2017

Evaluation of Math Block-Scheduling for Low Performing At-Risk and Economically-Disadvantaged Students

Toni Trice
Walden University

Walden University

College of Education

This is to certify that the doctoral study by

Toni Trice

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Charles McElroy, Committee Chairperson, Education Faculty
Dr. Katherine Hayes Fondation, Committee Member, Education Faculty
Dr. Andrea Wilson, University Reviewer, Education Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2017

Abstract

Evaluation of Math Block-Scheduling for Low Performing At-Risk and Economically-

Disadvantaged Students

by

Toni M. Trice

M.Ed., Tarleton State University, 2011

B.S, American Military University, 2004

Doctoral Study Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Education

Curriculum, Instruction, and Assessment

Walden University

August 2017

Abstract

Research shows a math achievement gap for at-risk and economically-disadvantaged students in the United States. To address this issue, a Texas school district implemented a 90-minute math block-scheduling program with 8th grade students. Shaped by the academic learning time and social justice theories, the purpose of this quantitative program evaluation was to determine if students in 3 key subgroups (low performing, low performing and at-risk, and low performing and economically-disadvantaged) who participated on the 90-minute block-schedule performed significantly better on the math state standardized assessment than did students in the same subgroups who remained on the traditional schedule. This retrospective causal-comparative design compared existing scores from the 2013 Math State of Texas Assessments of Academic Readiness (STAAR) of 109 8th grade students (n = 49 block-schedule; n = 60 traditional schedule) for each of the 3 key subgroups. Mann-Whitney U tests indicated no significant differences in Math STAAR scores for the 90-minute block-schedule groups versus the traditional schedule groups for any of the 3 key subgroups studied. Results suggest the 90-minute block-scheduling program was not effective in producing better math assessment scores compared to the traditional schedule for these students. Findings were presented to district decision makers in an evaluation report, which may motivate district stakeholders to reevaluate current educational practices and funding allocations to improve math achievement of low performing students and produce positive social change.

Evaluation of Math Block-Scheduling for Low Performing At-Risk and Economically-Disadvantaged Students

by

Toni M. Trice

M.Ed., Tarleton State University, 2013

B.S, American Military University, 2004

Doctoral Study Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Education

Curriculum, Instruction, and Assessment

Walden University

August 2017

Dedication

To my mother, Catherine Curry, the most amazing woman I know. Your strength, grace, and unconditional love have been a constant inspiration. You taught me to love life and all that God has blessed me with. Your belief in my ability to succeed was a continuous source of fuel during this long journey.

To my sons, TJ and Nick. I pray that my success inspires and encourages you to soar and reach for your dreams. Nothing is impossible with faith, focus, and the fight to finish.

Acknowledgements

First and foremost, I thank my heavenly Father. I know He was my source of strength and sanity during this endeavor. He placed my feet on this doctoral path several years ago and a seed was planted. Even when I didn’t understand, I knew enough to trust Him, stay the course, and allow Him to order my steps. I don’t know exactly what He has in store for me yet, but I know obtaining this degree was part of a bigger plan that will soon reveal itself.

With gratitude and humility I acknowledge and recognize all who have helped me arrive at this juncture. James, my husband, thank you for your patience and understanding during this daunting process. Family, friends, and colleagues, thank you for all of the support, positive encouragement, and prayers.

My deepest appreciation to my doctoral committee members. Dr. Charles McElroy, my doctoral study chair, thank you for your patience, guidance, and words of wisdom. You helped me keep my head in the game. Dr. Hayes and Dr. Wilson, thank you for the meticulous feedback and attention to detail. Your comments pushed me to excel.

Table of Contents

List of Tables

List of Figures

Section 1: The Problem

Introduction

Two major goals of the No Child Left Behind Act (NCLB; 2001) were to raise overall student achievement and close racial and socioeconomic achievement gaps in U.S. public schools. Over 15 years have passed since NCLB went into effect and the nation's students did not meet the 100% proficiency levels by the 2013–2014 deadline outlined by NCLB; in fact, the achievement gaps persist. Even though there was documented growth in math achievement scores for some subgroups of students, there was still a distinctive gap between economically-disadvantaged and at-risk students and their peers (Slavin & Lake, 2008).

The Every Student Succeeds Act (ESSA; 2015) was signed into law by President Obama in December 2015 as an attempt to placate rising sentiment against NCLB (Black, 2016). ESSA reauthorized the Elementary and Secondary Education Act (ESEA; n.d.). ESSA did not eliminate the accountability associated with NCLB but rather provided states with flexibility in establishing their accountability measures (Peterson, 2016; Rothman, 2016).

Students in the state of Texas have been struggling with mathematics for several years. The Texas Education Agency (TEA; 2013) reported that over 75,000 or 24% of eighth grade students did not score satisfactory on the math section of the state summative assessment in school year 2012 (SY12). This number decreased slightly in school year 2013 (SY13) to approximately 23% (TEA, 2014). The numbers were even more alarming when considering economically-disadvantaged and at-risk students. Students are classified as economically-disadvantaged if their family meets the eligibility requirements to receive free or reduced-cost meals under the National School Lunch and Child Nutrition Program (TEA, 2014). The unsatisfactory ratings for the economically-disadvantaged students were 32% and 29% for those same school years (TEA, 2013, 2014). At-risk students, those who met the state's criterion placing them at greater risk of dropping out, presented an even greater challenge with reported unsatisfactory performances of 45% and 39% (TEA, 2013, 2014). Summative assessments are given at the end of each grade level, but eighth grade math is significant because Texas state law requires that students pass the eighth grade summative math assessment as a prerequisite for promotion to the ninth grade (TEA, 2014).

In SY12, the Texas Assessment of Knowledge and Skills (TAKS) test was replaced with the State of Texas Assessments of Academic Readiness (STAAR) assessment (TEA, 2013). The STAAR assessment items were intentionally designed to measure increased cognitive complexity (TEA, 2011b). To meet the level of rigor needed for success on the STAAR, the new math curriculum included the Texas Essential Knowledge and Skills (TEKS) adopted to identify the skills all Texas students needed to master (State Board of Education [SBOE], 2012). The pace and depth of the new math curriculum using the TEKS caused great controversy (SBOE, 2012), so much so that the Education Commissioner, Michael Williams, temporarily suspended the Texas law requiring students to pass the STAAR test as a prerequisite for promotion (TEA, 2013).

It is unclear how the passage of ESSA will impact this current policy as TEA is currently in the process of gathering information from stakeholders. TEA plans to submit its final state plan to the U.S. Department of Education (USDE) in July 2017 (TEA, 2016).

I conducted this quantitative study in one Texas school district, referred to as K-Town ISD (KTISD) throughout the study. KTISD was not immune to the math achievement gaps prevalent nationwide and throughout Texas. The purpose of the study was to evaluate the 90-minute math block-scheduling program that was implemented to increase the math scores of low performing students and to decrease the achievement gap of students classified as at-risk and economically-disadvantaged. In Section 1, I will define the local problem, the significance of the problem, the rationale for the study, and evidence of the problem from professional literature. Additionally, in the section I will provide the justification for the guiding research questions, the theoretical framework behind the study, as well as possible implications.

Definition of the Problem

The satisfactory passing rate of 73% on the math section of the STAAR summative assessment in KTISD was below the state passing rate of 77% for SY12. KTISD is a large school district located in Central Texas. Its position, near a military instillation, results in a diverse student population. Currently more than 43,000 students are being served at six high schools, 14 middle schools, 35 elementary schools and a career center. The diverse student population consists of the following ethnic groups: 34% African American, 28% Hispanic, 26% Caucasian, 7% two or more races, 3%

Asian, 1% American Indian, and 1% Pacific Islander. Students classified as economically-disadvantaged account for 57% of the population, while 52% of the students are identified as at-risk. Approximately half of the students are military dependents, which contributes to a student mobility rate of 34%.

In an effort to address the low scores and the needs of a diverse population, the district implemented a pilot 90-minute math block-scheduling program at one middle school campus for SY13. During a personal interview, KTISD's director of curriculum shared that "this decision was based on the idea that block-scheduling would provide economically-disadvantaged and at-risk students with more time on task, and therefore, increasing the opportunities for more practice and concept mastery under the supervision of educators." The problem is the math block-scheduling program was never evaluated to determine its effectiveness. The findings of Martinez and Holland (2011) supported 90-minute block-scheduling in math class in schools with demographics similar to those in KTISD. They reported block-scheduling could increase the achievement of English language learners (ELL), economically-disadvantaged, at-risk, and special education students. The district leadership asked me to perform an evaluation of the program's outcomes to gauge the impact of extended class time on the achievement of the eighth grade students at the pilot campus.

Rationale

My rationale for this study was to provide the district decision makers with information regarding the effectiveness of the 90-minute math block-scheduling program in meeting the established program goals of increasing the math scores of low performing students and decreasing the achievement gap of at-risk and economically-disadvantaged students. Faced with decreased budgets, the district cannot afford to fund programs that are not effective. If the outcomes evaluation substantiates the program's effectiveness in increasing the math scores of low performing students and decreasing the achievement gap of the economically-disadvantaged and at-risk students, stakeholders may consider expanding the program throughout the district.

Evidence of the Problem at the Local Level

KTISD's curriculum director shared that "scores falling 4% below the state average along with KTISD's large at-risk and economically-disadvantaged population, 57% and 52% respectively, prompted the decision to adopt the 90-minute block-scheduling." District leaders viewed this decision as a positive intervention based on three assumptions. The first assumption was extended class time would allow teachers more time to deliver lessons and have students practice the concepts learned under the supervision of the teacher. Second, decision-makers believed that more hands-on activities could be incorporated to help differentiate for multiple learning styles. Third, the extended class time eliminated the need for daily homework which the majority of the students did not complete. The goals of the block-scheduling program were to increase the math scores of low achieving students and decrease the achievement gaps between the economically-disadvantaged and at-risk subpopulations on the eighth grade math portion of the STAAR assessment.

KTISD's SY12 STAAR data indicated that 31% of the students labeled as economically-disadvantaged performed below satisfactory level compared to 22% of those not classified as economically-disadvantaged (TEA, 2013). This represents a 9% achievement gap. The data concerning at-risk student performance presents an even more alarming situation as a 20 % gap was reported for this subpopulation in SY12. Only 64% of the at-risk students received satisfactory ratings in contrast to an 84% for students not considered to be at-risk (TEA, 2013). Additionally, KTISD's overall scores are 4% below the state passing score of 77%. Failure to reduce the achievement gaps would mean KTISD could be identified for targeted interventions imposed by the state. District leaders were eagerly seeking out options to address the issue; however, the implementation of block-scheduling came at a significant opportunity cost.

Adopting a block-schedule for eighth grade math classes increased the number of math teachers needed; however, due to budget constraints, the district initially was not able to hire additional teachers. Instead, elective courses were removed from the master schedule and elective teachers terminated to make room for the math teachers needed to support the block-scheduling model. As a result, parents and students complained about the lack of electives available for eighth grade students. They claimed the new schedule did not allow students to earn high school level credits while in middle school. Spanish, Advanced Theater Arts, Advanced Band, and Advanced Choir were all removed from the master calendar for SY13. This means students desiring to graduate on the distinguished plan, Texas's most prestigious public school diploma, would be forced to attend and pay for a summer school courses in order to earn the necessary credits for this honor.

These issues resulted in the filing of several grievances by teachers and parents. KTISD prides itself in satisfying the needs of its diverse population. District leaders need to know if the increased time in math class had an impact on the scores of low achieving students and has reduced the achievement gap between the subpopulations as measured on the math section of the summative STAAR assessment. KTISD leadership planned to utilize the findings of this outcomes evaluation to determine whether to reinstate the program, continue with the current 50-minute class scheduling, or explore other alternatives.

Evidence of the Problem from the Professional Literature

Mathematical concepts, such as money management, understanding of time, and critical problem solving, are fundamental skills needed for the development of productive citizens (Codding, Chan-Iannetta, Palmer, & Lukito, 2009). Consequently, math instruction receives a substantial amount of focus and funding (Smith, Marchand-Martella, & Martella, 2011). Despite all of the focus and attention placed on it by schools, many students continue to struggle with math. In the United States, approximately 50% of students in Grades 8 through 12 lack basic skills in mathematics (Duncan, Easton, & Kerachsky, 2009).

In addition to significant numbers of U.S. students being rated below proficiency in math, the achievement gap between students labeled as at-risk or economically-disadvantaged and their peers continues to grow. Paige and Witty (2010) believed the solution to this phenomenon would benefit civil rights movements. Robinson (2010) argued that reducing the achievement gap would be a significant factor in promoting equality in the United States. Despite the amount of attention being directed at closing the gap, it continues as schools are given mandates but not much guidance concerning how to address the persistent gap (Paige & Witty, 2010). Slavin and Lake (2008) reported that the gap between at-risk and economically-disadvantaged students and their counterparts remains significant. Memon, Joubish, and Khurram (2010) believed this gap was linked to the absence of resources available to families classified as low socio-economic. Riegle-Crumb and Grodsky (2010) revealed little headway is being made in eliminating the gap because they believed a deeper look is needed into factors outside of the schoolhouse that could impact student achievement. Eddy and Brooks (2011) have called for educational leaders to address this issue. However, none offer viable solutions to help schools effectively overcome the problem. More focus needs to be directed at solutions and not just awareness of the problem (Robinson, 2010). The purpose of this study was to evaluate one local program to address low student performance and the achievement gap.

Definitions

Academic learning time (ALT): The amount of time a student is actively, successfully, and productively engaged in learning. In its original conception by Carroll (1963), the formula for determining ALT included involvement, content overlap, and overall success. ALT is the span of time when a student is involved in instructional activities which are specifically aligned with the student's intuitiveness towards learning (Fisher & Berliner, 1985).

Achievement gap: The difference between the achievement scores of students, specifically those labeled as at-risk and economically-disadvantaged as measured on the STAAR standardized tests (see Madyun, 2011).

At-risk students: Students are classified as at-risk if they meet at least one of the 13 state identified criterion that place them at a higher risk of not completing high school (TEA, 2011a).

Block-scheduling: Block-scheduling involves the rearrangement of the school day to lengthen class periods. A regular school day may consist of six to eight class periods meeting less than 60 minutes each day. Block-scheduling reduces the number of classes each day, allowing for longer class periods (Zelkowski, 2010).

Economically-disadvantaged students: Students who qualify for free or reduced-cost meals. This determination is based on family size and income requirements established by the Texas Department of Agriculture Food and Nutrition Division (TEA, 2011a).

Elementary and Secondary Education Act (ESEA): Law signed by President Lyndon Baines Johnson in 1965. ESEA distributed grants to increase education opportunities available to low-income students (ESSA, n.d.).

Every Student Succeeds Act (ESSA): New legislation signed by President Barack Obama in 2015. ESSA is a reauthorization of ESEA (ESSA, n.d.).

Improvement Required State Accountability Rating: In 2013 Texas Legislature passed House Bill (5) mandating the creation of a new accountability system. Schools and districts failing to meet minimum standards on any of the four indexes (student progress, student achievement, postsecondary readiness, and closing performance gaps) receive an improvement required rating (TEA, 2013).

Low performing students: Students whose scaled scores on the STAAR assessment fall into the lowest category, Level 1. Scores in Level 1 are classified as unsatisfactory (TEA, 2013)

No Child Left Behind (NCLB): Federal law P. L. 107–110 passed in 2002. The legislation seeks to ensure every child is educated to his or her full potential through greater accountability (NCLB, 2001).

Scaled scores: The conversion of a raw score on an assessment to a common scale. The scaled score facilitates comparison between students and test years. Scaled scores are used to comparison slight variations from one version of the test to the next (TEA, 2013).

Social justice theory: Social justice theory revolves around the concept of reallocating goods and resources to improve the situations of the disadvantaged (Adams, Bell, & Griffin, 2007).

Targeted intervention and support: One of the consequences assigned to schools that consistently have low performing subpopulations. The prescribed interventions and level of support have yet to be defined (ESSA, n.d.)

Significance

The results of the study are significant for KTISD and to any other district struggling to address low math performance and the achievement gap existing between economically-disadvantaged and at-risk students. The primary goal of the block-scheduling program was to provide extended instructional time as a means to improve student performance, and therefore, reduce the achievement gaps present within the district. KTISD's SY12 STAAR data indicated that 31% of the students labeled as economically-disadvantaged performed below satisfactory level compared to 22% of those not classified as economically-disadvantaged (TEA, 2013). This represents a 9% achievement gap between the two subpopulations. The data concerning at-risk student performance present a dire situation. A 20% gap was reported for this subpopulation for SY12. Only 64% of the at-risk students received satisfactory ratings in contrast to an 84% for students not considered to be at-risk (TEA, 2013). District leaders were aware that the continuance of achievement gaps of this significance would result in KTISD failing to meet state standards.

Researchers offer several explanations regarding the achievement gap. Several (see Condron, 2011; Goodman & Burton, 2012; Reardon, 2013) focused on economic disparity, social class, and school racial segregation as possible associations. Others concluded that the achievement gap was a result of the differing levels of preparation a student acquired before entering the school system, affirming that families who were considered at-risk or economically-disadvantaged had less resources available to adequately prepare their students for formal education (Memon et al., 2010; Robinson, 2010). Huang and Sebastian (2014) and Ullucci (2011) claimed that the achievement gap is perpetuated by the lack of teacher experience in dealing with ethnic minority and disadvantaged students. Regardless of the possible reason for the gap, KTISD leaders knew the stakes were high and they strived to find solutions.

Guiding Research Questions

The purpose of this study was to evaluate the 90-minute block-scheduling program (independent variable measured on a nominal scale) on the student achievement (dependent variable measured on an interval scale) of low performing students. The block-scheduling math program was implemented to increase overall performance of low achieving students and decrease the achievement gap existing between at-risk and economically-disadvantaged students and their counterparts. I developed the following research questions (RQs) to guide this study:

RQ1: Is there a significant difference between the math scores of low performing students who participated in the 90-minute math block-scheduling program and those who did not?

$H_0$1: There is no significant difference in the math scores of low performing students who participated in the 90-minute math block-scheduling program and those who did not.

$H_1$1: There is a significant difference in the math scores of low performing students who participated in the 90-minute math block-scheduling program and those who did not.

RQ2: Is there a significant difference in the math scores of low performing at-risk students who participated in the 90-minute math block-scheduling program and those who did not?

$H_0$2: There is no significant difference in the math scores of low performing at-risk students who participated in the 90-minute math block-scheduling program and those who did not.

$H_1$2: There is a significant difference in the math scores of low performing at-risk students who participated in the 90-minute math block-scheduling program and those who did not.

RQ3: Is there a significant difference between the math scores of low performing economically-disadvantaged students who participated in the 90-minute math block-scheduling program and those who did not?

$H_0$3: There is no significant difference in the math scores of low performing economically-disadvantaged students who participated in the 90-minute math block-scheduling program and those who did not.

$H_1$3: There is a significant difference in the math scores of low performing economically-disadvantaged students who participated in the 90-minute math block-scheduling program and those who did not.

Review of the Literature

In this subsection, I will provide a literature review of the ALT theory that served as the theoretical framework and foundation of this study. Also included will be a review of literature involving social justice as it served as an underlying foundation for the study. My primary sources included peer-reviewed journal articles found in the Walden Library. I enhanced my search by including books and credible Internet sources located through Google Scholar. I identified key search terms such as *academic learning time theory*, *math achievement gap*, *block-scheduling*, and *social justice theory* to enhance the depth of resources. My initial search resulted in over 1,100 articles. To ensure the sources remained relevant and current, I periodically revisited the literature review to add sources published during the past 3–5 years.

Theoretical Framework

ALT Theory

The relationship of extended instructional time and increased student performance is rooted in the ALT theory. Carroll (1963) coined the term "time and learning" in a theory that focused on the time students spend engaged in learning. According to this theory, the amount of time a student applies to actively learning influences the degree of leaning (Carroll, 1963). This theory accounts for the fact that students are all individuals and learn at different speeds and levels of mastery based on experience, ability, and instructional quality (Carroll, 1963). However, learning time remains a factor that affects all students' performance; Grave's (2011) findings supported this theory and indicated a positive correlation with student's success and the time spent in class.

The time and learning theory was extended by Fisher and Berliner (1985) into what would eventually become the ALT theory. Fisher and Berliner described ALT as the time students are actively, successfully, and productively engaged in the learning process. In addition to simply time spent learning, this theory proposed that academic learning time must be instructional, engaging, have an appropriate difficulty level, and be aligned with desired outcomes (Fisher & Berliner, 1985). Schools extending their class periods and school days are, in effect, increasing the number of opportunities for ALT (Redd et al., 2012). Therefore, extended time and instruction are not mutually exclusive; both play an important role in student learning.

In its original conception by Carroll (1963), the formula for determining ALT included content overlap and successful student involvement. Content overlap involves the amount of tested material that is actually reviewed by students while in the classroom (Brady, Clinton, Sweeney, Peterson, & Poynor, 1977). Time on target is another phrase used to describe content overlap. The last component of Carroll's formula for success involves the extent a student correctly completes the assigned tasks (Fisher & Berliner, 1985). Therefore, to be effective, ALT requires (a) students cover assigned material, (b) students stay engaged during instruction, and (c) students are successful on those tasks (Farbman, 2011a). ALT is deemed the quantity of time and quality of education that a particular student engages in studying related academic assignments for performing future tasks with a high rate of success (Fisher & Berliner, 1985). In other words, ALT is the period in which a student engages in instructional activities that build toward future goals and align with desired outcomes (Buzza & Dol, 2015; Farbman, 2010).

ALT theory was developed further in consideration of the amount of time available to instruct students and its effect on achievement. Thomas and Grimes (1995) determined that four variables contributed to ALT: (a) time used for instruction, (b) allocated time, (c) engaged time, and (d) academic success. In researching this theory, social scientists moved from just looking at the quantity of instruction to considering the time available and time used successfully to improve academic achievement (Mega, Ronconi, & De Beni, 2014; Thomas & Grimes, 1995; Wagner, 2013). When increasing the time for instruction it is important to not slight the quality of instruction (Heitin, 2015; Patall, Cooper, & Allen, 2010; Tanner, 2009; Wagner, 2013).

The names associated with the theory have changed but the foundational principals remain steadfast (Gersten et al., 2012; Sams & Bergmann, 2013). Over the years, research has continued to focus on the importance of time spent on learning as it relates to student achievement. Jez and Wassmer (2013) conducted a study of low performing students at a California middle school. Their study showed an increase in the academic performance of disadvantaged students who participated in extended learning. Wang, Brinkworth, and Eccles (2013) reported similar results. Their findings showed students who participated in extended math classes experienced higher academic achievement and increased student motivation. Another study conducted in 2013 further showed support for ALT theory. Participants in an extended math program focused on engagement and time on task showed increased math scores on end of course assessments (Klanderman, Webster-Moore, Maxwell, & Robert, 2013).

Recently, Biddle and Mette (2016) published a report that reiterated the need for the relationship between engagement, quality instructional activities, and desired outcomes. They presented a report to the Maine Education Policy Research Institute outlining the advantages of extended learning time utilized properly. They reported extended programs that incorporated sequenced, active, focused, and explicit elements showed increased academic, emotional, and social gains for the students involved.

Not all researchers support the theory of ALT and its effects on academic achievement. Farbman (2011a) presented findings after reviewing instructional practices at three urban school districts attempting to use ALT. Farbman found that students participated in a curriculum that did not differentiate instruction or provide technology integration. While observing classrooms, the researcher found a significant number of students off task during instruction. One conclusion that can be drawn from this project is that ALT is difficult to attain without engaging lessons that are built around students' needs and change as the students grow.

Hackney (2013) conducted a study of extended learning time on math performance at an Illinois high school. The researcher found teacher's perceptions favored the traditional schedule vs. block-scheduling. Teachers reported it was difficult to keep students engaged in block-scheduling. Moreover, analysis of student outcomes on the math portion of the state standardized assessment did not show a significant difference in the percent of students meeting or exceeding standards between the students on the block schedule and those on the traditional schedule (Hackney, 2013).

While the evidence is not conclusive regarding ALT and its effects on academic achievement, it was the basis for this study. The theory relates to this study because the 90-minute math block schedule implemented by KTISD was developed around the hypothesis that extended time in the math class would increase the performance of low performing students. The extended class periods would allow teachers additional time to provide differentiation while supervising practice and exploration assignments and diminishing the need for take home assignments. Kim (2012) reported the success of students who were able to explore new concepts in a positive learning environment. The 90-minute math block schedule was an attempt to address some of the external factors that impact KTISD's at-risk and economically-disadvantaged students. District leaders believed applying an intervention along with positive feedback and continued remediation would help negate external factors students could not control. James and Folorunso (2012), Midkiff and Cohen-Vogel (2015), and Doykos (2015) all reported results that indicated increased performance of low performing students who participated in extended learning time in math classes.

Social Justice Theory

Social justice theory served as an underlying framework for this study. Identifying members of disadvantaged populations and removing boundaries preventing them from academic success is one of the founding principles of the social justice framework (Mullen, 2010). The social justice framework also addresses the hidden and blatant injustices in the lives of at-risk populations and vulnerable populations (Duncan & Murnane, 2014; Sue et al., 2007). Vulnerable populations include marginalized groups of people and ethnic minorities (Fitts & Weisman, 2010) and neglected children who lack stable, dependable, and nurturing relationships with positive care providers (Fleming-May, Mays, & Random, 2015; Gerdes, Segal, Jackson, & Mullins, 2011). Middle class students look to their parents to serve as their advocates, while disadvantaged students rely on educators to step in and act as advocates on their behalf (Jensen, 2013).

Social justice revolves around implementing programs and policies that will serve as equalizers for vulnerable populations (Nicotera & Walls, 2010). The 90-minute math block-scheduling program is one example of a supplemental service designed to equalize learning opportunities by allowing more time to foster student-teacher relationships, provide educational support that may not be available at home, and increase the opportunities for student reflection (Stebleton & Soria, 2012). These three outcomes have been reported to increase student academic performance and decrease discipline issues (Arce, 2005; Koyanna, 2011).

In order to establish relationships, there needs to be an understanding and respect of the students' background (Cohen & Schuchter, 2013; Van Soest & Garcia, 2003). The backgrounds of disadvantaged students vary and may be very different from that of the teacher. The background of a student plays an important role in how the student learns and processes information (Howe & Covell, 2013; Sahabudin & Ali, 2013). Teachers who are cognizant of a student's background have an opportunity to meet those students where they are socially and academically. This awareness is one of the first steps toward establishing social justice (Crosnoe & Leventhal, 2013; Rojas & Liou, 2017).

Bennett, Lutz, and Jayaram (2012) highlighted the important role educational institutions play in providing the structural environment that may be lacking at home. Extended time in class helps reinforce structure and allows for students to develop bonds with teachers (Feldman &Tyson, 2014). Francis, Mills, and Lupton (2017) believed these bonds provide students with an opportunity to learn how to accept productive and positive reinforcement designed to foster educational resilience. Educational resilience is a foundational element necessary for disadvantaged students to embrace as they realize high school graduation and the pursuit of secondary education are not only realistic but attainable (Foran, 2015).

Schools are full of students who may not be able to grasp the grade-level concepts being taught due to their personal and socioeconomic situations outside the classroom. These students will ultimately face long term disadvantages related to obtaining the benefits given to more advantaged counterparts if some intervention is not implemented (Cadiero-Kaplan, 2004; Power & Taylor, 2013). In the field of education, the social justice framework supports the concept that all students deserve equal educational opportunities; however, equal opportunity does not always mean identical treatment (Knudsen, 2009; Miller, 1976). Extra programs and services may be needed to support struggling students (Crosnoe & Leventhal, 2013; Kaufman & Blewett, 2012).

One example of extra programs and services is extended learning time. Additional time in class can provide students with an opportunity to complete homework assignments under the supervision of a teacher in a positive learning environment (Fan, Xu, Cai, He, & Fan, 2017). Often parents of disadvantaged students are not available or capable to assist students with homework assignments. Additionally, the home environment may not provide a suitable atmosphere that allows the student focus and concentrate on assignments (Gonida & Cortina, 2014). Either or both of these situations can result in the student failing to complete assignments or completing them incorrectly.

Regardless of the reason, the student has missed out on an opportunity to practice concepts. This continual pattern can lead to the student falling behind in class as they have not practiced enough to master necessary skills before moving onto new concepts (Bradley & Corwyn, 2016; O'Malley, Voight, Renshaw, & Eklund, 2015).

Advocates for social justice argue standardized testing does not promote social equity because a significant portion of at-risk and disadvantaged students struggle in the classroom to meet mandated proficiency standards mandated (Luginbuhl, McWhirter, & McWhirter, 2016; Power, 2012). Knudsen (2009) advocated supplemental services designed to equalize learning opportunities for those students who may be disserved by such broad decision-making. It is unclear how and if new educational policies (ESSA) will impact standardized testing. However, the need to safeguard the role of education in the pursuit of social justice should not be overlooked (Darling-Hammond, Zielezinski & Goldman, 2014; Morris & Perry, 2016).

Math Achievement Gap

Diligent efforts to close the achievement gap have fallen short. In fact, some believe the gap is widening (Goodman & Burton, 2012; Johnson & Kritsonis, 2010; Reardon, 2013) and educational leaders need to focus on programs to diminish the gap. Jackson and Wilson (2012) reported that teachers felt unprepared to effectively teach and reach students of diversity. They argued that schools needed to provide additional professional development to address best practices to assist teachers in addressing the unique needs of students classified as economically-disadvantaged as well as those who are at risk for high school completion. It is critical for educators to be conscious of the different types of knowledge and experiences students bring with them into the classroom (Delpit, 2012; Huang & Sebastian, 2014; McDonald, Polnick, & Robles-Pina, 2013) and incorporate these experiences into meaningful lessons (Marzano, 2013).

The possible reasons for the achievement gap are plentiful. Simms (2012) identified three potential reasons for the achievement gap: peer pressure, family background, and school effects. Perry and McConney's (2010) report from a 12-year longitudinal study identified the family's socio-economic status as a key factor influencing student performance and a student's potential for on-time high school completion. Gut, Reimann, and Grob (2013) showed students from disadvantaged populations had an increased association with behavioral problems. They concluded that these behavioral problems may be contributing factors to the achievement gap. Some of the problems related to restlessness, defiance, lack of self-control, disruptive classroom behavior, and hyperactivity (Reinke, Herman, & Stormont, 2013). Recently Ratcliff et al. (2016) studied teacher perceptions regarding the achievement gap. They reported teachers felt solutions to the achievement gap required: (a) more teacher professional development, (b) more time with students, (c) additional training for administrators regarding discipline, and (d) better communications with parents and family members.

Gregory, Skiba, and Noguera (2010) believed the achievement gap was attributed to discipline problems. The researchers associate the use of exclusionary practices, as a means of discipline, as a possible reason for the achievement gap. Morris and Perry

(2016) reported findings that supported Greagory, Skiba, & Noguera. They found disparities in school suspension practices between disadvantaged students and their peers. They claimed the high incidence of school suspension was a contributing factor in the achievement gap. Suspensions (both in and out of school) results in missed classroom instruction. In turn, missed instruction time can lead to achievement gaps (Morris & Perry, 2016).

Physical and psychological reasons have also been explored as a means to understand or justify the achievement gap. A neuroanatomical study was performed by Mackey et al. in 2015. The study showed students from higher income families had thicker cortical gray matter compared to the cortical gray matter of lower income students. Gray matter is associated with critical thinking and intelligence (Cheema & Kitsantas, 2014; Takeuchi et al., 2014).

In contrast, Wang (2010) argued that race and socioeconomic status have the greatest impact on math achievement. Instead of placing the blame on socioeconomic status and families, Naraian and Brown (2011) advocating research to determine what practices and programs were successful in helping economically-disadvantaged and at-risk students close the achievement gap. Young and Young (2016) conducted a study in Texas regarding the reason or reasons for the continued achievement gap. Their findings showed the reasons varied from region to region. They recommend local researchers conduct studies to address the issues present in that area.

Block-scheduling

My review of literature surrounding 90-minute math block-scheduling resulted in conflicting findings regarding the effectiveness of block-scheduling. President Obama and former Secretary of Education, Arnie Duncan, both promoted longer learning time as one of the core approaches to help turn around low performing schools (Farbman, 2011b; Kolbe, Partridge, & O'Reilly, 2011). They felt so strongly about this concept they fought to fund government programs especially designed to support additional learning time for students (Farbman, 2011b). In addition, Wright (2010) performed an extensive evaluation of 90-minute math block-scheduling that spanned 10 years. This study was conducted in South Carolina and showed a statistically significant improvement in math scores during the period the 90-minute math block-scheduling model was in place.

More support of 90-minute math block-scheduling is provided by Martinez and Holland (2011). Their study measured the performance of English Language Learners in Texas. Students who participated in the 90-minute math block schedule received higher math and reading scores compared to those in the traditional 45-minute class (Harvey, 2013; Joyner & Molina, 2012; Martinez & Holland, 2011). Although Gill (2011) was not able to find a notable difference in the overall scores for the populations he studied, higher scores were noted for the minority and at-risk students in the 90-minute math block schedule compared to those in the regular class.

Additional support for block-scheduling can be found in the findings of recent studies. Multiple researchers reported math block-scheduling presented positive relationships to increased student performance, decreased disciplinary issues, and overall student engagement of low performing students (Cortes, Goodman, & Nomi, 2015; Taylor, 2014). These findings further supported the implementation of 90-minute math block-scheduling as an option to address the achievement gap of economically-disadvantaged and at-risk populations in KTISD.

Researchers argued that extended class time provided educators with the opportunity to strengthen relationships with students. These relationships can lead to improved self-confidence, reduced classroom stress, and impact student learning (Benken, Ramirez, Li, & Wetendorf, 2015; Flocco, 2012; Harvey, 2013; McInerney, Cheng, Mok, & Lam, 2012). Sokal and Katz (2015) examined the engagement levels of middle school students enrolled in block-math classes for 1 year. Reports showed the students felt the extended time in class not only helped improve their math skills but it also promoted a more positive attitude about math in general.

In contrast, when Norton (2010) completed a study of South Carolina schools that utilized a block schedule, he found no significant differences in the scores of the students who participated in math block-scheduling and those that did not. Norton suggested continued professional development was needed for teachers. In that same vein, some researchers argue that extended class time has proven to be unyielding and labor intensive for teachers (Chute, 2012; Forman 2009). The intent to increase accountability has resulted in educators becoming frustrated with having to find ways to re-teach and re-work concepts they believe they have already taught (Brill, 2011; Martinez & Young, 2011). This opened the possibility professional development was warranted to help educators learn what to do with extended time (Hachey, 2013, Nomi & Raudenbush, 2016).

However, Gabrieli (2010) did not believe professional development for teachers was the solution. The researcher argued that teachers are already well trained in innovative and engaging teaching methods. Additionally, it was not believe extended classes would make a difference in student scores because middle school students have a restricted attention span lasting approximately 35-45 minutes. The findings suggested any positive results associated with block-scheduling could be attributed to the collection of school wide reforms implemented simultaneously with block-scheduling. Patall, Cooper, and Allen (2010) reported further doubts about the value of 90-minute math block-scheduling. Their study pointed out that it was difficult to isolate the reason for the increase in student achievement and posited that other researchers used weak casual inferences to support their findings. A later study echoed the opinion of Patall, Cooper, and Allen. Preston, Goldring, Guthrie, Ramsey, and Huff (2016) held that block-scheduling played only a small role in increased student achievement. They alleged several campus and family factors had to be present to bring about increased performance. Although several researchers observed increased performance of students enrolled in block math scheduling they cautioned that these increases were short-lived. Once students reverted back to a regular math schedule, summative assessments showed they could not maintain the levels of success reported while enrolled in the block schedule (Henry, Barrett, & Marder, 2016; Sheridan, Smith, & Pleggenkuhle-Miles, 2017).

Clearly, research is available to support both sides of the debate regarding 90-minute math block-scheduling. Leaders of KTISD were aware of both the positive and negative reports concerning block-scheduling. They carefully weighed the current research and decided to implement an extended math schedule for a 1 year test period during SY13. The district leadership has asked me to perform an evaluation of the program's outcomes to gauge the impact of extended class time on the performance of the eighth grade students at the pilot campus.

Implications

Tight budget constraints in SY14 did not afford the district with the option of considering the continuation of the extended math program and the program was never evaluated. It was recently announced that the district may be eligible for a substantial grant. District leaders are considering using these funds to cover the costs associated with the math block-scheduling program. The findings of this evaluation were presented to district leaders in the form of an evaluation report (See Appendix A). This summative report, my project study, detailed the findings of the evaluation along with recommendations concerning the program. District policy-makers will be considering the following options: (a) reinstate the math program, (b) continue with the 50-minute class schedule, or (c) explore other alternatives. The findings of one study cannot possibly provide a clear solution to address the achievement gap; however, this study may serve as catalyst for further research into block-scheduling as a means to address the math achievement gap in this school district.

Summary

The purpose of this project study was to evaluate the effectiveness of the 90-minute math block schedule implemented to decrease the achievement gap between low achieving at-risk students and economically-disadvantaged students and their counterparts at KTISD. In Section 1, I provided the outline of the problem and the justification for the study. In Section 2, I will provide details concerning the methodology for this project study.

Section 2: Methodology

Introduction

The purpose of this project study was to evaluate the 90-minute math block-scheduling program implemented at KTISD to increase the achievement of low performing students on the math section of the STAAR assessment and decrease the achievement gap between at-risk and economically-disadvantaged students and their counterparts. In this section, I will describe the research design and approach. Additionally, I will outline the research setting, the sample, and the methods of data collection and analysis. This section will also include the assumptions and limitations of the study along with the protection of participants' rights.

Research Design and Approach

To address the RQs concerning the effectiveness of 90-minute math block-scheduling program on the achievement of low performing at-risk and economically-disadvantaged students, I used a quantitative project study design in the form of an outcome-based evaluation. Quantitative research design methods allowed me to address the RQs in numerical terms (see Creswell, 2012). The eighth grade math STAAR scores for SY13 were used to evaluate the effectiveness of the 90-minute math block-scheduling program. In the study, I followed a retrospective casual comparative research approach. A retrospective casual-comparative approach was chosen because I sought to examine the effectiveness of a 90-minute math block-scheduling program that had already occurred (see Gay, Mills, & Airasian, 2008). I did not select an experimental research design because the participants are already grouped in relation to the independent variable (extended time in math class). School district leaders in KTISD selected the treatment group based on the SY12 STAAR achievement scores. The treatment group consisted of the eighth grade students who participated in the 90-minute math block-scheduling program for the entire SY13 and completed the math portion of the STAAR assessment while residing in KTISD. The comparison group was comprised of students who did not take part in the 90-minute math block-scheduling program for the entire SY13 and completed the math portion of the STAAR assessment while residing in KTISD.

With the call for increased accountability, program evaluations are an important tool used by educators to adequately evaluate programs and interventions to measure their effectiveness (Stufflebeam & Shinkfield, 2007). Program evaluations can serve as a means to measure whether or not a program's objectives were met. In addition to measuring goal attainment, program evaluations can help educational leaders make suggestions concerning program modifications or improvements (Lodico, Spaulding, & Voegtle, 2006).

This outcome-based evaluation was summative as outlined by Spaulding (2014). District leaders at KTISD sought to determine the future of the 90-minute math block-scheduling program. Outcome-based evaluations are usually conducted following the conclusion of the program and provide information to determine if the program's stated objectives were met (Spaulding, 2014). In this study, those objectives included (a) increasing the overall achievement scores of low performing students and (b) reducing the achievement gap of low performing at-risk and economically-disadvantaged students and their peers.

Setting and Sample

Two middle schools within the KTISD served as the setting for this study. KTISD is a large public school district situated in Central Texas. Currently, more than 43,000 students are being served at six high schools, 14 middle schools, 35 elementary schools, three alternative schools, and a Career Center. KTISD's diverse student population is 34% African American, 28% Hispanic, 26% Caucasian, 7% mixed race, 3% Asian, 1% American Indian, and 1% Pacific Islander. Fifty-seven percent of the student population is classified as economically-disadvantaged, while 52% of the students are identified as at-risk. District reports show approximately half of the students are military dependents, which contributes to a student mobility rate of 34%. The eighth grade students on one campus, the treatment group, participated in the 90-minute math block-scheduling program. The eighth grade students on the other campus, the comparison group, did not participate in the extended math program. The campus used as the comparison group was selected because it had a similar demographic profile to treatment group as reported by TEA (2014) for SY13. The at-risk and economically-disadvantaged populations were relatively equal to the treatment group. Additional similarities included overall campus size and the number of students performing at an unsatisfactory level.

The population included eighth grade students (N = 109) from the treatment and comparison campuses whose performance was classified as unsatisfactory on the math portion of the STAAR assessment. I used archival data in the form of math scores reported by TEA (2014) for the SY13 administration of the STAAR assessment. Only data as they related to the specific groups of low performing students were used; therefore, a purposeful sampling technique was suitable. Total population sampling, a form of purposeful sampling, involves collecting data on every individual in the identified population (Creswell, 2012). Total population sampling was appropriate because the populations being studied were small, the data were easily obtained, and there were no additional costs incurred by including the entire population (Creswell, 2012). Furthermore, the use of total population sampling eliminated the sampling error (Creswell, 2012).

I identified three subgroups from the population of the comparison and treatment campuses to address the RQs. The first subgroup included all of the students who were classified low performing as determined by the scores identified by TEA. For the Spring 2014 administration of the STAAR, students scoring less than 1583 on the math portion were classified as low performing (TEA, 2014). The treatment group (n = 49) and comparison group were identified (n = 60). The second subgroup consisted of students classified as low performing and identified as at-risk. TEA (2011a) classifies students as at-risk if they meet one of the 13 identified criteria that place them at a higher risk of not finishing high school. The treatment group in this subpopulation (n = 49) and the comparison group (n = 60) was the same as for the first RQ. The final subgroup included students who were classified as low performing and identified as economically-disadvantaged. Students who qualify for free or reduced meals based on the criterion established by the Texas Department of Agriculture Food and Nutrition Division are considered economically-disadvantaged (TEA, 2011a). In this case, the treatment group (n = 36) and the comparison group (n = 44) differed from those selected for the previous research questions.

The focus of this study was targeted at a specific population--low performing students. This narrow focus meant it was possible for a student to be included in all three subgroups. For example, the fact a student was classified as low performing meant they were automatically classified as at-risk. Low performance on a standardized test is one of the 13 indicators used by Texas to identify at-risk students (TEA, 2013). Therefore, the sample I used to address RQ1 and the sample for RQ2 were identical.

In contrast, not all students classified as low performing are economically-disadvantaged. Within the target population, 16 students on the comparison campus along with 13 on the treatment campus were not identified as economically-disadvantaged. The individual scores from each of the treatment groups are not related to any of the individual scores in the comparison group. The use of statistical analysis for independent measures was still appropriate as the RQs were targeted at three different populations.

Instrumentation and Materials

The STAAR served as the data collection instrument in this study. The STAAR assessment is given to all students in Texas to measure their achievement in relation to the TEKS (TEA, 2013). Each spring, eighth grade students are required to obtain a passing score on the math portion of the STAAR in order to be promoted to ninth grade (TEA, 2013). The STAAR is a criterion-referenced test, and a criterion-referenced test measures the achievement of an individual student against a predetermined standard (Creswell, 2012). TEA reports student performance as percentage passing, scaled scores, and commended performance. I used the scaled scores for this study because multiple versions of the assessment are administered. The use of scaled scores provides comparability (Educational Testing Service, 2010). The STAAR Technical Assistance Center (2010) reported that the use of scaled scores serves as a common reference from year to year. This common reference helps when researchers want to analyze data from different versions of the assessment. TEA (2013) conducts ongoing research to ensure the reliability and validity of the STAAR assessment annually.

In this study, I focused on students who received an unsatisfactory rating on the math section of the STAAR assessment administered during the Spring of 2014. This test was designed to measure a student's achievement at the end of SY13. A scaled score below 1583 resulted in a classification of performing unsatisfactorily (TEA, 2014).

Data Collection and Analysis

In accordance with the research guidelines set forth by KTISD, the district would not develop or prepare any reports to be used in this study. Therefore, I collected the data TEA in the form of STAAR assessment results for the treatment campus and the comparison campus. A written request was made to the TEA Student Assessment

Division to release the individual scaled mathematic scores of eighth grade students at the treatment and comparison campuses for SY13 (see Appendix B). Additionally, I requested TEA identify students by socioeconomic and at-risk status. Further, the data were requested to be deidentified with pseudo-codes attached to the students' scores.

I used SPSS to perform inferential statistics to analyze the data. Inferential statistics are appropriate when seeking to "compare two or more groups on the independent variable in terms of the dependent variable" (Creswell, 2012, p. 182). The independent variable was enrollment in the 90-minute math class, which was measured categorically. The dependent variable was the achievement scores from the STAAR assessments. The groups were independent and the participants only took one test. There was one independent and one dependent variable to be measured. Initial analysis determined the data were not normally distributed therefore a Mann-Whitney *U* test was selected to test the hypotheses. I selected the Mann-Whitney *U* test because it would allow me to compare two independent groups on the same continuous, dependent variable (see Creswell, 2012). The Mann-Whitney *U* test is similar to a *t* test but does not require the dependent variable to be normally distributed (Corder & Foreman, 2014). The goal of the data analysis was to determine if there was a significant difference between the scores of the treatment group and the comparison group. The results were used to determine the effectiveness of the 90-minute math block-scheduling program.

Assumptions, Limitations, Scope, and Delimitations

For the purposes of this study, I made several assumptions. First, it was assumed that the STAAR assessment was a valid and reliable instrument. Next, it was assumed that the students put forth their best effort on the STAAR assessment. Another assumption was that the extended time in class was utilized to provide academic enrichment. This assumption was based on the fact that teachers were required to turn in lesson plans outlining how the extended time would be used, and campus administrators performed walkthroughs to verify adherence to the lesson plans.

There were limitations associated with this study that could affect external validity. Threats to external validity involved issues that could threaten the capability to draw accurate inferences from the sample data (Creswell, 2012). The sample of students from only two campuses limited the ability to make any generalizations beyond the two schools involved in the study. Furthermore, the post-hoc power analysis showed the statistical power for all three Mann-Whitney *U* tests fell short of 80%. This could be attributed to the small sample population being tested (Sullivan & Feinn, 2012). The use of archival records could have also been a limitation if there were errors or changes in record-keeping procedures. Additionally, it is important to be sensitive to the possibility of spurious relationships that may arise when using archival data (Creswell, 2012). The analysis of archival data may overlook possible correlations between variables not being measured. For example, campus culture or level of teacher experience could have an influence on test scores.

Threats to internal validity cannot be overlooked when using a casual comparative research design. The groups being studied may differ in some way, not covered in the scope of this study, which may influence the results. In an attempt to control the extraneous variables, the comparison campus was selected because its demographic profile was similar to that of the treatment campus. Both campuses were similar in overall student enrollment, number of students ranked as low achieving, and in the percentages of students identified as at-risk and economically-disadvantaged.

The scope of this study focused on the extended time provided through the implementation of a 90-minute math block-scheduling program and its impact on the STAAR scores of low performing students in KTISD. The curriculum or teaching strategies employed during the program were not addressed, only the extended time in class. Further, the level of teacher experience was not considered in this study. The use of a casual comparative research design, along with the possibility of uncontrolled extraneous variables places limitations on the evaluation report. KTISD leaders plan to utilize the findings of the evaluation to make future budgetary decisions. However, they were advised to refrain from trying to over generalize the findings.

Protection of Participants' Rights

During the proposal stage of this study, permission to conduct the study within KTISD was obtained by submitting a written request to the district's executive officer (see Appendix C). The executive officer approved the request and provided district guidelines regarding KTISD's research policy (see Appendix D). Walden University

Internal Review Board (IRB) granted final approval to conduct the study on January 21, 2016 (01-05-16-0379390).

The use of pre-existing STARR Assessment data gathered from the Texas Education Agency (TEA) eliminated the need for informed consent forms. Once IRB and district approval were gained, a written request for the data was sent to TEA's public information coordinator of the Student Assessment Division (see Appendix B). The data received from TEA contained no student names and ID numbers in order to ensure the privacy of the participants. The data were saved on a password protected thumb drive and stored in a locked file cabinet located within an office that was not shared with others. The data will remain on file for 5 years after the conclusion of the study and will be shredded upon the established termination date.

Although I was previously employed as a math teacher in KTISD, I have not taught any of the students that could be involved in the study. I left the classroom several years before the implementation of the block schedule pilot program. Additionally, I have no direct leadership influence over any of the teachers or district personnel involved with the study.

Data Analysis

The purpose of this quantitative study was to evaluate the 90-minute math block-scheduling program implemented at one middle school in KTISD to determine if it improved the achievement of low performing students on the math section of the STAAR assessment and decreased the achievement gap between at-risk and economically-disadvantaged students and their counterparts. Data regarding student mathematical achievement on the eighth grade STAAR assessment were collected from students exposed to the 90-minute math block program, and those who were not exposed to the program. In the remaining portion of Section 2, I will discuss the summary statistics, data preparation of the dataset, hypothesis testing and finally present a conclusion based on the results.

All data analysis was conducted in IBM SPSS 23. I initially performed the data analysis myself, however to ensure accuracy and validity, I felt it was better to seek a second opinion. I utilized the expertise of a local statistics professor to review my input, analysis, and findings. The data provided to the professor contained no identifying campus or student information so there was no need for a confidentiality agreement.

Summary Statistics

Mathematics scores were collected from 109 low performing students, 60 of which were the comparison subjects and 49 who had been exposed to the 90-minute math block-scheduling program. A score of 1583 on the STAAR was used as a cut off for students who were considered low performing, as per the STAAR guidelines (TEA, 2014) and all students who scored above 1583 were removed from the analysis leaving only low performing students in the data set. Because the research questions were directed at specific groups within those identified as low performing, three sub-populations were established (low performing, low performing and at-risk, and low performing and economically-disadvantaged). During the development stage of this study, I decided to analyze the entire target population for each research question. This decision was based on information gained from the examination of archival data for the treatment and comparison campuses. The examination indicated that (a) the data would be relatively easy to retrieve and analyze, (b) there would be no additional time or costs associated with analyzing the entire population, and (c) there would be less than 100 students in each sub-population. The inclusion of the entire target population for each research question made it likely there would be unequal sample sizes. The populations for research questions RQ1 and RQ2 were $n = 49$ for the treatment campus and $n = 60$ for the comparison campus. Research question RQ3 was tested using a population of $n = 36$ for the treatment campus and $n = 44$ for the comparison campus. The unequal sample sizes did not preclude the use of Mann-Whitney *U*-tests. It only limited the statistical power to that of the smaller sample (Cohen, Manion, & Morrison, 2013; Corder & Foreman, 2014). Additionally, it was noted that statistical power decreased as sample sizes became more unequal (Creswell, 2012).

In order to run a Mann-Whitney *U* test, four assumptions must be satisfied. The first assumption is there is only one dependent variable measured on a continuous or ordinal scale. The second assumption is the dependent variable has two categorical independent groups. The third assumption is there is no relationship between the observations of each group; they are independent of each other. The fourth and last assumption is the distribution of the scores has to be determine to be either the same shape or different shapes, which will help determine if the results will be interpreted with differences in the median or mean scores (Corder & Foreman, 2014).

Equal sample size is not one of the four assumptions (Divine, Norton, Hunt, & Dienemann, 2013; Fay & Proschan, 2010). The variables in this study adhered to the stated assumptions; therefore, a Mann-Whitney *U* test was appropriate. Post-hoc analyses were performed to test the power of each test and the results are reported in the upcoming section along with the findings for each research question. Post-hoc power analyses are used to show if the population was adequate to detect the effect reported (Cohen, 1992; Sullivan & Feinn, 2012).

Summary statistics are presented in Table 1. The table depicts the number of subjects within each sub-population, variable level, and the mean and standard deviation calculations. Generally, students who had participated in the 90-minute math block-scheduling program had a lower mean score than students in the comparison group, with the exception of students who were not considered economically-disadvantaged.

Table 1 Summary Statistics of Math Scores for Comparison and Treatment Campuses

		Comparison			Treatment		
Variable	Level	Count	Mean	Standard Deviation	Count	Mean	Standard Deviation
Low Performing		60	1532.50	30.07	49	1522.51	44.48
At Risk	No	0	-	-	0	-	-
	Yes	60	1532.50	30.07	49	1522.51	44.48
Economically-disadvantaged	No	16	1520.50	34.83	13	1528.30	31.58
	Yes	44	1536.86	27.29	36	1520.41	48.52

The distribution of the mathematics scores for both the comparison and treatment group are presented in Figure 1. Most students scored in the range of 1500 to 1550. The comparison group had slightly higher mean mathematics scores than the treatment group. The treatment group appeared to have less variation in scores across the three sub-populations.

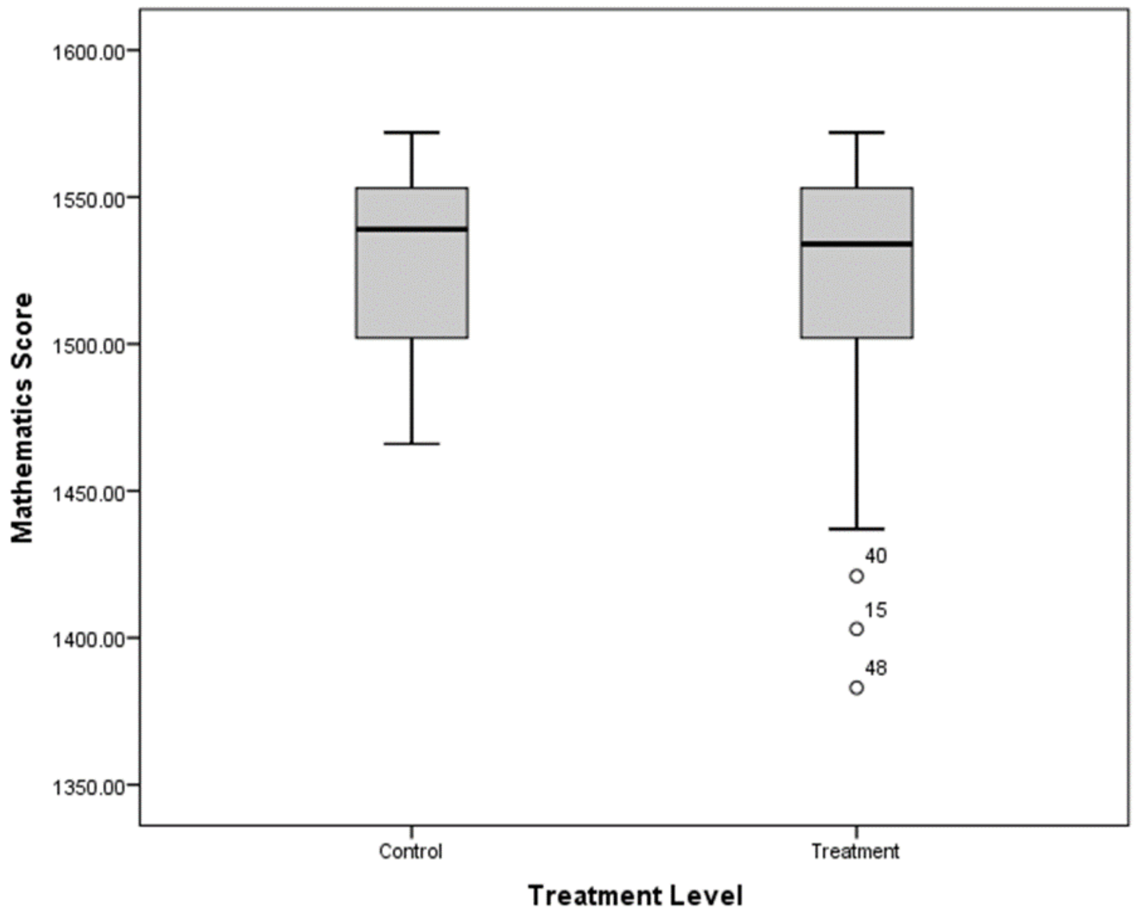


Figure 1. Distribution of math scores for comparison and treatment groups.

Hypothesis Testing

Before analyzing the data, it was necessary to check for the assumptions of normality and variance. Tests to check the assumptions of normality were conducted and revealed the data were not normally distributed. These findings resulted in electing a Mann-Whitney U test that did not require normal distribution of data. The reasons for the non-normal distribution of data and best practices concerning how to address the non-normal distribution will be discussed after the analysis of the research questions.

Research Question 1 (RQ1)

RQ1 was designed to compare the math scores of low performing students who had completed the 90-minute math block program with those who had not. First, the assumptions of normality and equality of variances were checked. Normality was checked using a Shapiro Wilks test and was found to be significantly non-normal (0.918, *df* = 60, $p = 0.00$ for the comparison group and 0.863, *df* = 49, $p = 0.00$ for the treatment group). As the assumption of normality was violated, a t-test was not appropriate and a Mann-Whitney *U* test was used for analysis to compare means as it did not require an assumption of normality. There are no assumptions regarding equal variances for a Mann-Whitney *U* test so these were not conducted.

Table 1 and Figure 1 show that the treatment group had a lower mean score than the comparison group, however this difference was found to be non-significant ($U = 1330.00$, $p = 0.39$). The hypotheses for RQ1 were:

> $H_0$1: There is no significant difference in the math scores of low performing students who participated in the 90-minute math block-scheduling program and those who did not.
>
> $H_1$1: There is a significant difference in the math scores of low performing students who participated in the 90-minute math block-scheduling program and those who did not.

Based on the findings, the null hypothesis failed to be rejected. There is no significant difference between the math scores for low performing students who participated in the

90-minute math block-scheduling program and those who did not. The post-hoc power analysis was performed using SPSS. The .26 effect size (ES) was considered to be small according to Cohen's (1992) criteria. With an alpha = .05 and *n*=49, the power was found to be .38. Meaning there was a 38% probability of getting statistically significant results.

Research Question 2 (RQ2)

RQ2 was designed to determine if there was a difference in the mathematics scores of low performing students who were also considered at risk who participated in the 90-minute math block-scheduling program and those who did not. The fact every student classified as low performing would be simultaneously identified as at-risk meant all data included in this set were exactly the same as for RQ1. Therefore, it was not a surprise that the findings for this research question were the same as the first research question.

Normality was again checked using a Shapiro Wilks test and was found to be significantly non-normal (0.918, $df = 60$, $p = 0.00$ for the comparison group and 0.863, $df = 49$, $p = 0.00$ for the treatment group). As the assumption of normality was violated, a *t*-test was not appropriate and a Mann-Whitney *U* test was used for analysis to compare means as it did not require an assumption of normality. There are no assumptions regarding equal variances for a Mann-Whitney U test so these were not conducted. The hypotheses for RQ2 were:

$H_0$2 There is no significant difference in the math scores of low performing at-risk students who participated in the 90-minute math block-scheduling program and those who did not.

$H_1$2 There is a significant difference in the math scores of low performing at-risk students who participated in the 90-minute math block-scheduling program and those who did not.

Based on the findings of RQ2, the null hypothesis failed to be rejected. There was no significant difference in the math scores of low performing at-risk students who participated in the 90-minute math block program and those who did not. A post-hoc power analysis was performed using SPSS. The .26 effect size was considered to be small according to Cohen's (1992) criteria. With an alpha = .05 and *n*=49 the power was found to be .38. Meaning there was a 38% probability of getting statistically significant results.

Research Question 3 (RQ3)

RQ3 was designed to compare the math scores of low performing students who were also economically-disadvantaged. Figure 2 shows a boxplot of mathematics scores for economically-disadvantaged students. The mean mathematics score for economically-disadvantaged students in the comparison group was 1536.86 (as per Table 1) and the mean mathematics score for economically-disadvantaged students in the treatment group was 1520.41 (as per Table 1).

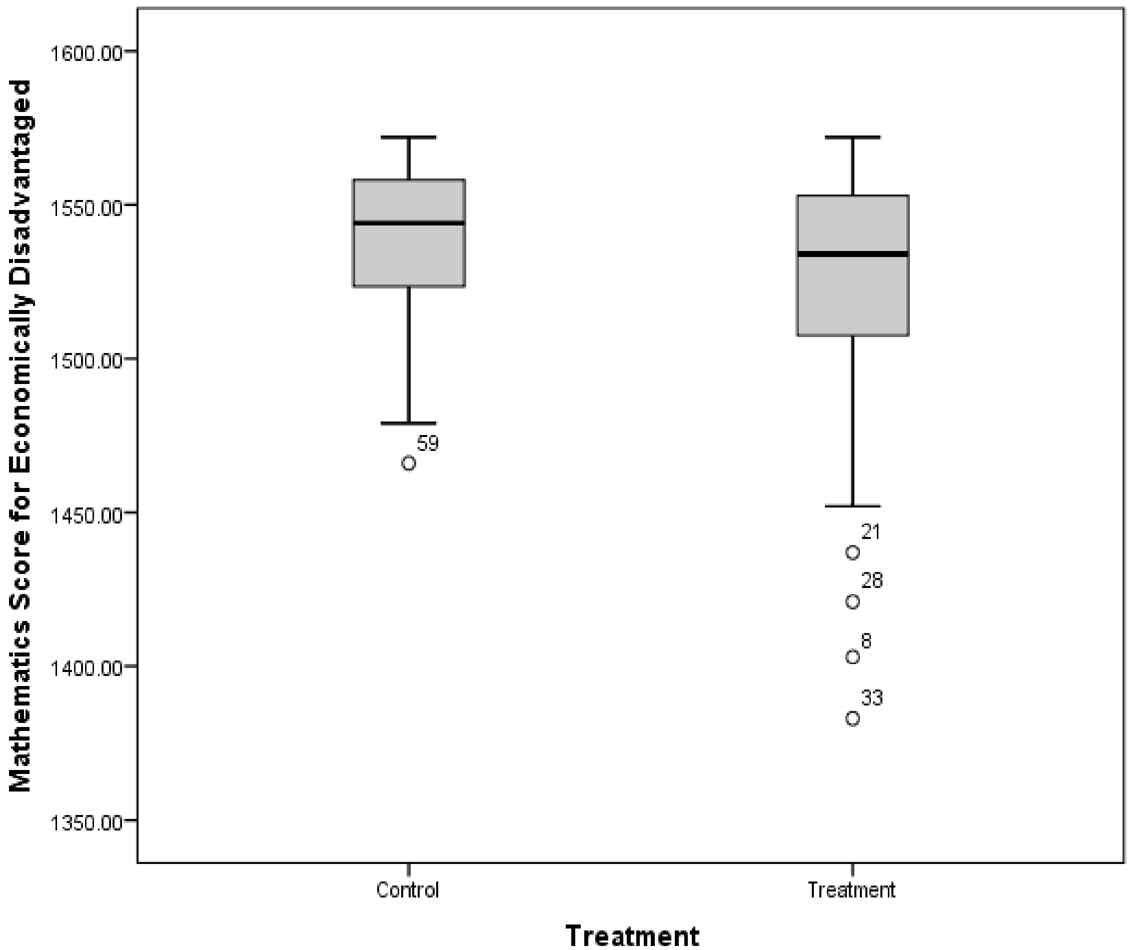


Figure 2. *Box plot of scores for economically-disadvantaged students*

Following the methods for RQ1, first, the assumption of normality was tested using Shapiro-Wilks tests and it was concluded that there was significant non-normality for both the comparison and treatment groups (0.908, *df* = 44, *p* = 0.00, and 0.851, *df* = 36, *p* = 0.00 respectively). Therefore, the Mann-Whitney *U*-test was once again used, this required no testing of the assumption of equal variances. Table 1 shows that the treatment group had a lower mean mathematics score than the comparison group for students who were considered economically-disadvantaged. This difference, however, was not significant (*U* = 662.50, *p* = 0.20). Therefore, I conclude there is no significant difference between the means of the two groups. The hypotheses for RQ3 were:

> $H_0$3 There is no significant difference in the math scores of low performing economically-disadvantaged students who participated in the 90-minute math block-scheduling program and those who did not.
>
> $H_1$3 There is a significant difference in the math scores of low performing economically-disadvantaged students who participated in the 90-minute math block-scheduling program and those who did not.

Based on the above findings, the null hypothesis failed to be rejected. There is no significant difference in the math scores of low performing economically-disadvantaged students who participated in the 90-minute math block program and those who did not. A post-hoc power analysis was performed using SPSS. The .45 effect size was considered to be moderate according to Cohen's (1992) criteria. With an alpha = .05 and *n*=36 the power was found to be .61. Meaning there was a 61 % probability of getting statistically significant results.

Non-Normal Distribution of Data

Use of a *t*-test for data analysis requires the data to be normally distributed. Normality was checked using a Shapiro Wilks test and found to be significantly non-normal. At this point it was essential to try to identify the reason(s) the data were not normally distributed and determine if there was another test that did not require normal distribution. According to Ghasemi and Zahediasl (2012), there are six common reasons data may not be normally distributed: (a) outliers or extreme values in the data, (b) by nature some data are not normally distributed, (c) too many values close to zero, (d) data collected from more than one shift or process can cause overlap (e) the way the data were sorted, and (f) insufficient data discrimination caused by rounding errors.

Examination of the data revealed two possible explanations for the non-normal distribution of the data used for this study. First, several scores could be considered outliers, as seen in Figures 1 and 2. While theses scores fell within the category of unsatisfactory math scores, they were below the majority of the other scores. Additionally, since the data were sorted to only include students who performed unsatisfactory this could affect normal distribution (Aguinis, Gottfredson, & Joo, 2013). It was determined the outliers were not erroneous and needed to be included in the data set. Furthermore, the data sorting was necessary to address the research questions. Therefore, a Mann-Whitney *U* test was selected to analyze the data. Table 2 shows several options available to analyze non-normal distributed data. The data collected in this study involved one independent variable (extended time in math class) and one dependent variable (student achievement). These factors supported the selection of a Mann-Whitney *U* test as the most appropriate statistical test to address the RQ's (Creswell, 2012; Fay & Proschan, 2010).

Table 2 Statistical Analysis Tools for Normally and Non-Normally Distributed Data

Tools for Normally Distributed Data	Equivalent Tools for Non-Normally Distributed Data	Distribution Required
***t*-test**	Mann-Whitney test; Mood's median test; Kruskal-Wallis test	Any
ANOVA	Mood's median test; Kruskal-Wallis test	Any
Paired *t*-test	One-sample sign test	Any
***F*-test; Bartlett's test**	Levene's test	Any
Individuals Comparison Chart	Run Chart	Any
C_p/C_{pk} analysis	C_p/C_{pk} analysis	Weibull; log-normal; largest extreme value; Poisson; exponential; binomial

Note: Reprinted from *Dealing with non-normal data: Strategies and tools,* by Arne Buthmann. Retrieved from https://www.isixsigma.com/tools-templates/normality/dealing-non-normal-data-strategies-and-tools/#comments.

Summary

The purpose of this study was to determine if there was a difference in the math scores of students who participated in a math block-scheduling program and those who did not. The RQ's specifically addressed low performing students, as well as low performing students considered either ‘at-risk’ or economically-disadvantaged, as this was an area of weakness for KTISD. This was done by comparing the means mathematics scores of 60 low performing students at a comparison school with 49 low performing students at a school that implemented the 90-minute block-scheduling program. Mann-Whitney *U* tests were utilized for data analysis, as the initially proposed *t*-tests could not be used because it was found the data were not normally distributed. The differences for all three RQ’s were found to be non-significant; therefore, it was concluded that there was no significant difference between the achievement of the low performing students who participated in the 90-minute blocking schedule program and those that did not. These findings will be discussed in the next chapter.

Conclusion

Section 2 explained the methodology of this study. The justification for the selection of a quantitative research approach and an outcome-based program evaluation design were presented. The research setting, the sample, and the instrumentation for this study were outlined. This section also described the data collection and analysis processes utilized for the study. The limitations of the study were discussed along with the precautions utilized to protect the rights of participants. In Section 3, I will elaborate on the details and the findings of the program evaluation.

Section 3: The Project

Introduction

The project I developed in association with this study was an evaluation report generated from the findings of the research project. In the project, I identified the problem of low performing, economically-disadvantaged and at-risk eighth grade students in the state of Texas's KTISD school district. These students received unsatisfactory scores in the math section of the STAAR. These unsatisfactory scores jeopardized the students' eligibility for promotion into the ninth grade and their schools' eligibility to receive federal funding. In this section, I will describe the project's goals and rationale and will provide a literature review that justifies the use of an outcome-based evaluation. In this section, I will also discuss and assess the roles and responsibilities of the program's stakeholders and the implications of the math block-scheduling program.

Description of Goals

The objective of this project study was to provide stakeholders with data to make an informed and educated decision regarding the math block schedule's future implementation or rejection in KTISD. The math block-scheduling program was never evaluated to determine what, if any, impact it had on the test population. The overarching goal of this evaluation was to provide key stakeholders with the answers to three RQs. The first question was aimed at comparing the math scores of low performing students who completed the 90-minute math block schedule program with those who did not. With the second RQ, I aimed to determine if there was a difference in the mathematics scores of low performing students who were also considered to be at risk who participated in the 90-minute math block-scheduling program and those who did not. The final question was aimed at comparing the math scores of low performing students who were also economically-disadvantaged. These questions were evaluated at the request of the program's key stakeholders.

Rationale

An evaluation report was appropriate for this project because my goal was to evaluate the math block-scheduling program to determine if the program met the pre-established program goals and report the findings to KTISD stakeholders. These findings would be used to make future budgetary decisions. The other project study genre options did not support the needs of the stakeholders as they related to evaluating the program's outcomes. Developing a curriculum plan, designing professional development, or making recommendations for policy changes would not have served the immediate needs of KTISD stakeholders. The decision makers only wanted to know if the program was successful in accomplishing the goals it was implemented to address (see Tam, 2014). This project served as the presentation of the background, data analysis, findings, recommendations, and supporting materials to assist stakeholders with the decision making process. I will further clarify my justification for selecting an evaluation report, specifically an outcome-based evaluation, in the following review of the literature subsection.

Review of the Literature

A program evaluation is a research method used to assess the results of a scheduled plan (Creswell, 2012). The block-scheduling program is the project being evaluated. The purpose of a program evaluation is to compare a program's performance to its original intent in order to determine whether it was effective according to initial performance standards (Dunsworth & Billings, 2012). The degree to which the evaluation finds the program successful may result in changes to the existing program. Often, pilot programs are evaluated using outcome-based evaluations; the program is initiated temporarily and then evaluated to determine its effectiveness prior to permanent implementation (Killion, 2013).

A program evaluation requires a program director and an evaluator. The director conducts the project and the evaluator observes and reports the findings or outcomes of the project's subjects over a specified period of time (Gargani & Miller, 2016; Linnell, 2014). The prepared report serves as an explanation of the nature of the program, the reason for its implementation, and its outcomes and future implications (Posavac, 2015). A program evaluation differs from plain research because a program evaluation is used to make decisions, while research is used to enhance topical knowledge (Spaulding, 2014). In other words, evaluation programs measure value, while research assesses the theory that undergirds a program (Cook & Lowe, 2012).

I gathered the research used to support this evaluation project using resources from the databases available through the Walden University Library. The research consisted of scholarly books, peer-reviewed articles, and journals. The key terms I used to identify these sources included: *outcome-based evaluation*, *program evaluations in schools, state standardized testing, STAAR, types of evaluation styles, importance of evaluation reports,* and *structure of evaluation reports*. Government organizations and affiliated websites also served as sources to further strengthen the literature review. Each search yielded thousands of results. I limited the results to sources in the discipline of education with a focus on those published within the last 3 to 5 years (2012–2017). Throughout the progression of this study, current literature was continually reviewed and incorporated to ensure the study was supported by timely and relevant sources.

Purpose of the Evaluation Report

To understand the purpose of an evaluation report, a person must first understand the meaning of a program and an evaluation in this context. According to Spaulding (2014), a program is “a set of specific activities designed for an intended purpose, with quantifiable goals and objectives” (p. 33). Prior to a program’s implementation, the program evaluator identified its intent and terms of purpose. The program evaluation can be quantitative, qualitative, or both (Creswell, 2012). The program evaluation report, which is referred to as the “project,” contains preprogram information that guides stakeholders through the purpose, execution, and outcome of the project (Connelly, 2015).

Evaluation is the assessment of information that is gathered through a set of preplanned strategies (Frye & Hemmer, 2012). Evaluation projects are context-specific, in that inferences are made about a certain subject within a certain setting (Chyung, 2015). The specificity determines a program's success in the context in which it is being evaluated; this success may not necessarily transmit to projects that are similar in scope, as the evaluation project is precisely tailored to fit each program's context. Evaluations offer limitless possibilities for stakeholders and evaluators alike, as they can be conducted in any industry. When evaluations are conducted in an academic setting, like this particular project, the evaluation is predominantly used to collect information about the performance of a program (see Frye & Hemmer, 2012). The evaluation ultimately serves to create a program that is linked to the greatest benefits by uncovering its flaws in order to eliminate them (Ahmady, Lakeh, Esmaeilpoor, & Yaghmaei, 2014).

According to Bogg (2012), an evaluation report should begin with the assessment of the project's purpose, needs, stakeholders, and scope. The role of the project report is not to provide a thorough list of the information that is associated with the project but to present factual material to be used for executive decision-making purposes (Bogg, 2012). The evaluation report essentially serves as a program's voice, demonstrating its level of success and or room for improvement (Connelly, 2015). In addition to this information, an evaluation report relates the project to prior research, such as theoretical frameworks, predecessors, and the reasons for selecting this form of evaluation (Leech, 2012). These sources, which are presented as a literature review, provide the project with additional context (Leech, 2012).

Ultimately, an evaluation determines whether a program's performance meets previously set standards (Killion, 2013). The evaluator conducts the evaluation and provides a detailed written account in the evaluation report, and then the evaluation report is presented to stakeholders, not only to understand the outcome, but also to provide context for the project (Smith, & Ory, 2014). The evaluation report informs the project stakeholders why the program was conducted, why a specific evaluation style was utilized, and the predicted implications, using data and material referenced in the literature review to serve as foundational material (Posavac, 2015). In this project, the evaluation report included: (a) a comparison of the math scores on the STAAR assessment of low performing, economically-disadvantaged, and at-risk eighth grade students attending KTISD before and after the math block-scheduling program; (b) data-driven recommendations about whether the math block-scheduling program was successful or unsuccessful; and (c) a written report to stakeholders to help justify their decision to support or reject the math block-scheduling program in KTISD. All three components of the report were needed to provide a comprehensive overview of the program (Lawton, Brandon, Cicchinelli, & Kekahio, 2014).

Types of Evaluations

The type of evaluation depends on the style of the project. There are several types of evaluations, but there are four that are commonly used: (a) formative evaluation, (b) process/implementation evaluation, (c) outcome-based evaluation, and (d) impact evaluation (Smith, & Ory, 2014;Yarbrough, Shulha, Hopson, & Caruthers, 2011). The study's subject, its participants, and what the evaluation seeks to ascertain determine the evaluation style (Yarbrough et al., 2011).

A formative evaluation is used to determine if a program fits properly within an institution's scope by evaluating during the process or program implementation (Smith, & Ory, 2014). For example, it may seek to verify whether a program fits within budget constraints or is conducted in a manner that adequately represents the institution prior to its permanent implementation. It may be argued that a formative evaluation could have been utilized for this project. However, a formative evaluation takes place while the program is ongoing and strives to assess whether a program fits within the scope of an institution, rather than determine the quantitative outcome of a program (Nsibande & Garraway, 2011). In this project, the success of the program was to be measured by the differences in the achievement scores of the students.

A process or implementation evaluation determines whether a program is carried out according to its set standards (Yarbrough et al., 2011). For example, if a kindergarten teacher is expected to read 10 children's books to students each week for one academic year as part of a new reading program, a process or implementation evaluation may determine whether or not these 10 books are being read. Alternately, it may determine how many teachers were able to follow this reading program in a set school district, as opposed to those who were not. The complexity of material assessed by this evaluation method may vary (Yarbrough et al., 2011). However, in this example, it would be used to determine whether the new reading program was successful in having teachers read 10 books to their students each week.

An impact evaluation report determines the success level of a program (Yarbrough et al., 2011). In other words, it assesses how well a program meets its preset standards. For example, if the same kindergarten class mentioned above encouraged its students to read five children's books outside of the classroom setting as part of the new reading program, this may assess how many students met this standard, the variation in how close students came to meeting this standard, and whether this reading program was successful overall in encouraging young readers to read more often. Perhaps most important, however, these questions may be revisited after the program and evaluation to determine whether the program played a role in long-term academic development and reading frequency of young readers. An impact evaluation shares a similar end-goal with an outcome-based evaluation; however, rather than solely assessing the results of a program, it is conducted after a program is finished to determine the degree of success and any lasting effects (Urban Reproductive Health Initiatives, [URHI], 2013).

My examination of the evaluation styles resulted in the selection of an outcome-based evaluation for this project. The study's goals and scope made an outcome-based evaluation the most appropriate. Outcome-based evaluations seek to assess a program's performance compared to previously set standards to uncover whether there was a change (Shek & Yu, 2012; URHI, 2013). For this project, the outcome-based evaluation helped me assess whether the block schedule program improved students' math scores to meet school district and state standards. In this situation, the only measure identified was the outcome measures, not the impact of the evaluation on individual participants or participants' perceptions of the program in relation to future academic experiences. This evaluation was entirely quantitative in nature.

Program Evaluation Design

Before an evaluation report can be produced, the proper evaluation design must be selected. According to the Harvard Family Research Project (HFRP), there are eight outcome models of program evaluation (Penna & Phillips, 2005). The HFRP described the outcome-based evaluation model as assessing the entire program from design, implementation and evaluation (Penna & Phillips, 2005). This model, in particular, evaluates a program to determine whether it meets target milestones or achievements. For these reasons, this model was selected for use with this project.

If possible, program evaluations should be conducted prior to a program's permanent implementation. The program design should provide a plan of action aligned with the project's goals that are easily comprehensible to individuals aside from the evaluator (Chacon, Sanduvete, Portell, & Anguera, 2013). Conducting the evaluation process during a program's earlier stages (formative) assists in identify problems and provides time to adjust the program (Puet, 2000).

This project used a casual-comparative research model, also known as ex post facto research. Casual comparative designs are often used to evaluate educational-based programs when the program has already been completed and the participants were not randomly assigned (Dunning, 2012; Ellis-O'Quinn, 2012). The casual comparative model involves comparing two groups in regards to the dependent variable and focuses solely on the program's outcomes (Hall, Freeman, & Roulston, 2014; Kravitz, n.d.)

The data for this program evaluation were collected from archival data in the form of math test scores on the eighth grade STAAR assessment. Since the study did not include all subjects in the population it was helpful to utilize measurement processes. Measurement processes involve the collection of specific data pertaining to the subject of evaluation (Frye & Hammer, 2012). In this study, it involved participants' demographics, as well as their academic performance. The data collection for a project can also be conducted using different techniques, including secondary data and active involvement (Chacón et al., 2013). Secondary data involve the collection of data that does not necessarily entail interaction with study participants (Chacón et al., 2013). This may require the evaluator to analyze material that is directly or indirectly supplied by study participants, but he or she does not interact directly with them. Active involvement, on the other hand, requires a level of interaction with the program participants through a means of one-on-one or group discussions (Chacón et al., 2013). There is also a middle ground between these two techniques, which allows strict result-focused interaction with participants in the form of data-recording assessments, such as tests or questionnaires (Chacón et al., 2013). However, this evaluation project required no interaction with its participants. The participants' academic performance was analyzed but there was no direct interaction with them.

The evaluation was a participant-based program. This means that the project focused on whether the behaviors of its participants, the students, yielded favorable results in terms of what the project aimed to study. The participant-based approach is a collective effort toward communal improvement (Hall et al., 2014). In addition to being participant-based, this project also included an intact-group design, which meant the students were not randomly assigned to membership in the groups (Frye & Hammer, 2012). The treatment group was selected by district leaders based on the previous year's math achievement scores. All eighth grade students on the treatment campus were included in the study. The comparison campus was specifically chosen based on performance status and demographic similarities, in order to ensure the most accurate outcomes possible. The students involved in the program were unaware of the study. There was no need for the participants to interact with the stakeholders or the evaluator, but an unspoken partnership existed and was necessary to evaluate the program in an effort to benefit all parties (Hall et al., 2014). The scaled test scores, measured by the STAAR assessment administered at the completion of the block-scheduling program, were collected without infringement, in order to produce reliable data for stakeholders.

A program's design must be based on strategic planning, which is the process of "matching the activities of an organization (or program) to its environment and resource capabilities" (Kim, 2011, p. 304). In other words, this assesses whether the program is feasible, financially or otherwise, in the environment in which it is being evaluated. For this project, the math block-scheduling program was evaluated prior to its permanent implementation because the costs of hiring additional math teachers to support the program are costly, and school officials did not wish to waste financial resources on an unsuccessful program. However, stakeholders knew that if this evaluation yielded successful results, the math block-scheduling program would be attainable via the district's annual budget or future grant awards.

Program Evaluation Standards

A program is successful if it presents desired results. However, a set of guidelines must dictate an evaluation program, especially in an academic setting. The Joint Committee on Standards for Educational Evaluation (JCSEE), established in 1975, is an organization of professionals accredited by the American National Standards Institute (ANSI), and it focuses on the quality of evaluation projects (JCSEE, 2016). JCSEE recognizes five categories as criteria for evaluating the quality of evaluation programs: (a) accuracy, (b) feasibility, (c) propriety, (d) utility, and (e) accountability standards (JCSEE, 2016). These categories each serve to assess different aspects of the evaluation. According to JCSEE (2016), the utility standards measure the significance of the evaluation program for stakeholders; the feasibility standards involve the reliability, sensibility, tactfulness and economical aspects of the evaluation, the accuracy standards ensure that the evaluation is based on reliable material; the evaluation accountability standards encourage documentation of the evaluation; and, finally, the propriety standards cover what is legally and ethically right. Prior to JCSEE, evaluation programs were often criticized for as being arbitrary, subjective, and biased (Hopkins, 2013).

JCSEE's standards are now requirements of evaluations in all educational settings. With these standards in place, the evaluation project must comply with the above criteria to be deemed reliable for stakeholders.

Advantages of Evaluation Reports

Evaluation reports provide a gateway to policy change. Oftentimes in institutions, such as educational settings, the stakeholders are not physically present to determine whether program changes are necessary. These individuals, while responsible for budgetary matters, may only be made aware of a program's need for improvement by employees or other individuals. The role of an evaluation report is to provide greater insight into programs to provide stakeholders with updated performance information. The quality of program improvement is arguably a direct result of the material relayed to its stakeholders (Dunsworth & Billings, 2012).

In addition to providing the background and outcome information of an evaluation project, the report also provides information about the limitations associated with the project (Mertens & Wilson, 2012). In other words, the evaluation report provides the full scope of a project to include its implementation, goals, expectations, performance, participants, and finally its constraints. Rather than solely providing the description of a program, it explores the practicality and feasibility of a particular program in a particular context. The evaluator, in this sense, acts as a communication medium, while the report serves as the catalyst for change (Mertens & Wilson, 2012).

In the context of this project, the evaluation report functioned similarly to student report cards (Shawer, 2013). The report provided stakeholders with an updated account of the performance of the math block schedule program, as it was the stakeholders' responsibility to make final decisions regarding the program's incorporation into the school system (Kim, 2011). The report also provided a written account for stakeholders to share with others, including state officials and policymakers (Shawer, 2013). The stakeholders reviewed this evaluation report prior to their decision regarding the implementation of a math block schedule program, so they knew the thorough scope of the project.

Importance of Stakeholders in Program Evaluations

Part of developing the program design is identifying the key stakeholders. Stakeholders and those around them can be impacted by the decisions made by district decision makers (Chyung, 2015; O'Sullivan, 2012). Stakeholders can be individuals or groups who have a vested interest in the situation. This study involved two types of stakeholders, upstream stakeholders and downstream stakeholders. This project included both. Upstream stakeholders are the individuals who make the executive decisions about the program in terms of its structure, resources, completion, or consent (Chyung, 2015). Downstream stakeholders are those directly affected by the evaluation, such as its participants (Chyung, 2015).

In this program, the upstream stakeholders were school officials, including the superintendent and other decision-making executives, who must approve the program.

The downstream stakeholders were the students, particularly those who participated in the treatment group of this evaluation. Others impacted by the program were; teachers, who would have to incorporate the math block-scheduling program into their daily curriculum; other school employees, including bus drivers, who may have to accommodate a longer school day by providing after-school transportation to students; and parents/guardians, as they are the executive decision-makers of the participating students' affairs until they become legal adults.

Evaluation Report Presentation

At the completion of the project, stakeholders were presented with an evaluation report (see Appendix A). This report served as the explanation of the project and its results. The report was needed, not only because it provided a permanent record that the evaluation was conducted, but also because it assisted stakeholders in making an executive decision about whether to incorporate block-scheduling into the school district's curriculum. The purpose of an evaluation project, after all, is to provide stakeholders with materials to justify, dissuade, or change their decisions regarding a program (Johnson, Hall, Greene, & Ahn, 2013). Although the presentation generally should be crafted in a manner that will provoke analysis and discussion, the nature of this project does not necessarily call for such a presentation (Johnson et al., 2013). The quantitative data gathered throughout the project served as the sole determinant of whether or not the program was effective, and presentation design cannot change the outcome of the project.

Although this project did not present the data in an alternative manner, the report must still be written for the specific target audience. In this case, the target audience included school district officials and potentially state federal officials involved with education reforms. According to the International Academy of Education, a nonprofit scientific organization that supports educational research, since the presentation may be distributed to individuals who are unfamiliar with terms associated with the education system and/or initially may be unaware of the program's context, it is important for the written report to provide detail about how the project was conducted (Anderson & Postlethwaite, 2007). The intent was not to confuse stakeholders or other individuals who may oversee the project, but to provide the information in the most clear and concise way as possible (Fitzgerald, 2014).

The Use of Evaluation Reports in Education

Evaluation reports are used in education to make changes to existing programs (Andawei, 2015). Arguably, the greatest objective in an academic setting is to offer programs that yield the highest academic success rate. In addition to meeting the demands of state and national assessment standards, schools are trusted by parents, guardians, and students alike with the responsibility of educating youth. These schools must offer innovating and intriguing programs that captivate students' attention while providing them with credible knowledge. In order to determine whether these programs are successful, an evaluation and its corresponding report provide information about the pros and cons of a program. An evaluation report can also indicate whether stakeholders should seek alternative means that are potentially more beneficial than an existing program and evaluate this new program.

In addition to providing stakeholders with an opportunity to assess new or existing programs, it also provides a justification for including programs in the annual budget. Each year, public schools receive an allotment to spend on school programs, extracurricular activities, supplies, etc. Stakeholders are responsible for allocating that budget. Programs, such as the one evaluated in this study, are not without cost, and before permanently implementing a new program into a school's curricula, stakeholders must assess whether it is both successful and feasible. This evaluation report served as a source of justification for the stakeholders' decision to fund or reject the program, not only between each other, but to outside officials, as well. The report also serves as a predecessor and point of reference for similar projects that may be conducted in the future.

Project Description

The purpose of this outcome-based evaluation was to collect information about the math block-scheduling program in order to appraise the success of said program. An evaluation report, referred to as the "project," is a written account of the evaluation that describes its objective, outcome, and implications. The main resource informing this evaluation report were scores from participating students of the appropriate demographic and academic level. Stakeholders had previously identified the treatment campus based on previous math STAAR scores. Additional resources included the faculty responsible for executing the math block schedule program, along with the required academic materials needed by participating students. This, too, was previously configured and approved by stakeholders. This project did not have any outstanding potential barriers, as the school district requested this evaluation to be conducted.

My role and responsibility as the evaluator was to collect and analyze data; compile the evaluation report; and be available to address stakeholder's inquiries. This report provided the stakeholders with information necessary to assist in deciding whether to approve or reject the permanent implementation of a math block schedule program in the school district. The evaluation report served as the sole point of reference, as a similar project had not been conducted in the school district. I presented the evaluation report to the stakeholders during a board meeting. The stakeholders were provided with the report prior to their annual budgetary meeting, allowing them time to review the material. The stakeholders were school officials with decision-making responsibilities for the school district. Although this project assessed students' scores, the students themselves were not involved in the project and held no role or responsibilities. My responsibilities included accurately analyzing the data and appropriately reporting the findings in the evaluation repot. When the evaluation report was initially presented, the district decision-makers bore the sole responsibility of reviewing the findings and making budgetary decisions regarding the math block-scheduling program. Consequently, the decision-making responsibility was modified and will be addressed in Section 4.

Project Implications

Nationwide, thousands of schools are struggling to meet the Federal standards along with individual state standards. For many schools, especially those in low-income areas, government funding is crucial, and government standards add pressure to improve. Although students differ in terms of demographics, socioeconomic status, intelligence level, emphasis on education, and drive for success, there are ways that schools can alter their traditional procedures to accommodate students of varying academic levels. By structuring classes to meet the needs of students more effectively, students may have greater chances to succeed, as well as remain interested in learning.

The outcome of this program proved ineffective in regards to raising eighth grade students' math scores. Following the completion of the math block schedule program, neither a positive nor a negative impact on math scores between the comparison group and the treatment group could be proven. Although the study did not prove the math program successful for eighth grade students in KTISD, this is not a telltale determinant of whether the block schedule program could be successful for other classes, other school districts in the state of Texas, or across the country. This program may be utilized on other populations, such as different grade levels or core content, and have a different result. Because the evaluation was content-specific, it cannot serve as the ultimate determinant of student behavior in relation to extended learning time.

Local Community

Walker, Clancy, and Cheney (2013) reported well-designed program evaluations help stakeholders meet the needs of at-risk students. While this study did not provide the immediate solution to address closing the achievement gap between eighth grade low performing, at-risk and economically- disadvantaged students and their peers in KTISD, it did open discussions about alternate and additional options to address the issue. Although the findings did not show that extended time in math class increased student performance on the STAAR assessment, the decision-makers in KTISD still favored the extended time in class for students classified as low performing, at-risk and economically-disadvantaged. District leaders believed the extended time offered students an opportunity to succeed or improve in a subject that is difficult.

Instead of rejecting the ALT theory, KTISD elected to adopt the math block-scheduling program along with recommendations suggested in the evaluation plan. KTISD will launch comprehensive professional development and mentoring programs for the teachers, support staff, and administrators to facilitate program implementation. The professional development and mentoring programs will focus on building relationships with students and families while learning how to make lessons relevant, challenging, and realistic. ALT in itself may not be enough to address the achievement gap in KTISD; however, it may serve as a supplemental services designed to promote social justice and reduce educational inequality. KTISD leaders understand this requires professional development targeted at training educators to recognize oppression and adapt teaching strategies (Voss & Rickards, 2016).

Beyond the Local Community

When considering the key stakeholders of this project, it is obvious that the students, teachers, and school officials of the KTISD district were impacted. However, this study has the potential to influence neighboring school districts, states, and even nationwide schools who struggle with similar performance issues and achievement status. When viewed from a wider perspective, this study has the potential to impact all students, whether at-risk, economically-disadvantaged, or not, who struggle in particular subject areas that are regarded heavily by state standardized testing policies. As districts search for more effective ways to reach their students academically and socially, this study provides a viable starting place for other districts' searching for solutions.

Conclusion

Section 3 outlined the project, described the project goals, and the scholarly rationale for selecting a program evaluation. A review of the literature further supported the use of an evaluation report as the appropriate means to address the problem and guide the project development. The final portion of the section is devoted to the implications of the project as they relate to the local community and beyond. Section 4 discusses my personal reflections and conclusions concerning this study.

Section 4: Reflections and Conclusions

Introduction

This project study served as the program evaluation of a 90-minute math block-scheduling program implemented at KTISD. In this section, I will address the project study's strengths; weaknesses; possible alternative approaches to the problem; and personal reflections about myself as a scholar, practitioner, and program developer. The importance of the project study, its implications, applications, and possible directions for further research will also be included.

Project Strengths

I identified several strengths associated with this project study. First, there was no need to develop a data collection tool. The STAAR assessment administered by the state of Texas served as the data collection tool. These assessments undergo a thorough process to check for validity and reliability (TEA, 2014). Another strength of the study was that the data collection process was relatively quick and inexpensive since the data I analyzed were archival and easily accessed. The use of statistical software (SPSS) to perform the data analysis provided concise and consistent numerical data to address the RQs was also a strength. Finally, the possibility of personal bias associated with me as the researcher was minimized by the use of numerical data (see Fitzpatrick, Sanders, & Worthen, 2011). The use of archival data collected for state accountability purposes further strengthened the project's validity and reliability.

Project Limitations

A significant limitation to this study was the fact it was restricted to students from only two middle schools located within one Central Texas school district. The small samples, (n = 49 and n = 60 for RQ1 and RQ2; n = 36 and n = 44 for RQ3) along with the very specific scope of the study meant the findings could not be generalized to a larger population. Further, I did not randomly assign participants to the treatment and comparison groups. They were grouped based on the independent variable (time in math class). Although there was an attempt to control for extraneous variables by selecting a comparison campus with a similar demographic profile to the treatment campus, there was no way to determine if factors outside the scope of this study had an influence on the findings. For example, I did not assess the curriculum to determine if it was viable. The experience levels of participating teachers and how they spent the extra time in class were also not examined. Additionally, home factors and campus culture were not factored into this study.

Recommendations for Alternative Approaches

This study could have been strengthened by the use of a mixed-methods approach to evaluate the program. A mixed-methods approach would have supplied the necessary statistical data along with qualitative data concerning teacher and parent opinions of the program. The qualitative portion of the study could have provided insight regarding some of the limitations previously discussed. For example, surveys and interviews could have been conducted to determine how the teachers felt about the extended class time and how the time was used.

Measuring the quality of the instruction should not be overlooked in a study of this type. Incorporating formative assessments designed to measure the quality of instruction throughout the program's implementation could have provided valuable feedback. Because quality of instruction can be difficult to measure, future studies should consider including multiple tools of measurement. High-stakes standardized assessments should not be the sole instrument used to gauge quality of instruction. Teacher self-reports, administrative walk-throughs, peer-evaluations, and parent surveys could all provide useful data to measure quality of instruction.

Adhering to a true experimental design as opposed to the quasi-experimental design I selected is also an alternate approach for this type of study. The random assignment of students along with the administration of a pretests and posttests could limit threats to internal validity (see Creswell, 2012). Random assignment helps control the extraneous variables that can occur naturally. Because the students were already assigned to treatment and comparison groups, a true experiment could not be conducted for this study (see Dunning, 2012)

Scholarship, Project Development, Leadership, and Change

Kriner, Coffman, Adkisson, Putnam, and Monaghan (2015) described involvement in a doctoral program as a transformative process capable of altering the character of the learner. This process was intense and required self-discipline, time-management, and perseverance. During this transformative process, I discovered it was necessary to step back and conduct self-reflections that were both personal and honest. I needed to contemplate: (a) my goals, (b) why it was important to me, and (c) what I was going to do about it.

Reflection as a Scholar and Practitioner

Scholarship is a process that encompasses critical thinking, discovering, integrating, applying, and teaching (McLay, 2013). During this project study, I learned the significance of each of these functions and gained a better perspective regarding how they work together to develop a well-rounded scholar. Critical thinking and discovery were key skills that helped me navigate through the extensive literature review process. Searching for relevant sources, analyzing their content, investigating frameworks, looking for gaps in the literature, and constructing the foundation for the project pushed me to transform from a student to a scholar (see Simms, 2013). The more I searched, the more excited I became about finding a resolution to the problem being addressed in the study. I identified connections to other disciplines and began contextualizing the information gathered. This progression allowed me to gain insight into the importance of integrating concepts and ideas when searching for solutions. As my project developed, I understood the relevance of a scholar applying new knowledge and sharing it with other scholars. My doctoral journey has shaped me into a practitioner who is ready and prepared to investigate relevant issues in education, construct meaningful hypotheses, design appropriate research, analyze data, publish accurate results, and collaborate with other practitioners to initiate desired changes (see Flessner, 2012). However, this study has also helped me realize that initiating a change should not be the final step in the process. Program monitoring and following up are key steps in the evaluation process that cannot be overlooked.

Reflection as a Project Developer

This project study was the result of a personal interview conducted with a KTISD district leader as part of a Walden course assignment. I learned a previously piloted math block-scheduling program had not been evaluated to measure its effectiveness. District leaders were seeking statistical data to assist in making informed decisions regarding the program. The needs of the stakeholders led me to select a program evaluation for my project study. Developing a curriculum plan, designing appropriate professional development, or making policy recommendations would not have met the needs of the stakeholders. My passion for math and my desire to support educational equality as a means to social change made evaluating the math block-scheduling program both relevant and well-timed. These two factors helped me muster the stamina to push through and stay the course. If I had not been enthusiastic about pursing my topic, I would have succumbed to the challenges, frustrations, and obstacles that seemingly lurked around every corner (see Segol, 2014).

As I navigated through the various stages of the project study, I eventually learned to defer to the rubrics and checklists provided by Walden to assist in the development of the project. It took me a while to grasp the concept that I needed to rely on these resources, as I felt I could "figure" things out on my own. A blueprint was being provided and I needed to follow it. Careful examination of the resources helped me understand the requirements and expectations for developing the project study. All the pieces were finally starting to fall into place and make logical sense. Furthermore, I have come to appreciate the emphasis placed on locating and identifying scholarly resources to justify and serve as the foundation for the project development. I realize as a scholar and practitioner, my reputation and creditability are contingent with the sources I choose to reference and support.

Leadership and Change

Several years ago I completed a course focused on the role of an effective leader in the change process. At that time, I was not in a leadership position and was unable to draw from authentic experiences to solidify my understanding of the change process. I knew an effective leader: (a) has a firm understanding of the change process, (b) understands individual reactions to change, and (c) uses collaborative strategies to support the process (Adler & Sfard, 2016). This study reminded me that teamwork is a key factor in navigating change. The era of principals and superintendents serving as the solitary instructional leaders of organizations is long gone. Teamwork cultivates an organization's leadership capacity and promotes sustainable school improvement (Jarvis, Bell, & Kelly, 2016).

During this evaluation process I had an opportunity to witness how classroom teachers, principals, support staff, parents, and the district's decision makers were able to work together toward a common goal. Unfortunately, the findings of the evaluation I conducted did not provide the district leaders with the clear answers they desired. Nonetheless, the findings and recommendations provided the catalyst to encourage academic conversations between the stakeholders to explore viable, data-driven options for the district.

Undoubtedly, KTISD still faces challenges concerning low math achievement scores and the achievement gap. However, the team is passionate, focused, and committed to developing effective solutions to increase student achievement throughout the district. It was my honor to serve a small role in this process. I come away with an enriched understanding of the educational practitioner's role in supporting effective leadership and the change process.

Reflection on the Importance of the Work

The importance of this project study revolved around the evaluation of a math block-scheduling program implemented as an intervention to increase the scores of low performing students. Because the majority of the low performing students are also classified as at-risk and economically-disadvantaged, the findings of this evaluation serve as a means to encourage awareness of social inequalities and foster conversations about strategies to minimize these inequalities. Equality in education encompasses not only protecting the rights and opportunities of marginalized students but also the support of programs and practices to assist these students (Bartell, 2013). The math block-scheduling program evaluated for this project study serves as an example of one such program.

In this instance, the findings of the program evaluation did not show that extended class time had a significant impact on the math scores of the three target populations: (a) low performing, (b) at-risk, and (c) economically-disadvantaged students. However, pertinent and worthwhile information was still extracted from the findings and recommendations. The acceptance or rejection of the hypothesis does not negate the importance of the work.

As a practitioner, I have learned it is important for my focus to remain on conducting a quality evaluation from start to finish regardless of the findings. The information gained from a properly-executed evaluation has the potential to impact local funding decisions, determine staffing and professional development needs, and serve as a resource to help other scholars design and develop programs to achieve their goals (Wellington, 2015).

Implications, Applications, and Directions for Future Research

At the onset of this study, both the KTISD leadership, requesting the evaluation, and I were aware the findings could not be generalized to a larger population. In other words, there could be no implications for social change outside the district based on the findings of the study. The study was designed with a specific scope: (a) low performing, (b) at-risk, and (c) economically-disadvantaged eighth grade students at two middle schools in KTISD. However, this does not mean the study lacks significance.

The organizational implication of this study was KTISD leaders embracing the importance of working together to develop and implement intervention programs. The initial pilot math program was reviewed, approved, and mandated from the top down without any input from those involved in facilitating the new program. Failure to gain participant buy-in, provide effective training and support, along with the lack of feedback throughout the program implementation could be the main reasons students in the treatment group did not show a significant difference in math achievement (see Adler & Sfard, 2016). In contrast, individuals sharing a common interest and committed to working together toward a common goal are a powerful force capable of perpetuating educational equality and social justice. KTISD leaders have the ability to transcend the physical brick and mortar borders of their environment by learning to investigate, develop, and share alternative practices and programs designed to address some of the barriers faced by underserved populations.

Academic Learning Time (ALT) theory was the theoretical foundation for this project study. According to ALT, spending additional time learning should increase student achievement (Thomas & Grimes, 1995). The findings of this study did not support this hypothesis. For KTISD, solely providing extended time in class proved to be insufficient in accomplishing the goals of the program. However, before dismissing the theory of extended class time as an option to increase student achievement in KTISD, an examination of how that time was utilized should be conducted.

The scope of this study was very specific; however, several factors outside that scope provide potential areas of focus for future studies. For example, the experience levels of the teachers involved were not taken into account for this study. The experience level of a teacher could have a significant influence on what takes place in and outside the classroom. Additional research concerning the impact of professional development programs focused on training teachers how to utilize extended class time would also be beneficial. Research regarding the timing of implementation is another area that could be examined. In this study, the program was implemented to eighth grade students. The results may differ if the program was administered to a different grade level.

Conclusion

Performing a program evaluation on the math block-scheduling program implemented at KTISD was both an educational and personal experience I will never forget. As an educator, I am a dedicated advocate constantly seeking out resources to assist all students. Personally, I possess a strong desire to level the playing field for those students who may be academically disadvantaged due to factors beyond their comparison. While this study did not show there was a statistically significant difference between the scores of those who participated in the program and those that did not, it did facilitate open conversations about future ideas to address the achievement gaps in the KTISD.

District leaders decided not only to give the program another chance but to expand it to all 13 middle schools in the district. This time, they plan to implement a comprehensive professional development program designed to support the teachers and staff involved in the program. The professional development will focus on helping teachers: (a) develop a deeper understanding of the curriculum and the resource available, (b) cultivate student and parent relationships, (c) share best classroom practices, (d) undergo culturally relevant training, and (e) discover how to analyze and use data to drive instruction and remediation. Additionally, formative evaluations will be conducted throughout the program instead of waiting until the program has ended to perform an evaluation. As a future curriculum, instruction, and assessment professional, it would be an honor to join and work with KTISD's leadership team to increase student achievement throughout KTISD.

References

Adams, M., Bell, L. A., & Griffin, P. (2007). *Teaching for diversity and social justice*. New York, NY: Routledge.

Adler, J., & Sfard, A. (2016). *Research for educational change: Transforming researchers' insights into improvement in mathematics teaching and learning*. New York, NY: Routledge.

Aguinis, H., Gottfredson, R., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods 16*(2), 270–301. doi:10.1177/1094428112470848

Ahmady, S., Lakeh, M. A., Esmaeilpoor, S., Arab, M., & Yaghmaei, M. (2014). Educational program evaluation model, from the perspective of the new theories. *Research and Development in Medical Education, 3*(1), 5–8. doi:10.5681/rdme.2014.003

Andawei, M. (2015). Project evaluation techniques of small and medium enterprises: Case study of selected firms in Bayelsa state. *Journal of Chemical, Biological and Physical Sciences, 5*(2), 2182–2190. Retrieved from http://www.jcbsc.org/issueenv.php?volume=5&issue=2

Anderson, L. W., & Postlethwaite, T. N. (2007). Program evaluation: Large-scale and small-scale studies. *Education Policy Series: Education Policy Series,* 1–24. Retrieved from http://www.unesco.org/iiep/PDF/Edpol8.pdf

Arce, J., Luna, D., Borjian, A., & Conrad, M. (2005). No Child Left Behind: Who wins? Who loses? *Social Justice*, *32*(3), 56–71.

Bartell, T. G. (2013). Learning to teach mathematics for social justice: Negotiating social justice and mathematical goals. Journal for Research in Mathematics Education, *44*(1), 129–163.

Benken, B., Ramirez, J., Li, X., & Wetendorf, S. (2015). Developmental mathematics success: Impact of students' knowledge and attitudes. *Journal of Developmental Education, 38*(2), 14-22.

Bennett, P. R., Lutz, A. C., & Jayaram, L. (2012). Beyond the schoolyard the role of parenting logics, financial resources, and social institutions in the social class gap in structured activity participation. *Sociology of Education, 85*(2), 131-157.

Biddle, C., & Mette, I. (2016*). Successful strategies for extended learning opportunities: A literature review of academic benefits and beyond*. A report of the Maine Education Policy Research Institute. Orono, ME: University of Maine.

Black, D. W. (2016). Abandoning the federal role in education: The Every Student Succeeds Act. *California Law Review*, forthcoming. Retrieved from https://ssrn.com/abstract=2848415

Bogg, D. (2012). *Report writing* (1st ed.). Maidenhead, England: McGraw-Hill Education.

Bradley, R., & Corwyn, R. (2016). Home life and the development of competence in mathematics: Implications of research with the HOME inventory. *Early Childhood Mathematics Skill Development in the Home Environment*, 29–49. doi: 10.1007/978-3-319-43974-7_3

Brady, M., Clinton, D., Sweeney, J., Peterson, M., & Poynor, H. (1977). *Instructional dimensions study*. Washington, DC: Kirschner Associates, Inc.

Brill, S. (2011). *Class warfare: Inside the fight to fix American's schools*. New York, NY: Simon and Schuster.

Buthmann, A. (2016). *Dealing with non-normal data: Strategies and tools.* iSixSigma. Retrieved from https://www.isixsigma.com/tools-templates/normality/dealing-non-normal-data-strategies-and-tools/

Buzza, D. C., & Dol, M. (2015). Goal-setting support, motivation, and engagement in alternative math classes. *Exceptionality Education International*, *25*(1), 35–66.

Cadiero-Kaplan, K. (2004). *The literacy curriculum and bilingual education*. New York, NY: Peter Lang Publishing Inc.

Carroll, J. (1963). A model of school learning. *Teachers College Record*, *64*, 723–733.

Chacón Moscoso, S., Sanduvete Chaves, S., Portell Vidal, M., & Anguera Argilaga, M. T. (2013). Reporting a program evaluation: Needs, program plan, intervention, and decisions. *International Journal of Clinical Health & Psychology, 13*(1), 58–66.

Cheema, J., & Kitsantas, A. (2014). Influences of disciplinary classroom climate on high school student self-efficacy and mathematics achievement: A look at gender and racial–ethnic differences. I*nternational Journal of Science and Mathematics Education*, *12*(5), 1261–1279.

Chute, E. (2012, September 24). Changes coming for standardized tests will make AYP less feasible. *Pittsburgh Post-Gazette (PA).* Retrieved from http://www.post-gazette.com/news/education/2012/09/24/Changes-coming-for-standardized-tests-will-make-AYP-less-feasible/stories/201209240134

Chyung, S. Y. (2015). Foundational concepts for conducting program evaluations. *Performance Improvement Quarterly*, *27*(4), 77–96. doi:10.1002/piq.21181

Codding, R. S., Chan-Iannetta, L., Palmer, M., & Lukito, G. (2009). Examining a class wide application of cover-copy-compare with and without goal setting to enhance mathematics fluency. *School Psychology Quarterly*, *24*, 173–185.

Cohen, A. K., & Schuchter, J. W. (2013). Revitalizing communities together: The shared values, goals, and work of education, urban planning, and public health. *Journal of Urban Health*, *90*(2), 187–196. doi:10.1007/s11524-012-9733-3

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, *1*(3), 98 –101.

Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education*. New York, NY: Routledge.

Condron, D. (2011). Egalitarianism and educational excellence. *Educational Research, 40*(2), 47–55.

Connelly, L. M. (2015). Evaluation research. *Medsurg Nursing: Official Journal of the Academy of Medical-Surgical Nurses, 24*(1), 34–60.

Cook, P. F., & Lowe, N. K. (2012). Differentiating the scientific endeavors of research, program evaluation, and quality improvement studies. *Journal of Obstetric, Gynecologic, and Neonatal Nursing, 41*(1), 1.

Corder, G. W., & Foreman, D. I. (2014). *Nonparametric statistics for non-statisticians: A step-by-step approach*. Hoboken, NJ: Wiley.

Cortes, K., Goodman, J., & Nomi, T. (2015). Intensive math instruction and educational attainment: Long-run impacts of double-dose algebra. *Journal of Human Resources, 50*(1), 108–158.

Creswell, J. (2012). *Educational research: Planning conducting, and evaluating quantitative and qualitative research.* Boston, MA: Pearson Education.

Crosnoe, R., & Leventhal, T. (2013). School and neighborhood based interventions to improve the lives of disadvantaged children. *Societal Contexts of Child Development: Pathways of Influence and Implications for Practice and Policy, 44*(2), 156.

Darling-Hammond, L., Zielezinski, M., & Goldman, S. (2014). Using technology to support at-risk students' learning. Stanford Center for Opportunity Policy in Education. Retrieved from https://edpolicy. stanford. edu/publications/pubs/1241

Delpit, L. (2012). *Multiplication is for white people: Raising expectations for other people's children.* New York, NY: The New Press.

Divine, G., Norton, H. J., Hunt, R., & Dienemann, J. (2013). A review of analysis and sample size calculation considerations for Wilcoxon tests. *Anesthesia & Analgesia*, *117*(3), 699–710.

Doykos, B. (2015). A review of research evidence of impacts of summer and extended learning opportunities. A report of the Maine Education Policy Research Institute (MEPRI). Gorham, ME: University of Southern Maine. Retrieved from http://maine.gov/doe/mlti/about/research/

Maine Education Policy Research Institute.

Duncan, A., Easton, J. Q., & Kerachsky, S. (2009). *The nation's report card: Mathematics 2009. Assessment of Educational Progress at Grades 4 and 8.* Washington, DC: U.S. Department of Education.

Duncan, G. J., Murnane, R. J. (2014). Meeting the educational challenges of income inequality. *Phi Delta Kappan, 95*(7), 50–54. doi:10.1177/003172171409500712

Dunning, T. (2012). *Natural experiments in the social sciences: A design-based approach*. New York, NY: Cambridge University Press.

Dunsworth, M., Billings, D. L., & Ebrary, I. (2012). *Effective program evaluation* (2nd ed.). Bloomington, IN: Solution Tree Press.

Eddy, C. M., & Brooks, D. E. (2011). Ethnic matching, school placement, and mathematics achievement of African American students from kindergarten through fifth grade. *Urban Education, 46*(4), 1280–1299. doi:10.1177/004 2085911413149

Educational Testing Service. (2010). *Glossary of standardized testing terms.* Retrieved from http://www.ets.org/understanding_testing/glossary/

Ellis-O'Quinn, D. (2012). An ex post facto study of first-year student orientation as an indicator of student success at a community college. *Inquiry: The Journal of the Virginia Community Colleges, 17*(1), 51–57. Retrieved from http://commons.vccs.edu/inquiry/vol17/iss1/6

Every Student Succeeds Act of 2015, Pub. L. No. 114–95 § 114 Stat. 1177 (2015–2016)

Fan, H., Xu, J., Cai, Z., He, J., & Fan, X. (2017). Homework and students' achievement in math and science: A 30-year meta-analysis, 1986–2015. *Educational Research Review, 20*, 35–54. doi: 10.1016/j.edurev.2016.11.003

Farbman, D. (2010). Tracking an emerging movement: A report on expanded-time schools in *America. Education Digest*, *75*(6), 17–19.

Farbman, D. (2011a). *Harnessing the power of expanded time to improve schools: A study of three districts.* Boston, MA: National Center on Time & Learning. Retrieved from http://www.timeandlearning.org/?q=learning-time-america

Farbman, D. (2011b). *Trends to reform the American school calendar: A snapshot of federal, state, and local action.* Boston, MA: National Center on Time & Learning. Retrieved from http://www.timeandlearning.org/?q=learning-time-america

Fay, M., & Proschan, M. (2010). Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Statistics Surveys, 4,* 1–39. doi:10.1214/09-SS051

Feldman, S. B., & Tyson, K. (2014). Clarifying conceptual foundations for social justice in education. In *International Handbook of Educational Leadership and Social (In) Justice*, *29,* 1105–1124. doi:10.1007/978-94-007-6555-9_55

Fisher, C., & Berliner, D. (1985). *Perspectives on instructional time*. New York, NY: Longman.

Fitts, S., & Weisman, E. M. (2010). Exploring questions of social justice in Bilingual/Bicultural teacher education: Towards a parity of participation. *Urban Review, 42*(5), 373–393.

Fitzgerald, E. (2014). Show and tell: Disseminating evaluation results. Retrieved from Grant Professional Association website: http://grantprofessionals.org/site-map/articles/77-gpa/254-show-and-telldisseminating-evaluation-results

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed). Upper Saddle River, NJ: Pearson.

Fleming-May, R. A., Mays, R., & Radom, R. (2015). I never had to use the library in high school: A library instruction program for at-risk students. *Libraries and the Academy, 15*(3), 433–456.

Flessner, R. (2012). Addressing the research/practice divide in teacher education. *Action In Teacher Education, 34*, 159–171.

Flocco, D. C. (2012). Deeper learning, reduced stress. *Independent School, 71*(4), 62–68.

Foran, M. (2015). Creating opportunity for struggling students. *Education Digest, 81*(2), 4–11.

Francis, B., Mills, M., & Lupton, R. (2017). Towards social justice in education: contradictions and dilemmas. *Journal of Education Policy, 32*(4), 414–431. doi:10.1080/02680939.2016.1276218

Frye, A., & Hammer, A. (2012). Program evaluation models and related theories: *AMEE Guide No. 67. Medical Teacher, 34*(5), 288–299. doi:10.3109/0142159X.2012.668637

Gabrieli, C. (2010). More time, more learning. *Educational Leadership, 67*(7), 38–47.

Gargani, J., & Miller, R. L. (2016). What is program evaluation? *American Journal of Public Health, 106*(6), e13.

Gay, M., Mills, G. & Afrasian, P. (2008). *Educational research competencies for analysis and applications* (9th ed.). Upper Sadler River, NJ: Pearson.

Gerdes, K. E., Segal, E. A., Jackson, K. F., & Mullins, J. (2011). Teaching empathy: A model rooted in social cognitive neuroscience and social justice. *Journal of Social Work Education, 47*(1), 109–131. doi:10.5175/JSWE.2011.200900085

Gersten, R., Clarke, B., Jordan, N. C., Newman-Gonchar, R., Haymond, K., & Wilkins, C. (2012). Universal screening in mathematics for the primary grades: Beginnings of a research base. *Exceptional Children, 78*(4), 423–445.

Ghasemi, A., & Zahedials, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism, 10*(2), 486–486. doi:10.5812/ijem.3505

Gill, W. (2011). Middle school A/B block and traditional scheduling: An analysis of math and reading performance by race. *NASSP Bulletin, 95*(4), 281–301.

Gonida, E. N., & Cortina, K. S. (2014). Parental involvement in homework: Relations with parent and student achievement-related motivational beliefs and achievement. *British Journal of Educational Psychology*, *84*(3), 376–396.

Goodman, R., & Burton, D. (2012). What is the nature of the achievement gap? Why does it persist and are government goals sufficient to create social justice in the education system? *Education*, *40*(5), 500–514. doi:10.1080/03004279.2010.550586

Grave, B. S. (2011). The effect of student time allocation on academic achievement. *Educational Economics*, *19*(3), 291–301. doi:10.1080/09645292.2011.585794

Gregory, A., Sciba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher*, *39*(1), 59–68. doi:10.3102/0013189X09357621

Gut, J., Reimann, G., & Grob, A. (2013). A contextualized view on long-term predictors of academic performance. *Journal of Educational Psychology, 105*(2), 436.

Hachey, A. C. (2013). Early childhood mathematics education: The critical issue is change. *Early Education and Development*, *24*(4), 443–445.

Hackney, J. (2013). *The impact of high school schedule type on instructional effectiveness and student achievement in mathematics* (Doctoral dissertation, Indiana State University). Retrieved from ProQuest Dissertations & Theses. (UMI No. 3589509)

Hall, J., Freeman, M., & Roulston, K. (2014). Right timing in formative program evaluation. *Evaluation and Program Planning*, *45*, 151–156. doi:10.1016/j.evalprogplan.2014.04.007

Harvey, D. (2013). *A case study of the impact of an intervention-scheduling model on academic achievement in English and Mathematics by students in need of remediation* (Doctoral dissertation). Available from http://digitalcommons.georgiasouthern.edu/etd/800

Heitin, L. (2015). Four-day school week linked to gains in math; 'Does shortening the school week impact student performance? Evidence from the four-day school week. *Education Week, 35*(4). Retrieved from http://edweek.org/ew/articles/2015/09/16/four-day-school-week-linked-to-gains-in.html

Henry, G., Barrett, N., Marder, C. (2016). Double-dosing in math in North Carolina public schools (REL 2016–140). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. Retrieved from http://ies.ed.gov/ncee/edlabs.

Hopkins, P. T. (2013). *Teacher perspectives of the use of student performance data in teacher evaluations* (Doctoral dissertation, The College of William and Mary). Retrieved from ERIC/Proquest. (ERIC No. ED556569)

Howe, R. B., & Covell, K. (2013). *Education in the best interests of the child: A children's rights perspective on closing the achievement gap*. Toronto, Canada: University of Toronto Press.

Huang, H., & Sebastian, J. (2014). The role of schools in bridging within-school achievement gaps based on socioeconomic status: a cross-national comparative study. *Compare: A Journal of Comparative and International Education, 45*(4), 501–525.

Jackson, K., & Wilson, J. (2012). Supporting African American students’ learning of mathematics: A problem of practice. *Urban Education*, *47*(2), 354–398. doi:10.1177/0042085911429083

Jarvis, J., Bell, M., & Sharp, K. (2016). Leadership for differentiation: An appreciative inquiry of how educational leadership shapes pedagogical change. *Leading and Managing, 22*(1), 75–91.

James, A. O., & Folorunso, A. M. (2012). Effect of feedback and remediation on students' achievement in junior secondary school mathematics. *International Education Studies, 5*(5), 153–162. doi:10.5539/ies.v5n5p153

Jensen, E. (2013). How poverty affects classroom engagement. *Educational Leadership, 70*(8), 24–30.

Jez, S., & Wassmer, R. (2013). The impact of learning time on academic achievement. *Education and Urban Society*, *12*(2), 1–23. doi:10.1177/0013124513495275

Johnson, C., & Kritsonis, W. (2010). The achievement gap in mathematics: A significant problem for African American students. *Doctoral Forum, 7*(1), 1–12.

Johnson, J., Hall, J., Greene, J. C., & Ahn, J. (2013). Exploring alternative approaches for presenting evaluation results. *American Journal of Evaluation*, *34*(4), 486–503.

Joint Committee on Standards for Educational Evaluation. (2016). Program evaluation standards statement. Retrieved from http://www.jcsee.org/program-evaluation-standards-statements

Joyner, S., & Molina, C. (2012). Class time and student learning. *Texas Comprehensive Center (TXCC)*. Retrieved from http://txcc.sedl.org/resources/briefs/number6/

Kaufman, A., & Blewett, E. (2012). When good enough is no longer good enough: How the high stakes nature of the No Child Left Behind Act supplanted the Rowley definition of a free appropriate public education. *Journal of Law and Education*, *41*(1), 5–23.

Killion, J. (2013). Comprehensive professional learning system: A workbook for states and districts. Oxford, OH: Learning Forward.

Kim, C. (2012). The role of affective and motivational factors in designing personalized learning environments. *Educational Technology Research and Development*, *60*(4), 563–584. doi:10.1007/s11423-012-9253-6

Kim, Y. (2011). Program evaluation for strategic planning and resource management: An approach for school effectiveness evaluation. *KEDI Journal of Educational Policy*, *8*(2), 303–322.

Klanderman, D., Webster-Moore, M., Maxwell, M., & Robbert, S. (2013). Creating problems and their solutions: Service-learning through trinity mathematics triathlons, math nights, and math centers. *PRIMUS, 23*(6), 563–571. doi.org/10.1080/10511970.2013.764366

Knudsen, A. (2009). The importance of student centered democratic education & the effects on placement of English language learners. *Middle Grades Research Journal*, *4(*1), 7–97.

Kolbe, T., Partridge, M, & O'Reilly, F. (2011). *Time and learning in schools: A national profile.* Boston, MA: National Center on Time & Learning. Retrieved from http://www.timeandlearning.org/research/Time%20and%20Learning%20in%20Schools%20-%20A%20National%20Profile.pdf

Koyama, J. (2011). Principals, power, and policy: Enacting "supplemental educational services." *Anthropology & Education Quarterly, 42*(1), 20–36.

Kravitz, L. (n.d.). Understanding and enjoying research. Retrieved from https://www.unm.edu/~lkravitz/Article%20folder/understandres.html

Kriner, B., Coffman, K., Adkisson, A., Putman, P., & Monaghan, C. (2015). From students to scholars the transformative power of communities of practice. *Adult Learning, 26*(2), 73–80. doi: 10.1177/1045159515573021

Lawton, B., Brandon, P. R., Cicchinelli, L., & Kekahio, W. (2014). Logic models: A tool for designing and monitoring program evaluations. REL 2014-007. Washington, DC: US Department of Education, Institute of Education Sciences. National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Pacific. Retrieved from http://eric. ed. gov.

Leech, N. L. (2012). Writing mixed research reports. *American Behavioral Scientist, 56*(6), 866–881. doi:10.1177/0002764211433800

Linnell, D. (2014). Process evaluation vs. outcome evaluation. Retrieved from http://tsne.org/blog/process-evaluation-vs-outcome-evaluation

Lodico, M. G., Spaulding, D. T., & Voegtle, K. H. (2006). *Methods in educational research*. San Fransisco, CA: Jossey-Bass.

Luginbuhl, P., McWhirter, E., & McWhirter, B., (2016). Sociopolitical development, autonomous motivation, and education outcomes: Implications for low-income Latina/o adolescents. *Journal of Latina/o Psychology*, *4*(1), 43.

Mackey, A., Finn, A., Leonard, J., Jacoby-Senghor, D., West, M., Gabrieli, C., & Gabrieli, J. (2015). Neuroanatomical correlates of the income-achievement gap. *Psychological Science, 26*(6), 925–933. doi:10.1177/0956797615572233

Madyun, N. (2011). Connecting social disorganization theory to African-American outcomes to explain the achievement gap. *Educational Foundations, 25*(3), 21–35.

Martinez, P., & Holland, G. (2011). Impact of ninety-minute math block schedule on math achievement of middle school students. *American National Journal of Contemporary Research, 1*(2), 44–52.

Martinez, R., & Young, A. (2011). Response to intervention: How is it practiced and perceived? *International Journal of Special Education*, *26*(1), 44–52.

Marzano, R. J. (2013). Ask yourself: Are students engaged? *Educational Leadership, 70*(6), 81–82.

McDonald, B., Polnick, B., & Robles-Pina, R. (2013). Impact of instructional practices on students' mathematics achievement in urban middle schools. *Delta Kappa Gamma Bulletin, 79*(2), 52–65. Retrieved from http://www.deltakappagamma.org/NH/bulletin_vol79-2_2013winter.pdf#page=53

McInerney, D. M., Cheng, R., Mok, M., & Lam, A. (2012). Academic self-concept and learning strategies: Direction of effect on student academic achievement. *Journal of Advanced Academics*, *23*(3), 249–269.

McLay, A. (2013). Developing the parameters of scholarship in postgraduate coursework studies. *Journal of University Teaching & Learning Practice, 10*(2), 1–14. Retrieved from http://ro.uow.edu.au/jutlp/vol10/iss2/7

Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology*, *106*(1), 121.

Memon, G.R. Joubish, M.F., & Khurram, M.A. (2010). Impact of parental socio-economic status on student's educational achievements and secondary schools of district Malir. *International Journal of Academic Research and Reflection, 1*(3), 25–33.

Mertens, D. M., & Wilson, A. T. (2012). *Program evaluation theory and practice: A comprehensive guide*. New York, NY: The Guilford Press. Retrieved from http://www.ebrary.com.libproxy.clemson.edu

Midkiff, B., & Cohen-Vogel, L. (2015). Understanding local instructional responses to federal and state accountability mandates: A typology of extended learning time. *Peabody Journal of Education, 90*(1), 9–26.

Miller, D. (1976). *Social justice.* Oxford, England: Oxford University Press.

Morris, E. & Perry, B. (2016). The punishment gap: School suspension and racial disparities in achievement. *Social Problems, 63*(1), 68–86. doi:10.1093/socpro/spv026

Mullen, C. A. (2010). 21st-Century priorities for leadership education and prospective school leaders. *Scholar-Practitioner Quarterly, 4*(4), 331–333.

Naraian, S., & Brown, K. S. (2011). Readiness for change: Appropriation of external resources in two urban schools. *Education and Urban Society, 43*(1), 42–72. doi:10.1177/0013124510379824

Nicotera, N., & Walls, N. (2010). Challenging perceptions of academic research as bias free: Promoting a social justice framework in social work research methods courses. *Journal of Teaching in Social Work, 30*(3), 334–350.

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).

Nomi, T., & Raudenbush, S. W. (2016). Making a success of “Algebra for All”. The impact of extended instructional time and classroom peer skill in Chicago. *Educational Evaluation and Policy Analysis, 38*(2), 431–451. doi:10.3102/0162373716643756

Norton, M. (2010). *A study of the impact of block-scheduling on student academic achievement in public high schools* (Doctoral dissertation). Available from ProQuest Digital Dissertations and Theses. (UMI No. 3397428)

Nsibande, R., & Garraway, J. (2011). Professional development through formative evaluation. *International Journal for Academic Development, 16(*2), 97–107. doi:10.1080/1360144X.2011.568279

O'Malley, M., Voight, A., Renshaw, T. L., & Eklund, K. (2015). School climate, family structure, and academic achievement: a study of moderation effects. *American Psychological Association, 30*(1), 142–157. doi:10.1037/spq0000076

O'Sullivan, R. G. (2012). Collaborative evaluation within a framework of stakeholder-oriented evaluation approaches. *Evaluation and Program Planning, 35*(4), 518–522. doi.10.1016/j.evalprogplan.2011.12.005

Paige, R., & Witty, E. (2010). *The black-white achievement gap*. New York: NY: Amacon.

Patall, E., Cooper, H., & Allen, A. (2010). Extending the school day or year: A systematic review of research (1985-2009). *Review of Educational Research, 98*(5), 299–309. doi:10.3102/0034654310377086

Penna, R. M., & Phillips, W. J. (2005). *Eight outcome models*. Retrieved from http://www.hfrp.org/evaluation/the-evaluation-exchange/issue-archive/evaluation-methodology/eight-outcome-models

Perry, L & McConney, A. (2010). School socioeconomic composition and student outcomes in Australia: Implications for educational policy. *Australian Journal of Education, 54*(1), 72–85. doi:10.1177/000494411005400106

Peterson, P. E. (2016). The end of the Bush-Obama regulatory approach to school reform. *Education Next, 16*(3), 22–32.

Posavac, E. (2015). *Program evaluation: Methods and case studies*. New York, NY: Routledge.

Power, S. (2012). From redistribution to recognition to representation: Social injustice and the changing politics of education. *Globalization, Societies and Education, 10*(4), 473–492. doi:10.1080/14767724.2012.735154

Power, S., & Taylor, C. (2013). Social justice and education in the public and private spheres. *Oxford Review of Education, 39*(4), 464–479. doi:10.1080/03054985.2013.821854

Preston, C., Goldring, E., Guthrie, J. E., Ramsey, R., & Huff, J. (2016). Conceptualizing essential components of effective high schools. *Leadership and Policy in Schools,* 1–38. doi.org/10.1080/15700763.2016.1205198

Puet, R. (2000). *Program evaluation 101*. Retrieved from https://mainweb-v.musc.edu/vawprevention/research/programeval.shtml

Ratcliff, N., Costner, R., Carroll, K., Jones, C., Sheehan, H., & Hunt, G. (2016). Causes of and solutions to the achievement gap: Teachers' perceptions. *Teacher Educators' Journal, 9,* 97–111. Retrieved from https://eric.ed.gov/?id=EJ1095644

Reardon, S. F. (2013). The widening income achievement gap. *Educational Leadership, 70*(8), 10–16.

Redd, Z., Boccanfuso, C., Walker, K., Princiotta, D., Knewstub, D., & Moore, K. (2012). Expanding time for learning both inside and outside of the classroom: a review of evidence base. Retrieved from http://www.childtrends.org/Files/Child_Trends2012_08_16_RB_TimeForLearning.pdf

Riegle-Crumb, C., & Grodsky, E. (2010). Racial-ethnic differences at the intersection of math course-taking and achievement. *Sociology of Education, 83*(3), 248–270. doi:10.1177/0038040710375689

Reinke, W., Herman, K., & Stormont, M. (2013). Classroom-level positive behavior supports in schools implementing SW-PBIS: Identifying areas for enhancement. *Journal of Positive Behavior Interventions, 15*(1), 39–50. doi:10.1177/1098300712459079

Robinson, K. (2010). Black-white inequality in reading and math across K-12 schooling: A synthetic cohort perspective. *Review of African American Political Economy, 37*(3-4), 263–273. doi:10.1007/121140109074

Rojas, L., & Liou, D. D. (2017). Social justice teaching through the sympathetic touch of caring and high expectations for students of color. *Journal of Teacher Education, 68*(1), 28–40. doi:10.1177/0022487116676314

Rothman, R. (2016). Accountability for what matters. *State Education Standard, 16*(1), 10–13. Retrieved from https://eric.ed.gov/?id=EJ1087556

Sahabudin, N. A., & Ali, M. B. (2013). Personalized learning and learning style among upper secondary school students. *Procedia-social and Behavioral Sciences,103*, 710–716. doi./10.1016/j.sbspro.2013.10.391

Sams, A., & Bergmann, J. (2013). Flip your students' learning. *Educational Leadership*, *70*(6), 16–20. Retrieved from https://eric.ed.gov/?id=EJ1015329

Segol, G. (2014). Choosing a dissertation topic: Additional topics. *College Student Journal*, *48*(1), 108–113.

Shawer, S. F. (2013). Accreditation and standards-driven program evaluation: Implications for program quality assurance and stakeholder professional development. *Quality & Quantity, 47*(5), 2883–2913. doi:10.1007/s11135-012-9696-1

Shek, D., & Yu, L. (2012). Subjective outcome evaluation of the project PATHS (extension phase) based on the perspective of program implementer. *Scientific World Journal*, 589257-8. doi:10.1100/2012/589257

Sheridan, B., Smith, B., Pleggenkuhle-Miles, E. (2017) Short vs. long: Cognitive load, retention and changing class structures. *Education Economics*, 101–112.

Simms, K. (2012). Is the black-white achievement gap a public sector effect? An examination of student achievement in the third grade. *Journal of At-Risk Issues*, *17*(1), 23–29. Retrieved from EBSCOhost.

Simms, M. (2013). A teacher-educator uses action research to develop culturally conscious curriculum planners. *Democracy & Education, 21*(2). Retrieved from http://democracyeducationjournal.org/home/vol21/Iss2/3.

Slavin, R., & Lake, C. (2008). Effective programs in elementary mathematics: A best-evidence synthesis. *Review of Educational Research, 78*(3), 427–515. doi:10.3102/0034654308317473

Smith, C. R., Marchand-Martella, N. E., & Martella, R. C. (2011). Assessing the effects of the "Rocket Math" program with a primary elementary school student at risk for school failure: A case study. *Education and Treatment of Children, 34*(2), 247–258. Retrieved from http://eric.ed.gov

Smith, M. L., & Ory, M. G. (2014). Measuring success: Evaluation article types for the public health education and promotion section of frontiers in public health. *Frontiers in Public Health, 2*, 111. doi:10.3389/fpubh.2014.00111

Sokal, L., & Katz, J. (2015). Effects of the three-block model of universal design for learning on early and late middle school students' engagement. *Middle Grades Research Journal, 10*(2), 65–82.

Spaulding, D. T. (2014). *Program evaluation in practice: Core concepts and example for discussion and analysis* (2nd ed.). San Francisco, CA: John Wiley & Sons, Inc.

Stebleton, M. J., & Soria, K. M. (2012). Breaking down barriers: academic obstacles of first-generation students at research universities. *Learning Assistance Review (TLAR), 17*(2), 7–19.

STAAR Technical Assistance Center. (2010). *Post-test guide technical information for STAAR District and test site coordinators and research specialists*. Retrieved from http://www.startest.org/pdfs/STAAR.post-test_guide.2010.pdf

State Board of Education. (2012). Commissioner's draft of the Texas mathematics standards. Retrieved from http://www.tea.state.tx.us/index2.aspx?id=2147499971

Sullivan, G. M., & Feinn, R. (2012). Using effect size, or why the p value is not enough. *Journal of Graduate Medical Education, 4*(3), 279–282.

Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. San Francisco, CA: Jossey-Bass.

Sue, D. W, Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A. M., Nadal, K. L., & Esquilin, M. (2007). Racial micro-aggressions in everyday life: Implications for clinical practice. *American Psychologist, 62*(4), 271–286.

Takeuchi, H., Taki, Y., Nouchi, R., Sekiguchi, A., Kotozaki, Y., Miyauchi, C., & Kunitoki, K. (2014). Regional gray matter density is associated with achievement motivation: evidence from voxel-based morphometry. *Brain Structure and Function*, *219*(1), 71–83. doi:10.1007/s00429-012-0485-3

Tam, M. (2014). Outcomes-based approach to quality assessment and curriculum improvement in higher education. *Quality Assurance in Education 22*(2), 158–168. doi:10.1108/QAE-09-2011-0059

Tanner, C. K. (2009). Effects of school design on student outcomes. *Journal of Educational Administration, 47*(3), 381–399. Retrieved from http://eric.ed.gov

Taylor, E. S. (2014). Spending more of the school day in math class: Evidence from a regression discontinuity in middle school. *Journal of Public Economics*, *117*(1), 162–181. doi:10.1016/j.jpubeco.2014.06.002

Texas Education Agency (2011a). *Glossary for the Academic Excellence Indicator System.* Retrieved from http://ritter.tea.state.tx.us/perfreport/aeis/2011/glossary.html

Texas Education Agency. (2011b). *State of Texas assessments of academic readiness (STAAR) resources*. Retrieved from http://www.tea.state.tx.us/student.assessment/staar/.

Texas Education Agency. (2013*). State of Texas assessments of academic readiness summary report*. Retrieved from http://www.tea.state.tx.us/staar/rpt/sum/yr12/

Texas Education Agency. (2014). *State of Texas assessments of academic readiness summary report.* Retrieved from http://www.tea.state.tx.us/star/rpt/sum/yr13/

Texas Education Agency. (2016). *Every Student Succeeds Act (ESSA).* Retrieved from http://tea.texas.gov/About_TEA/Laws_and_Rules/ESSA/Every_Student_Succeeds_Act_(ESSA)/

Thomas, A., & Grimes, J. (1995). *Best practices in school psychology III*. Washington, DC: National Association of School Psychologist.

Ullucci, K. (2011). Learning to see: The development of race and class-consciousness in white teachers, *Race Ethnicity and Education, 14*(4), 561–577. doi:10.1080/13613324.2010.519982

Urban Reproductive Health Initiatives. (2013). Measurement, Learning, & Evaluation Project. Outcome and impact evaluation. Retrieved from https://www.urbanreproductivehealth.org/toolkits/measuring-success/outcome-and impact-evaluation

Van Soest, D., & Garcia, B. (2003). *Diversity education for social justice: Mastering teaching skills.* Alexandria, VA: Council on Social Work Education.

Voss, R. & Rickards, A. (2016). Using social justice pedagogies to improve student numeracy in secondary school education. *Journal of Education and Practice, 7*(15) 40–47. Retrieved from https://eric.ed.gov/?id=EJ1103148

Wagner, K. M. (2013). *The effects of an extended-day online math program on math achievement* (Doctoral dissertation). Retrieved from proquest.com/openview/d5e05a31671406b9278325e09644b43c/1?pq-origsite=gscholar&cbl=18750&diss=y

Walker, B., Clancy, M., Tsai, S. F., & Cheney, D. (2013). Bridging the research-to practice gap: empowering staff to implement meaningful program evaluation and improvement to better serve students with emotional or behavioral disorders. *Beyond Behavior, 13*(2) 3–14. doi:10.1177/107429561302200302

Wang, A. (2010). Optimizing early mathematics experiences for children from low income families: A study on opportunity to learn mathematics. *Early Childhood Education Journal, 37*(4), 295–302. doi:10.1007/10643009103539

Wang, M., Brinkworth, M., & Eccles, J. (2013). Moderating effects of teacher–student relationship in adolescent trajectories of emotional and behavioral adjustment. *Developmental Psychology, 49*(4), 690–705. doi:10.1037/a0027916

Wellington, J. (2015). *Educational research: Contemporary issues and practical approaches.* London, England: Bloomsbury Academic.

Wright, M. (2010). A longitudinal study of block-scheduling in one South Carolina high school: *A descriptive twenty-five year case study from traditional to block* (Doctoral dissertation). Available from ProQuest Dissertations & Theses. (UMI No. 500017648)

Yarbrough, D., Shulha, L., Hopson, R. & Caruthers, F. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

Young, J., & Young, J. (2016). Closing the gap or reaching the ceiling: An exploratory trend analysis of the black white achievement gap in Texas. *Journal of Multicultural Affairs, 1*(1), 4. Retrieved from http://scholarworks.sfasu.edu/jma/vol1/iss1/

Zelkowski, J. (2010). Secondary mathematics: Four credits, block schedules, continuous enrollment? What maximizes college readiness? *Mathematics Educator*, *20*(1), 8–21.

Appendix A: Summative Evaluation Report

Executive Summary

An outcomes-based summative program evaluation was conducted to determine if the 90-minute math block-scheduling program implemented at one local middle school accomplished its preset goals. The goals of the math block-scheduling program were to increase the overall achievement of low performing eighth grade students on the math portion of the STAAR assessment and decrease the achievement gaps between at-risk and economically-disadvantaged students and their peers. KTISD leaders requested this evaluation because the program had not been previously assessed. The findings of this evaluation will assist district leaders in making future budgetary decisions regarding district wide interventions.

A quantitative study was performed to evaluate the scores of low performing eighth grade students who participated in the math block-scheduling program and those that did not. The study included 109 low performing students, 49 from the treatment campus, and 60 from the comparison campus. Mann-Whitney *U* tests were performed on three subgroups to address the targeted research questions concerning: (a) low performing, (b) low performing and at-risk, and (c) low performing and economically-disadvantaged students. The samples identified for RQ1 and RQ3 included the entire population, $n = 49$ and $n = 60$. The sample for RQ3 was $n = 36$ and $n = 44$. The eighth grade math STAAR assessment scores from both campuses were retrieved from Texas Educational Agency (TEA). Because the block-scheduling program had already occurred a retrospective casual-comparative approach was chosen. Mann-Whitney *U* tests were utilized to analyze the data to address three guiding research questions.

The results of all three tests showed there was no significant difference between the scores of students who participated in the block-scheduling program and those who did not. Therefore, the program did not accomplish its pre-established goals. The three measured outcomes of the program did not show a significant difference in the scores.

The findings of the evaluation were conveyed in the form of a summative evaluation report. The report included an executive summary, an evaluation report recapping the key components of the study, and a PowerPoint presentation. A formal presentation of the findings can be arranged if district stakeholders so desire.

The results of this study showed extended learning time alone was not sufficient in addressing the issues associated with the low performing students within KTISD. This in itself does not mean extended learning time does not have value. Because this was a retrospective program evaluation there was no way to control for extenuating factors. The results of this study indicated there could be multiple intervening and potentially confounding variables that were not addressed. It is my opinion that an examination of how the extended time was utilized should be explored and this information used to develop a professional development program to support a math block-scheduling program.

Recommendations

I recommend the math block-scheduling program be piloted again, after the teachers and staff have taken part in a comprehensive professional development program designed to increase the understanding of:

- The curriculum and available resources to support differentiation and student engagement;
- The importance of using data from formative assessments to drive instruction and remediation;
- The importance of building relationships with students and families; and
- Culturally Relevant Teaching (CRT) practices.

Once the teachers and campus administrators have been provided with the support and resources to effectively utilize the extended class time they will be better equipped to capitalize on the benefits associated with block-scheduling.

An unexpected outcome from this study was KTISD's leadership acknowledging the importance of including teachers and staff members when developing intervention programs. The math block-scheduling program was mandated by district leaders without any input from those responsible for the daily administration of the intervention. While the reports from other campuses boasted positive results concerning math block-scheduling programs, each campus is unique and what works for one entity may have no impact on other. KTISD leadership decided in the future, stakeholders from all levels would be included in the decision making processes associated with items directly related to classroom instruction. This was a win-win situation for all associated with KTISD.

Evaluation Report

Introduction

Low scores on the math portion of the eighth grade STAAR assessment prompted KTISD leaders to pilot a 90-minute block-scheduling math program during SY13. Not only were the district's math scores below the state average; there were also significant achievement gaps noted between at-risk and economically-disadvantaged students compared to their peers. To address the needs of the diverse population within KTISD, district leaders sought out an intervention that would provide more time on task, thus increasing the opportunities for practice and mastery of skills. They opted for extended learning time as the intervention based on the findings of Martinez and Holland (2011) who reported 90-minute math block-scheduling increased the student achievement in schools with demographics similar to those in KTSD. The math block-scheduling program was never evaluated to determine if it met the program goals. The goals of the program were to increase the math scores of low performing students and reduce the achievement gaps between at-risk and economically-disadvantaged students compared to their peers.

Background and Rationale

KTISD faced several problems regarding the performance of students on the math portion of the eighth grade STAAR assessment for SY12. First, students performed 4% below the state average. Second, the same data indicated that 31% of the students labeled as economically-disadvantaged performed below satisfactory level compared to 22% of those not classified as economically-disadvantaged (TEA, 2013). This represents a 9%

achievement gap. Third, a 20 % achievement gap was reported for the at-risk subpopulation in SY12. Only 64% of the at-risk students received satisfactory ratings in contrast to an 84% for students not considered to be at-risk (TEA, 2013). Lastly, failure to reduce the achievement gaps could mean KTISD would be identified for *targeted support and interventions* under ESSA.

The curriculum director at KTISD explained the rationale behind piloting the math block-scheduling program was based on three assumptions. The first assumption was extended class time would allow teachers more time to deliver lessons and have students practice the concepts learned under the supervision of the teacher. Second, decision-makers believed that more hands-on activities could be incorporated to help differentiate for multiple learning styles. Third, the extended class time eliminated the need for daily homework which the majority of the students did not complete

Purpose and Evaluation Questions

The purpose of this study was to evaluate the 90-minute block-scheduling program (independent variable) on the student achievement (dependent variable) of low performing students. The independent variable was measured on a nominal scale while an interval scale was used to measure the dependent variable. The block-scheduling math program was implemented to increase overall performance of low achieving students and decrease the achievement gaps existing between at-risk and economically-disadvantaged students and their counterparts. The following research questions guided the study:

Research Question 1 (RQ1)

Is there a significant difference between the math scores of low performing students who participated in the 90-minute math block-scheduling program and those who did not?

$H_0$1: There is no significant difference in the math scores of low performing students who participated in the 90-minute math block-scheduling program and those who did not.

$H_1$1: There is a significant difference in the math scores of low performing students who participated in the 90-minute math block-scheduling program and those who did not.

Research Question 2 (RQ2)

Is there a significant difference in the math scores of low performing at-risk students who participated in the 90-minute math block-scheduling program and those who did not?

H_0 2: There is no significant difference in the math scores of low performing at-risk students who participated in the 90-minute math block-scheduling program and those who did not.

$H_1$2: There is a significant difference in the math scores of low performing at-risk students who participated in the 90-minute math block-scheduling program and those who did not.

Research Question 3 (RQ3)

Is there a significant difference between the math scores of low performing economically-disadvantaged students who participated in the 90-minute math block-scheduling program and those who did not?

$H_0$3: There is no significant difference in the math scores of low performing economically-disadvantaged students who participated in the 90-minute math block-scheduling program and those who did not.

$H_1$3: There is a significant difference in the math scores of low performing economically-disadvantaged students who participated in the 90-minute math block-scheduling program and those who did not.

Methodology

To address the research questions concerning the effectiveness of 90-minute math block-scheduling program on the achievement of low performing at-risk and economically-disadvantaged students, a quantitative project study design was utilized in the form of an outcomes based evaluation. Quantitative research design methods allowed me to address the research questions in numerical terms (Creswell, 2012). The eighth grade math STAAR scores for SY13 were used to evaluate the effectiveness of the 90-minute math block-scheduling program. The study followed a retrospective casual comparative research approach. A retrospective casual-comparative approach was chosen because the research questions seek to examine the effectiveness of a 90-minute math block-scheduling program that had already occurred (Gay, Mills, & Airasian, 2008). An experimental research design was not selected because the participants were already grouped in relation to the independent variable (extended time in math class).

School district leaders in KTISD selected the treatment group based on the SY12 STAAR achievement scores. The treatment group consisted of the 49 eighth grade students who participated in the 90-minute math block-scheduling program for the entire SY13 and completed the math portion of the STAAR assessment while residing in KTISD. The comparison group consisted of 60 students who did not participate in the 90-minute math block-scheduling program for the entire SY13 and completed the math portion of the STAAR assessment while residing in KTISD.

With the call for increased accountability, program evaluations are an important tool used by educators to adequately evaluate programs and interventions to measure their effectiveness (Stufflebeam & Shinkfield, 2007). Program evaluations can serve as a means to measure if a program's objectives were met. In addition to measuring goal attainment, program evaluations can help educational leaders make suggestions concerning program modifications or improvements (Lodico, Spaulding, & Voegtle, 2006). District leaders at KTISD sought to determine the future of the 90-minute math block-scheduling program. Outcome-based evaluations are usually conducted following the conclusion of the program. This outcome-based evaluation was summative as outlined by Spaulding (2014). An outcome-based evaluation provides information to determine if the programs stated objectives were met. Those objectives included (a) increasing the overall achievement scores of low performing students and (b) reducing the achievement gap of low performing at-risk and economically-disadvantaged students and their peers.

Results

All data analysis was conducted in IBM SPSS 23. Mathematics scores were collected from 109 students, 60 of which were the comparison subjects and 49 who had been exposed to the 90-minute math block-scheduling program. A score of 1583 on the STAAR was used as a cut off for students who were considered low performing, as per the STAAR guidelines (TEA, 2014) and all students who scored above 1583 were removed from the analysis leaving only low performing students in the data set. This resulted in a sample of 60 students in the comparison group and 49 in the treatment group. Summary statistics are presented in Table 1 showing the number of subjects within each subpopulation and variable level as well as mean and standard deviation. Generally, students who had participated in the 90-minute math block-scheduling program had a lower mean score than students in the comparison group, with the exception of students who were not considered economically-disadvantaged.

Table 1

Summary Statistics of Mathematics Scores for Low Performing Students Who Have Not Been Exposed to a 90-minute Math Block Program (Comparison) and Those Who Have (Treatment).

		Comparison (Regular Schedule)			**Treatment (90-minute Schedule)**		
Variable	Level	Count	Mean	Standard Deviation	Count	Mean	Standard Deviation
Low Performing		60	1532.50	30.07	49	1522.51	44.48
At Risk	No	0	-	-	0	-	-
	Yes	60	1532.50	30.07	49	1522.51	44.48
Economically-disadvantaged	No	16	1520.50	34.83	13	1528.30	31.58
	Yes	44	1536.86	27.29	36	1520.41	48.52

The distribution of the mathematics scores for both the comparison and treatment group are presented in Figure 1. Most students scored in the range of 1500 to 1550. The comparison group had slightly higher mean mathematics scores than the treatment group. The treatment group appeared to have less variation in scores across the three sub-populations.

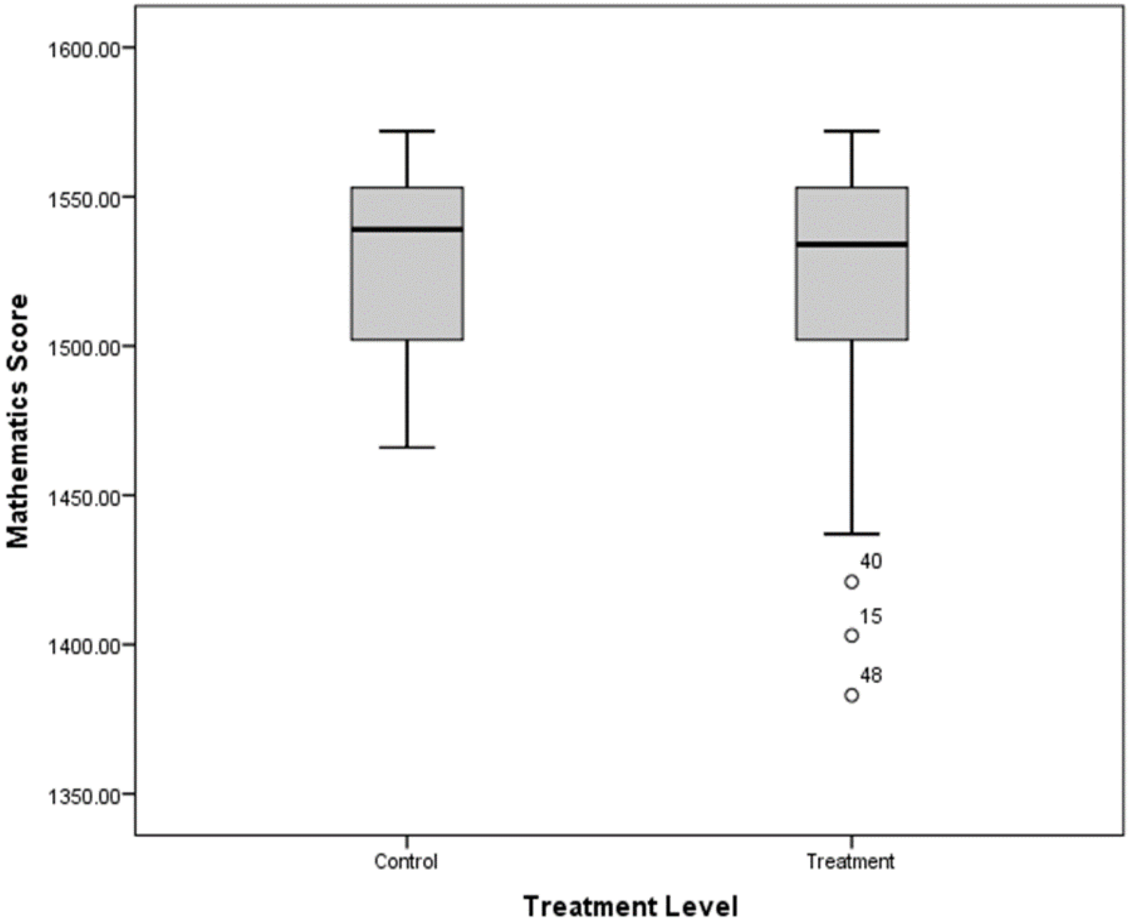


Figure 1. Distribution of mathematics scores for students in the comparison and treatment groups.

Figure 2 shows a boxplot of mathematics scores for economically-disadvantaged students. The mean mathematics score for economically-disadvantaged students in the comparison group was 1536.86 (as per Table 1) and the mean mathematics score for economically-disadvantaged students in the treatment group was 1520.41 (as per Table

1). Again, the scores were lower for the treatment group. Also, note the number of scores in the treatment group falling well below median of the majority of the scores.

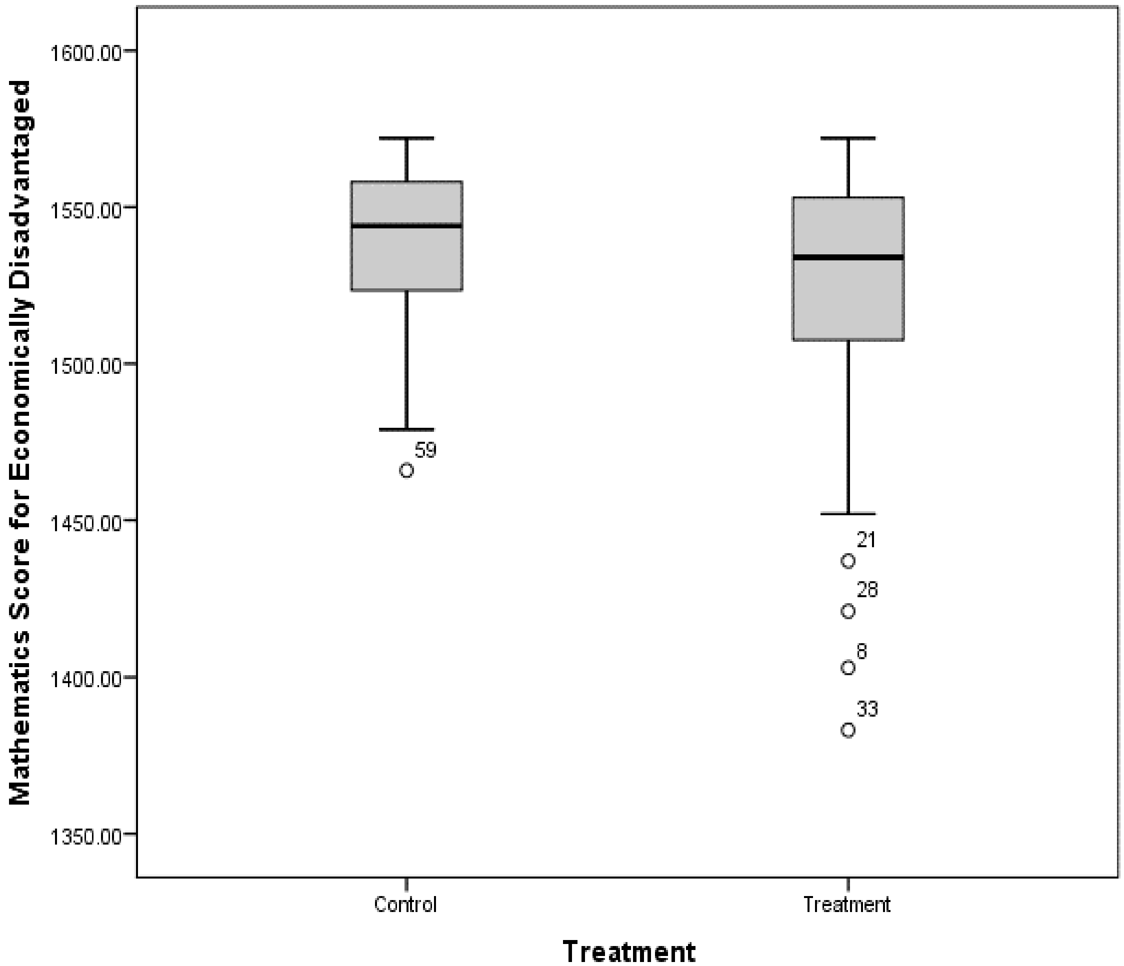


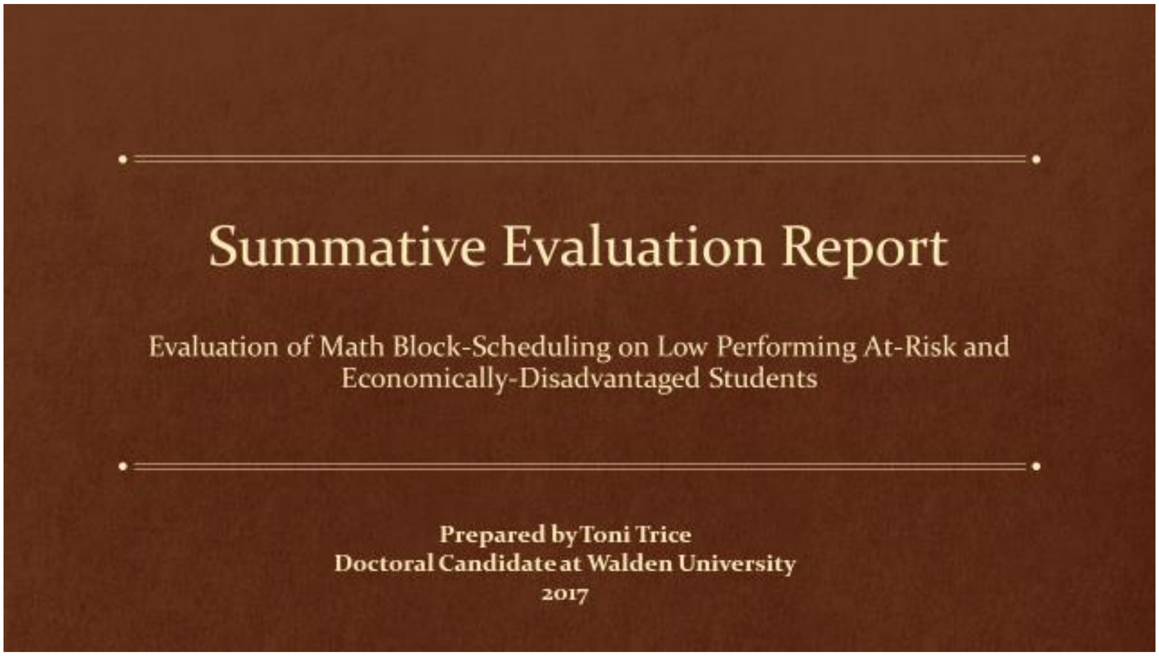
Figure 2. Box plot of mathematics scores for economically-disadvantaged students in the comparison and treatment groups.

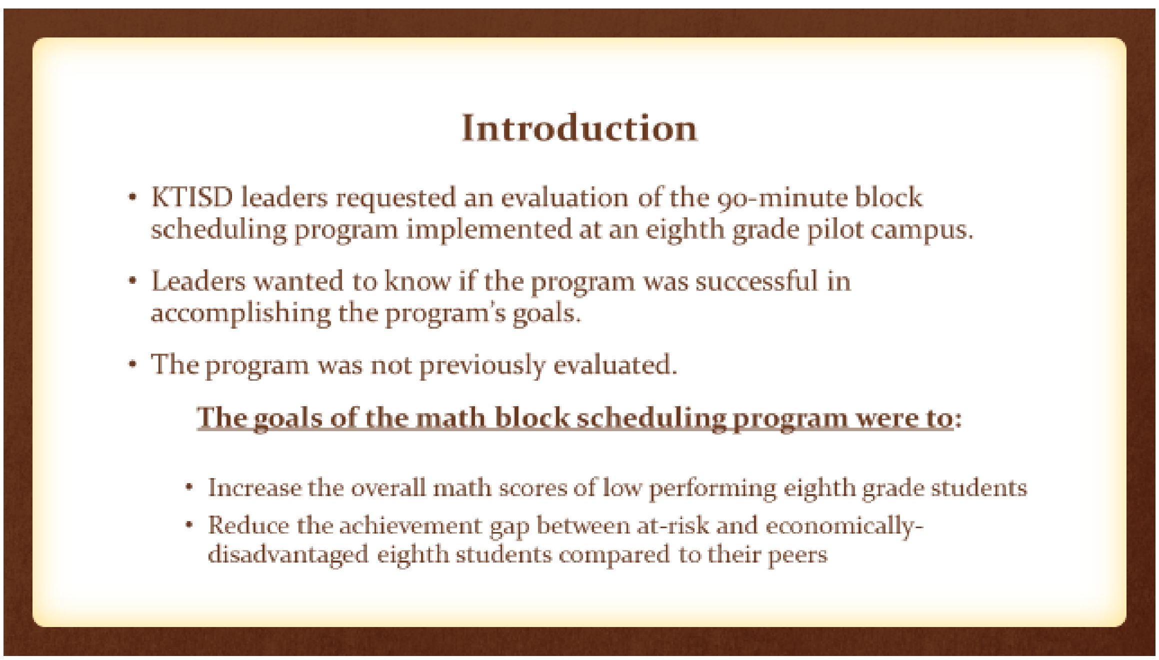
Conclusion and Recommendations

Although the results of this evaluation did not show that participation in the math block-scheduling program increased the achievement of low performing students, I believe there is merit behind extended class time in addressing the issues KTISD is attempting to address. However, the initial program was not well structured, implemented, or monitored. I recommend the program be implemented again. This time supported by a comprehensive professional development program. The professional development should focus on helping teachers and staff:

- Gain a better understanding of the curriculum and the resource available to support differentiation and increase student engagement;
- Develop and sustain student and parent relationships;
- Share best classroom practices;
- Embrace culturally relevant teaching strategies;
- Discover ways to make the learning applicable and transferable;
- Learn not to fear formative assessments and work collaboratively to increase student achievement; and
- Discover the importance of analyzing and using data to drive instruction and remediation.

PowerPoint Presentation





Definition of the Problem

- Eighth grade math scores in KTISD are 4% below the state average.
- There is a 9% achievement gap between economically-disadvantaged students and their peers.
- There is a 20% achievement gap between at-risk students and their peers.
- Continued scores below the state average and failure to close the achievement gaps between subpopulations would cause KTISD to be targeted for state "interventions and support."

Rationale

District leaders believed extended time in math class would:

1. Allow teachers more time to explain lessons and have students practice new concepts under the supervision of the teacher.
2. Provide more time for hands-on activities to assist with differentiation and making connections to the real world.
3. Eliminate the need for daily homework because the students would be able to get the necessary practice in class.

District leaders piloted the extended math program as an intervention to increase the math achievement of low-performing students and reduce the achievement gaps between subpopulations.

Purpose of the Study

- The purpose of this study was to evaluated a previously unevaluated, 90-minute math block scheduling program to determine if it accomplished its goals.
- The findings of the study would be used to help district leaders determine if they wanted to invest in expanding the program throughout the district's 13 middle schools.

Research Questions

- A quantitative design was utilized to address three research questions:

RQ1: Is there a significant difference between the math scores of low performing students who participated in the 90 minute math block scheduling program and those who did not?

RQ2: Is there a significant difference in the math scores of low performing at-risk students who participated in the 90 minute math block scheduling program and those who did not?

RQ3: Is there a significant difference between the math scores of low performing economically-disadvantaged students who participated in the 90 minute math block scheduling program and those who did not?

Data Collection and Analysis

- The individual scaled math scores for the spring 2014 administration of the STARR assessment were used.
- Texas Education Agency supplied a report of the eighth grade math scores for the treatment campus and the comparison campus.
- Pseudo-codes were assigned to the students to protect privacy.
- SPSS software was used to perform statistical analysis.
- Mann-Whitney *U* tests were used to test the hypothesis.

Overall Results

Table 1

Summary Statistics of Mathematics Scores for Low Performing Students who have not been exposed to a 90 Minute Math Block Program (Comparison) and those who have (Treatment).

		Comparison (Regular Schedule)			Treatment (90 Minute Schedule)		
Variable	Level	Count	Mean	Standard Deviation	Count	Mean	Standard Deviation
Low Performing		60	1532.50	30.07	49	1522.51	44.48
At Risk	No	0	-	-	0	-	-
	Yes	60	1532.50	30.07	49	1522.51	44.48
Economically Disadvantaged	No	16	1520.50	34.83	13	1528.30	31.58
	Yes	44	1536.86	27.29	36	1520.41	48.52

Results

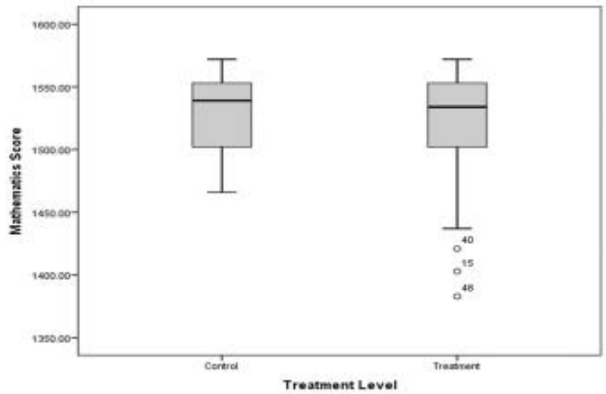


Figure 1. Distribution of mathematics scores for students in the comparison and treatment groups.

Recommendations

- Implement math block scheduling program again.
- This time include a comprehensive professional development program to support teachers and staff to:
  - Gain a better understanding of the curriculum and the resource available to support differentiation and increase student engagement;
  - Develop and sustain student and parent relationships;
  - Share best classroom practices;


Recommendations
- Embrace culturally relevant teaching strategies;
- Discover ways to make the learning applicable and transferable;
- Learn not to fear formative assessments and work collaboratively to increase student achievement; and
- Discover the importance of analyzing and using data to drive instruction and remediation.

References

Creswell, J. (2012). *Educational research: Planning conducting, and evaluating quantitative and qualitative research.* Boston, MA: Pearson Education.

Gay, M., Mills, G., & Airasian, P. (2008). *Educational research competencies for analysis and applications* (9th ed.). Upper Sadler River, NJ: Pearson.

Lodico, M. G., Spaulding, D. T., & Voegtle, K. H. (2006). *Methods in educational research*. San Fransisco, CA: Jossey-Bass.

Martinez, P., & Holland, G. (2011). Impact of ninety-minute math block schedule on math achievement of middle school students. *American National Journal of Contemporary Research, 1*(2), 44–52.

Spaulding, D. T. (2008). *Program evaluation in practice: Core concepts and examples for discussion and analysis*. San Francisco, CA: John Wiley & Sons, Inc.

Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. San Francisco, CA: Jossey-Bass.

Texas Education Agency. (2013*). State of Texas assessments of academic readiness summary report*. Retrieved from http://www.tea.state.tx.us/staar/rpt/sum/yr12/

Texas Education Agency. (2014). *State of Texas assessments of academic readiness summary report.* Retrieved from http://www.tea.state.tx.us/star/rpt/sum/yr13/

Appendix B**:** Data Request for Texas Education Agency (TEA)

Date:

Public Information Coordinator

Dear Public Information Coordinator, Student Assessment Division:

I am also a doctoral student at Walden University, where I am currently working on my final project study. The title of my study is The Evaluation of a Block-scheduling Program on Math Achievement Scores of Low Performing At-risk and Economically-disadvantaged Students. I am interested in determining if the extended math classes had an impact on the achievement scores of students on the STAAR assessment given in **April 2014**. To protect the privacy of the students I am requesting that names be removed and replaced with codes.

Request # 1

- The scaled math scores of eighth grade students classified as Level I at District XXXX campus XXX
- The scaled math scores of eighth grade students classified as Level I at District XXXXX campus XXX

Request #2

- The scaled math scores of eighth grade students classified as Level I and at-risk at District XXXXX, campus XXX
- The scaled math scores of eighth grade students classified as Level I and at-risk at District XXXXX, campus XXX

Request #3

- The scaled math scores of eighth grade students classified as Level I and economically-disadvantaged at District XXXXX, campus XXX
- The scaled math scores of eighth grade students classified as Level I and economically-disadvantaged at District XXXXX, campus XXX

Thank you in advance for your assistance. If you have any questions regarding this request, please contact:

Appendix C: Request to Conduct Study

The following information is provided in order to request permission to conduct research with the XXXXXXXXXX

1. Who is conducting the research? (the person, university, business, or agency)

This is an individual research project to be conducted by Toni Trice, a former XXXXX employee and current XXXXXXXXXX

2. The purpose of the research.

The purpose of the research is for the Partial Fulfillment of the Requirements for the Degree of Doctor of Education Curriculum, Instruction, and Assessment (CIA) at Walden University.

The researcher will conduct an evaluation of the math block-scheduling currently in place at middle school campuses in the district. The title of the study is “An Evaluation of a Block-scheduling Program on Math Achievement Scores of At-Risk and Economically-disadvantaged Students.”

3. How the information collected will be used.

The information collected will be use to evaluate the effectiveness of block math scheduling to determine if the extended class time has a positive influence on the math scores of at-risk and economically-disadvantaged students. The finding of the research will be presented to district leaders.

4. How the research will be conducted; that is, by electronic or manual survey, interviews, etc.

This is a quantitative study using archival data. The data collection instrument for this study will be the State of Texas Assessments of Academic Readiness (STAAR) eighth grade mathematics results as reported on the Academic Excellence Indicator System (AEIS) SY 2013. The researcher will gain access to the data from Texas Education Agency (TEA).

5. When the research is to be done (beginning and estimated end date).

Sincerely,

XXXXXXXXXXXXXXXXXXXXXXXXXXX

Appendix D: Permission to Conduct Study Within District

Greetings XXXXX

Your request to conduct research in XXXXX is approved provided you strictly adhere to the research guidelines sent in an earlier email message. **As a reminder**:

- Students may not be involved without parental permission. I understand that you plan to obtain parental permission during the 1st grading period and conduct the study during the 2nd.
- Student or employee names or other individual identification information may not be used.
- If you are planning to use an electronic survey (via email) to collect data, no survey may be sent to all XXXXX employees, that is, no widely distributed survey may be used.
- XXXXX internal email system or campus/department hard-copy mail distribution system may not be used by a district employee to "internally" survey staff or collect information. Researcher must use his/her own personal computer and home email to send surveys or requests for information from employees or use regular U.S. mail. Researchers may send information requests/surveys to district email addresses but employee responses must be done after work hours, if the employee chooses to participate.
- Researcher may not interview or survey XXXXX teachers and/or administrators during campus or department work hours.
- Participation of teachers or administrators must be entirely voluntary on their part. The district will not direct participation.
- Students may not be involved without parental permission.
- Student or employee names or other individual identification information may not be used.
- If you are planning to use an electronic survey (via email) to collect data, no survey may be sent to all XXXX employees, that is, no widely distributed survey may be used.
- XXXXX internal email system or campus/department hard-copy mail distribution system may not be used by a district employee to "internally" survey staff or collect information. Researcher must use his/her own personal computer and home email to send surveys or requests for information from employees or use regular U.S. mail. Researchers may send information requests/surveys to district email addresses but employee responses must be done after work hours, if the employee chooses to participate.
- XXXXX will not develop/prepare special reports for a researcher. Data / information requested must come from existing reports.
- Audio or video recordings may not be made of staff or students.
- XXXXX must be provided a copy of the research findings.

This is a standard message I send once approved however I realize that most of the above does not pertain to you since you will pull your info from TEA and your research wont impact students or staff ☺

Best of luck to you
Executive Officer
Office of the Superintendent