). A matrix to evaluate technology implementations contrasted this "unacceptable use" with "ideal use" wherein the ILS is used "as a tool for regularly accomplishing classroom instructional objectives" (Mills & Ragan, 2000, p. 28). Because of such variation, Slavin (1987) urged that time spent using the program be a factor in determining the effectiveness of an ILS.

SuccessMaker is an ILS for which a historical review may be necessary to identify relevant research. The program is rooted in the work of Suppes and Zancotti at

Stanford University in the late 1960s (Kulik, 1994; Wood, 2004). Out of their work came the Computer Curriculum Corporation (CCC) and, ultimately, this program. The company was purchased by Simon & Schuster in 1990 (Manning, 2004). Pearson acquired Simon & Schuster and its holdings, including SuccessMaker, in 1998 (Pearson Digital Learning, n.d.). Previous research with the program has identified it as Stanford-CCC, SuccessMaker, SuccessMaker Enterprise, or even by a portion of the product such as Math Concepts and Skills (Manning, 2004). Given the changes in ownership and name, it is doubted that all previous relevant studies were identified in previous ILS meta-analyses.

A discussion about the nature of SuccessMaker Mathematics (SMM) is helpful for identifying an underlying theoretical framework. Students begin their use of SMM with an initial placement assessment designed to identify grade level skills. This process may take up to three hours (Pearson Digital Learning, 2012) or approximately 300 questions (Wood, 2004). Students may begin this initial placement at either their enrolled grade level or a level determined by the teacher managing the student's use of the program. Students are presented with questions that increase or decrease in difficulty depending on the accuracy of student responses. A branching algorithm is used to work through various skill strands and grade levels (Svoboda, Jones, van Vulpen, & Harrington, 2012). Students may work on skills at their ability level, in 15 strands of content (Pearson Digital Learning, 2004), with difficulty contingent on student success. Additionally, teachers may assign specific skill units to students instead of having students work only on grade-level skills. SMM, as anticipated by Bailey's (1992) description of an ILS, maintains an ongoing record of student skill capabilities and program usage, allowing the teacher to produce up-to-date records of student use and progress when needed. SMM also incorporates a regular review of previously mastered skills into student work to ensure continued understanding (Wood, 2004).

SMM is an interactive program within a multimedia environment. Students are provided with audio and video material regarding a particular concept or skill. Students have access to virtual tools such as a highlighter and sticky notes to keep students active during learning (Pearson Education, 2013). No research studies were found that examined these particular tools for effectiveness. SMM provides immediate feedback for student responses. A "cognitive coach who offers hints and insights" (p. 6) is provided when a student answers incorrectly. This use of a multimedia environment for learning has been found to improve student comprehension during instruction (Bransford, Sherwood, Hasselbring, Kinzer, & Williams, 1990).

Theoretical Considerations

The behaviorist definition of learning is the acquisition of a new behavior. A person learns what is practiced, and learning prepares the student to demonstrate "specific responses to particular stimuli rather than general responses to vague stimuli" (Schiro, 2013, p. 63). The learner is considered an active participant in the learning process, and exhibition of learned behaviors is necessary for continued learning (Ormrod, 2014). Shaping occurs as increasingly complex or difficult behaviors are presented to the learner. Schiro (2013) noted that even the most complex tasks are considered by behaviorists as compositions of discrete simpler skills that can be taught.

Immediate feedback is necessary, and technology increases that immediacy. Learning is self-paced; not all learners will acquire the same skill at the same speed or in the same number of discrete trials.

SMM has its foundations in behaviorism through programmed instruction. Programmed instruction, as developed by Skinner (1986), is a specific application of behaviorist principles built on the early work of Thorndike and Pressey. Material to be learned should be presented in small increments to reduce the likelihood of error. Material is arranged by complexity, and learners enter at the highest level at which they can demonstrate mastery (Svoboda et al., 2012). The learner is presented with a question in response to some stimulus, and the teacher (or, for SMM, the program) awaits a response. The student is provided differential feedback based on the response. Failure to respond correctly in SMM may result in continued exposure to the same skill with additional support from the "cognitive coach" or a change in skill or skill level following multiple failures suggesting frustration. Students experiencing consistent success may experience an increase in the grade level of skills presented through a process known as branching (Joyce, Weil, & Calhoun, 2009). The present level of student ability is identified as the skill level where the student's performance plateaus, and instruction is provided at that level.

Programmed instruction has changed significantly as technology has changed. The rise and fall in favor with programmed instruction has been directly linked to these technological changes (Svoboda et al., 2012). In early years, programmed instruction led to an over-reliance on technology which, coupled with a limited range of stimulating media, led to student boredom (McDonald, Yanchar, & Osguthorpe, 2005). Rigid application of the principles of programmed instruction identified above has relaxed in later years (McDonald et al., 2005), and later programs and versions have been more interactive and student-directed (Cruthirds & Hanna, 1997). Current iterations of SMM have retained core principles of programmed instruction – success-driven increases in complexity, immediate feedback, and active participation – while sprinkling in tools more consistent with cognitive and constructivist frameworks.

Programmed instruction works, though research findings are inconsistent. Early meta-analytic research found that programmed instruction yielded an effect size of just over $d = .20$ (Kulik, Kulik, & Cohen, 1980), at the low end of Cohen's (1988) bracket for a small effect. Two years later, another meta-analysis determined that programmed instruction was no better than traditional instruction (Kulik, Schwalb, & Kulik, 1982), with an effect size for mathematics of $d = -.01$. Another early estimate of the effectiveness of computer-aided instruction, to include systems utilizing programmed instruction, yielded an effect size of $d = .57$ (Schmidt, Weinstein, Niemiec, & Walberg, 1985). Ormrod (2014) contends that programmed instruction remains viable for students with little previous success, including students with learning or behavior difficulties, as well as those for whom previous attempts at teaching and learning have proven unsuccessful. Behaviorist principles are well-established, though their application may be time-intensive and less than enjoyable.

Behaviorist strategies have demonstrated success with learning-disabled students (Zafiropoulou & Karmba-Schina, 2005). The reason may be attributed to the ability of computer-based interventions, such as SMM, to provide immediate feedback (Burton, Moore, & Magliare, 2008). Cooley (2007) proposed that students with mathematics disabilities be provided with step-by-step modeling of solving problems, frequent monitoring of progress, and the use of work sessions that are more frequent but less intense. Drill-and-practice models have been recommended (Pellegrino & Goldman, 1987) as a step towards building automaticity of skills (Cummings & Elkins, 1999). "Those who lack automaticity at the basic skills level exhaust their cognitive resources trying to recall math facts and, therefore, have few resources left for solving problems" (Wendling & Mather, 2009, p. 173). SuccessMaker Mathematics incorporates these recommendations and behaviorist principles, and it is anticipated that its use with students with learning and behavior disabilities should prove effective in increasing achievement levels.

Constructivist principles may also be seen in more recent iterations of SMM. By providing incremental increases in skills under review, SMM incorporates a mechanical version of Vygotsky's (1978) zone of proximal development. According to Vygotsky, students learn best when challenged with skills at or slightly above their current ability level. By reinforcing previously learned skills, SMM also provides instructional scaffolds on an individual basis. Though the interpersonal contact and communication are absent from a true sociocultural position, communication via the cognitive coach and program use facilitated by the teacher may serve as surrogates. The communication provided by SMM during its instruction is a version of math dialogue akin to Richards' (1996) "school math" characterized by rigidity and computational focus. This style is further characterized by an invitation-reply-discourse sequence; SMM provides a prompt-response-feedback communication loop. Mills and Ragan (2000) noted that the teacher should not be supplanted by any coaching provided through the ILS, and their ideal use of the ILS includes the teacher as an ongoing participant in the teaching process.

This author assumes a pragmatist position (Creswell, 2011; Creswell & Plano Clark, 2011) that avoids the discontinuities between the various theoretical frameworks above. Instead, pragmatism takes a "what works" approach and considers the question asked as more important than the underlying theory (Creswell, 2011; Creswell & Plano Clark, 2011; Tashakkori & Teddlie, 2003). This leads to a philosophical pluralism that allows for the inclusion of both behaviorist understandings of learning as well as constructivist epistemologies. Practicality, a focus on the outcomes and consequences of choices, is most valued (Cherryholmes, 1992; Tashakkori & Teddlie, 2003). The question being asked here is whether or not SuccessMaker is effective for improving mathematical learning for students with disabilities, not by what means it may do so.

Previous Research Findings

Though the research on instructional learning systems is rich, a historical review of SMM was more difficult. Possibly due to the variety of names by which the product has been called over the years, few primary source documents were found. Many studies that were identified had not been submitted to peer review through the journal publication process. A review of existing meta-analyses and research syntheses was undertaken. These studies are presented in Table 2.1, including selected details and effect sizes.

The studies presented in Table 2.1 are not without concern. Only six of the studies in Table 2.1 (Crawford, 1970; Delon, 1970; Mendelsohn, 1972; Ragosta, 1983; Suppes & Morningstar, 1969; Underwood, Cavendish, Dowling, Fogelman, & Lawson, 1996) have been subject to peer review. This increases the possibility that design flaws and inaccurate reporting may have led to erroneous results. Slavin and Lake (2008) identified design flaws in eleven studies, including Kirk (2003) and Underwood et al. (1996) presented here. A frequent design issue cited by Slavin and Lake (2008) was the lack of an adequate control group, though inadequate outcome measures and group equivalence were also noted as concerns among their excluded studies. Table 2.1 includes four institutional reports, and the most recent report included (Gatti, 2009) should be interpreted with caution as it appears to be research sponsored by the vendor for SMM.

The lack of recent research regarding SMM is of concern. No peer-reviewed research was found that was been conducted in the past twenty years. The most recent research studies located were conducted by doctoral students as part of their dissertations (Gee, 2008; Kirk, 2003; Manning, 2004; Mintz, 2000). Though the research has investigated the same program, that program has doubtlessly changed over time to leverage new technological capabilities. At present, Pearson (2015) is advertising SuccessMaker 8 as the newest version of their software. It is unclear if differences between this version and previous versions are cosmetic, functional, or instructional.

Table 2.1

Previous SuccessMaker Research

Study	Type of Publication	Location	Grade	Number of Subjects	Effect Size (*d*)
†Cranford (1976)	Dissertation	Mississippi	$5^{th} – 6^{th}$		.64
†Crawford (1970)	Journal Article	California	7^{th}	2 classrooms, 36 students	.10
†Davies (1972)	Dissertation	California	$3^{rd} – 6^{th}$	240 students	.34
†Delon (1970)	Journal Article	Mississippi	1^{st}	5 classrooms, 99 students	1.08
Gatti (2009)	Institutional Report	4 states (AZ, FL, MA, NJ)	3^{rd}, 5^{th}	8 schools, 792 students	.14 (for 3^{rd}) .50 (for 5^{th})
Gee (2008)	Dissertation	Georgia	$3^{rd} – 5^{th}$	1 school, 180 students	.61
*Hotard & Cortez (1983)	Institutional Report	Louisiana	$3^{rd} – 6^{th}$	2 schools, 190 students	.39
†Jamison, Fletcher, Suppes, & Atkinson (1976)	Book Chapter	Mississippi	$1^{st} – 6^{th}$	12 schools, 600 students	.40
Kirk (2003)	Dissertation	Tennessee	$2^{nd} – 5^{th}$	4 schools, 348 students	.84 (.93 for 5^{th})
Laub (1995)	Dissertation	Pennsylvania	4^{th}-5^{th}	2 schools, 314 students	.56
Manning (2004)	Dissertation	Florida	6^{th}	1 school, 64 students	.75
Manuel (1987)	Dissertation	Nebraska	3^{rd}-6^{th}	3 schools, 165 students	.06
†Mendelsohn (1972)	Journal Article	New York	$2^{nd} – 6^{th}$	20 schools, 3,282 students	.49

Table 2.1 Continued

Study	Type of Publication	Location	Grade	Number of Subjects	Effect Size (*d*)
†Miller (1984)	Dissertation	Oregon	$5^{th} – 8^{th}$	15 schools, 577 students	.38
Mintz (2000)	Dissertation	Alabama	$4^{th} – 5^{th}$	8 schools, 487 students	-.06
†Palmer (1973)	Institutional Report	California	$4^{th} – 6^{th}$	3 schools, 171 students	.36
†Prince (1969)	Institutional Report	Mississippi	$1^{st} – 6^{th}$	12 schools, 544 students	.64
*Ragosta (1983)	Journal Article	California	$1^{st} – 6^{th}$	4 schools	.77
†Suppes & Morningstar (1969)	Journal Article	California	$1^{st} – 6^{th}$	7 schools, 1896 students	.28
Underwood, Cavendish, Dowling, Fogelman, & Lawson (1996)	Journal Article	United Kingdom	primary & secondary	9 schools, 173 students	.40
†Vincent (1977)	Dissertation	Ohio	$9^{th} – 12^{th}$	2 schools, 35 students	.34

Notes: †Included in Kulik (1994) meta-analysis. *Included in Slavin and Lake (2008).

Given the ages of the studies listed in Table 2.1, it is reasonable to assume that the underlying theoretical framework relied heavily on programmed instruction (Svoboda et al., 2012).

An average effect size was found for the studies provided in Table 2.1, though certain assumptions were required. It was assumed that the sample in Gatti (2009) was equally split into two groups. The low effect size for Kirk (2003) was used as representative of her study given the concerns presented by Slavin and Lake (2008). The simple mean effect size found for studies in Table 2.1 was $d = .46$ (95%CI [.34, .57]). Using Cohen's (1988) suggestions regarding the interpretation of effect sizes, this result would be considered small. Removal of two significant outliers (Delon, 1970; Mintz, 2000) yielded a similar though slightly lower simple mean effect size of $d = .41$ (95%CI [.32, .50]). Notably, three of the highest effect sizes from these studies were from studies conducted in Mississippi nearly forty years ago (Cranford, 1976; Delon, 1970; Prince, 1969). Restricting this process to only studies conducted since 2000 did not result in significantly different results.

An additional evaluation of SMM research was conducted by Becker (1992). Results from 11 studies conducted during the 1980s were included, though citations for these studies were omitted by the author. As a consequence, locating Becker's original sources is unlikely. Becker's (1992) studies are described in Table 2.2. Becker included both sample sizes and effect sizes for the included studies, and a weighted mean effect size can be calculated. It is assumed that the sample size from the Calvert Co., Maryland study was equal for all three groups. The weighted mean effect size was $d = .30$ (95%CI

Table 2.2

Studies Included in Becker (1992) Meta-Analysis

Study	Design	Location	Grade	Number of Subjects	Effect Size (*d*)
1988-89	Individual Change vs. Test Norms	Ft. Worth, TX	$1^{st} – 7^{th}$	120 students, ~25 hours use	1.60
1988-89	Individual Change vs. Test Norms	Omaha, NE	$2^{nd} – 6^{th}$	170 students, ~20 hours use	1.30
1987-88	Individual Change vs. Test Norms	Milwaukee, WI	$2^{nd} – 9^{th}$	600 students, ~40 hours use	.80
1987-88	Individual Change vs. Test Norms	Aiken Co., SC	$2^{nd} – 8^{th}$	600 students, ~30 hours use	.70
1983-88	Cohort Change to Statewide Change	Calvert Co., MD	3^{rd}, 5^{th}, 8^{th}	1,500 students, ~35 hours use	.10 (3^{rd}) .25 (5^{th}) .50 (8^{th})
1983-86	Individual Change vs. Test Norms	Calvert CO., MD	$4^{th} – 6^{th}$	653 students	.35
1977-80	Random Assignment	Los Angeles, CA	$1^{st} – 6^{th}$	750 students, ~50 hours use	.26
1980-81	Random Assignment	Lafayette Parish, LA	$3^{rd} – 6^{th}$	94 students, ~25 hours use	.19
1981-82	Comparison Group	Portland, OR	$5^{th} – 8^{th}$	80 students, ~25 hours use	.30
1984-86	Comparison Group	Rochester, NY	$4^{th} – 6^{th}$	2,600 students, 19 schools	.00
1984-85	Comparison Group	Atlanta, GA	Elementary, Middle	700 students, 7 schools ~25 hours use	.40

Note. Becker (1992) failed to provide authors for any of the studies included in his meta-analysis. Consequently, these studies are only descriptions of studies rather than identifications of studies. Most sample sizes are approximate.

[.12, .47]). This small effect size was statistically significant. However, the New York study contained nearly one-third of the cumulative sample in Becker's presentation, and the effect size for that study was a statistical outlier. Removal of this study and recalculation of the weighted mean effect size yielded an effect size of d = .45 (95%CI [.28, .63]). Studies done most recently generated effect sizes greater than the confidence interval for the revised mean effect size, suggesting a time-based effect perhaps tied to technology innovations.

A number of studies have been identified by previous authors but rejected for various reasons. Table 2.3 provides an overview of these studies. Many of the studies were rejected by Slavin and Lake (2008) for various reasons, though Pearson (2002) provided a collection of summaries for these. All of the studies in Pearson (2002) failed to provide sufficient statistical information from which to derive effect size information. Instead, percentiles and percentage passing rates appeared more frequently. None of the original studies could be found, though most appeared to be reports produced by either Pearson (vendor for SMM) or the school districts in which the product was used. None were submitted for peer review, and the likelihood of corporate authorship casts doubts as to the replicability of the studies. None of the studies were conducted in the past ten years.

Previous research has suggested that SMM produces a small but significant effect on student achievement. Findings were inconsistent across types of studies (journal article vs. dissertation, etc.) as noted above. Study location may have even impacted findings. Research efforts regarding SMM may be tapering off; the last peer-reviewed

Table 2.3

Documents Not Included in Meta-Analytic Comparisons

Study	Type of Publication	Location	Grade	Number of Subjects	Data Provided
Crenshaw (1982)	Dissertation				(a)
Donnelly (2004)	Presentation				(b)
Humphries (1997)	Institutional Report	North Carolina	3rd – 8th	11 classrooms	percentiles
Laub & Wildasin (1998)	Institutional Report	Pennsylvania	2nd – 6th	6 schools, 522 students	percentiles, grade equivalents (a)
McWhirt, Mentavlos, Rose-Baele, & Donnelly, (2003)	Institutional Report				(a)
Office of Research, Loudoun Co. Public Schools (1998)	Institutional Report	Virginia	3rd – 5th	3 schools, 254 students	qualitative overview
Phillips (2001)	Dissertation				(c)
Simon & Tingey (2001)	Institutional Report	Florida	4th – 5th	12 schools, 459 students	FCAT results
Tingey & Simon (2001)	Institutional Report	California	4th – 5th	9 schools, 597 students	mean gains, normal curve equivalents (a)
Tingey & Thrall (2000)	Institutional Report	Florida	4th – 5th	12 schools	percentage comparisons (a)
Tuscher (1998)	Institutional Report	Pennsylvania	3rd – 5th	4 schools	SAT-9 percentiles (a)
Wildasin (1984)	Institutional Report				(a)

Note. All deficiency comments from Slavin & Lake (2008).
(a) Lack of an adequate control group. (b) Insufficient control group matching. (c) Inadequate outcome measure.

article was published twenty years ago. Previous research has also focused on elementary mathematics performance. Only eight studies included students in 7th or 8th grades (traditional junior high or middle school grades). It is telling that the What Works Clearinghouse provides no judgment of the evidence-based effectiveness of SMM. More research is needed to determine if SMM truly yields an effect on students' mathematics achievement.

Purpose of This Study

National standards have been set through No Child Left Behind and Race to the Top by which schools are expected to demonstrate adequate yearly progress in mathematics. Students with disabilities have historically underperformed on these assessments relative to their non-disabled peers. As the number of students with disabilities grows, it becomes increasingly important to provide adequate supports for these students in order to meet state and national standards (Manning, 2004). Students with disabilities generally only make small achievement gains, especially during the middle school years (Graham, Bellert, Thomas, & Pegg, 2007). Pressures for students with disabilities, especially learning disabilities, to succeed are increasing (Martindale, Pearson, Curda, & Pilcher, 2005) while the gap between high and low achievers grows wider every year (Cawley, Parmar, Yan, & Miller, 1998).

Despite the research base for SMM outlined above, limited research exists to support its effectiveness for students with disabilities (Wood, 2004). Vockell and Mihail (1993) suggested that consistent computer-based instruction may provide students with disabilities a greater chance of success through development of automaticity and overlearning of concepts. It has also been suggested that technology should be integrated into mathematics instruction for all at-risk learners (Li & Edmonds, 2005). The aim of this study is to determine if SMM effectively improves mathematics achievement for students with disabilities.

Methods

SuccessMaker Mathematics was purchased by a central Texas school district at the beginning of the 2010-2011 school year by the Special Education department. Consequently, schools were instructed that only students eligible for special educations services were to use the program. Licenses were purchased and given to all 12 middle schools in the district. Identification of specific students and development of a campus implementation plan was left to the campuses. Vendor recommendations to the district regarding yearly usage totals suggested that 20-25 hours of use per student should produce measurable achievement gains. Those recommendations are consistent with those currently provided by vendor representatives (D. Wayland, personal communication, January 28, 2016). A matrix of time usage estimates based on IP level and expected gain provided by the vendor (Pearson Education, 2012) was not available to the district at the start of their implementation. The array considers homogeneous clusters of students grouped by their IP level. Based on desired gain levels, usage levels are provided at three incremental levels of student success. The publication reads, in part, "Achieving the time in the 50^{th} percentile column will result in approximately one-half of students reaching at least that gain; achieving the time in the 75^{th} percentile will result in approximately three-fourths of students reaching at least that gain" (Pearson

Education, 2012). Given the wide range of achievement levels for students using SuccessMaker both district-wide and at each campus, the matrix was condensed to a yearly usage recommendation of approximately 20-25 hours consistent with on-site vendor recommendations. For students with an IP level of 3.0 or greater, the matrix provided indicates that usage at these recommended levels is capable of yielding at least 1.0 years of growth. For students with an IP level of 4.5 or greater, the matrix indicates that usage at these recommended levels is capable of yielding 1.5 years of growth. Data for this research spans 5 years beginning with the 2010-2011 school year.

Participants

Each year the program has been available, students with disabilities have had access to the program contingent on campus implementation plans. Consequently, some students have received multiple years of program usage. There is limited research available (McKissick, 2016) to suggest that multiple years of program use might affect program effectiveness. Each student-year of program use, then, will be considered unaffected by use in previous years.

The State of Texas has developed a number of end-of-year high-stakes examinations for its students. Prior to 2012, students took the Texas Assessment of Knowledge and Skills (TAKS). Five versions of that test were available to students: TAKS, a grade-level assessment identical to that taken by non-disabled students; TAKS-Accommodated, a grade-level assessment with additional allowable accommodations not believed to influence the rigor of the assessment; TAKS-Modified, testing grade-level concepts using simplified vocabulary, reduced answer choices, and a simplified format;

TAKS-Alternate, for students with severe cognitive disabilities interfering with administration of paper-and-pencil examinations; and LAT, for students requiring linguistic accommodations. Beginning in 2012, students took the State of Texas Assessment of Academic Readiness (STAAR). Four versions of the STAAR were originally available, mirroring the versions available with TAKS, with the exception of a STAAR-Accommodated version. The STAAR-Modified test was replaced during the 2014-2015 school year with the STAAR-Accommodated version, an online assessment utilizing virtual tools such as a highlighter and sticky notes. State testing expectations are considered annually as part of the development of Individualized Education Plan for each student with disabilities.

During the five years of SMM use in the district, 2,441 student-years of data were collected. Of these, 156 were removed because prior-year (baseline) or current-year state testing data included the Alternate or linguistically accommodated version of the state assessment. Some students were introduced to SMM but did not complete initial placement. The reporting of state testing data for the previous year was taken as evidence that the student began the year in the district, and reporting of state testing data for the year of SMM was taken as evidence that the student ended the year in the district. Thus, an additional 668 were removed for lack of current- or prior-year state test data or SMM usage data indicative of either lack of treatment exposure or limited use due to partial-year enrollment. An additional 15 student-years of data were removed because no special education eligibility could be verified. Of the resultant 1,603 student-years of data, 398 included current- and prior-year state testing data at the different levels (grade-level or modified). These were removed for lack of adequate techniques to compare scores between various levels of the state assessments. The resultant dataset included 1,204 student-years of data from 920 unique students. There were 673 students who used the program for one year, 210 in two different years, and 36 students in three different years.

Materials and Procedure

SMM was made available for all middle school campuses in the district for use with students with disabilities. Campuses assumed responsibility for implementation of the program, including which students would access the program during various times of the day. Students at most of the campuses were provided opportunities to use the program before and after school as time and access allowed. Students were also able to access the program from home. Campus plans have undergone revision and refinement in subsequent years, and some campuses have integrated SMM use as part of the curriculum for resource mathematics classes (McKissick, 2016). Variations in campus implementation plans have not changed the specific intervention, namely SMM.

The district provided two measures of student achievement. First, SMM cumulative usage reports by student for each year were reviewed. These reports included an initial placement score, a grade level placement identified by SMM based on an initial evaluation of student abilities. A final grade placement score was also included so that a measure of math achievement gain during program use could be calculated. Because students from multiple grade levels were to have their performance analyzed simultaneously, it was determined that a measure of previous learning was needed. It was expected that students beginning a grade level should have an initial placement score equal to that grade level, indicative of achieving one academic grade level for each prior year of school. Thus, an average rate of growth was calculated by dividing the initial placement score by the grade. Additionally, state testing results from the previous year were made available. As mentioned above, changes in state testing have been frequent. Though scaled scores were made available, changes in scales between test versions and across years have made comparisons nearly impossible. Using district means and standard deviations, these scores were transformed to *z*-scores by test type and year. The design for this study is modeled in the diagram below, where *O1* and *O2* represent state testing results and SMM grade placement results respectively:

NR	$\{O_{1A}, O_{2A}\}$	$X_{FULL\ (>20\ Hours)}$	$\{O_{1B}, O_{2B}\}$
NR	$\{O_{1A}, O_{2A}\}$	$X_{LIMITED\ (15\text{-}20\ Hours)}$	$\{O_{1B}, O_{2B}\}$
NR	$\{O_{1A}, O_{2A}\}$	$X_{LIMITED\ (10\text{-}15\ Hours)}$	$\{O_{1B}, O_{2B}\}$
NR	$\{O_{1A}, O_{2A}\}$	$X_{LIMITED\ (5\text{-}10\ Hours)}$	$\{O_{1B}, O_{2B}\}$
NR	$\{O_{1A}, O_{2A}\}$	$X_{LIMITED\ (0\text{-}5\ Hours)}$	$\{O_{1B}, O_{2B}\}$

Students were classified by their level of program use. Group A used SMM for 0-5 hours during a year, Group B used the program for 5-10 hours during a year, Group C used the program for 10-15 hours, Group D used the program for 15-20 hours, and Group E used the program for more than 20 hours. Two revisions were made to the dataset. First, all students with an average rate of prior growth greater than 1.0 were removed. Though these 18 students had identified disabilities, it was not apparent that the disabilities had impacted their mathematics achievement. Second, it was determined that the unbounded upper end of Group E allowed for the inclusion of "super-users" who had accumulated well over 25 hours of program use (maximum use reported was 81.4 hours in a year). Consequently, Group E was amended to include students with 20-25 hours of program use, resulting in the exclusion of 194 "super-users." This resultant range coincides with vendor recommendations to the district regarding target usage levels.

A primary concern in the absence of random assignment is the establishment of between-group homogeneity. An analysis of variance identified no significant variations between groups regarding their prior year state testing performance, $F(4, 885) = 1.56$, $p = .1817$. Similar analyses were conducted between groups for all disability areas. A significant difference was found only among students with an intellectual disability, though the result may be due to a small number of students in the sample with that disability. An analysis of variance was conducted to determine if there were any differences between usage groups regarding the average rate of growth. Again, no statistically significant between-group differences were found, $F(4, 885) = 1.14$, $p =$

.3375. Analyses for between-group differences in average rate of growth were conducted by disability area. Between-group differences existed for students with autism, likely due to small sample sizes. Summary information for theses analyses are provided in Table 2.4. Analyses of both variables were extended to grade, gender, ethnicity, and school year. All tests identified homogeneity of groups except for prior state testing in 2013 and average rate of growth in 2014. Both may indicate refinement of campus implementation plans, though it should also be noted that the state test changed from TAKS to STAAR for the 2013 school year. Based on these analyses, the usage groups demonstrate sufficient homogeneity to proceed with further analysis. Additional group description, including demographic information, is provided in Tables 2.5 and 2.6.

Results

Two outcome measures were identified that were consistent with the variables used to determine between-group equivalence. Prior rate of learning was subtracted from SMM-reported achievement gain to determine a change in learning rate. State testing scores from the year of program use and the year prior to program use were transformed to *z*-scores, and a *z*-score difference was derived by subtracting the two. The use of both measures was indicated by the dual expectations of program used – improvement in state testing performance and growth in student achievement rates.

Multivariate analysis of variance (MANOVA) was conducted to determine if student usage significantly affected these achievement measures. Attention was given to the assumptions of MANOVA prior to analysis. Assumptions regarding sample size,

Table 2.4

Tests for Group Homogeneity

Dependent Variable 1: Average Rate of Growth Prior to SuccessMaker Use

		df	SS	MS	*F*	*p*
All Disabilities	Group	4	.0863	.0216	1.137	.3375
	Error	885	16.793	.0190		
Autism	Group	4	.218	.0544	3.025	.0280
	Error	42	.756	.0180		
Emotional Disturbance	Group	4	.0257	.0064	.267	.8979
	Error	44	1.059	.0241		
Learning Disability	Group	4	.0565	.0141	.825	.5093
	Error	540	9.244	.0171		
Intellectual Disability	Group	4	.0272	.0091	1.084	.3861
	Error	15	.1255	.0084		
Other Health Impairment	Group	4	.0299	.0075	.388	.8173
	Error	123	2.374	.0192		

Dependent Variable 2: State Testing *z*-Score for Year Before SuccessMaker Use

		df	SS	MS	*F*	*p*
All Disabilities	Group	4	3.942	.9855	1.565	.1817
	Error	885	557.45	.6299		
Autism	Group	4	2.049	.5121	.991	.4230
	Error	42	21.704	.5168		
Emotional Disturbance	Group	4	1.689	.4223	.618	.6518
	Error	44	30.048	.6829		
Learning Disability	Group	4	2.996	.7490	1.145	.3346
	Error	540	353.35	.6544		
Intellectual Disability	Group	3	8.033	2.678	4.051	.0270
	Error	15	9.916	.6610		
Other Health Impairment	Group	4	1.535	.3838	.6763	.6097
	Error	123	69.799	.5675		

Table 2.5

Usage Group Demographics

	Group A (0-5 hours)	Group B (5-10 hours)	Group C (10-15 hours)	Group D (15-20 hours)	Group E (20-25 hours)	Group F* (>25 hours)
N	227	292	190	102	79	194
Male/Female	137 / 90	188 / 104	116 / 74	67 / 35	45 / 32	121 / 73
Afr.-Amer.	91	144	75	39	26	80
Hispanic	57	63	51	31	27	58
White	68	72	54	26	25	48
Other	11	13	10	6	1	8
Autism	16	10	8	7	6	15
Emotional Disturbance	16	11	13	6	3	7
Learning Disability	128	180	125	66	46	121
Intellectual Disability	5	7	5	2	0	6
Other Health Impairment	34	45	22	11	16	20
Other Disabilities**	7	7	3	4	1	6
Multiple Disability Codes†	21	32	14	6	7	19

Notes: *Group F was not included in the MANOVA and follow-up ANOVAs. **This category includes students who have auditory, visual, or orthopedic impairments. †Students may have disabilities in multiple areas. They are grouped separately here as the impact of multiple disabilities is not known.

Table 2.6

Usage Statistics (Means and Standard Deviations) per Usage Group

	Group A (0-5 hours)	Group B (5-10 hours)	Group C (10-15 hours)	Group D (15-20 hours)	Group E (20-25 hours)	Group F* (>25 hours)
IP Level	4.44 (1.02)	4.27 (.94)	4.36 (1.07)	4.36 (.91)	4.37 (.89)	3.84 (1.04)
Avg. Growth Rate	.64 (.15)	.62 (.13)	.63 (.15)	.63 (.13)	.61 (.12)	.56 (.14)
Gain	.06 (.06)	.20 (.11)	.38 (.16)	.52 (.20)	.59 (.20)	1.07 (.51)
Prior Year State Testing *z*-Score	-.64 (.84)	-.70 (.81)	-.76 (.73)	-.70 (.79)	-.51 (.76)	-.34 (.91)
Current Year State Testing *z*-Score	-.52 (.82)	-.60 (.84)	-.57 (.73)	-.48 (.80)	-.25 (.84)	-.23 (.91)
Accuracy	.62 (.16)	.65 (.09)	.65 (.08)	.64 (.09)	.63 (.07)	.62 (.07)
Questions per Session	8.85 (6.71)	14.62 (7.27)	18.96 (8.17)	20.93 (9.87)	21.15 (9.52)	23.38 (9.27)
Questions per Hour of Use	37.76 (25.70)	61.58 (26.76)	75.08 (25.83)	76.93 (23.52)	75.85 (25.34)	83.13 (27.96)
Session Length (in minutes)	14.4 (4.2)	14.4 (3.0)	15.0 (3.6)	16.2 (4.8)	16.8 (5.4)	16.8 (4.2)

Note: Group F was not included in the MANOVA or follow-up ANOVAs.

independence of observations, and types of variables used in the analysis appeared to be met. Analysis of univariate distributions for the dependent variables resulted in the removal of 49 outliers. Analysis of multivariate distributions, resulting in Mahalanobis distances, resulted in the removal of 53 outliers. Multivariate normality was determined by examination of the normality of each dependent variable, inspection of *Q-Q* plots, and review of residuals from a generalized linear model. For each usage level for each dependent variable, the Shapiro-Wilk *W* was not significant. These are provided in Table 2.7. The generalized linear model yielded a measure of overdispersion of 0.4328, the ratio of deviance to degrees of freedom. Overdispersion rates greater than 1 are problematic (Carruthers, Lewis, McCue, & Westley, 2008), so the assumption regarding multivariate normality was resolved. A comparison of linear and quadratic fit lines between the two dependent variables resulted in fractional increases to R^2, suggesting that a linear relationship between variables existed. The Levene statistic identified no variance concerns for the change in state testing *z*-scores. Comparison of group variances for the change in growth rate involved a comparison of the highest and lowest group variances. This yielded an F_{MAX} = 2.048, and the greatest ratio of sample sizes was 3.696. According to Tabachnick and Fidell (2001), "F_{MAX} is the ratio of the largest cell variance to the smallest. If sample sizes are relatively equal (with a ratio of 4 to 1 or less for largest to smallest cell size, an F_{MAX} as great as 10 is acceptable" (p. 80). To assess multicollinearity, the correlation between dependent variables was found to be low yet significant based on the sample size, r = .082 (95% CI [.016, .147]). The sample appears

Table 2.7

Shapiro-Wilk Values for DV Univariate Normality

	Change in Growth Rate	Change in State Testing *z*-Score
Group A (0-5 hours) *N* = 227	*W* = .9888 *p* = .0734	*W* = .9908 *p* = .1619
Group B (5-10 hours) *N* = 292	*W* = .9908 *p* = .0655	*W* = .9952 *p* = .5002
Group C (10-15 hours) *N* = 190	*W* = .9896 *p* = .1807	*W* = .9917 *p* = .3443
Group D (15-20 hours) *N* = 102	*W* = .9852 *p* = .3128	*W* = .9862 *p* = .3707
Group E (20-25 hours) *N* = 79	*W* = .9832 *p* = .3866	*W* = .9832 *p* = .3878

to meet all assumptions for the MANOVA. The MANOVA yielded a Wilks' $\Lambda = .5161$, $F(8, 1768) = 86.63$, $p < .0001$.

Univariate analysis of variance was conducted with each dependent variable. The analysis for change in state testing *z*-score was not significant, $F(4, 885) = 1.497$, $p = .2012$. Between groups *t*-tests found no usage groups to be statistically different for this outcome measure. The analysis of variance (ANOVA) for change in growth rate was significant, $F(4, 885) = 206.57$, $p < .0001$. All usage groups were statistically different from each other. Results for these analyses can be found in Table 2.8. The greatest change in growth rate was found for Group E, $\bar{x} = -.0213$ (95% CI [-.064, .021]). ANOVAs were also conducted to determine if there were any differences in both dependent variables for gender, ethnic, and disability groups; no group differences were found.

Because the analysis of state testing *z*-scores was found to be not significant, attention was focused on the analysis of growth rate. All users, except for those with an average rate of growth before SMM use greater than 1.0, were considered for inclusion. This sample of 1186 included the 194 "super-users" excluded from previous analyses. In preparation for an ANOVA to determine if any variations existed between the six usage groups (previous five plus Group F, those who used the program for more than 25 hours) regarding a change in growth rate, the variable was analyzed for univariate normality. This resulted in the removal of 25 univariate outliers, resulting in a sample of 1161 student-years of usage. Subsequent Shapiro-Wilk *W* tests failed to confirm normality for 4 of the 6 groups on the dependent variable. A logarithmic transformation of the

Table 2.8

Results of ANOVAs for Each Outcome Measure for Groups A-E

Change in Growth Rate

Source	df	SS	MS	*F*	*p*
Usage Group	4	30.457	7.614	206.57	<.0001
Error	885	32.621	.039		
Total	889	63.078			

Group	*N*	Mean	Lower 95%CI	Upper 95%CI
A	227	-.5819	-.6069	-.5569
B	292	-.4230	-.4451	-.4010
C	190	-.2514	-.2788	-.2241
D	102	-.1010	-.1383	-.0637
E	79	-.0213	-.0637	.0211

Change in State Testing *z*-Score

Source	df	SS	MS	*F*	*p*
Usage Group	4	2.538	.634	1.497	.2012
Error	885	375.153	.424		
Total	889	377.691			

dependent variable was tested for univariate normality, and all groups demonstrated normality on the variable. A significant difference was found between groups, $F(5,1155) = 431.51$, $p < .0001$. Subsequent *t*-tests found significant differences ($p < .0001$) between all group pairings except Groups D and E (15-20 hours of use and 202-25 hours of use, respectively). Values for the means and confidence intervals of each group, converted into units of years change in growth rate, are provided in Table 2.9. The inclusion of previously excluded multivariate outliers resulted in minimal changes to the means for Groups A-D. The mean for Group E increased from the first to second ANOVA, though the 95% confidence interval still contained zero. The mean and confidence interval for Group F suggest that students with disabilities who use SMM for more than 25 hours are likely to realize significant changes in their rate of mathematics achievement.

To determine if different student populations received differential benefit from program use, ANOVAs were conducted to determine variations existed within each usage group. No differences were found for gender or ethnicity groups. Small samples of students with intellectual disabilities and “other” impairments (not those with an OHI eligibility) were removed prior to analysis. No differences were found within usage groups to indicate differential impact of similar usage for students with different disabilities. ANOVAs were conducted across usage groups for each disability group. These analyses mirrored the combined ANOVA conducted above that indicated significant differences between all levels of usage. Results can be found in Table 2.10.

Variation in usage patterns between campuses was identified. Fidelity of implementation has been identified as a reason why interventions fail (Mills & Ragan,

Table 2.9

Results of ANOVA for Change in Growth Rate for All Usage Groups

Source	df	SS	MS	*F*	*p*
Usage Group	5	39.578	7.916	431.51	<.0001
Error	1155	21.187	.018		
Total	1160	60.765			

Group	*N*	Mean	Lower 95%CI	Upper 95%CI
A	241	-.5910	-.6149	-.5667
B	312	-.4350	-.4584	-.4113
C	202	-.2499	-.2823	-.2169
D	121	-.0428	-.0895	.0051
E	94	.0115	-.0429	.0674
F	191	.4387	-.3922	.4860

Note: Means and confidence intervals have been converted from logarithmic values used in ANOVA to years of growth.

2000). A Chi-Square analysis of implementation variations between campuses, reflecting comparable number of students at each usage level, was significant, $\chi^2(44) = 245.77$, $p < .0001$. Students in Groups E and F, those who received the recommended usage and those who exceeded usage recommendations, were included in the same group for this analysis. Table 2.11 presents the percent of students from each campus that received or exceeded the recommended usage levels for each campus. The percentage of students in the current sample receiving or exceeding usage recommendations was 24.62%.

Variations in usage patterns between usage level groups were also identified. Table 6 presents information regarding performance variables for each usage group. Accuracy is defined as the percent of exercises completed correctly. To achieve normality for this variable, 20 outliers were removed and an exponential transformation was applied. Six users were removed who had 0% accuracy (each attempted fewer than 12 questions), and an additional 5 users with 100% accuracy were removed (each attempted fewer than 5 questions). The resultant ANOVA identified a significant variation in accuracy between usage groups, $F(5, 1150) = 6.372$, $p < .0001$. Post-hoc *t*-testing identified that users in Group F had a significantly lower accuracy rate than users in Groups A-D (all $p < .0002$). Session length was calculated as the total usage time divided by the number of sessions (included in the SMM usage report). Attempts to normalize the variable were unsuccessful, so a non-parametric test was used to determine group differences. A Kruskal-Wallis analysis of variance by ranks found significant differences between groups on this variable ($H[5] = 98.107$, $p < .0001$).

Table 2.10

Results of ANOVA for Change in Growth Rate for Disability Groups

	Autism	Emotional Disturbance	Learning Disabilities	Other Health Impairment
F	*F*(5,59) = 19.053	*F*(5, 54) = 10.783	*F*(5, 713) = 236.73	*F*(5, 153) = 32.054
p	< .0001	< .0001	< .0001	< .0001
N	65	60	719	159
Usage Group	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
A	-.63 (-.86, -.40)	-.58 (-.73, -.43)	-.57 (-.61, -.53)	-.58 (-.69, -.46)
B	-.49 (-.77, -.20)	-.44 (-.63, -.26)	-.42 (-.46, -.38)	-.41 (-.51, -.31)
C	-.13 (-.43, .18)	-.11 (-.28, .06)	-.23 (-.28, -.19)	-31 (-.45, -.17)
D	-.09 (-.42, 22)	-.16 (-.40, .08)	-.04 (-.10, .02)	-.01 (-.19, .17)
E	.09 (-.25, .44)	-.05 (-.42, .32)	.06 (-.01, .13)	.02 (-.14, .18)
F	.84 (.61, 1.08)	.37 (.13, .61)	.43 (.38, .47)	.51 (.36, .67)

Table 2.11

Campus Fidelity of Use

Campus	Total *N* for Campus	Percentage of Users Receiving or Exceeding Usage Recommendations
A	89	14.61%
B	145	31.03%
C	154	20.13%
D	86	17.44%
E	72	45.83%
F	82	54.88%
G	139	17.99%
H	25	16.00%
I	147	13.61%
J	80	53.75%
K	76	1.32%
L	91	18.68%
Total	1186	24.62%

Note: Totals cover the five years of usage for this review, and includes only students whose data was used in the analyses conducted.

Two measures of efficiency of use were identified. The number of questions per session provides a measure of the student's effort during each session of program use. To achieve normality for this variable, 16 outliers were removed and a square-root transformation was applied. Three users were removed who had 0% efficiency. All usage groups demonstrated normality except Group C (Shapiro-Wilk $W = .9832$, $p = .0154$), so interpretation of the resultant ANOVA should consider this normality concern. The ANOVA identified a significant variation in questions per session between usage groups, $F(5, 1161) = 126.52$, $p < .0001$. Post-hoc *t*-tests identified differences between all groups (all $p < .02$) except Groups D, E, and F. A second measure of efficiency, the number of questions per hour of use, was identified that removed the impact of session length differences between usage groups. Again, a square-root transformation was applied to achieve normality for each group level. Four outliers were removed, and three students with 0% efficiency were excluded from the analysis. The ANOVA identified a significant variation in the number of questions per hour between usage groups, $F(5, 1173) = 102.84$, $p < .0001$. Post-hoc *t*-tests identified difference between all pairings of Groups A and B with Groups C-F.

Each of these four performance variables was reviewed for differences between demographic groups. ANOVAs were conducted using the three transformed variables (accuracy, questions per session, and questions per hour), and a nonparametric test was conducted using session length. No differences for gender or ethnicity were found. Differences were found among disability groups for questions per session ($F[3, 999] = 3.475$, $p = .0156$) and session length ($H[3] = 9.626$, $p = .022$). Students with autism were found to answer more questions per session despite spending less time per session than students in other disability groups.

To determine the predictive capacity of these usage pattern variables regarding gain in achievement rates, a regression analysis was conducted. Since the amount of usage time has already been identified as having a significant impact on change in growth rates, this analysis was restricted to those students who had received or exceeded the usage recommendations (N = 292). A logarithmic transformation of time was required to achieve normality for this variable. The regression analysis identified time, accuracy, and questions per hour of program use as significant predictors of change in growth rate. Parameter estimates may be found in Table 2.12. A model including these three predictor variables accounted for 84% of the variance in student change in growth rate among students receiving or exceeding program usage recommendations (R^2 = .8411). Using the mean accuracy and mean number of questions per hour for these students, it was found that 25 hours of program use would result in growth rates commensurate with previous years of schooling. Increasing the use to 42 hours, holding the other two parameters constant, is predicted to yield a growth rate change of .5. This level of program use – nearly double the recommendations – may lead to closing the math achievement gap by half of a school year. To close the math achievement gap by a full school year, nearly 70 hours of program use is predicted to be necessary.

Regression analysis was also conducted for disability groups for those students receiving or exceeding usage recommendations. Small samples sizes prohibit generalizations for students with autism, emotional disturbances, and intellectual disabilities. Regression equations for students with learning disabilities and other health impairments (often, ADHD) identified the same parameters as significant. As the parameter estimates do not overlap, their differential impact may be of predictive value. Estimates for these parameters are also found in Table 2.12.

There are multiple ways to determine the effect size for the treatment used. When students are re-grouped dichotomously as to whether or not they received the treatment with fidelity, the impact on the outcome variable (logarithmic transformation in change in growth rate) is significant, $F(1, 1159) = 907.42$, $p < .0001$, with an accompanying $R^2 = .439$. Using Kabacoff's (2014) formula below for using R^2 to find effect size, $f^2 = .78$.

$$f^2 = \frac{R^2}{1-R^2} \quad (1)$$

Cohen's (1988) recommendations for interpreting this statistic consider .35 to be a large effect. Using Cohen's (1988) formulae for converting between effect sizes, this effect size is equivalent to $d = 1.77$, large by Cohen's standards. Problematically, this calculation involves the use of SMM data for students who used the program sparingly (consider those with 0-5 hours of use). Information from SMM regarding yearly growth rates may be limited to a portion of the reporting year due to the limited use, therefore creating validity concerns regarding this interpretation.

Alternately, students receiving the program with fidelity might have their rate of growth during treatment use compared to their rate of growth prior to SMM use. Students in Groups E and F, who met or exceeded usage recommendations ($N = 292$), had a combined mean growth during treatment of .93 ($SD = .48$). Their annual rate of growth prior to SMM use was .57 ($SD = .13$). Using formulae (2) and (3) below from

Table 2.12

Regression Analysis Results

	Intercept	Time (Log-Transformed)	Accuracy (Exponential-Transformed)	Questions per Hour (Root-Transformed)
All students receiving or exceeding usage recommendations	-7.11* (-7.74, -6.48)	1.02* (.95, 1.10)	1.24* (1.04, 1.43)	.164* (.07, .258)
Students with Learning Disabilties	-4.81* (-5.34, -4.29)	.90* (.82, .98)	1.99* (1.62, 2.36)	.009* (.004, .014)
Students with Other Health Impairments	-6.35* (-7.69, -5.00)	1.07* (.85, 1.29)	3.60* (2.42, 4.79)	.008** (.001, .016)

Note: *Significant at $p < .001$. **Significant at $p < .03$

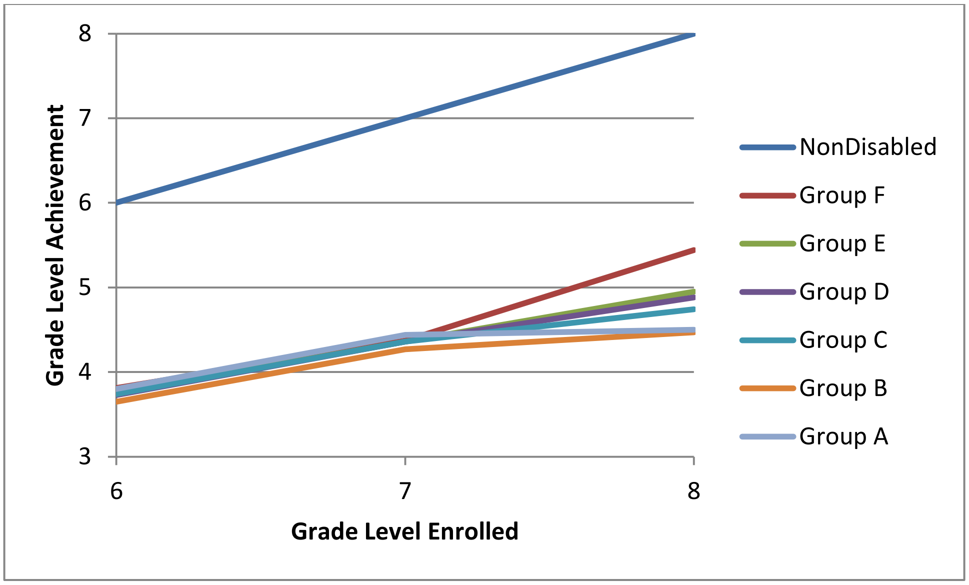
Ellis (2010), an effect size was found, Cohen's *d* = 1.02. Cohen's (1988) benchmarks for evaluating effect sizes identify .80 as a large effect for this statistic. Similar comparisons for state testing performance utilize a prior mean *z*-score of -.387 (*SD* = .884) and end-of-treatment *z*-score of -.240 (*SD* = .889), yielding an insignificant effect size of *d* = .01.

$$d = \frac{\bar{M}_1 - \bar{M}_2}{SD_{pooled}} \tag{2}$$

$$SD_{pooled} = \sqrt{\frac{(n_A - 1)SD_A^2 + (n_B - 1)SD_B^2}{n_A + n_B - 2}} \tag{3}$$

Discussion

Regarding the effectiveness of SuccessMaker Mathematics for students with disabilities, the research conducted demonstrates the potential of the program for closing mathematics achievement gaps. The regression analyses identified that usage patterns regarding accuracy and efficiency (number of questions attempted per hour of program use), in addition to usage time, are useful predictors of changes in achievement growth rate. Though gender and ethnicity did not lead to group differences, variations between disability groups were present in various analyses. Figure 2.1 compares the changes in achievement growth rates for the six usage groups in this study to a hypothetical non-disabled student. Students are expected to experience one year of achievement growth for each year of school. Figure 1 illustrates that this has not historically happened for the disabled students using the program. Though the recommended use of SMM yields a learning trajectory similar to non-disabled students, much greater use would be needed to close the existing gaps.



Notes: The figure utilizes average growth rates and gains from Table 6. Data from 6th, 7th, and 8th grade students were consolidated into representative trend lines for 7th grade comparison. A hypothetical, non-disabled peer is provided as reference.

Figure 2.1. Learning trajectories of disabled students by usage group.

The use of outcome measures for this study present a variety of problems for interpreting the findings. State testing scores, the score of greatest concern to school districts, present significant comparison issues across years. Though equated scores may be useful for comparing across STAAR tests, no bridge was created to compare TAKS scores to STAAR scores. The issue is exponentially worse when addressing students with disabilities as the possible test versions and levels expands. This study has considered only those students whose state testing level (modified or on-level) remained constant from the previous year through the year of treatment. The use of *z*-scores for performance comparisons is less than desirable since students are compared to each other rather to an objective benchmark. Until the State of Texas provides a standardized and consistent measure of achievement, such poor comparison methods are likely to continue.

The consequence of poor state testing data is the need for measurement within SMM itself. Though the program provides an initial placement score, it is unable to assess student effort during the process. Consequently, students who are less motivated may intentionally perform poorly on the initial placement in an effort to meet a teacher's expectation for completion. It is believed that several students whose data were used in this study fall in this category of initial placement responding, though the large sample size and removal of outliers is believed to have reduced or eliminated their impact on analyses.

Further, use of treatment-provided achievement data as an outcome variable is not ideal. Identification and use of additional assessment instruments would be of assistance, and correlational analysis between those instruments and SMM would be useful. As with initial placement testing, performance on any other assessment instrument including state tests is subject to student motivational issues. A design that employs periodic evaluations of student motivation in addition to pre- and post-testing of achievement would improve upon these findings.

The quasi-experimental nature of this research also presents concerns. Though efforts were made to demonstrate homogeneity of usage groups on a host of factors, there is no good substitute for true random assignment. In the school setting, however, true randomization presents possible ethical and practical difficulties. Withholding access to a treatment believed to have benefit, especially for students with disabilities, may be ill-advised. Delaying access to treatment, as might be done in a design involving switching replications (Shadish, Cook, & Campbell, 2002), is difficult to implement for a year-long intervention. The use of a within-subjects design, as has been conducted here, may be necessary. Many interventions, such as SMM, are expensive purchases for school districts. In the absence of available funds or grants, researchers may be forced to utilize existing data. Forward-thinking districts are encouraged to develop an implementation plan that allows for appropriate data collection from the beginning to analyze program effectiveness.

This analysis considers effectiveness of SMM from a treatment dosage perspective. Students who received SMM with fidelity produced significantly higher mathematics achievement gains than students who did not receive the recommended usage of the treatment. When students who exceeded treatment usage recommendations are considered, those gains in achievement are even greater. Future research regarding SMM should consider implementing usage groups for greater usage levels than were considered for this project. Excessive use of the treatment was beyond the scope of this research. It is not yet known if use of SMM well beyond usage recommendations will result in continued linear growth or potential diminishing returns.

Though this paper has taken a pragmatist position, there is reason to believe that behaviorist instructional methods are helpful for students with disabilities. The behaviorist roots of SMM were reviewed above, and the effectiveness of the program for student with disabilities has been shown. This study did not investigate the use and perceptions of features more in line with cognitivist or constructivist theories. Instead, the repeated skill repetition and branching algorithms that serve as a foundation for skill presentation and assessment have yielded usage data consistent with this theoretical position. Further research that addresses the various components of the program is needed to determine what combined and individual effects these components have.

Previous research regarding SMM has included few studies in Texas. Most recently, Tucker (2008) found that SMM provided no benefit to 5^{th} grade students using district passing rates as an outcome measure. This study has focused on the individual student, but has identified a similar lack of state testing differences following program use. Additionally, the current study has opted to address only those students with disabilities. Findings and conclusions from this study may not be generalizable to other student groups or school districts.

The need for effective remediation tools for students with disabilities is clear and ongoing. SuccessMaker has demonstrated an ability to assist struggling learners, but only if minimum usage recommendations are followed. Even then, these learners may not achieve learning gains commensurate with their non-disabled peers. Schools using SMM are encouraged to develop a clear plan for implementation that will allow students to meet targeted usage levels. Ongoing monitoring of student performance during program use is recommended so motivational issues discussed above may be addressed early. A discussion of implementation concerns is presented in McKissick (2016), though users are encouraged to identify the needs and target population for their campus.

CHAPTER III

PERCEPTIONS AND OBSTACLES ENCOUNTERED DURING SUCCESSMAKER IMPLEMENTATION

Overview

Technology implementations face a variety of barriers that affect implementation fidelity. These barriers include both extrinsic concerns (time, equipment, and training) and intrinsic concerns (beliefs, values, and practices). Curriculum considerations and clarity of program use also contribute to implementation fidelity issues. An instrumental case study approach was employed to determine user perceptions of SuccessMaker following five years of use in a central Texas school district as well as reflections on the implementation process. Twelve participants – 6 teachers and 6 campus administrators – participated in a structured interview and review of a technology implementation matrix. Barriers to implementation were identified, as were strategies employed by participants to overcome these. Five themes emerged from the interviews to guide future technology implementations: advanced preparation, development of a campus plan, making time for program use, development of a core team, and the need for feedback for teachers and students.

Introduction

In 2011, as a result of consistent underachievement by students with disabilities, the Special Education department of a central Texas school district purchased access to SuccessMaker for this specific population. SuccessMaker is an instructional learning system with a significant research history, with much of the research conducted during the 1980s and 1990s and included in later meta-analyses (Kulik, 1994; 2003). Secondary campuses were provided with access to the program and a directive to use the program. The district contracted with Pearson, the current owner of SuccessMaker, to have a company representative train staff and generate student profiles to facilitate an expedient start to the implementation process. However, specific details regarding the implementation were largely left to the campuses. Consequently, a variety of implementation patterns have been seen in the years since the program was purchased.

SuccessMaker is an instructional learning system built on behaviorist principles. Instruction is available in both reading and mathematics, and students are presented with instruction at their identified academic level. Program use begins with an initial placement during which the student is presented with problems of increasing or decreasing difficulty, depending on successful completion of presented problems. A branching algorithm is used to work through various learning strands and levels of difficulty (Svoboda, Jones, van Vulpen, & Harrington, 2012). Regular program use consists of sessions of approximately 15-20 minutes in length, and total usage recommendations based on initial placement level and desired growth have been made by the vendor (Pearson Education, 2012). The program presents students with questions at or slightly above their current level with opportunities for frequent drill and practice. Successful performance is rewarded immediately through sound and visuals as well as cumulatively through games incorporating previously mastered skills. The integration of a “cognitive coach” in recent versions of the program (Pearson Education, 2013) has incorporated aspects of cognitive learning theories with an existing behaviorist foundation. SuccessMaker utilizes current technological capabilities to provide a multimedia learning environment complete with audio and visual supports including animation and engaging graphics.

Literature Review

Technology and curriculum implementation research has consistently identified that the implementation process is key to the success of the intervention. Implementation issues often limit the potential effects of instructional learning systems (Mageau, 1992; Mills & Ragan, 2000), and a meta-analysis has suggested that this problem was consistent across studies (Durlak & DuPre, 2008). Technology complexity may be the cause of poor implementation as well as a reason why poor implementation failed to yield the intended results (Mills & Ragan, 2000). Clarity regarding implementation specifics, as well as ongoing monitoring and feedback, were recommended as tools to improve implementation fidelity (O'Donnell, 2008). Even so, Durlak and DuPre (2008) suggested that high implementation variability should be expected and that perfect implementation was unrealistic.

Implementation diffusions take time, and judgments regarding their efficacy should be delayed. Fullan (2001) identified three phases of change regarding innovation adoptions. The initial adoption stage is followed by a two to three year implementation stage. A continuation (or institutionalization) stage follows thereafter once an institution has worked out any issues and integrated the innovation. Educational change may therefore take three to six years (Baylor & Ritchie, 2002; Felner et al., 2001; Fullan, 2001; Hall & Hurd, 2001) before consistent effects may be observed. Measuring the effects of an innovation or technology prior to a completed implementation may lead to inaccurate findings.

On the one hand, the role of the teacher in the success of technology implementation has been repeatedly identified. Zhao and Frank (2003), summarizing the findings of dozens of studies, indicated that the teacher's attitude toward technology is a key factor in the technology diffusion literature. Teacher beliefs regarding the perceived value and usefulness of the technology were critical in adopting and maintaining the innovation (Agarwal & Prasad, 1999; Barnes, 2005; Davis, 1989). These beliefs, including their pedagogical beliefs, influenced their professional practice (Haney & Lumpe, 1995) and may have limited the full implementation of the innovation. Hadley and Sheingold (1993) found that student accomplishment and independent use of technology by students were the two greatest incentives for teachers as they integrated technologies.

On the other hand, teacher openness to change has been identified as an obstacle to technology implementation (Baylor & Ritchie, 2002; Lapointe & Rivard, 2005). Willingness to change was also seen as an important precursor to actual change (Jazzar & Algozzine, 2006; Mills & Ragan, 2000). The lack of willingness may be due to fear (Stigler & Hiebert, 1999), though it could just as easily be the result of a lack of personal experience. Teacher confidence in the technology and its use may also have limited full implementation (Sobol et al., 1989; Weston, 2005).

Administrative support is also essential in successful innovation implementations. The administrators influenced school structure and culture, including institutional willingness to change (Weston, 2005), through their allocation of resources (Earle, 2002). Administrators may need to reallocate time to allow room for technology implementation as teachers have found that integrating technology in the fixed class period time to be impractical (Bowman, Newman, & Masterson, 2001; Cuban, Kirkpatrick, & Peck, 2001). Training, specifically pre-implementation and ongoing professional development regarding the use of the technology, has been identified as a critical component (Carroll et al., 2007; Durlak & DuPre, 2008; Dusenbury, Brannigan, Falco, & Hansen, 2003). Multiple researchers have identified the need for a resource person or program champion as a valuable support for implementation (Berman & McLaughlin, 1976; Durlak & DuPre, 2008). Early research (Berman & McLaughlin, 1976) did not find that outside consultants were useful in program implementations, though later research has found that full-time technology support is needed for curricular integration (Granger, Morbey, Lotherington, Owston, & Wideman, 2002).

Most research has focused on the adult contributions to technology implementation. However, student attitudes and beliefs may also serve to shape the implementation process (McGhee & Kozma, 2003). Their enjoyment of the technology, motivation to use the program, and perceived importance of the tools were found to be important in the implementation process (Liu & Johnson, 1998).

Numerous barriers to implementation have been identified by previous researchers. Ertmer (1999) posited that implementers encountered first-order extrinsic barriers to implementation. These include time (Dusenbury et al., 2003; Granger et al., 2002; Hadley & Sheingold, 1993; Pelgrum, 2001; Weston, 2005), availability of needed equipment (Granger et al., 2002; Hadley & Sheingold, 1993; Pelgrum, 2001; Weston, 2005), and access to needed support and training (Hadley & Sheingold, 1993; Pelgrum, 2001). These may be augmented by insufficient infrastructure that impeded use of available equipment such as internet access and installation issues (Groff & Mouza, 2008). Second-order intrinsic barriers identified by Ertmer (1999) included teacher beliefs, values, and practices. These barriers were consistent with curriculum concerns (Bowman et al., 2001; Pelgrum, 2001; Weston, 2005), lack of a plan or vision (O'Donnell, 2008; Pelgrum, 2001), and personal attitudes or resistance (Bowman et al., 2001; Lapointe & Rivard, 2005). Additionally, emergent technologies are notoriously unreliable, and this may lead to teachers failing to fully integrate it into lessons (Zhao, Pugh, Sheldon, & Byers, 2002).

Various models for measuring technology implementations have been proposed. Groff and Mouza (2008) proposed a model that considers the context of the implementation and characteristics of the teacher, student, and innovation. Their model incorporates beliefs and attitudes of the teacher and student, infrastructure and systems, and distance between the innovation and current practices. The Concerns-Based Adoption Model [CBAM] (Hall, 1979) "generally approaches change as a mandate from an administrator or other leader position then diffused to the teachers as the ultimate consumer of the innovation" (Straub, 2009, p. 636). Using the CBAM as their theoretical guide, Mills and Ragan (2000) generated an Instructional Learning System Configuration Matrix (ILSCM) to measure implementation of SuccessMaker. While there are thematic overlaps between these two models, the ILSCM better captures observable behaviors whereas that of Groff and Mouza (2008) considers attitudes and existing capabilities. The ILSCM (Mills & Ragan, 2000) will be reviewed as part of this study of the campus perceptions regarding implementation of SuccessMaker for students with disabilities.

Purpose

The proposed study seeks to capture campus perceptions regarding the implementation of SuccessMaker for middle school students with disabilities. Teachers and campus administrators were asked to describe their implementation process, including barriers encountered as well as their strategies for program use. Their perceptions of the program, including their perceptions of the students using the program were reviewed. Feedback from campus staff regarding the ILSCM (Mills & Ragan, 2000) was integrated with these processes and perceptions.

Methods

Role of the Researcher

During the period covered by this investigation (2011-2014), I served as a special education coordinator for a central Texas school district. This position afforded me numerous opportunities to collaborate with many teachers and administrators across the district, including all of the individuals approached for participation in this study. My responsibilities did not include appraisal or evaluative feedback for any of the potential participants. I also was not tasked with providing support of any kind for SuccessMaker during this time; another special education coordinator was responsible for the rollout and ongoing supervision of the program.

Recognizing the potential conflicts that might arise as a result of my professional relationships with the potential participants, multiple steps were taken to prevent such issues. Permission was sought from the superintendent of the school district to conduct the research, including soliciting and interviewing campus administrators and teachers. Research guidelines were provided by the school district to govern the conduct of the research and designed to minimize the impact of the existing professional relationship. Permission was also sought from the Institutional Review Board of Texas A&M University, where this potential conflict of interest was also reviewed before the research proposal was accepted. Communication with potential participants identified me as a student rather than by my professional position. The start of each interview included a clarification of my purpose and role during the interview, and I occasionally deflected questions from participants that sought approval for either their implementation or perceptions of the program. Once a draft of this article was developed, I asked the participants to review the text and provide feedback. Despite these efforts, it is possible that responses from the participants may include a positive spin as they discussed the use and perceptions of SuccessMaker with a district administrator from whose department the program came. In order to reduce this limitation, I intentionally identified myself as a student rather than as a district employee and, when necessary, assured the participants that their responses would have no impact on my perceptions regarding the quality of their teaching or campus administration performance.

Study Context

Access to SuccessMaker was purchased by the special education department of a central Texas school district. Initially, the program was provided to secondary campuses, including twelve middle schools. Use of the program was limited to special education students as a result of rules regarding the source of funding for the intervention. The program provided support for both mathematics and reading, and many students participated in both content areas. A fixed amount of licenses, approximately 15-20 each year, were provided to each campus. These licenses limited the number of students who were able to concurrently access the program from each campus. Campuses were instructed to develop a plan of implemention for the program, leaving complete discretion to the campuses to develop a schedule for program use and identify students for participation.

SuccessMaker provides an evaluative component as well as instructional support. The initial placement evaluation presents students with exercises that vary in difficulty depending on the accuracy of response. The process may take up to three hours (Pearson, 2012) or approximately 300 questions (Wood, 2004). Once the program identified the student's current level of academic performance in a subject, exercises were presented at or slightly above that level in an instructional format. The program tracks the student's mastery of skills, increases their difficulty as appropriate, and maintains a record of student performance. Teachers may also direct the program to present a specific set of information to a student through the use of custom courses. These are generally shortened instructional sets with a limited number of embedded exercises for students to complete. Consequently, the student may receive either remedial instruction based on initial placement results to fill in instructional gaps or targeted instruction, often for grade-level skills, at the discretion of the teacher. Discretionary use of the program regarding targeted instruction was left to the campuses.

Cost estimates for SuccessMaker are dependent on the number of licenses purchased as well as any on-site vendor support desired. Licenses are described as concurrent, meaning that the number of students who may simultaneously access the program is contingent on the number of licenses purchased. Though five licenses would allow five students to access SuccessMaker at the same time, those five licenses might allow fifty or more students access to program over the course of the day. The cost of one license is approximately $1,000 (EdSurge, n.d.), with additional yearly fees for continued maintenance of the license as well as any costs associated with program upgrades. The district purchased at least 15 licenses for each middle school as well as an unspecified number of licenses for all high schools and selected elementary schools. Documents from the target district indicate that for each of the first four years of program use, on-site vendor support for 140 days was purchased at a cost of $139,500 per year. Approximately 400-450 middle school students with disabilities have used the program each year it has been available in the district.

Access to SuccessMaker requires significant technology support. The program is computer based, so the campus must have sufficient access to computers for their targeted population. Software, including a current web browser and Java (until recently) must also be upgraded and functional. The program requires internet access, and the increased internet usage may drain available bandwidth from other users resulting in a campus-wide internet slowdown. Recommended supports include: a) mice for page navigation, b) keyboards for response entry, and c) headphones for students to take advantage of the multimedia environment.

Participants

Because the foci of this investigation were the perceptions of implementation and use of SuccessMaker at the middle school level, several criteria for inclusion in the study were developed. Only staff who had three or more years of consecutive years of experience with SuccessMaker were considered for inclusion. Those staff members were to be in an administrative role with responsibilities for program oversight or regular facilitators of the program. Staff must also have been currently employed by the district at the time of the research. Only middle schools were considered for this research as that was the initial focus of the SuccessMaker intervention by the district.

Given those constraints, I was able to identify 20 staff members who met those criteria. This included six principals, seven assistant principals, and seven special education teachers, representing nine middle school campuses. I received affirmative responses from twelve of these staff members, representing a 60% response rate. The final interview sample consisted of three principals, three assistant principals, and six teachers, representing seven middle school campuses.

Materials and Procedure

An instrumental case study approach (Creswell, 2007; Stake, 2005) was employed to determine campus perceptions of SuccessMaker and the implementation process. Instrumental case studies examine a specific case to develop insight into a particular issue. The perceptions of teachers and administrators regarding the implementation of SuccessMaker in middle schools in a central Texas school district will be considered the instrumental case which is believed to be illustrative of SuccessMaker implementations elsewhere and technology-based interventions collectively. A structured interview was conducted with each of the participants following their provision of informed consent for participation in the study. The following questions were asked of each participant:

1. In your own words, please describe SuccessMaker Mathematics (SMM). What is it, and how is it used?

2. What was your role selection and implementation of SMM for your campus? How did your participation in these processes affect your perception of the program and your use of the program?

3. Describe your implementation process. Where did you find time during the school day? How did you establish buy-in from teachers? How did you establish buy-in from the students? How long did it take before implementation felt like a normal part of the school day? What obstacles to implementation did you encounter?

4. Describe your results. How successful have your students been? Does this level of success lead you to believe that the program is useful? How have students responded to the program?

5. Do you believe that ongoing use of the program would be beneficial? Why?

6. What advice would you provide to other campuses considering implementation of SMM?

7. What concerns do you have about future instructional learning system technology implementations?

The interview also included a review of the matrix developed by Mills and Ragan (2002) regarding the implementation of instructional learning systems. Participants graciously allowed me to record these interviews for later transcription; each interview lasted approximately 45 minutes. I transcribed the recorded interviews and analyzed them using the constant comparative method (Glaser & Strauss, 1967). I employed open, axial, and selective coding (Creswell, 2007). Open coding involves reading the transcriptions and generating codes. Axial coding involves re-reading the transcriptions, re-wording the codes, and organizing them into categories. Those categories included perceptions of the program and student response to the program, identified obstacles to implementation, implementation strategies, and recommendations regarding future instructional learning system technology interventions. Responses regarding the implementation matrix and the open coding were reassembled into themes that demonstrated consistency across participants. Selective coding involves reading the transcriptions for a third time to delimit and organize the main theme and categories. As mentioned above, participants were provided an opportunity to review my assembled findings and offer corrections and feedback. These were integrated into the final manuscript. As much as possible, I have used the participants' own words to describe their experiences and perceptions.

Data regarding the historical use of the program, as well as aggregated results, were compiled for analysis elsewhere (McKissick, 2016). These campus data profiles provided a window through which interview responses were viewed. Usage patterns may highlight, compliment, or even demonstrate a contrast between participant perceptions and actual student use. The usage pattern for the district is provided in Table 3.1, and illustrative patterns of implementation at the campus level are provided in Tables 3.2 – 3.4.

Perceptions of the Program

Almost unanimously, participants had a strong positive perception of SuccessMaker. Though they frequently pointed to the specific capabilities of the program, their perceptions were consistently tied to their perceptions of student success as a result of program use. Only two participants had a negative perception of both the program and the level of student success. Tom, a special education teacher at Campus E, explained simply: “The results were outstanding. It sold me.” Nubian, a special education teacher at Campus G, added: “I think is a very, very, very [sic] good program for our kids because the implementation and the success rate are very good, and the kids can see it right away.”

Participants consistently identified five features of SuccessMaker that affected their perceptions of the program. Assessment, primarily done through the initial placement process, was seen as a positive feature of the program. Phil, a special education teacher at Campus F, was impressed by the accuracy of the assessment, saying: “If I look at paperwork, and it comes to us with a 3rd grade ability in math

Table 3.1

SuccessMaker Usage Profile for the District

	2010-11	2011-12	2012-13	2013-14	2014-15
Number of Students Using SuccessMaker Mathematics	351	527	438	445	378
Student Usage Levels					
0-5 Hours	135	98	78	88	64
5-10 Hours	125	147	126	93	72
10-15 Hours	65	101	84	49	77
15-20 Hours	20	65	52	49	59
20-25 Hours	5	37	48	38	39
25 or more Hours	1	79	50	128	67
For Student with 5 or more Hours of Usage	Mean (Standard Deviation)				
Number of Years Behind at the Start of Year[a]	3.08 (1.47)	2.65 (1.17)	2.83 (1.18)	2.94 (1.23)	2.75 (2.45)
Gain During Usage	.31 (.22)	.55 (1.28)	.50 (.38)	.68 (.57)	.54 (.42)
Time Using SuccessMaker Mathematics	10:10 (4:06)	15:56 (9:31)	15:51 (9:45)	24:05 (17:23)	19:19 (13:52)
Percent of Users Who Passed End-of-Year State Test	62.9%	-----[b]	57.6%	59.4%	13.5%[c]

Note. Students with less than 5 hours of program usage were removed from aggregate data because low usage may indicate that the student moved during the school year or that usage was restricted to initial placement only.
[a]Mean number of years behind a start of year reflects the difference between enrolled grade level and initial placement identified grade level.
[b]State testing results for mathematics were not reported for 2011-12 as a result of changes in the assessment instrument.
[c]The STAAR-Modified version of the state assessment was no longer available, and most students who took the STAAR-Modified during the previous year took an on-grade-level assessment this year.

Table 3.2

SuccessMaker Usage Profile – High Implementation

	2010-11	2011-12	2012-13	2013-14	2014-15[c]
Number of Students Using SuccessMaker Mathematics	25	31	24	34	-----
Student Usage Levels					
0-5 Hours	0	8	4	1	-----
5-10 Hours	4	13	2	4	-----
10-15 Hours	16	8	1	0	-----
15-20 Hours	5	2	0	1	-----
20-25 Hours	0	0	1	2	-----
25 or more Hours	0	0	16	26	-----
For Student with 5 or more Hours of Usage	Mean (Standard Deviation)				
Number of Years Behind at the Start of Year[a]	3.23 (1.32)	1.90 (1.15)	3.00 (1.07)	3.58 (1.27)	-----
Gain During Usage	.41 (.20)	.27 (.18)	.78 (.39)	1.16 (.57)	-----
Time Using SuccessMaker Mathematics	12:26 (2:43)	9:50 (3:36)	29:43 (11:03)	41:47 (17:25)	-----
Percent of Users Who Passed End-of-Year State Test	88.0%	-----[b]	55.0%	69.7%	-----

Note. Students with less than 5 hours of program usage were removed from aggregate data because low usage may indicate that the student moved during the school year or that usage was restricted to initial placement only.
[a]Mean number of years behind a start of year reflects the difference between enrolled grade level and initial placement identified grade level.
[b]State testing results for mathematics were not reported for 2011-12 as a result of changes in the assessment instrument.
[c]The campus did not use SuccessMaker during this school year.

Table 3.3

SuccessMaker Usage Profile – Still Developing

	2010-11	2011-12	2012-13	2013-14[c]	2014-15
Number of Students Using SuccessMaker Mathematics	23	27	37	-----	55
Student Usage Levels					
0-5 Hours	18	17	13	-----	8
5-10 Hours	5	10	10	-----	7
10-15 Hours	0	0	11	-----	14
15-20 Hours	0	0	3	-----	11
20-25 Hours	0	0	0	-----	11
25 or more Hours	0	0	0	-----	4
For Student with 5 or more Hours of Usage	Mean (Standard Deviation)				
Number of Years Behind at the Start of Year[a]	1.65 (.86)	1.88 (.74)	2.96 (1.09)	-----	2.49 (.81)
Gain During Usage	.23 (.09)	.19 (.08)	.40 (.29)	-----	.57 (.30)
Time Using SuccessMaker Mathematics	7:53 (1:32)	6:45 (1:18)	11:08 (3:10)	-----	16:36 (6:21)
Percent of Users Who Passed End-of-Year State Test	80%	-----[b]	70.8%	-----	21.3%[d]

Note. Students with less than 5 hours of program usage were removed from aggregate data because low usage may indicate that the student moved during the school year or that usage was restricted to initial placement only.

[a]Mean number of years behind a start of year reflects the difference between enrolled grade level and initial placement identified grade level.

[b]State testing results for mathematics were not reported for 2011-12 as a result of changes in the assessment instrument.

[c]The campus did not use SuccessMaker during this school year.

[d]The STAAR-Modified version of the state assessment was no longer available, and most students who took the STAAR-Modified during the previous year took an on-grade-level assessment this year.

Table 3.4

SuccessMaker Usage Profile – Peaked and Falling

	2010-11	2011-12	2012-13	2013-14	2014-15
Number of Students Using SuccessMaker Mathematics	23	78	18	41	52
Student Usage Levels					
0-5 Hours	14	22	6	0	19
5-10 Hours	8	34	11	5	5
10-15 Hours	1	16	1	5	10
15-20 Hours	0	5	0	6	14
20-25 Hours	0	0	0	4	4
25 or more Hours	0	1	0	21	0
For Student with 5 or more Hours of Usage	Mean (Standard Deviation)				
Number of Years Behind at the Start of Year[a]	4.34 (2.17)	2.73 (1.12)	2.08 (1.48)	3.28 (1.14)	3.21 (.90)
Gain During Usage	.33 (.21)	.33 (.22)	.32 (.12)	1.05 (.70)	.39 (.25)
Time Using SuccessMaker Mathematics	7:50 (2:18)	10:11 (4:52)	8:22 (1:40)	31:07 (17:48)	14:52 (4:19)
Percent of Users Who Passed End-of-Year State Test	77.8%	-----[b]	50%	70.7%	9.1%[c]

Note. Students with less than 5 hours of program usage were removed from aggregate data because low usage may indicate that the student moved during the school year or that usage was restricted to initial placement only.

[a]Mean number of years behind a start of year reflects the difference between enrolled grade level and initial placement identified grade level.

[b]State testing results for mathematics were not reported for 2011-12 as a result of changes in the assessment instrument.

[c]The STAAR-Modified version of the state assessment was no longer available, and most students who took the STAAR-Modified during the previous year took an on-grade-level assessment this year.

according to previous testing and previous teachers, 95% of the time SuccessMaker is going to put them within half a grade of that. If it doesn't, I worry about previous teachers." This function moves seamlessly into the instructional feature of the program. Kelly, a special education teacher at Campus L, described these processes: "Initially, they go into the program and they're tested for probably the first week that they are on it. Then it develops a pathway for them, depending on their needs." Instruction varied as students demonstrated success or difficulty, though the program also incorporated periodic review of previously mastered skills. This spiraling of instruction was frequently identified as a key feature of the program. Paula, an Assistant Principal at Campus F, noted, "It's spiraling the stuff they are weak on and then introducing new concepts at the same time. And, the benefit of that is [that] the kids can see where those prerequisite skills are coming into play." The program provided feedback to the students, both regarding their immediate performance on an exercise as well as their cumulative performance and usage. Additionally, the program provided usable reports to staff regarding student needs and performance. Allan, an Assistant Principal at Campus L, described the individual student report: "It was a nice report that we got from SuccessMaker which would show their usage, the skills that they worked on, and their progress toward them. So it was something which, in my opinion, parents enjoyed looking at and was tangible for them to be able to see progress." Michael, the Principal at Campus L, noted that students frequently asked for color printouts of their reports to show to parents. "They will ask you to copy these reports because there's a lot of green [indicating skill mastery] and parents know that green is good."

The perception of student gains was important to most participants in identifying whether or not the program was useful or should be continued. Betty, the Principal at Campus G, observed: “In the past, the [state testing, TAKS and STAAR] with the kids before SuccessMaker was not that successful, about 25% of our SPED kids passed. Once we implemented SuccessMaker, it got much better. It probably went up to about 50%, which was a great increase for those kids who’ve never passed before.” Her special education teacher, Nubian, agreed. “For the first time, you see kids crying because they actually passed the STAAR test.” Students from Paula’s campus have averaged high usage levels for the past few years. She noted: “We have had kids that have been on for three years that in a year’s time, they have gained three years. We have had a huge amount of progress. We have kids that have been able to work out of resource and go into a regular classroom. I’m not talking about the inclusion classroom. I’m talking about the regular classroom. They are so excited because they came to us in sixth grade and they were barely on the third grade, and they are on grade level, and being successful.”

Not everyone had a high opinion of the program. From the onset of implementation, Carmen, the Principal at Campus K, was concerned about the remedial instruction: “What I found was that it was not necessarily teaching students at grade level. It was actually going down and trying to take them…it was like the students would never get to grade level.” Kelly, a special education teacher at Campus L, did not consider the gains from the program useful. She stated: “I would say maybe ten percent have made at least a year’s growth. Probably more like 80 percent have stayed [at the same rate of growth]. The last ten percent have done the other way." Both of these educators were quick to note the value of a live person over reliance on a machine. Carmen said, "I still think people are better than a program. If you hire well, then that would be way better than a program because there is going to be a relationship there. The kids are not going to have a relationship with the program." Allan, though positive in his perceptions of the program, agreed: "I think it's the relationships that teachers have with their kids that contributed as much to the success of the program as the program itself."

Perceptions of Students Using the Program

Perceptions of student engagement and program usage were generally positive. When asked about student engagement, only Kelly provided a consistently negative response. She offered: "Of course they say, 'This is baby stuff; I know all this.' But if you test students on it, they don't know it. They don't want to get on the computer and see these little cartoons." The principal at Kelly's campus, Michael, disagreed: "The kids are getting more enthusiastic about learning, trying things that they wouldn't have tried before. It really did have a positive effect on kids that they would put forth more effort. All of our kids that showed growth felt better about themselves. That's going to make them feel better about what they're taking [such as] the test that's in front of them." Paula identified the same description of the program, but found the student response to be very different: "Now we get some that say this is too babyish. But when you turn away, they are the ones that are just as involved as the others. They just don't want to admit it." Tom, a special education teacher at Campus E, suggested that motivation may not be universal. However, “the ones that are motivated, there’s definite improvement. They want to come and work on it. They think it’s fun. Once they get to the end of the module, they get a little game.”

Opportunities to personalize one’s avatar and play embedded games stood out as specifically engaging to numerous participants. As Nubian described, “I thought they were goofy-looking, but the kids loved their avatars. When they did something right, the little avatar would jump or do something crazy. They thought it was kind of cool.” Tom considered the games to be not only a reward for success within the program but also educational in their own right. The games required the students to use previously learned skills, adding a speed dimension to their knowledge. For other participants, the simple use of technology was sufficient to spark student interest. Carmen, whose campus did not emphasize the use of the program, noted: “They like anything with technology. They liked games and things like any program has.”

Several participants noted that boredom was a real possibility. For Phil, the issue was the repeated use of similar visual stimuli. He also suggested that student boredom may set in sooner unless the teacher maintains an active presence in the learning environment. “Stay in touch with your kids while they’re working on it. They’re going to get bored if you let them.” Kelly suggested that the repetitive practice itself may be the key to student boredom: “There was a lot of repetition of skills that they were having a hard time with. They kept harping on and harping on, and the kids were getting frustrated and…they were reluctant to get on it. Students became bored of it.” Susan, a special education teacher at Campus C, recognized student boredom, but found that introducing incentives and rewards was an effective solution to this issue.

Themes

Through the coding process, five themes emerged as consistent and relevant to the implementation and perceptions of SuccessMaker. These five themes share some overlap, though they may also serve as discrete issues to be considered. The discussion below considered the issues and obstacles identified during the implementation process as well as strategies used by the campus to overcome those obstacles. A discussion of ILS integration components from the ILSCM developed by Mills and Ragan (2002), as they relate to this specific program implementationis included within each theme. Where applicable, advice from the participants was included within the theme.

Theme 1 - Be Prepared!

SuccessMaker was and is a technology-dependent intervention program. Early versions, including the version provided for the campuses in this study, leveraged software on the computer to support internet-based applications. Issues with technology were identified by all twelve participants. Phil was quick to identify the limitations of the software being utilized. He noted that SuccessMaker was Java-based, and issues with Java presented frequent interruptions for students at his campus. Tom noted that SuccessMaker did not run well on web browsers other than Internet Explorer. Limited bandwidth became a problem as large groups of students attempted to simultaneously access the internet to use SuccessMaker. Michael identified the fundamental relationship between these two: “Usage went up, and speed went down.” Slow internet speed is made worse by power outages. Carmen was quick to point out that technology can fail, and that can lead to frustration. Teachers must have a backup plan. Susan agreed, noting that usage at her campus dipped as a result of the technology being frequently unusable or unavailable. A limited number of student usage licenses were purchased for each campus. These licenses limited the number of students who were able to simultaneously access SuccessMaker. This limitation was paired with the lack of available computers at the campuses.

The need for initial and ongoing training was identified as a key aspect of a successful implementation. All teachers, as well as four administrators, acknowledged that they had participated in the initial training as SuccessMaker was being brought to the district. Nubian indicated that this initial training was full day in length and provided by a company representative. Phil noted that several times a year, the representative would visit with him and his campus administrator, Paula. Ongoing training was provided, as needed, during such on-site visits, though additional training was available at the vendor's web page. Not all participants appeared aware that training regarding program updates was available from a source other than the on-site representative. Nubian identified that she would take some time as the school year was starting to walk through the program again to see what changes had been made. Component 1 from the ILS matrix from Mills and Ragan (2002) concerns the design of the courseware. All participants in this study except one affirmed that SuccessMaker describes enrollment levels and mastery of skills. Seven also noted that the program provides individually prescriptive programming for users. Variance in training may account for an inconsistent understanding of the program's capabilities. Component 2 from this matrix concerned courseware training, and all participants (except for four administrators) affirmed their receipt of initial and ongoing training.

A clear recognition of the need for training and additional technology was evident as participants provided advice for subsequent program implementations. As Allan put it, "If you're thinking about implementing it, first you need to make sure that your teachers will have the resources, whether it be the tools for making the program work or the necessary training associated with it." Sarai, an assistant principal at Campus E, received no training. She identified that training was important, but it was more important for her that the teachers using the program be well-trained. Michael and Kelly both noted that Campus L invested in additional computers so that students would have access to SuccessMaker as needed. Component 15 of the Mills and Ragan (2012) matrix concerns the degree of teacher knowledge regarding courseware capabilities. Five teachers and one administrator self-reported skills consistent with the highest level of familiarity; Four others, including three administrators, indicated little to no familiarity with the content or resources available.

Theme 2 - Have a Plan!

In addition to equipping staff with the necessary equipment and training, there is a need for clarity in the purpose and use of the program. The district where the participants work provided SuccessMaker for the exclusive use of students with identified disabilities served under the Individuals with Disabilities in Education Act (IDEA). Phil identified the challenges with this group of students: "Their disabilities affect everything they do and every piece of information they process. No two are alike. If two come to you with learning disabilities and they're both working math at the 4th grade level, I guarantee you – give me a couple of days and I'm going to see big differences between those two kids, as similar as they may sound." In addition to specific identified disabilities and learning styles, the achievement inconsistencies among this group present challenges. Tom noted that "It's like, go on Christmas break and you have them at this level, and they get back on [two weeks later] – all of the sudden, they're regressing." Identifying a targeted level of usage is also important. Eight of the twelve participants in this study identified a common target for use identified by the district – 15 to 20 minutes of daily use for a total of at least 20 hours during the school year. Tables 3.2 – 3.4 provide evidence that such consistency is difficult.

One strategy that facilitated regular use of the program was integrating SuccessMaker into the resource mathematics curriculum. Students in the resource classroom were often multiple years below grade level in the content area. SuccessMaker's ability to provide instruction at the student's grade level or to target particular skills was seen as advantageous by the users. Five of the seven campuses represented identified this as an effective strategy. Susan identified that her campus' decision to move towards an all-inclusion schedule interfered with her ability to have students access the program. Components 5 and 6 from the matrix developed by Mills and Ragan (2002) were concerned with integration of the program within classroom instruction and curriculum. Participant responses identified integration with classroom instruction at Campuses E, F, and L. All participants noted that SuccessMaker supplemented district curriculum, and five participants believed that the program aligned with or supported the curriculum.

Teacher facilitation of program use was also identified as a key strategy for improving student performance. Component 11 of the ILS matrix from Mills and Ragan (2002) considered acceptable use to include occasional facilitation of program use. All participants except Susan identified at least this level of facilitation. Because her campus had moved away from a resource classroom model, students used the program in the inclusion classroom or in a content mastery area where non-disabled students also received support. The opportunities for monitoring student computer use dropped as the number of students needing support grew. Others, including Phil, recognized ongoing facilitation as essential to the learning process. He said, “You’ve got to stay involved with the kids. You’ve got to continue to teach during SuccessMaker. Now the kid maybe has a little bit better understanding, realizes I’m involved with it, and is maybe just a little bit more interested in the next question.”

The development of program usage goals is also a useful tool to guide implementation. Mills and Ragan (2002) identified this as Component 4 of their ILS implementation matrix. Participants from Campuses F and L, where participation was consistently higher, identified that goals were established and celebrated. Michael, the principal at Campus L, provided this statement of purpose: “Our goal was to give kids the opportunity to see their own growth, celebrate their growth as they went, and then to use data to inform us on kids’ growth and abilities.” Campuses with less consistent usage may have only identified the district usage expectations or did not reference any campus goals for the program. Chris, a special education teacher at Campus D, stated, “The goal was for everyone to make a year’s gain.”

Several participants provided advice regarding the speed of implementation. Michael advocated a trial run for a select group of students. Others, including Paula and Nubian, suggested that campuses looking to implement the program should do it wholeheartedly from the start. In Paula’s words, “My recommendation is not to take baby steps. My recommendation is to jump straight in.” She added that the sooner regular use can be established, the sooner the campus will see gains in student improvement. Sustained over time, students may make significant academic gains. Most participants were able to remember at least one student who made at least a one-year gain in mathematics while using the program or who had been moved out of a resource setting because of demonstrated abilities.

Theme 3 - Make Time for Use

For a program to be well-implemented and regularly used, there must be time in the day for students to use the program. The challenge, identified by participants in this study, was that the school schedule is often too rigid to allow the inclusion of any program requiring a fixed daily usage. Tom summed it up like this: “You’ve got to give up some instructional time or find some time during the day to do it.” Chris, whose campus has had few regular users in the years since the program was rolled out, added this: “Initially, when it first came out, I wasn’t very happy with the program due to the fact that we didn’t have the time allotted for the program. This past year, I made sure I implemented the program and I do see a lot of gains in it.”

Most users identified a philosophical shift at the district level that facilitated program use. In the third year of program availability (2012-2013), the district adjusted the schedule for all middle schools to introduce a double-blocked mathematics period as a strategy for improving mathematics achievement scores. Borrowing a few minutes from each class period, an extra class period was built into the school day. Doubling the instructional time allotted to mathematics allowed campuses and teachers to make better use of the program. Chris noted, “They had two fifty-minute math classes. It was a lot easier to let them off the last twenty minutes of the second period versus twenty minutes of a fifty-minute class. That wasn’t effective.”

Finding additional time during the day has resulted in a variety of strategies. Students at many campuses have the option of coming in before school or after school to use SuccessMaker. Students are also able to use the program from home through the district’s website using personalized log-in credentials. From Tom’s experience: “If you get the parents’ buy-in, they could use it at home. They had home access. If they were going to be absent or something, and I knew about it in advance, [the parents] would email me. I could say, what he can do is log on to SuccessMaker every day for 20 minutes.” Betty, recalling her students’ likelihood of using the program from home, said, “If you had twenty [students], maybe two out of the twenty would use it at home.”

The most effective and common strategy for establishing regular student use was the introduction of a routine. Mills and Ragan (2002) identified this need in Components 8 and 10 of their implementation matrix. Participants in this study consistently rated their scheduling as at least acceptable, indicating that some students were scheduled for regular use. The restricted implementation of SuccessMaker in this district – allowing only students receiving special education support to access the program – caused confusion for respondents as they graded their implementation by the standards in the matrix. With the exception of two principals, all participants believed that students received at least 30 minutes of instruction per week using the courseware. Usage patterns in Tables 3.2 – 3.4 demonstrate a discrepancy between this perceived usage level and actual usage for three campuses. The use of SuccessMaker as part of the resource curriculum for mathematics, as well as the introduction of a double-blocked mathematics instructional period, did yield usage improvements. Table 3.2 demonstrates the sudden increase in student usage attributable to these strategies.

Theme 4 - Use a Team Approach

The choice to use SuccessMaker in this district was a top-down decision. Campuses were not included in the selection of the program; rather, the special education department identified a program believed to be beneficial for its students in need of academic support. That top-down decision making was also seen at the campus level in several instances. When asked about her role in the program selection process, Kelly replied, "I wasn't a part of that. It was a done deal. This is coming to your campus." Allan, an administrator at Kelly's campus, discussed the challenges presented by teachers reluctant to release students from class to use the program: "The problem was that they [the teachers] just didn't want to send them. I would get involved and say, 'This is non-negotiable; the kid has to go. This is an expectation.' They didn't really have too much of a choice."

Despite the limited roles afforded to the campuses in the program selection process, many participants in this study considered their roles in the implementation process to be substantial. Many self-identified as the program leader for their campus. Those who identified their role in implementation to be minimal tended to have lower perceptions of the program. Campuses F and L, where usage per student reached high levels during the 2013-2014 school year, had administrators and teachers who both identified themselves as program leaders.

Of particular interest across the participants' responses was the high value placed on the role of the vendor representative. Betty had this to say about the representative: "She worked really hard to help us design custom courses. She worked one-on-one. She came in on weekends when we had camps to work with teachers with those kids. She put a lot of time and dedication into it, and she made that program very successful. She was instrumental in making sure that we did a good job with the program." Phil, Tom, and Nubian agreed regarding the representative's support. They cited frequent meetings with the representative as well as the ability to contact her whenever necessary to work through technical issues and explore program capabilities. When the program changed its representative for the district, participants found the change to be counterproductive. In Michael's assessment, "She is not as positively overwhelming. The program is only as good as the person selling it." Nubian has also felt the change: "Things have changed drastically. I would like to go back to that [representative], because the production, in my opinion, dropped off."

Participants also identified that the campus implementation team needed to include a variety of members. Campus technologists were considered critical because the program was rooted in technology. Several participants cited the technologist's role in resolving issues such as out-of-date software, internet access issues, and computer availability. The campus administrator was seen by teachers as a necessary ally in implementing the program. Nubian pointed to her need to involve administrators to pull students from class to use the program. The principal at Campus E was perceived as an obstacle by Tom. Despite an increasingly high level of implementation through the 2013-2014 school year, the principal discontinued use of the program the following year following a disagreement with the vendor's representative. As he recounted, "The biggest thing is teachers can't do it alone. They need to have administrator buy-in. If they don't have the administration buy-in, it's not going to be a successful program." Participants also identified the parent as a helpful member of the implementation team for the individual students. Several recalled conversations with parents about the use of SuccessMaker at home, though they also noted that follow-through was lacking. The use of SuccessMaker reports to guide the development of Individualized Educational Plans for special education students necessarily involved the parent in the implementation process. Teachers noted that these reports clearly identified student needs and abilities for parents. Component 3 from the implementation matrix developed by Mills and Ragan (2002) considers the ongoing meetings to discuss courseware. Participants routinely noted that building level meetings, often with the vendor, were regularly held during the early implementation years.

Theme 5 - Provide Feedback

As students use SuccessMaker, they receive feedback in a variety of ways. The program provides immediate feedback as each exercise is completed. Students may also receive periodic updates via the program regarding their overall improvement. Mills and Ragan (2002) included three components in their ILS implementation matrix regarding the use of reinforcement and student reports. The use of group motivational strategies was perceived by most participants to be inconsistent with student privacy concerns, especially for students with disabilities. However, the use of individual feedback was common. Though not part of any district directive, five of six teachers indicated that students received some form of grade for their work on SuccessMaker. Four of six administrators responded that courseware was not included in student grades, representing an area of disconnect between teachers and administrators.

Group and individual recognition has been provided at several campuses in ways that minimize privacy violations. Often, students receive recognition within their resource class. As all students in the class have some disability and are well below grade level, privacy issues are not seen as obstacles to praise and reward strategies. Participants from several campuses recalled an award given by Pearson to the district and, ultimately, the campuses, for high levels of student usage. Michael recalled a specific celebration held to honor the students with the greatest growth while using SuccessMaker. Parents and teachers of these students were invited to attend, and certificates of achievement were given to the students as their achievements were announced. Michael noted that these students and their parents “had never had that ceremony before. Most of those kids were lucky to get B's and C's, so for somebody to say, 'You are the top performer in math in SuccessMaker,' that was the type of thing that those parents had been wanting forever."

Discussion

Participants in this study identified many of the barriers identified by other researchers. Rather than allow those issues to remain obstacles, schools where implementation fidelity improved over time found ways to respond to those obstacles productively. In some instances, obstacles were removed as a result of district planning. The move to a double-blocked mathematics period began as a structural change at one campus but has now spread to all middle schools in the district. Technology limitations were overcome through targeted resource acquisition, improvement in district capacity (internet bandwidth), and support from campus technologists. At only one campus has an obstacle persisted and inhibited implementation – the preference for human support over technological support for providing remedial and on-level instruction for students with disabilities.

The five themes developed through these interviews are clearly interdependent. The call to "be prepared" presupposes "having a plan" for implementation in addition to securing needed resources and training. "Having a plan" may require changes to the master schedule that "make time for use" easier to accomplish. The development of a core team and processes for feedback should be considered as part of the plan.

At almost all campuses identified in this study, the campus level of implementation is consistent with the perceptions of the users from that campus. At

Campus L, low teacher perceptions of the program were overcome by positive administrator perspectives and top-down directives regarding program use. Positive teacher perceptions of the program at Campus D have to date yielded minimal implementation, though there is a developing clarity about future implementation steps for that campus. As Straub (2009) noted, "Successful facilitation of adoption is most likely to occur at the intersection of the cognitive, affective, and contextual factors" (p. 644). Results from this study appear to validate this position.

Whereas other implementation frameworks might suggest that implementation proceeds through a series of stages (Straub, 2009), the experiences of the participants in this study are centered in two activities. The CBAM framework includes room for teachers to engage in information-seeking and determination of the personal cost of an implementation. The nature of the implementation under review here did not allow room for this personal buy-in to be developed. Rather, teachers were thrust into the logistics management stage wherein the details of integration were of primary concern. Teachers also entered the consequence stage of the CBAM model as they considered the effects of SuccessMaker on their students. Teachers found little room and need to compare their work with peers, though refining the implementation occurred for some participants. Earle (2002) posited that "change starts with the individual teacher" (p. 11). In the context of this district's implementation, change started with the district purchasing the program for campus use. Teachers – end users of the technology – were directed to implement the technology. Early versions of the ILSCM (Mills & Ragan, 2002) included a component for participation in ILS selection; it was dropped from their final version because no teachers reported participating in the selection process. The current implementation likewise found a lack of campus personnel participating in the selection process.

This study identifies the limitations of the ILSCM as developed by Mills and Ragan (2002), even for SuccessMaker implementations. Though some campuses (such as Campus F) had a teacher and an administrator who rated themselves highly on all ILSCM components, other campuses had perhaps one or the other. Some components, such as the development of clear rules for program use, may be considered part of a school district's acceptable use policy for all technologies. Formal grade or department meetings, part of a component for an ongoing support system, may need refinement in implementations where limited staff will be involved. Several components considered aspects of training and demonstration of knowledge regarding the use of the program; these may instead represent a single research construct. Consideration of the parameters of each specific technology implementation ("having a plan") might result in development of different, targeted configuration matrices that account for context, target audience, usage expectations, teacher capacities, and identified barriers.

Bowman et al. (2001) identified that previous research has included little investigation of the impact of technology implementations on teachers. Almost all participants in this study indicated that SuccessMaker had positively impacted their students. Having their students with disabilities grow academically, and in some cases succeed on state assessments, was sufficient evidence of program effectiveness and a reason to continue using the program. Recollections of specific students who had experienced significant achievement gains led participants to get animated during interviews. Though the foci of this study have been implementation issues and program perceptions, teachers who have used the program and experience success are eager to continue using the program and improving their fidelity. They felt a need for targeted intervention for these struggling learners, and they expressed clarity regarding the capability of SuccessMaker to fill that gap.

CHAPTER IV

EFFECTIVENESS OF LONGITUDINAL ILS USAGE FOR STUDENTS WITH DISABILITIES

Overview

Longitudinal research regarding mathematics interventions is limited, as is longitudinal research investigating the impact of instructional learning systems on the mathematics achievement of students with disabilities. Previous research has been inconsistent regarding achievement trajectories for students with disabilities. This study used longitudinal and multilevel modeling to determine what factors related to the use of SuccessMaker Mathematics might account for changes in student achievement over time. A sample of two-year program users (N = 183) was found, and four usage patterns were identified. Between-group homogeneity was established or accounted for in analyses. No student-level or group-level factors were identified to account for more variance than a baseline model without factors. Subsequent ANOVAs found that student achievement trajectories, regardless of program use, trended consistently downward and demonstrated an increase in the achievement gap over time. SuccessMaker Mathematics was not found to impact this decline in student performance relative to non-disabled peers.

Introduction

State and national standards have been set through legislation such as No Child Left Behind and Race to the Top which hold schools accountable for demonstrating yearly student achievement gains in mathematics. Students with disabilities have historically underperformed on almost all state assessments relative to their non-disabled peers. Numerous states have adopted achievement benchmarks for various sub-populations, including students with disabilities, and district performance ratings are tied to the performance of these sub-populations. It has become increasingly important, then, for school districts to identify and provide appropriate supports for these students (Manning, 2004). Students with disabilities often make only small achievement gains, especially during the middle school years (Graham, Bellert, Thomas, & Pegg, 2007). Pressure to assist students with disabilities, especially those with learning disabilities, has increased (Martindale, Pearson, Curda, & Pilcher, 2005) while the gap between high and low achievers has grown wider each year (Cawley, Parmar, Yan, & Miller, 1998).

SuccessMaker Mathematics (SMM) provides a possible solution for addressing the disability achievement gap. Early work on the program began in the late 1960s (Suppes & Morningstar, 1969), and sufficient research enabled meta-analytic reviews by the 1990s (Kulik, 1994). Current versions have incorporated modern computing capabilities related to audiovisual interfaces and program speed. SuccessMaker Mathematics is an integrated learning system (Bailey, 1992) that provides targeted computer-based instruction at the student's measured achievement level and allows teachers to track the progress of multiple students across multiple subjects. Student progression through the program is contingent upon continued successful performance using a branching algorithm (Svoboda, Jones, van Vulpen, & Harrington, 2012). Student performance on discrete skills or skill groups can be measured, and SMM provides a recurrent check of mastered skills. As students demonstrate mastery or failure of presented skills, the difficulty level of subsequent material is adjusted by the program. Student usage data is collected along the way to include the numbers of questions encountered and answered correctly, percent of skills mastered, time spent using the program, and initial and current grade placement levels.

Reviews of SMM, individually or alongside other instructional learning systems, have been numerous. Though meta-analyses (Becker, 1992; Kulik, 1994) have generally found small but significant effects for SMM usage, previous research has been plagued with methodological issues. Slavin and Lake (2008) identified design flaws in eleven studies, and McKissick (2016) identified that only six studies had been subjected to peer review. Most original research was done throughout the 1980s, leaving the impact of more recent versions of SMM untested. Corporate-sponsored research (Gatti, 2009), institutional reports (e.g., Laub & Wildasin, 1998) and doctoral dissertations (Gee, 2008; Kirk, 2003; Manning, 2004) comprise most of the recent corpus of available research findings.

Two recent investigations regarding the effectiveness of SMM cause some concern regarding the use of program. Tucker (2009) compared the achievement scores for Texas districts where SMM was implemented to achievement scores from a set of control districts not using the program. No significant difference in state assessment passing rates was found between these groups. McKissick (2016) noted that state testing data were problematic given the almost yearly changes in format and scale, and no differences between students at varying program usage levels was found using state testing data. A second study of note was a review of interventions found to be effective for remediating student mathematics achievement (Slavin, Lake, & Groff, 2009). To be considered for inclusion in this study, an intervention must have been at least 12 weeks in length, conducted at the middle or high school level, and met design quality considerations. No studies reviewing SMM were found to qualify for inclusion in the study. This lack of review regarding SMM research was consistent with that done by the What Works Clearinghouse, which has thus far been unable to find sufficient evidence on which to base a judgment regarding the effectiveness of SMM.

The findings regarding the effectiveness of SMM for disabled students are also questionable. A meta-analysis reviewing components of technology interventions found that students with disabilities realized greater gains when using technology than nondisabled students (Li & Ma, 2010). This disparate benefit accounted for 11% of the variation in effect sizes between studies used in their review. For comparison, publication date (before or after 1999) accounted for 13% of the variance between effect sizes while experimental design (true experiment versus quasi-experiment) accounted for 0% of the variance. Their review included a study with a sample size of 6 (Irish, 2002) that led Li and Ma (2010) to conclude that "CAI [computer-assisted instruction] could be an effective mechanism for teaching these special needs students" (p. 220).

The long-term achievement trajectories for students with disabilities are not well-understood. The commonly held belief regarding these students was identified by Judge and Watson (2011) as the "ever-widening mathematics achievement gap with the passage of time for students with LD [learning disabilities]" (p. 154), discovered through longitudinal review of achievement scores. Their review spanned grades 1 through 5, and they found that the gap was present from the start and grew despite the point at which a learning disability was identified. This starting achievement gap was also identified by Jordan, Hanich, and Kaplan (2003), though they found no differences in growth rate between normal students and students with mathematics disabilities.

Although a thorough review of the longitudinal research regarding instructional learning systems is beyond the scope of this study, the limited research is presented in Table 4.1. All longitudinal studies regarding SuccessMaker were conducted prior to 1992, nearly 25 years ago. Advances in computing suggest that the findings from these studies may no longer remain true for modern program versions. Several reviews have discovered that achievement has not been improved despite technological advances, as evidenced by a lack of positive trend in outcomes from more recent studies (Cheung & Slavin, 2013; Christmann & Badgett, 2003; Liao, 1998). An unweighted average of the effect sizes in Table 4.1 is $d = .23$, slightly above Cohen's (1988) threshold for a small effect. The marked discrepancy between the two schools in Fitzgerald and Fitzgerald (2002) suggested that campus implementation or other school factor may have impacted program effectiveness.

The length of an intervention has been linked to its effectiveness. Definitions of brief studies range from as little as four weeks (Kulik, 1994) to one term or semester (Li & Ma, 2010). These brief interventions have resulted in larger effects than longer interventions (Cheung & Slavin, 2013; Li & Ma, 2010). Interventions lasting longer than 1 year showed no advantage over interventions lasting one semester (Li & Ma, 2010), perhaps due to novelty effects wearing off over time (Kulik, Schwalb, & Kulik, 1982;

Table 4.1

Previous Instructional Learning System Longitudinal Research

Study	ILS	Duration	Effect Size
Birch (2002)*	Lightspan	2 years	+.28
Estep, McInerney, Vockell, & Kosmoski (2000)*	Jostens	1-5 years	+.02
Fitzgerald & Fitzgerald (2002)	SuccessMaker	3 years	+.55 for Campus 1 .00 for Campus 2
Nunnery & Ross (2007)*	Ace Math	2 years	+.20
Ragosta (1983)*	CCC/SuccessMaker	3 years	+.36
Resendez, Azrin, & Strobel (2009)*	EnVision	2 years	+.35
Schmidt (1991)*	Wasatch ILS	1-3 years	+.17
Spencer (1999)*	Jostens	5 years	+.40
Becker (1992)†			
Chicago	WICAT	3 years	+.10
Sandy, UT	Waterford	5 years	+.23
Calvert Co., MD	CCC	5 years	+.10
Los Angeles	CCC	3.5 years	+.27

Notes: CCC refers to the Computer Curriculum Corporation, original developer and vendor for SuccessMaker. *Included in Cheung and Slavin (2013). †Becker (1992) did not provide references for these studies.

Kozma, 2001). Program intensity appeared to have an inverse parabolic effect, with a medium level of intensity (between 30 to 75 minutes per week) yielding an effect size of .20 whereas low use (less than 30 minutes per week) and high use (more than 75 minutes per week) yielded effect sizes of .06 and .14, respectively.

The ongoing underperformance of students with disabilities suggests that lengthier interventions may be necessary to close mathematics achievement gaps. It is doubtful that brief interventions of less than one semester and a medium level of intensity would be sufficient to address these gaps. The issue becomes even greater when considering students at the secondary level. McKissick (2016) found that students with disabilities averaged a rate of growth of approximately .62 (compared to assumed nondisabled student growth of 1.0) per academic year, resulting in multiple grade-year academic deficits by the time students reach middle school. The need for interventions to close these gaps is great. This study considers whether multiple years of SuccessMaker Mathematics usage at or above recommended levels results in greater gains than single-year program usage.

Methods

SuccessMaker Mathematics has been purchased for a central Texas school district for the past 5 school years for use with students with disabilities served under the Individuals with Disabilities in Education Act. A fixed number of licenses were provided to each middle school campus, and campuses were given flexibility regarding their implementation approach. Campuses were provided with a vendor recommendation of yearly usage totals of 20-25 hours per student consistent with both current on-site vendor recommendations (D. Wayland, personal communication, January 28, 2016) as well as a published matrix of time recommendations based on initial placement level and anticipated gains (Pearson Education, 2012). Because the intervention was provided only to students with disabilities, random assignment and control group assignment were not permissible. Ethical issues regarding the withholding of a potentially effective intervention were also identified. As a result, a within-subjects design was developed to consider whether multiple years of program use at or above recommended levels yielded results that were significantly different than recommended usage for only one academic year.

Participants

Despite the program's availability for all students with disabilities at all middle school campuses, only 282 students were found that had received multiple years of program usage. Of these, 26 were excluded from this study as their usage was in nonconsecutive years. The remaining 256 students were analyzed for between-group homogeneity. Variations between disability groups were identified regarding their initial achievement gap (number of years behind grade level at the start). Consequently, only students with learning disabilities and "other health impairments" (often, attention deficit-hyperactivity disorder) were included. The final sample consisted of 183 students. Demographics for the sample are included in Table 4.2.

Between-group homogeneity was established prior to analysis. An analysis of variance found no difference in initial achievement gap between the four usage groups, $F(3, 179) = 2.14$, $p = .097$. No differences between usage groups were found for ethnicity, $\chi^2(9) = 8.626$, $p = .4725$. No differences between the usage groups were found for disability, $\chi^2(3) = 2.633$, $p = .4518$. Chi-square analysis identified variations in usage groups for gender, $\chi^2(3) = 8.407$, $p = .0383$. It was found that more males were in the Limited-Full group than expected, though more females than expected were in the Full-Full group. Previous research has found no gender impact on treatment outcome (McKissick, 2016), and the groups were considered sufficiently similar for further investigation.

Materials and Procedure

SuccessMaker Mathematics use begins with an initial placement of each student. Students complete a series of questions from a preset level of difficulty, increasing or reducing the difficulty of subsequent questions based on the accuracy of student responses. The process may take up to three hours (Pearson Education, 2012) and may include up to 300 questions (Wood, 2004). Reports available from SMM identify the starting grade level of the student as their IP, or initial placement, score. This score was taken to be the baseline achievement level for each student. The initial achievement gap was calculated by subtracting this IP score from the student's enrolled grade level. Students entering 6th grade who are on grade level, for example, should earn an IP of 6.0 and have a resultant gap of 0.0. An entering 6th grader with an IP of 4.5 would show an initial achievement gap of 1.5 years.

At the end of each school year, students are assumed to be ready for the next grade level. Thus, a 6th grader who is achieving on grade level should have a placement score of 7.0 at the end of the 6th grade. Usage reports provided by the district included a

Table 4.2

Student Sample Characteristics

	Sample	Limited-Limited	Full-Limited	Limited-Full	Full-Full
N	183	101	20	42	20
Male/Female	115 / 68	61 / 40	14 / 6	32 / 10	8 / 12
Black	46%	50%	55%	43%	25%
Hispanic	26%	24%	20%	21%	50%
White	22%	21%	20%	29%	20%
Other	5%	5%	5%	7%	5%
LD/OHI	145 / 38	83 / 18	15 / 5	30 / 12	17 / 3
Grade Levels Below at Time Points	Mean (Standard Deviation)				
Year 1 Start	-2.59 (1.01)	-2.59 (1.07)	-2.28 (0.84)	-2.54 (1.02)	-3.06 (0.65)
Year 1 End	-3.23 (1.02)	-3.33 (1.05)	-2.57 (0.83)	-3.30 (1.04)	-3.30 (0.79)
Year 2 Start	-3.13 (1.02)	-3.06 (1.09)	-2.83 (0.68)	-3.14 (0.91)	-3.76 (0.93)
Year 2 End	-3.64 (1.06)	-3.79 (1.12)	-3.54 (0.78)	-3.22 (1.00)	-3.91 (0.92)
Gain/Loss During Summer	0.10 (0.82)	0.27 (0.73)	-0.27 (0.74)	0.16 (0.92)	-0.46 (0.81)
Gain/Loss from Year 1 Start to Year 2 Start	-0.54 (0.84)	-0.47 (0.78)	-0.56 (0.91)	-0.60 (0.91)	-0.47 (0.78)

final achievement level for each user. A 6th grader with a final achievement level of 5.2 would be considered to be 1.8 years behind grade level at the end of that year.

Achievement gap measures were taken at four times. The first measurement, used in establishing group equivalence, was taken from the initial placement score from SMM at the start of the first usage year. The second measurement came at the end of the first usage year. The third measurement followed the summer break between years one and two. The third and fourth measures were taken from the initial placement and final achievement scores from SMM reports during the second year of usage. Whereas some researchers have used state testing data as a measure of achievement, McKissick (2016) found no differences in state testing outcomes for varying usage groups. Consequently, only SMM-measured achievement levels were considered for this study. The design for this study is modeled in the diagram below, where O_1 represents academic years behind as measured by SMM:

NR	O_{1A}	$X_{FULL\ (>20\ Hours)}$	O_{1B}	O_{1C}	$X_{FULL\ (>20\ Hours)}$	O_{1D}
NR	O_{1A}	$X_{FULL\ (>20\ Hours)}$	O_{1B}	O_{1C}	$X_{LIMITED\ (0\text{-}20\ Hours)}$	O_{1D}
NR	O_{1A}	$X_{LIMITED\ (0\text{-}20\ Hours)}$	O_{1B}	O_{1C}	$X_{FULL\ (>20\ Hours)}$	O_{1D}
NR	O_{1A}	$X_{LIMITED\ (0\text{-}20\ Hours)}$	O_{1B}	O_{1C}	$X_{LIMITED\ (0\text{-}20\ Hours)}$	O_{1D}

Given previous research findings that SMM has a small but significant effect on student achievement, it is predicted that multiple years of program usage at or above recommended levels should result in closing the achievement gap more than any partial-usage pattern identified above. This prediction is consistent with the longitudinal studies identified in Table 4.1, but it runs counter to the claims made by Li and Ma (2010) regarding lengthy interventions.

Results

A longitudinal and multilevel modeling approach was first used to determine what impact, if any, the treatment had on the achievement gap. In this approach, a Level 1 model was developed that accounted for the multiple achievement measurements without any predictor variables. This baseline model is effectively a regression model fitted to the individual student, with the following equation used:

$$Y_{ti} = \pi_{0i} + \pi_{1i}(Time\ Point)_{ti} + \epsilon_{ti} \quad (1)$$

where Y_{ti} is the achievement gap for student *i* at time *t*; "time point" is 0 for the first initial placement, 1 for the end of the first year, 2 for the second initial placement, and 3 for the end of the second year; π_{0i} is the initial achievement gap for student *i*; π_{1i} is the rate of change for student's achievement gap; and ϵ_{ti} is the residual associated with the student's achievement gap at a specific time point, assumed to be normally distributed with a mean of 0 and variance of σ^2.

Level 2 equations allow for student- and group-specific variables to be introduced into the model as predictors that might account for variations between students. For the baseline model, these equations are as follows:

$$\pi_{0i} = \beta_{00} + r_{0i} \tag{2}$$

$$\pi_{1i} = \beta_{10} + r_{1i} \tag{3}$$

where β_{00} represents the mean initial achievement gap; β_{10} represents the mean change in achievement gap; r_{0i} is the variation in initial achievement gap between students; and r_{1i} is the variation in the change in the achievement gap. Both r_{0i} and r_{1i} are assumed to be normally distributed with a mean of 0 with some variance (τ_{0i} and τ_{1i}, respectively) and covariance between them (τ_{01i}).

Parameter estimates from this baseline model are included in Table 4.3. Participants were found to have a mean starting mathematics achievement gap of 2.69 years, and each subsequent measurement of this gap found students an average of 0.30 years further behind. Both of these parameters are significant at $p < .001$. This model accounted for 84.8% of the variance between scores. The variance of the random components of this model, τ_{00}, was .945 and was significantly different from zero, Wald $Z = 8.474$, $p < .001$. Hayes (2006) identified that the use of a likelihood ratio test presents an alternative means of comparing various models. The deviance for this model, using a -2 Log Likelihood measure, is 1536.92. Future models may be compared to this baseline model to determine if additional model parameters significantly reduce this deviance.

Model 2 was developed to determine if any student level factors may have influenced student achievement while using the program. Three dichotomous variables were used to identify ethnicity, one dichotomous variable was used to identify gender, and one dichotomous variable was used to identify disability. The resultant model did

Table 4.3

Parameter Estimates for Multilevel Models with Years Behind as Outcome Variable

	Model 1 (Baseline)	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Level 1 Parameters								
Intercept	-2.69***	-2.44	-2.91***	-2.90***	-2.87***	-2.83***	-2.65***	-2.82***
Time	-0.30***	-0.30***	-0.30***	-0.32***	-0.33***	-0.30***	-0.30***	-0.35***
Level 2 Parameters								
Gender (Male=1)		0.29	0.35*	0.36*	.40*	0.27*		0.35*†
African-American		-0.10						
Hispanic		0.19						
White		-0.09						
Disability		-0.53						
Level 3 Parameters								
Fidelity				-0.35				
------Fidelity*Time				0.15				
BothYrsFull						-0.37*	-0.40*	
------BothYrsFull*Time						0.01	0.00	
Yr1Full					0.19			
------Yr1Full*Time					-0.01			
Yr2Full					-0.34*			-0.27†
------Yr2Full*Time					0.15*			0.12*
Intraclass Correlation (ICC)	0.848	0.737	0.825	0.754	0.745	0.842	0.854	0.841
Deviance (-2 Log Likelihood)	1536.92	1606.19	1527.69	1616.56	1584.71	1518.78	1526.77	1524.29

Note: *$p < .02$. ***$p < .001$. †Gender and Yr2Full did not have random components in Model 8.

not identify any of these variables as significant. Additionally, the intercept for this model was not significant. Measures of variance (intra-class correlation [ICC]) and deviance (-2 Log Likelihood) were less useful for this model than the baseline model, and this model was discarded. Results from homogeneity tests referenced above found no differences between disability and ethnicity groups, so these results were not unexpected. Model 3 considered only gender as a student-level variable, consistent with previous homogeneity tests. Model 3 identified a significant intercept, slope across measurements, and gender coefficient. However, the model reduced the measured variance between scores. The difference between Model 3 and Model 1, using a Chi-Square difference between deviance measures was not significant, $\chi^2(4) = 9.23$, $p = .056$.

Models that incorporated group variables were considered next. A measure of fidelity was identified for each campus-year when the program was used by identifying the percent of students using the program that year who used the program at least 20 hours. Dummy variables were developed to distinguish between the four treatment conditions. A model was developed that incorporated these variables in both level 2 equations, resulting in both main effects and interaction effects with the time-point variable. None of the dummy variable terms, including the interaction terms, were found to be significant. The intraclass correlation for this model was .711, and the deviance measure (-2 Log Likelihood) was 1669.12. Neither measure was an improvement over the baseline model. A model utilizing these dummy treatment variables was developed that did not include any interaction effects. Again, none of the dummy variable terms were found to be significant. The intraclass correlation for this model was .709, and the deviance measure was 1675.35. As before, neither measure was an improvement over the baseline model. These models are not presented in Table 4.3.

Dummy variables were developed for program usage for each of the student's two consecutive years (Yr1Full and Yr2Full) as well as for usage in both years of at least 20 hours (BothYrsFull). Model 4 did not identify fidelity as an influence in student achievement scores, yielding nonsignificant parameter estimates and lower overall model measures. Model 5 identified only usage during year two as significant, though the ICC and deviance measures did not improve on baseline model measures. Model 6 identified both years with full usage (at least 20 hours in each year) as significant, though the interaction with time was not significant. The model accounted for less variance than the baseline model (84.1% vs. 84.8%). The Chi-Square difference between Model 6 and the baseline model was not significant, $\chi^2(10) = 18.11$, $p = .054$. Model 7 considered the effects of having two years of full usage without the mediating effect of gender. This model accounted for more variance than the baseline model, though the Chi-Square difference between the models was not significant, $\chi^2(5) = 10.15$, $p = .071$.

Only one model was found to yield a significant improvement in deviance over the baseline model. This model considered the effects of gender and year two full usage only. Both of these variables included only fixed components, holding the effect of each variable constant across participants without considering random variations within each variable. Though this reduced the number of model parameters, it may have unnecessarily restricted the impact of each variable on the achievement measures. The full model is described below:

Level 1: $$Y_{ti} = \pi_{0i} + \pi_{1i}(Time\ Point)_{ti} + \epsilon_{ti} \quad (4)$$

Level 2: $$\pi_{0i} = \beta_{00} + \beta_{01}(Male)_i + \beta_{02}(Yr2Full)_i \quad (5)$$

$$\pi_{1i} = \beta_{10} + \beta_{1i}(Yr2Full) \quad (6)$$

The model yielded a slightly lower ability to explain variance between scores (ICC = 84.1%), but the Chi-Square difference in deviance between the models was significant, $\chi^2(3) = 12.63$, $p = .006$. Given the restrictions on the mediating variables, it is doubtful that this model accurately identified the impact and relationships between variables and student achievement gap measures.

A graph of the achievement gap measures for each group is provided in Figure 4.1. Error bars were removed from the graph to enhance clarity, and it should be noted that there were no significant differences between groups at the start of Year 1. The graph suggests that there are similarities between usage groups across measurement intervals. An analysis of variance identified that there was significant variation between groups in their final achievement gap, $F(3, 179) = 3.562$, $p = .015$. Between group *t*-tests identified that Group 3 (Partial-Full) was significantly different than Group 4 (Full-Full) and Group 1 (Partial-Partial) at $p < .02$ for each comparison. Total change in achievement gap, as measured by the difference between the start of year 1 and end of year 2, was found to be significantly different between groups, $F(3, 179) = 4.933$, $p < .003$. Group 3 was found to be significantly different than Group 1 and Group 2 (Full-Partial).

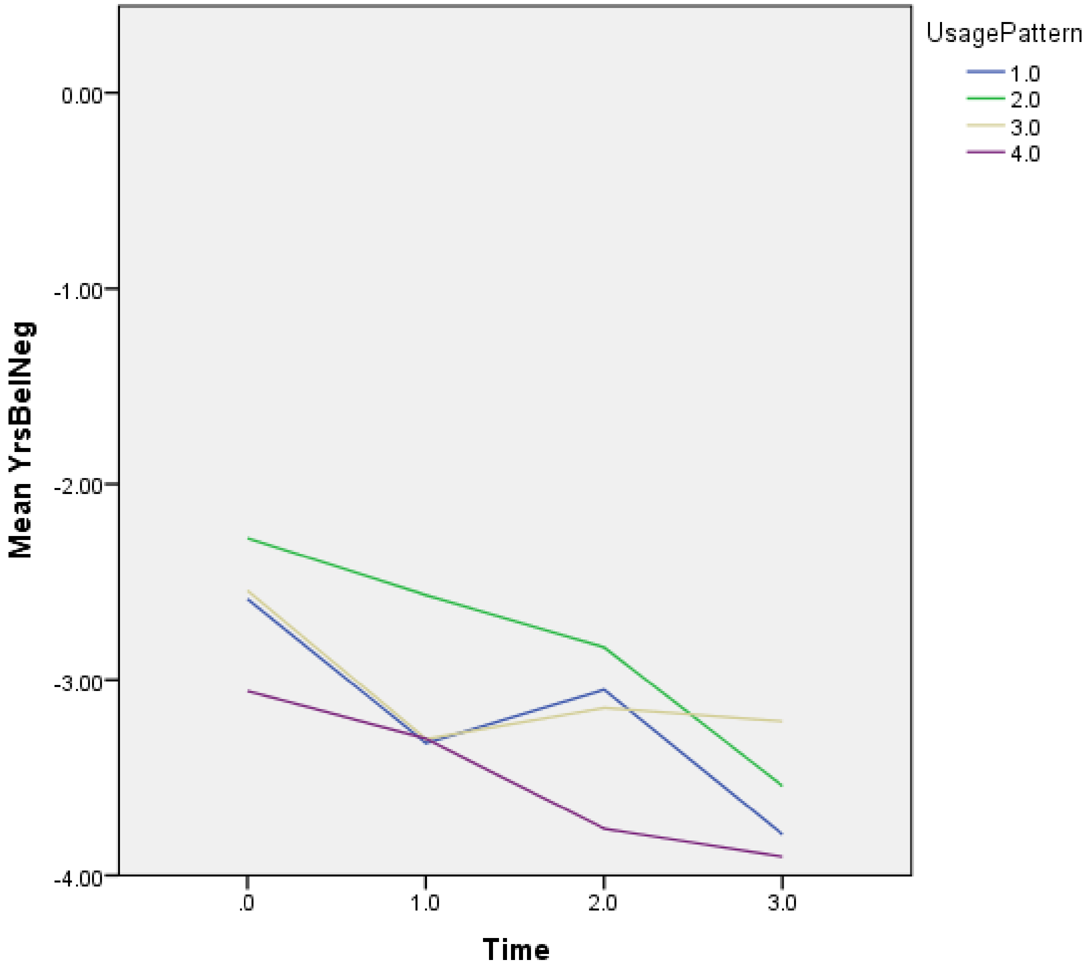


Figure 4.1. A visual comparison of the achievement gap between usage groups.

Group 1 received limited usage during both years; Group 2 received full-limited usage over the two years; Group 3 received limited-full usage over the two years; Group 4 received usage of least 20 hours each year. Students on grade level are assumed to remain at 0 at each measurement interval, indicating that all usage groups declined over time. The long term trajectories for each group show a widening of the mathematics achievement gap. Variations in summer patterns, from Time 1 to Time 2, may be a correction for poor measurement during the first school year.

Comparing measurement points during the year may be a means of determining if the program has a consistent effect across the year. An ANOVA comparing achievement gap at the start of each year (Time 0 and Time 2) found no differences between groups, $F(3, 179) = 0.544$, $p = .653$. However, an ANOVA comparing achievement gap at the end of each year (Time 1 and Time 3) found differences between usage patterns, $F(3, 179) = 10.323$, $p < .0001$. An ANOVA comparing achievement gap differences during the summer between years (Time 1 vs. Time 2) identified between group differences, $F(3, 179) = 6.336$, $p = .0004$. Since the usage groups were considered homogenous above, it is doubtful that there was significant variation between summer mediating factors such as summer school program attendance or impact. Rather, comparisons across starts suggest that end-of-year measurements (Time 1 and Time 3) may be erroneous for students in partial usage patterns.

Indeed, tests between full and partial users in each year found significant differences during each year and the summer break. Students who used the program at least 20 hours during the first year were found to have a significantly flatter rate of change during the school year, $t(181) = -11.018$, $p < .0001$. Similar results were found for students who used the program at least 20 hours during the second year, $t(181) = -13.349$, $p < .0001$. The unexpected increase in achievement scores for Year 1 Partial users is better accounted as incomplete year 1 data since no differences were found between usage patterns between the start of year 1 and the start of year 2. It is also worth noting that, despite usage pattern, the mathematics achievement gap between disabled and non-disabled students continued to grow at all measurement points.

Discussion

The need continues to exist for interventions to close the mathematics achievement gap for students with disabilities. Consistent with Judge and Watson (2011), this study identified that the achievement gap for students with learning and attention disorders continues to grow over time. This study extends their findings through the middle school years. Results from this study suggest that the use of SuccessMaker Mathematics for at least 20 hours each year does not yield results that persist over time. There were no additional benefits found for students who used the program at that recommended level for consecutive years.

Additionally, the results from this study failed to identify moderator variables that might account for some of the variance between students' achievement gap measures. Demographic variables that were included in multilevel models failed to yield models with substantially better variance and deviance measures than a baseline model void of moderator variables. Only gender was identified as a potential moderator, though initial variance between groups on this variable may rule out the impact of gender on achievement outcomes. Models that included usage variables, including the fidelity with which campuses implemented SMM, were also found to not be significantly different from the baseline model.

Analyses of usage patterns across measure points revealed interesting patterns of achievement differences. Regardless of usage level, all students in the sample remained on the same learning trajectory from the start of one year to the next. Changes over the summer break appear to be corrections from poor measurement during the previous school year. It is believed that end-of-year achievement levels identified in SMM reports may be invalid, and future researchers are advised to find more stable and accurate measures of achievement.

The sample size for this study may be considered rather limited given the five year span of the data collection. The use of multilevel models given this sample size is not advised (Hox, 2010). Other methods, including regression analysis or structural equation modeling, may provide better insights into the factors affecting student performance. That said, the failure for multiple years of program usage at or above recommended levels to yield positive outcomes in achievement is consistent with other studies (Cheung & Slavin, 2013; Christmann & Badgett, 2003). It is also conceivable that factors specific to the district or its implementation may have produced depressed effects that might not be found elsewhere. This conjecture is similar to that made by Fitzgerald and Fitzgerald (2002) as an explanation for their disparate findings between campuses. Indeed, the high mobility rate for this district likely had a substantial impact on student attrition that led to exclusion from this study.

The low threshold for "full" program usage may also be a limiting factor in this study. Previous research has found that "super-users" – those who use the program well beyond 25 hours during the school year – may realize greater results than those who use the program between 20-25 hours per year (McKissick, 2016). Only two students in the sample received more than 25 hours each year, and extrapolation from such a small sample would substantially increase the likelihood of error.

There may also be sub-patterns of program usage that deserve further exploration. Because of the campus-controlled schedule for program usage, it is conceivable that some users experienced spikes in usage time at various points rather than a continuous level of usage throughout the school year. Program-generated reports do not provide this level of usage, so researchers are advised that ongoing monitoring of student performance may be necessary to capture these variations. Students who garner significant hours of usage during a brief period may not experience long term benefits from that usage.

The summer decline experienced by all user groups, as identified by a comparison of year one and year two starting measures, may be of use to those working with students with disabilities. Intensive summer programs may allow students an opportunity to focus on a narrow set of skills (e.g., mathematics) rather than a 7-9 period course of study as expected during the regular school year. This narrow focus may nudge the starting levels for the subsequent year higher, perhaps adjusting the long-term achievement gaps identified in this study.

This study explored only one possible long-term intervention strategy to address the mathematics achievement gap for students with disabilities. The need remains for ongoing research to identify solutions to this persistent problem. It remains unclear if technology-based interventions, specifically SuccessMaker, will have any lasting effects. However, short-term interventions are unlikely to resolve this issue. Campuses implementing interventions are encouraged to ensure treatment fidelity and even increase dosage until desired results are seen.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Students with disabilities have academic deficiencies that present challenges to both researchers and practitioners. Previous research was inconsistent regarding the learning trajectories of students with disabilities. The research presented here favors the "ever-widening mathematics achievement gap with the passage of time" (Judge & Watson, 2011, p. 154) perspective regarding this issue. This suggests that, without intensive intervention, students with disabilities should be expected to fall ever further behind as they progress through school. The current research focused on students in middle school (6^{th} – 8^{th} grades). It is presumed, based on this trajectory pattern, that the gap between disabled and non-disabled students would be smallest at the earlier grades. Though intervention may be necessary at all grades, the time period where the greatest change would be expected is in the primary grades when the gap is the smallest. Some research exists to support the use of SuccessMaker for students in elementary school, though most are dissertations and not peer-reviewed (Gatti, 2009; Gee, 2008; Kirk, 2003; Laub, 1995). The current research has also identified a direct relationship between the amount of program usage and resultant achievement growth. Students with disabilities in later grades, including middle school, may need to use the program for much longer than expected to address their achievement gaps. This may present situations wherein students must forego some educational experiences such as electives, arts, or even other core subjects to provide sufficient time for the intervention. Alternately, parents have the opportunity to assist schools in remedying this underachievement by facilitating program usage at home and by enabling students' access to the program before and after school. Similar concerns may exist for reading, though that is beyond the scope of the present research. Rather, schools and parents must identify their educational priorities and use available resources in pursuit of those concerns which are most pressing.

Previous research identified that certain intervention strategies and approaches may be advantageous for students with disabilities. Early research was inconsistent regarding the effects of programmed instruction, a methodology rooted in behaviorist principles. Cohen's *d* effect sizes ranging from .57 (Schmidt, Weinstein, Niemiec, & Walberg, 1985) to -.01 (Kulik, Schwalb, & Kulik, 1982) were found through meta-analyses. The current research found that increased use of SMM, a program utilizing programmed instruction, yielded significantly greater outcomes. However, the current research was unable to demonstrate that these outcomes did much to address the achievement gap for students with disabilities. Sustained use of SMM at high levels (20 or more hours per year) did not demonstrate any impact on the learning trajectories of students with learning disabilities or other health impairments. Thus, the present research fails to support the proposition that behaviorist strategies are inherently beneficial for students with disabilities (Cooley, 2007; Cummings & Elkins, 1999; Ormrod, 2014; Zafiropoulou & Karmba-Schina, 2005). Because the present version of SMM has incorporated a number of features that are not behaviorist in nature (such as a cognitive tutor), it may be more pragmatic to say that the current research has not found SMM to provide sufficient benefit at recommended levels to remedy the achievement gap for students with disabilities.

A remarkable homogeneity of achievement trajectories was identified across disability groups. At various stages of the present research, students with various disabilities were considered as a single group because of similar achievement gaps. This includes students with learning disabilities, students with “other health impairments” such as attention deficit/hyperactivity disorder, autism, emotional disturbances, and even intellectual disabilities. The term “learning disabilities” itself represents a disparate group of students as students with disorders in mathematics, reading, and written expression. The specific areas of learning disability for the students in the present research were unknown, though this group demonstrated consistency in both achievement gap and response to the provided treatment. Future research considering the learning trajectories and impact of SuccessMaker Mathematics (SMM) for students with various learning disabilities is warranted. There is some evidence from the present research that the use of a common intervention may demonstrate success for student with various areas of disabilities, though the treatment dosage may vary between groups. That said, it is recommended that any intervention applied to students with disabilities be individualized to address the needs of the specific student rather than based on any disability classification.

A variety of obstacles, and accompanying solutions, were identified by staff as possibly impacting student performance and program usage. Student usage at levels needed to impact the achievement gap may necessitate campus infrastructure improvements. Staff identified that increases in the number of available computers as well as available internet bandwidth were needed to support the increase in student program usage. These upgrades may be costly for some schools or districts. Future technology-based intervention implementations should consider the cost of expected technology upgrades. Additionally, alterations to the daily schedule may be needed to facilitate implementation. Though campuses experienced limited success in reaching recommended usage levels, what success was experienced appears to have been aided by the introduction of a double-blocked period of mathematics instruction. The increase in available instruction time facilitated the introduction and use of the program. It is conceivable that a dedicated intervention period, as some elementary campuses presently use, may provide additional time during the school day for targeted intervention. Alternately, campuses may wish to consider individualized scheduling for students for whom program usage and academic remediation may be priorities. This is supported by the lack of student program usage outside of the school day despite opportunities for home use and program use before and after school hours.

The need for a team approach to implementation and ongoing use, supported by a clear plan for implementation, is also great. Implementation fidelity rates across campuses and academic years demonstrated significant fluctuations in usage patterns. Prior to implementation, the campus should identify those individuals who will secure the needed technological supports and provide ongoing monitoring of student usage. Campus administrators are advised to be aware of implementation resistance that may adversely impact students. Users and administrators must be equipped to address routine issues associated with technology, including software glitches and program upgrades. Whereas some campuses in the present study reported ongoing reliance on vendor representative support, campuses with more successful implementations developed staff capacity for addressing issues as they arose. Clear points of contact should be established to facilitate communication regarding the program. Staff should provide ongoing monitoring of student performance during program use, and periodic reviews of cumulative performance should be conducted.

Use of SuccessMaker at recommended levels (20-25 hours per year) yielded gains that were statistically significant but that failed to address the achievement gaps present for students with disabilities. While the present research identified that students receiving more program usage realized greater achievement gains, students receiving the treatment at recommended levels failed to achieve at levels that would address the achievement gap relative to their non-disabled peers. Rather, students receiving significantly more than the recommended dosage were more likely to experience achievement growth that brought them nearer to grade level performance. Additionally, students who received the recommended dosage for two consecutive years failed to demonstrate a change in learning trajectory relative to students receiving the treatment at recommended levels for just one year. Sustained program use was not found to be beneficial at recommended levels. More accurate and useful recommendations regarding program usage are needed from the vendor. This is further justified given the inconsistent implementation patterns identified in this research.

Other measures of student performance besides usage time were found to be useful guides to treatment effectiveness. Student accuracy during usage, as measured by the percent of questions answered correctly, was found to influence growth models. Several staff interviewed during the qualitative portion of the research identified that some students developed patterns of behavior that involved “quick-clicking” or “button-mashing” rather than engaged problem-solving. This pattern frequently resulted in low accuracy scores as a result of minimal effort and increased likelihood of erroneous responses. Student efficiency during program use was also identified as a significant mediator of achievement gains. The number of questions encountered per hour of program use influenced the student’s gains. Two distinct patterns were identified by staff, of which the above pattern of rapid and disengaged learning was but one. The other pattern identified those students with low engagement characterized by limited responses. These students may have the program open, though responses to presented problems are limited. It is believed that there is a subset of these low-efficiency students for whom response time is affected by disability or personality. In other words, engagement may be high though the time needed to solve presented problems is greater. This results in a low efficiency score, which may affect achievement gains. More research is needed to determine how student affective components, including engagement levels and perception of the program, impact student performance.

Despite the limited impact of SMM on student performance, staff perceptions of the program were overwhelmingly positive. Several campuses experienced the confluence of several factors – available time, positive staff attitudes, positive student engagement, and lack of technological obstacles – that allowed individual students to demonstrated marked growth during a single year of program usage. Vendor reports highlighted these individual cases as examples that the program was working and could work for all students. Perceptions of teachers who work with students with disabilities were buoyed by these individual cases, despite these cases representing usage outliers. Student performance on end-of-year state testing was not significantly impacted by student usage, though this was rarely mentioned by staff. The present research was hampered by variations in test versions and score reporting used in Texas, resulting in a less than ideal method of score comparison. However, it should be noted that the particular test to which a student was assigned (e.g, grade-level STAAR vs. off-level STAAR-Modified) may have been influenced by achievement gains resulting from program use.

Not all staff perceptions of SMM, either regarding the program or its effectiveness for students with disabilities, were positive. Some staff identified that the program was very effective for a small portion of the students using the program but rather unsuccessful for many students. Others identified obstacles related to technology, schedule, and leadership that hampered implementation and resulted in little change in achievement. Others identified that the reliance on technology to address student achievement issues overlooked the importance of personal relationships in the learning process. For these staff members, funds earmarked for technology interventions would be best spent on trained staff that can leverage interpersonal relationships with students to effect achievement gains.

The present research has sought to address a number of limitations with previous research. Most of the available research is over a decade old. Those studies within the past ten years have almost all been dissertations not subjected to peer review. The age of previous research also suggests that newer versions of SMM have not received much research attention. These newer versions have incorporated advances in technology to yield more engaging audio and video presentations as well as learning devices. These advancements include the use of a cognitive tutor and virtual devices such as a highlighter and sticky notes. No research was identified that considered the impact of these virtual devices on student achievement or program effectiveness. Though the data used was historical, the present study has demonstrated between-group equivalences that may offset the lack of random assignment to treatment usage groups. Ethical issues with assignment to treatment groups in educational settings, especially for students with disabilities, have been identified. The proposed submission of the three enclosed articles to peer-reviewed journals should yield a level of validity unachieved by recent doctoral students investigating SMM.

The research herein has sought to find a solution to the stubborn problem of an achievement gap for students with disabilities. The identification of an adequate comparison sample for the target population presented challenges during data analysis. However, the naturally occurring differentiation in the dose of SMM at each campus provided the opportunity to examine that factor on student learning. Comparison of student performance to individual prior achievement rates is a satisfactory means of demonstrating program impact, but it fails to satisfactorily address the achievement gap.

By keeping an eye on the expected achievement levels of these students, researchers may determine whether interventions are absolutely effective or simply relatively effective. Any intervention which fails to demonstrate a rate of achievement greater than 1.00 (one year of learning in one year's time) fails to impact the disability achievement gap. Effect sizes presented in similar research may be misleading, and an examination of the comparison groups used is strongly advised. Effect sizes may not represent the most useful data regarding interventions for students with disabilities given the need to demonstrate closure of the achievement gap.

The present research has also identified that students with disabilities may not have adequate access to resources to address their achievement deficits. Despite present technology resources and capabilities, the introduction of a technology-based intervention may necessitate the purchase of additional resources to facilitate successful implementation. The cost of these upgrades, as well as the cost of the intervention itself, may prove to be insurmountable obstacles without supplemental allocations or financial grants. Access to resources may also be interpreted as problems with existing instructional schedules. Students with disabilities may need increased opportunities to address fundamental skills that require changes to the school schedule or individualized program of instruction. Additional staff may also be needed to provide needed opportunities for these students.

The research presented here has identified several characteristics of students with disabilities that may be useful to not only future researchers but also current practitioners. There is an ongoing need to identify the educational needs and learning characteristics of students with disabilities in order to provide appropriate supports. Those who provide direct support to students with disabilities should benefit from the discussion presented above regarding the learning trajectories and homogeneity of their students despite varying areas of eligibility. Students with disabilities also represent an underrepresented group as researchers investigate student learning. Additional research that considers interventions to close the disability achievement gap is needed, and a pragmatist mindset that considers the end goal of student learning as most important is encouraged. The research conducted for this project was encouraged by various representatives of the vendor interfacing with the district. It is hoped that the findings herein demonstrate that recommendations for program usage need to be revised and that campuses need additional support to effectively implement the program. Results from this study will be presented to the district from which the data were generated, guiding their implementation of SMM as well as future intervention implementations.

Implementation of SuccessMaker Mathematics identified a number of concerns for teachers and administrators alike. Participants consistently identified obstacles regarding technological infrastructure, personnel support, and time as interfering with implementation efforts. Despite these obstacles, most users demonstrated a strong positive perception of the program and its results. Often, users identified obstacles and quickly identified how the campus worked to overcome them.

In spite of the overwhelming support for the program, implementation was inconsistent across and within campuses. Student usage was found to vary from year to year within the same campus, and some campuses have yet to identify any student who has used the program at or above recommended levels. The variety of campus implementation plans led to the identification of natural usage groups that facilitated statistical analysis. Interestingly, disparate usage did not favor any specific gender, race, or disability group. Instead, analyses consistently revealed between group homogeneity at the start of implementation. However, the findings from this research support previous findings that partial program implementations may end up with participants farther behind than they were before the intervention (cf. Capraro et al., 2016, in press).

Student academic achievement as a result of SMM usage was identified as the outcome variable of interest. Student growth using SMM was compared to previous achievement levels using a within-subjects design. Students who used the program at recommended levels remained unchanged while students at lower usage levels experienced negative learning gains. When compared to students, disabled or not, who are on grade level, students who received the program at usage levels below vendor recommendations experienced a widening of the disability achievement gap. Students receiving program usage at vendor recommendation levels demonstrated no change in their achievement gap. However, students who used the program at levels greater than vendor recommended levels experienced accelerated achievement that began to close the gap. Longitudinal usage of SMM at recommended levels did not result in achievement outcomes that were statistically different from usage below recommended levels over the same length of time.

Program cost is a potential obstacle that may affect program implementation. Licenses for SuccessMaker cost approximately $1,000 apiece (EdSurge, n.d.), though licenses may be used for both reading and mathematics. SuccessMaker licenses are also concurrent and perpetual, meaning that multiple students may use the available licenses and that the license becomes the property of the purchaser. The district in which the research was conducted purchased approximately 15 licenses per middle school campus. This allowed 15 students to simultaneously access the program. Annual fees to renew the license and provide program maintenance represent recurring fees associated with the program. On-site support from the vendor is available. Records from the participating district indicate that $139,500 was spent for at least the first four years of program implementation, and the district was provided with 140 days of on-site support from Pearson representatives.

Campuses or districts where SuccessMaker is being considered for adoption should consider whether sufficient resources are in place to ensure a productive and financially responsible implementation. Considerations regarding technological capabilities and needs should be identified prior to adoption. Personnel support for program adoption and eventual use are needed. An accountability system that included provision of achievement feedback to both students and campuses should be developed.

These recommendations follow from the discontinuity of voice and action from administrators who explained that they "really like the program" and the discordant level of program use within their building. The scope of the proposed implementation necessitates that districts consider their financial and technological resources. Districts should determine the size of their intervention sample and review their present academic schedule to determine how many licenses would be needed for a successful implementation. The concurrent nature of the licenses allows for multiple students to utilize a single license sequentially. Vendor recommendations for usage suggest that daily sessions of 15-20 minutes in length are appropriate. It is conceivable, then, that three students per hour could utilize the same license. Campus scheduling issues that restrict the amount of available time for program implementation should be identified to more effectively calculate the number of licenses needed for the intervention. Funding must also be secured to include startup costs as well as annual license and support costs. Current technology should also be reviewed prior to program adoption. Additional computers or increased bandwidth may be necessary to accommodate the increased internet usage resulting from the program. The costs of additional computers, software and operating system upgrades, and improved internet speed should be captured as part of a pre-adoption analysis of capabilities.

Administrator and teacher support for program usage should be sought prior to implementation. This support must necessarily extend from verbal support of the program to a commitment for its eventual implementation with fidelity. A clear plan is needed to identify which students will access the program at what times. Teachers who may be impacted by this plan must commit to ensuring that students have the access anticipated. The intervention plan should identify which staff will have oversight of the program, regularly reviewing student usage and growth reports, and on what schedule that oversight should occur. Teacher leaders, as well as campus administrators, will be needed to review student usage and maintain faithful adherence to the campus' usage plan. The identification of support personnel, including technologists and even learning specialists, is advised prior to program adoption. Ongoing support of the program may result in additional demands on limited time for these personnel, and opportunities to address these issues pre-adoption are encouraged.

Implementation may necessitate a review of existing campus and course schedules. Participants in the current study identified that working the program into existing scheduled provided difficult. Campuses where implementation was more successful found creative ways of working program use into existing courses without significantly interrupting core curriculum. Teachers providing resource (below-level) instruction perceived that integrating SuccessMaker with their remedial instruction resulted in achievement gains. During the second year of program implementation, the district implemented a double-block schedule for mathematics at the middle-school level. This resulted in all middle school students receiving approximately 100 minutes of mathematics daily. Teachers found that the use of 15-20 minutes for SuccessMaker use was more viable given the additional available time, and usage levels improved as a result. Similar changes to campus schedules may be needed to effectively implement this and other instructional interventions.

Accountability systems should be developed prior to program implementation. The use of public funds to purchase educational programs for students requires the ethical administration of those dollars. Consequently, districts are encouraged to identify key campus and district personnel who will maintain responsibility for ongoing program usage monitoring. These individuals are encouraged to identify periodic benchmarks by which program usage may be gauged to ensure implementation fidelity. In addition to student usage levels, indicators such as average session length and accuracy of response should be monitored for the individual students utilizing the program. The investment of large amounts of funds, personnel, and time that may be needed for successful program implementation suggests that senior administrators in the district be supportive of implementation plans.

The absence of a clear implementation plan that includes an analysis of available resources may result in an incomplete – and unsuccessful – program implementation. In five years of program implementation in the target district, only 17.9% of students used SuccessMaker at a level that resulted in any closure of the disability achievement gap. Inconsistency between campuses within the same district resulted in wide differences in the percent of users receiving or exceeding recommended usage levels, ranging from 1.32% to 54.88%. It is further recommended, then, that districts adopting the program for a group of campuses develop a plan for providing oversight and support across campuses to ensure implementation fidelity.

Though this research has added to the existing literature regarding SuccessMaker Mathematics as well as students with disabilities, much more research is needed in both domains. The present research identified that any closure of the mathematics achievement gap for middle school students with disabilities using SuccessMaker was achieved only when using the program at levels exceeding those recommended by the vendor. Additional research is still needed to identify issues related to program implementation so that obstacles that might interfere with both program effectiveness and student learning can be found and removed. The influence of affective variables for students with disabilities still deserves exploration, including their perceptions regarding the effectiveness of the intervention selected for them. Further research regarding the longitudinal outcomes of students with disabilities is encouraged.

REFERENCES

Agarwal, R., & Prasad, J. (1999). Are individual differences germane to the acceptance of new information technologies? *Decision Sciences, 30*, 361-391.

Bailey, G. D. (1992). Wanted: A road map for understanding integrated learning systems. *Educational Technology, 32*(9), 3-5.

Barnes, R. (2005). Moving towards technology education: Factors that facilitated teachers' implementation of a technology curriculum. *Journal of Technology Education, 17*, 6-18.

Baylor, A. L., & Ritchie, D. (2002). What factors facilitate teacher skill, teacher morale, and perceived student learning in technology-using classrooms? *Computers & Education, 39*, 395-414.

Becker, H. J. (1992). Computer-based integrated learning systems in the elementary and middle grades: A critical review and synthesis of evaluation reports. *Journal of Educational Computing Research, 8*(1), 1-41.

Berman, P., & McLaughlin, M. W. (1976). Implementation of educational innovation. *The Educational Forum, 40*, 345-370.

Birch, J. (2002). *The effects of the Delaware Challenge Grant Program on the standardized reading and mathematics test scores of second and third grade students in the Caesar Rodney School District* (Doctoral dissertation). Wilmington College, New Castle, DE.

Bowman, J., Newman, D. L., & Masterson, J. (2001). Adopting educational technology: Implications for designing interventions. *Journal of Educational Computing Research, 25*(1), 81-94.

Bransford, J. D., Sherwood, R. D., Hasselbring, T. S., Kinzer, C. K., & Williams, S. M. (1990). Anchored instruction: Why we need it and how technology can help. In D. Nix & R. Spiro (Eds.), *Cognition, education, and multimedia: Exploring ideas in high technology* (pp. 115-142). Hillsdale, NJ: Erlbaum.

Burton, J., Moore, D., & Magliare, S. (2008). Behaviorism and instructional technology. In D. H. Jonassen & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology: A project of the Association for Educational Communications and Technology* (2nd Ed.; pp. 3-36). Mahwah, NJ: Erlbaum.

Capraro, R. M., Capraro, M. M., Morgan, J., Scheurich, J., Jones, M., Huggins, K., Corlu, S. M., Han, S. Y., & Younes, R. (2016, in press). The impact of sustained professional development in STEM in a diverse urban district. *Journal of Educational Research*. DOI 10.1080/00220671.2014.936997.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science, 2*(40), 1-9.

Cawley, J., Parmar, R., Yan., W., & Miller, J. (1998). Arithmetic computation performance of students with learning disabilities: Implications for curriculum. *Learning Disabilities Research and Practice, 13*, 68-74.

Cherryholmes, C. H. (1992). Notes on pragmatism and scientific realism. *Educational Researcher, 14,* 13-17.

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9*, 88-113.

Christmann, E. P., & Badgett, J. L. (2003). A meta-analytic comparison of the effects of computer-assisted instruction on elementary students' academic achievement. *Information Technology in Childhood Education Annual, 1*, 91-104.

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

Cooley, M. L. (2007). *Teaching kids with mental health and learning disorders in the regular classroom: How to recognize, understand, and help challenged (and challenging) students succeed.* Minneapolis, MN: Free Spirit Publishing.

Crawford, A. N. (1970). A pilot study of computer assisted drill and practice in seventh grade remedial mathematics. *California Journal of Educational Research, 21*, 170-174.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W. (2011). Controversies in mixed methods research. In. N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (4th ed., pp. 269-283). Thousand Oaks, CA: Sage.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.

Cruthirds, J., & Hanna, M. S. (1997). Programmed instruction and interactive multimedia: A third consideration. Retrieved from http://files.eric.ed.gov/fulltext/ED439464.pdf

Cuban, L., Kirkpatrick, H., & Peck, C. (2001). High access and low use of technologies in high school classrooms: Explaining an apparent paradox. *American Educational Research Journal, 38*, 813-834.

Cummings, J. J., & Elkins, J. (1999). Lack of automaticity in the basic addition facts as a characteristic of arithmetic learning problems and instructional needs. *Mathematical Cognition, 5,* 149-180.

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*, 319-340.

Delon, F. G. (1970). A field test of computer assisted instruction in first grade mathematics. *Educational Leadership, 28*, 170-180.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18*(2), 237-256.

Earle, R. S. (2002). The integration of instructional technology into public education: Promises and challenges. *ET Magazine, 42*(1), 5-13. Retrieved from http://asianvu.com/digital-library/educational_technology/earle.pdf

EdSurge (n.d.). *Successmaker*. Retrieved from https://www.edsurge.com/product-reviews/successmaker

Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge, UK: Cambridge University Press.

Ertmer, P. (1999). Addressing first- and second-order barriers to change: Strategies for technology implementation. *Educational Technology Research and Development, 47*(4), 47-61.

Estep, S. G., McInerney, W. D., Vockell, E., & Kosmoski, G. (2000). An investigation of the relationship between integrated learning systems and academic achievement. *Journal of Educational Technology Systems, 28*(1), 5-19.

Felner, R. D., Favazza, A., Shim, M., Brand, S., Gu, K., & Noonan, N. (2001). Whole school improvement and restructuring as prevention and promotion. *Journal of School Psychology, 39*, 177-202.

Fitzgerald, D., & Fitzgerald, R. (2002). *The use of integrated learning systems in developing number and language concepts in primary school children: A longitudinal study of individual differences.* Retrieved from http://canberra.academia.edu/RobertFitzgerald/Books

Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York, NY: Teachers College Press.

Gatti, G. G. (2009). *Pearson SuccessMaker math pilot study.* Pittsburgh, PA: Gatti Evaluation, Inc.

Gee, A. P. (2008). *An investigation of the impact of SuccessMaker on reading and math achievement at an elementary school* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses. (UMI: 3326592)

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory.* Chicago, IL: Aldine.

Graham, L., Bellert, A., Thomas, J., & Pegg, J., (2007). A basic skills intervention for middle school students with learning difficulties. *Journal of Learning Disabilities, 40*, 410-419.

Granger, C. A., Morbey, M. L., Lotherington, H., Owston, R. D., & Wideman, H. H. (2002). Factors contributing to teachers' successful implementation of IT. *Journal of Computer Assisted Learning*, *18*, 480-488.

Groff, J., & Mouza, C. (2008). A framework for addressing challenges to classroom technology use. *AACE Journal, 16*(1), 21-46.

Hadley, M., & Sheingold, K. (1993). Commonalities and distinctive patterns in teachers' integration of computers. *American Journal of Education, 101*, 261-315.

Hall, G. E., & Hord, S. M. (2001). *Implementing change: Patterns, principles, and potholes.* Needham Heights, MA: Allyn & Bacon.

Haney, J., & Lumpe, A. (1995). A teacher professional development inventory guided by reform policies, teachers' needs, and research. *Journal of Science Teacher Education, 6*(4), 187-196.

Hayes, A. F. (2006). A primer on multilevel modeling. *Human Communication Research, 32*, 385-410.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd Ed.). New York, NY: Routledge.

Irish, C. (2002). Using peg- and keyword mnemonics and computer-assisted instruction to enhance basic multiplication performance in elementary students with learning and cognitive disabilities. *Journal of Special Education Technology, 17*(4), 29-40.

Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with comorbid mathematics and reading difficulties. *Child Development, 74*(3), 834-850.

Joyce, B., Weil, M., & Calhoun, E. (2009). *Models of teaching* (8th ed.). Upper Saddle River, NJ: Pearson.

Judge, S., & Watson, S. M. R. (2011). Longitudinal outcomes for mathematics achievement for students with learning disabilities. *The Journal of Educational Research, 104*(3), 147-157.

Kabacoff, R. I. (2014). *Power analysis*. Retrieved from http://www.statmethods.net/stats/power.html

Kirk, V. C. (2003). *Investigation of the impact of integrated learning system use on mathematics achievement of elementary student* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses. (ProQuest Document ID: 305324110).

Kozma, R. (2001). Learning with media. In D. P. Ely & T. Plomp (Eds.), *Classic writings on instructional technology* (pp. 155-158). Denver, CO: Libraries Unlimited, Inc.

Kulik, C.-L. C., Kulik, J. A., & Cohen, P. A. (1980). Instructional technology and college teaching. *Teaching of Psychology*, *7*(4), 199-205.

Kulik, C.-L. C., Schwalb, B. J., & Kulik, J. A. (1982). Programmed instruction in secondary education: A meta-analysis of evaluation findings. *The Journal of Educational Research*, *75*(3), 133-138.

Kulik, J. A. (1994). Meta-analytic studies of findings on computer-based instruction. In E. L. Baker & H. F. O'Neil (Eds.), *Technology assessment in education and training*, (pp. 9-33). Hillsdale, NJ: Erlbaum.

Kulik, J. A. (2002). *School mathematics and science programs benefit from instructional technology* (Info Brief NSF 03-301). Washington, DC: National Science Foundation.

Kulik, J. A., (2003). *Effects of using instructional technology in elementary and secondary schools: What controlled evaluation studies say.* Arlington, VA: SRI International.

Lapointe, L., & Rivard, S. (2005). A multilevel model of resistance to information technology implementation. *MIS Quarterly, 29*(3), 461-491.

Laub, C. M., & Wildasin, R. L. (1998). Student achievement in mathematics and the use of computer-based instruction in the Hempfield School District. In Pearson Digital Learning (Ed.), *SuccessMaker evidence of effectiveness: Selected evaluation studies* (p. 5). Retrieved from http://www.pearsoned.com/wp-content/uploads/dc4-successmaker-enterprise-evidence-of-effectiveness.pdf

Li, Q., & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review, 22*(3), 215-243.

Liao, Y. K. (1998). Effects of hypermedia versus traditional instruction on students' achievement: A meta-analysis. *Journal of Research on Computing in Education, 30*(4), 341-359.

Liu, L., & Johnson, D. (1998). A computer achievement model: Computer use, confidence, attitudes, and knowledge: A causal analysis. *Computers in Human Behavior, 14*(1), 125-146.

Mageau, T. (1992). Integrating and ILS: Two teaching models that work. *Electronic Learning, 11*(4), 16-22.

Manning, C. A. (2004). *The effect of the Math Concepts and Skills (MCS) computer program on standardized test scores at a middle school in east central Florida* (Doctoral dissertation). University of Central Florida, Orlando.

Martindale, T., Pearson, C., Curda, L., & Pilcher, J. (2005). Effects of an online instructional application on reading and mathematics standardized test scores. *Journal of Research on Technology in Education, 37*, 349-360.

McDonald, J., Yanchar, S., & Osguthorpe, R. (2005). Learning from programmed instruction: Examining implications for modern instructional technology. *Educational Technology Research and Development, 53*(2), 84-98.

McGhee, R., & Kozma, R. (2003, April). *New teacher and student roles in the technology-supported classroom.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

McKissick, S. K. (2016). *Effectiveness of Pearson's SuccessMaker for students with disabilities* (Unpublished doctoral dissertation, Chapter 2). Texas A&M University, College Station, TX.

McKissick, S. K. (2016). *Perceptions and obstacles encountered during SuccessMaker implementation* (Unpublished doctoral dissertation, Chapter 3). Texas A&M University, College Station, TX.

Mendelsohn, M. (1972). CAI in New York City: The slow learner in mathematics. *National Council of Teachers of Mathematics Yearbook*, 355-364.

Mills, S. C., & Ragan, T. J. (2000). A tool for analyzing implementation fidelity of an integrated learning system. *Educational Technology Research and Development*, *48*(4), 21-41.

Mintz, K. S. (2000). *A comparison of computerized and traditional instruction in elementary mathematics* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses. (ProQuest Document ID: 304577773).

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

Nunnery, J. A., & Ross, S. M. (2007). Effects of the school renaissance program on student achievement in reading and mathematics. *Research in the Schools, 14*(1), 40-59.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, *78*(1), 33-84.

Ormrod, J. (2014). *Human learning* (6th ed.). Essex, UK: Pearson.

Pearson Digital Learning (n.d.). *The Pearson timeline: Our history*. Retrieved from http://timeline.pearson.com

Pearson Digital Learning. (2002). *SuccessMaker evidence of effectiveness: Selected evaluation studies.* Retrieved from http://www.pearsoned.com/wp-content/uploads/dc4-successmaker-enterprise-evidence-of-effectiveness.pdf

Pearson Digital Learning. (2004). *Notes on SuccessMaker course levels, gain, and prescriptive scheduling*.

Pearson Education Inc. (2012, March). *SuccessMaker mathematics time/gain estimates for student scheduling: A guide for scheduling student sessions to achieve target scores.* Released publication from vendor provided by representative.

Pearson Education Inc. (2013). *SuccessMaker 6 math reference guide*. Retrieved from http://it.dadeschools.net/SuccessMakerNew/Teacher%20Resources/SuccessMaker_6_Math_Reference_Guide.pdf

Pelgrum, W. J. (2001). Obstacles to the integration of ICT in education: Results from a worldwide educational assessment. *Computers and Education, 37*(2), 163-178.

Pellegrino, J. W., & Goldman, S. R. (1987). Information processing and elementary mathematics. *Journal of Learning Disabilities,20,* 23-32.

Prince, J. D. (1969). *A practitioner's report results of two years of computer-assisted instruction in drill and practice mathematics*. McComb, MS: McComb Schools.

Ragosta, M. (1983). Computer-assisted instruction and compensatory education: A longitudinal analysis. *Machine-Mediated Learning, 1*(1), 97-127.

Resendez, M., Azrin, M., & Strobel, A. (2009). *A study on the effects of Pearson's 2009 enVisionMATH program: Final summative report*. Jackson, WY: PRES Associates.

Richards, J. (1996). Negotiating the negotiation of meaning: Comments on Voigt (1992) and Saxe and Bermudez (1992). In L. P. Steffe, P. Nesher, P. Cobb, G. A. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 69-75). Mahwah, NJ: Erlbaum.

Schiro, M. S. (2013). *Curriculum theory: Conflicting visions and enduring concerns* (2nd ed.). Thousand Oaks, CA: Sage.

Schmidt, M., Weinstein, T., Niemiec, R., & Walberg, H. J. (1985). Computer-assisted instruction with exceptional children. *Journal of Special Education, 19,* 493-501.

Schmidt, S. (1991). *Technology for the 21st century: The effects of an integrated distributive computer network system on student achievement* (Doctoral dissertation). University of La Verne, California.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth.

Skinner, B. (1986). Programmed instruction revisited. *The Phi Delta Kappan, 68*(2), 103-110.

Slavin, R. E. (1987). A theory of school and classroom organization. *Educational Psychologist*, *22*(2), 89-109.

Slavin, R. E., & Lake, C. (2008). Effective programs in elementary mathematics: A best-evidence synthesis. *Review of Educational Research, 78*(3), 427-515.

Slavin, R. E., Lake, C., & Groff, C. (2009). Effective programs in middle and high school mathematics: A best-evidence synthesis. *Review of Educational Research, 79*(2), 839-911.

Sobol, D. F., Rohrbach, L. A., Dent, C. W., Gleason, L., Brannon , B. R., Johnson, C. A., & Flay, B. R. (1989). The integrity of smoking prevention curriculum delivery. *Health Education Research, 4*, 59-67.

Spencer, T. M. (1999). *The systemic impact of integrated learning systems on mathematics achievement* (Doctoral dissertation). Eastern Michigan University, Ypsilanti, Michigan.

Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *SAGE handbook of qualitative research* (443-466). Thousand Oaks, CA: SAGE.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* New York, NY: Free Press.

Straub, E. T. (2009). Understanding technology adoption: Theory and future directions for informal learning. *Review of Educational Research, 79*(2), 625-649.

Strudler, N., & Wetzel, K. (2005). The diffusion of electronic portfolios in teacher education. *Journal of Research on Technology in Education, 37*(4), 411-433.

Suppes, P., & Morningstar, M. (1969). Computer-assisted instruction. *Science, 166,* 343-350.

Svoboda, D. S., Jones, A. L., van Vulpen, F., & Harrington, D. (2012). Programmed instruction. In J. Hattie & E. M. Andermann (Eds.), *International guide to student achievement* (392-395). New York, NY: Routledge.

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.

Tucker, T. H. (2009). *The relationship between computer-assisted instruction and alternative programs to enhance fifth-grade mathematics success on the annual Texas Assessment of Knowledge and Skills* (Doctoral dissertation). University of North Texas, Denton.

Underwood, J., Cavendish, S., Dowling, S., Fogelman, K., & Lawson, T. (1996). Are integrated learning systems effective learning support tools? *Computers & Education, 26*(1), 33-40.

Van Dusen, L. M., & Worthen, B. R. (1995). Can integrated instructional technology transform the classroom? *Educational Leadership, 53*(2), 28-33.

Vockell, E. L., & Mihail, T. (1993). Behind computerized instruction for students with exceptionalities. *Teaching Exceptional Children, 25*(3), 38-43.

Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Wendling, B. J., & Mather, N. (2009). *Essentials of evidence-based academic interventions*. Hoboken, NJ: John Wiley & Sons.

Weston, T. J. (2005). Why faculty did – and did not – integrate instructional software in their undergraduate classrooms. *Innovative Higher Education, 30*(2), 99-115.

Wood, K. M. (2004). Effects of SuccessMaker Math on students with learning disabilities in inclusive and special education classrooms. *Journal of Teacher Initiated Research, 1*. Retrieved from http://www.otterbein.edu/Files/pdf/Education/JTIR/VolumeI/woodfinal.pdf

Zafiropoulou, M., & Karmba-Schina, C. (2005). Applying cognitive-behavioral interventions in Greek mainstream school settings: The case of learning difficulties. *Learning Disabilities: A Contemporary Journal, 3*(2), 29-48.

Zhao, Y., & Frank, K. A. (2003). Factors affecting technology uses in schools: An ecological perspective. *American Educational Research Journal, 40*(4), 807-840.

Zhao, Y., Pugh, K., Sheldon, S., & Byers, J. (2002). Conditions for classroom technology innovations. *Teachers College Record, 104*(3), 482-515.