Summer 7-14-2023

THE EFFECT OF A PERSONALIZED LEARNING MODEL ON THE MATHEMATICAL ACHIEVEMENT OF ELEMENTARY STUDENTS

Jaclyn L. Pedersen
jpedersen@uttyler.edu

THE EFFECT OF A PERSONALIZED LEARNING MODEL ON THE MATHEMATICAL ACHIEVEMENT OF ELEMENTARY STUDENTS

by

JACLYN PEDERSEN

A dissertation submitted in partial fulfillment
of the requirements for the degree of
EdD in School Improvement
School of Education

Michael Odell, Ph.D., Committee Chair

College of Education and Psychology

The University of Texas at Tyler
August 2023

The University of Texas at Tyler
Tyler, Texas

This is to certify that the Doctoral Dissertation of

JACLYN PEDERSEN

has been approved for the dissertation requirement on
June 14, 2023
for the EdD in School Improvement

Approvals:

____
Dissertation Chair: Michael Odell, Ph.D.

____
Member: Yanira Oliveras-Ortiz, Ph.D.

____
Member: Dominick Fazarro, Ph.D.

____
Director, School of Education

Sarah Sass__________
Dean, College of Education and Psychology

ACKNOWLEDGEMENT

I would like to start by thanking my Committee Chair, who has also been my mentor for quite some time, Dr. Michael Odell, who serves as the Roosth Endowed Chair and a Professor of STEM Education for the School of Education at The University of Teas at Tyler. Words cannot express my deepest gratitude to Dr. Odell for the support, encouragement, and countless meetings over the years, continuously giving his expertise and feedback on improvement science as well as my dissertation. I would also like to thank my other committee members, Dr. Yanira Oliveras-Ortiz, Interim Director of the School of Education and Dr. Dominick Fazarro, Professor of Industrial Technology and Industrial Management, College of Business at The University of Texas at Tyler. I would also like to thank the entire faculty of the Doctorate of Education in School Improvement for your time, advice, support and expertise in this field.

Next, I would like to thank my colleagues at the University Academy, specifically the team of Instructional Coaches and Dr. Jo Ann Simmons, and each member of Cohort 1 of our program for your encouragement over the past three years. The moral support, help when needed, feedback sessions, and inspiration given by you all will not be forgotten.

Lastly, I would like to extend a heartfelt thank you and appreciation to my husband Chris for your daily support in my work and always being there to listen when I needed you to be. Without your help with our family during this time, this dissertation would not have been possible. Finally, a thank you to my dad who has instilled the value of education in me from a very early is due. You have always inspired me to go above and beyond in all that I do.

Abstract

THE EFFECT OF A PERSONALIZED LEARNING MODEL ON THE MATHEMAATICAL ACHEIVEMENT OF ELEMENTARY STUDENTS

Jaclyn Pedersen

Dissertation Chair: Michael Odell, Ph.D.

The University of Texas at Tyler

June 2023

This is a two-year mixed-methods study designed to discover the effects of a personalized learning model on the mathematical achievement in elementary. The setting is a university laboratory school that utilizes a distinct instructional math model based on mathematical best practices, problem-based learning (PrBL), and the use of personalized learning (PL) through stations, small-group pull-outs and a designated software program (IXL). The evaluation study was conducted over the course of a year. Teacher observations were used to understand the implementation of the model as well as the teacher's perceptions. Student data from the previous year's state assessment, State of Texas Assessment of Academic Readiness (STAAR), were used as baseline data for improvements as well as the students' mathematical functional level in IXL. The results of year one revealed that while some gains were seen academically, the model was not implemented to fidelity. Therefore, year two consisted of an intervention of a job embedded professional development plan for teachers in the PL model in order to determine the effect of mathematical achievement. The same measures as year one were used with the addition of a teacher survey given the middle and end of year. Overall, the study found that the job embedded professional development plan for teachers did have a positive impact on the achievement scores for both STAAR and functional levels in IXL. Subsequent years of data would need to be collected to be able to make recommendations for future studies as well as the impacts on the field at large.

Key words: Mathematics Achievement, School Improvement, Problem-based Learning, Improvement Science, Laboratory School, Job-embedded Professional Development, Blended Learning, Personalized Learning

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

The following is an improvement science dissertation in practice that includes five distinct chapters. Improvement science is a “methodology that disciplines inquiries to improve practice” (Bryk et al, p.10, 2017). The goal of improvement science is to ensure improvement efforts are based on evidence as well as best practices (Shojania & Grimshaw, 2005). Improvement science focuses on six distinct principles; the work is problem-specific and user-centered, focus on variation in performance, see the system that produces the current outcomes, measuring what is intended to improve, disciplined inquiry to drive improvement, and the use of networked communities (Bryk et al., 2017). The dissertation was designed using the overall philosophy of continuous improvement known as Plan-Do-Study-Act (PDSA) cycles. Continuous school improvement uses a similar model, Plan-Implement-Evaluate-Improve cycles, as an effort to improve schools on an ongoing basis (Bernhardt, 2018). The aim of this dissertation is to practice continuous school improvement methods while adhering to the principles of improvement science and engage in design-based implementation research (DBIR).

The study is based in one school setting, a public charter school in Texas, and includes two iterations of PDSA cycles. The first cycle is for the evaluation of an existing program and the second cycle is based on the introduction of an intervention to the program. The intervention is a result of using the improvement science principles listed above as well as the results of the first PDSA cycle. Each study was conducted over the course of a school year and were designed to be mixed method studies; using qualitative as well as quantitative data. The chapters in this dissertation are designed to give an overview of the setting as well as the problem of practice, a thorough review of the current literature, the evaluation of a program directly related to the problem of practice, the evaluation of the intervention to the program, and a final discussion and recommendations based on the findings from the two years of research.

Chapter 1: Problem of Practice

The achievement of students in mathematics in the United States (U.S.) has lagged behind other countries for years. According to some studies, the achievement of U.S. students in mathematics ranks around average (Desilver, 2017). In the most recent Program for International Student Assessment (PISA) in 2018, students in the United States performed below the Organization for Economic Cooperation and Development (OECD) member average in mathematics (OECD, 2019). Mathematics scores have remained stable since 2006 with no significant improvement in results. However, other studies which include results from the Trends in International Mathematics and Science Study (TIMSS) show U.S. students, particularly fourth-grade students, to be ranked eleventh out of 45 countries (Provasnik et al., 2016). When reviewing the TIMSS study, Provasnik et al. found that while U.S. fourth graders had shown average progress over the five administrations of TIMSS, they have shown little to no progress from the 2011 administration to the last administration in 2015. An examination of the results of the 2019 National Assessment of Education Progress (NAEP) results in mathematics shows that the average mathematics score for students in Texas was not statistically different from their average score in 2017. Only 30% of students in Texas performed at or above the NAEP proficient level (NAEP, 2019).

Setting

The setting of the study is an open-enrollment public charter school in Texas made up of two K-12 campuses and one 1-12 campus. The campuses are set in three distinct counties. The charter was written to be a lab school for a university and it is modeled after the Texas Science Technology Engineering and Mathematics (T-STEM) blueprint. The district implements project based, problem based and blended learning (BL) as the primary methods of instruction while also implementing Project Lead the Way (PLTW) Engineering and Biomedical pathways for students. Students take dual credit classes beginning their ninth-grade year and are able to graduate with 42+ hours of university credit.

The district has an enrollment of approximately 827 students and employed 57 teachers. The student population was 66.5% White, 17% Hispanic, 6.3% African American, 4.7% Asian, 49.3% female, and 50.7% male. The district is 38% economically disadvantaged across the three campuses, with 6.4% special education population and 10.3% Section 504 students. The teachers have various backgrounds ranging from novice probationary teachers to veteran teachers with thirty plus years of experience. Each campus has one Director (much like a principal) and one instructional coach whose main role is to help assist in the implementation of the instructional model as outlined in the charter.

The district is considered a high performing district in the state, earning an overall rating of an “A” on the 2019 accountability ratings, a rating only given to ten percent of districts. However, the district received a “B” rating in the domain related to student progress. The rating in student progress was largely related to math. To address the lack of academic growth in math, the Math Innovation Zones (MIZ) program was put into place for the 2019-2020 school year. The MIZ program is a BL grant awarded by the state with the purpose of improving mathematics achievement in students. The district chose to implement the MIZ program as a PL model. The math PL model was designed by stakeholders, which included the Director of Curriculum, three campus instructional coaches, and teacher representatives from each campus. The design team created the PL model to include four main components: data driven decisions, student reflection, targeted instruction, and integrated technology.

Problem of Practice

Student success for every student is the goal for classroom teachers from year to year. The state of Texas defines student success as meeting or mastering the grade level Texas Essential Knowledge and Skills (TEKS) as well as meeting set progress measures (specific to math and reading) from year to year on the STAAR. In some cases, these set progress measures can simply mean to maintain scores from the year before and in other cases it means students need to improve their score from the previous year. The measures are specific to students' scores from their prior math or reading test. This specific measure was put into the state's accountability system in 2017 to make sure students who were meeting state standards were not being neglected and therefore insured to be held accountable to their own individual growth from year to year.

In reviewing district data, a gap in student progress data was identified. In 2019, 22% of students in the district not only did not meet progress in math, reading, or both, but also declined from one standard to the next (ex: Masters in 2018 to Meets in 2019, or Masters in 2018 to Approaches in 2019, and so on) on the STAAR exam. Fourth-grade math students, in particular, had the highest rate of students not meeting progress. A problem of practice was identified: 73% of students in fourth grade did not meet progress on the math STAAR, causing students to fall further behind when entering the next grade level.

Purpose

The purpose of this study is to evaluate the effects of four components of a BL model (data driven decisions, student reflection, targeted instruction and integrated technology) integrated into one PL model on the mathematical achievement of fourth-grade students. For the purpose of this study, PL is defined as a station rotation model where students' individual needs are met through targeted instruction provided by the teacher in small group pull-outs as well as digital platforms. Students are given some choice in their pacing, content, and groupings as well as set individual academic goals, reflecting upon them often to address their own gaps and strengths.

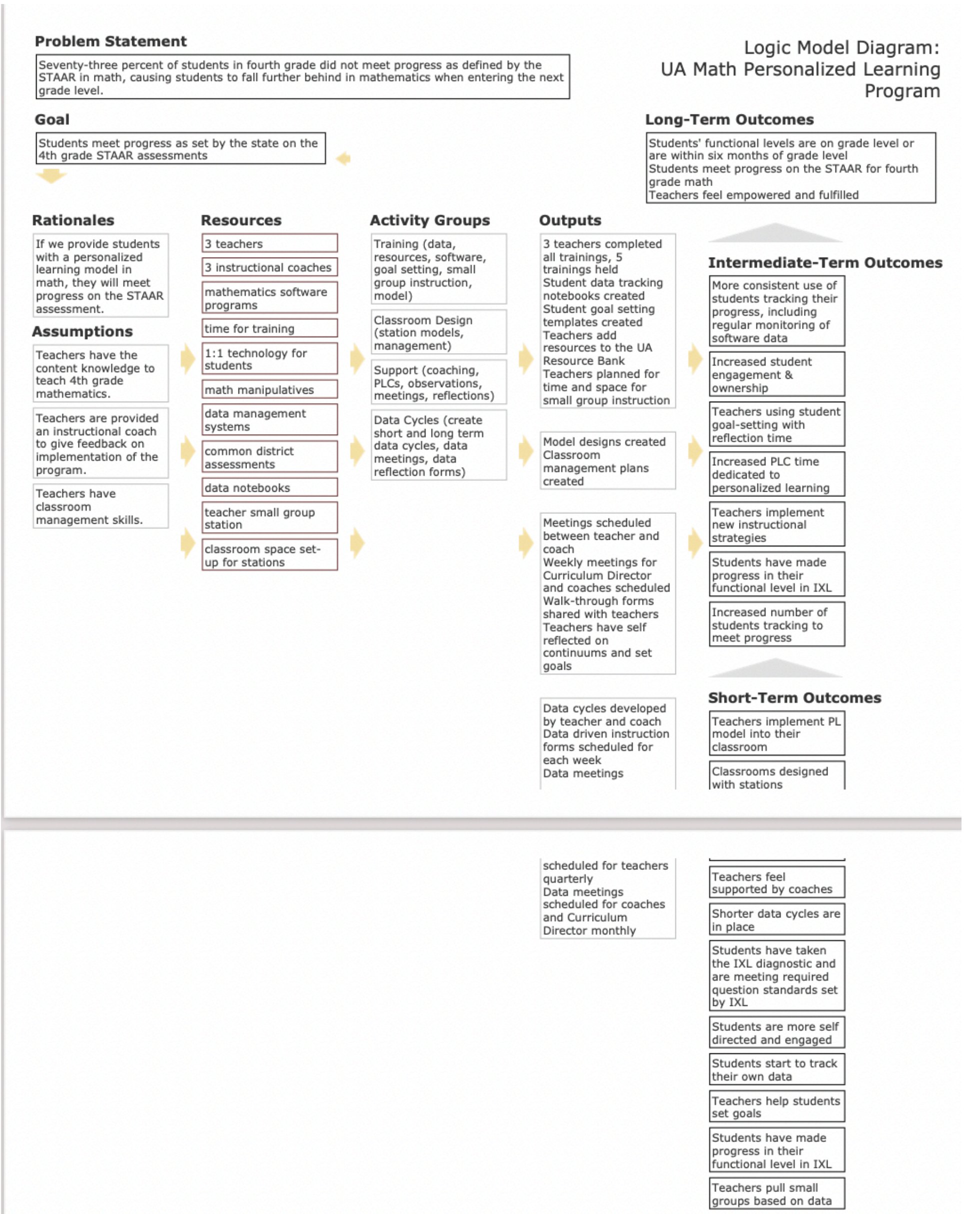
The study will seek to answer the following questions: (1) How have students' mathematical functional levels changed?, (2) How have students progressed as defined by the state assessment, STAAR?, and (3) How have teachers' views of their success in the classroom changed?

Theory of Change

The theory behind the PL model is based on three guiding principles: students met on their functional level will fill gaps more quickly, teachers making data driven decisions (in short data cycles) to inform their instruction will help each student progress, and students taking ownership in their own learning leads to academic growth. Figure 1 is a logic model that was created to serve as a visual representation of the major components and outcomes of the math PL model being evaluated.

Figure 1

Logic Model: Math Personalized Learning Program



Long-term goals

The ultimate goal of the program is for students to meet progress on the fourth-grade math STAAR. Additional goals of the program are (1) students' functional levels are on grade level or are within six months of grade level by the end of fourth grade and (2) teachers feel empowered and fulfilled under the new instructional model.

Intermediate goals

Intermediate goals of the program are aligned to the overall goals of the program. Students tracking their own data, goal-setting regularly based on their data, and their overall engagement in learning increasing will lead to the goal of students making academic growth from the prior year as well as contributes to the ultimate goal of students meeting progress on the STAAR. Teachers increasing their professional learning community (PLC) time to focus on PL as well as trying new strategies in their classroom with the support of the campus instructional coaches will lead to the overall goal of teachers feeling empowered and fulfilled. In addition, the success of their students showing progress in their functional levels and tracking to meet progress also contributes to the goal of teacher fulfillment. See Figure 2 below for all short, intermediate, and long-term objectives for the model.

Table 1

Short-Term, Intermediate, and Long-Term Objectives

	Logic Model Components	**Evaluation Questions**	**Indicators**	**Targets**
Early/Short-term Objectives	Teachers implement PL model into their classroom Students have taken the IXL diagnostic and are meeting required question	To what extent and how have teachers implemented PL into their classrooms? To what extent have students started IXL?	Increased teacher use of the PL model. IXL reports show that students have taken the diagnostic and are meeting an	By October 2020, all fourth grade math classrooms are implementing PL components as observed on walkthrough forms. By October 2020, all students will have taken the diagnostic and the class average of

	standards set by IXL.		average of 65 questions per week in IXL.	each fourth grade classroom for answered questions is 65/per week.
	Teachers help students set goals	How have teachers structured student goal setting in PL classrooms?	Increased number and percentage of students utilizing a data tracking notebook	By October 2020, all fourth grade math students will have set one or more academic goal based on standards.
	Teachers pull small groups based on data	To what extent and how are teachers pulling small groups during instruction?	Increased number of flexible groupings in PL classrooms	By October 2020, all fourth grade math teachers will have pulled a small group based on data. By October 2020, all fourth grade math teachers will have reflected on this with a DDI form.
	Classrooms designed with stations	How are stations being utilized in the classroom?	Increased number of classrooms using station models	By October 2020, all fourth grade math classrooms are utilizing a station model.
	Students are more self directed & engaged	To what extent are students self-directed?	Increased number and percentage of students can move through stations with little direction from the teacher.	By October 2020, 75%+ of students are engaged and can move through stations with little direction from the teacher.
	Teachers feel supported by coaches	To what extent to teachers feel supported by campus coaches?	Increased number of teachers feel supported	By October 2020, all fourth grade teachers will express support .
	Shorter data cycles are in place	How has the length of data cycles changed?	Increased number of weekly Data Driven	

	Students start to track their own data Students have made progress in their functional level in IXL	To what extent and how have students tracked their own data? To what extent have students' functional level increased in IXL?	Instruction (DDI) forms are being completed. Increased number and percentage of students utilizing a data tracking notebook Increased functional levels by student	By October 2020, all fourth grade math teachers will be completing a DDI form each week. By October 2020, all fourth grade math students will have a data tracking notebook and will have tracked progress on at least one standard or assessment. By October 2020, 90% of students will have made progress in their functional level in IXL.
Intermediate Objectives	More consistent use of students tracking their progress, including regular monitoring of software data Increased student engagement & ownership Teachers using student goal-	How are students tracking their progress by standard? To what extent and how did students interact in PL classrooms? To what extent and how do teachers utilize student goal	Increased number of student data tracking Increased number and percentage of student engagement Increased number of goal setting artifacts found in	By February 2021, all 4th grade classrooms are utilizing student data tracking notebooks. By February 2021, multiple trackers are present for each unit in student data tracking notebooks. By February 2021, 90%+ will be actively engaged in PL classrooms as noted on walkthrough forms.

	setting with reflection time Increased PLC time dedicated to PL Teachers implement new instructional strategies Students have made progress in their functional level in IXL Increased number of students tracking to meet progress	setting and reflection in the classroom? To what extent and how do teachers utilize PLCs? To what extent and how have teachers tried new instructional strategies? To what extent have students' functional level increased in IXL? To what extent are students tracking to meet progress on STAAR?	classroom. Increased number of student reflection times with teacher observed in classrooms through walkthrough forms. Increased number of PLCs dedicated to PL discussions and planning Increasing number of fourth grade teachers trying new instructional strategies in PL classrooms. Increasing number of teachers will express confidence in trying new strategies. Increased functional levels by student	By February 2021, 100% of student notebooks will contain student academic goals by unit. By February 2021, 3+ walkthrough forms per teacher will have noted teacher/student reflection time. By February 2021, two PLCs per month will be dedicated to the planning and the sharing of experiences and resources as seen on PLC documentation forms. By February 2021, all fourth grade teachers will have implemented new practices into their classroom. By February 2021, all fourth grade teachers will feel more confident in taking risks. By February 2021, 90% of students will have made progress in their functional level in

			Increasing number of students tracking to meet progress from benchmark one to benchmark two	IXL since October 2020. By March 2021, 40% of fourth grade students will be on track to meet progress on STAAR.
Long-term Objectives	Students' functional levels are on grade level or are within six months of grade level Students meet progress on STAAR for fourth grade math Teachers feel empowered and fulfilled	To what extent did students' functional level improve? To what extent did students meet progress on STAAR? How have teachers' views of their success in the classroom changed?	Students functional level of mathematics increased from third to fourth grade Increased number and percentage of students meeting or exceeding progress on STAAR Increased number of teachers feel their students made gains in the classroom	By May 2021, 100% of students' functional level will have increased from the beginning of the year. By May 2021, 50% of fourth grade students will meet progress on STAAR. By May 2022, 75% of fourth graders will meet progress on STAAR. By May 2021, 100% of fourth grade teachers will feel that students made gains in the classroom.

Assumptions and Justifications

Stakeholders believe that if students are provided a PL model in math designed to meet students on their functional level and address their individual needs, then they will achieve academic growth from one grade level to the next and will ultimately meet progress on the

STAAR. It is assumed that teachers have the appropriate content knowledge, as well as classroom management skills, and are provided an instructional coach to support their implementation of personalized learning.

The chosen software for the integrated technology component of the personalized learning program is IXL and each student will be issued a device to take home and use at school. The district will use a data management system to track data and teachers will be required to keep Student Progress Monitoring Spreadsheets (SPMS) that will house all the classroom summative assessments, scores from IXL, and scores from past STAAR exams, as well as current scores on district checkpoints and benchmarks. Participating teachers will need time for training as well as planning. Professional development days will consist of a week during the summer. Teachers will receive training on the model as well as have time to build designs and templates for use in their classroom. On-going training will also be provided throughout the school year by the campus instructional coaches. Teachers will need math manipualtives that match the fourth-grade standards, the TEKS, as well as supplies for data notebooks for students. Classrooms will need adequate space for stations, teacher-small group set-up, and a space for whole group instruction. Campus instructional coaches will schedule meetings, observe, provide feedback and necessary resources, as well as review data and help the teachers set goals. Teachers will also be supported through PLCs. Specific evaluation questions were made for each strategy used in the program Table 2 below.

Table 2

Activities and Evaluation Questions

Strategies and Activities	**Evaluation Questions**
Teacher professional development on personalized software platform IXL	To what extent did teachers receive professional development on how to utilize IXL in their classroom designs?

Teacher professional development on the UA Personalized Instructional Model (core four components – Data driven instruction, student reflection and ownership, targeted instruction, integrated digital content)	To what extent did teachers receive professional development on the core components of blended learning?
Classroom designs (station designs, management)	How did teachers receive professional development on personalized learning classroom designs and management?
Teacher support through coaching (observations, feedback, meetings, reflections)	How often did teachers receive feedback and reflect on their instruction?
Teacher support through PLCs (schedules, surveys, shared resources)	To what extent were teachers trained on productive PLCs?
Data Cycles (short and long cycles, data meetings, DDI forms, SPMS)	How were DDI forms introduced and implemented weekly? How were the SPMS implemented and communicated?

Activities included in the math PL program can be grouped into four categories: training, classroom design, support, and data cycles. First, training for teachers will include the four core components of the PL model; data driven decisions, integrated technology, targeted instruction, and student reflection and ownership. Each of the components contributes to the next. Integrated software (IXL) gives teachers real time data as well as meets students at their functional level. Data is then used to group students, target instruction, and monitor student progress. Students set goals based on their data and revise as needed. Indicators that the model is being implemented in respect to the training received include increased teacher use of the PL model, IXL reports indicating student progress in their functional level, increased usage of data tracking notebooks used by students, increased number of flexible groupings based on data, and increased number of goal setting artifacts found in classrooms as well as student reflection times with the teacher.

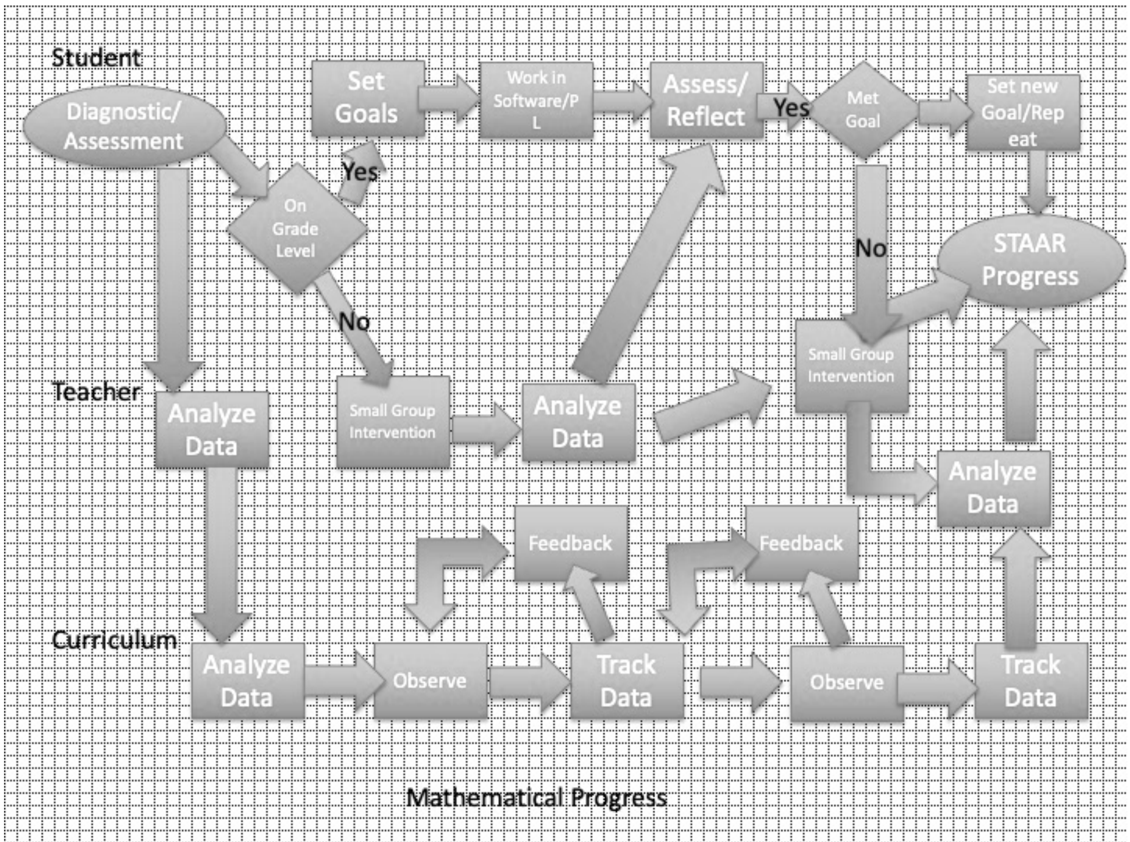
Second, classroom designs will need to include stations as well as classroom management strategies that promote student engagement. Teachers will be given design models as exemplars but ultimately be able to create their own designs as long as they include time and space for each of the four components. Indicators will include increased number of classrooms using station models and increased number of students being able to move through stations with little direction from the teacher.

Third, support will be given through campus instructional coaches as well as PLCs. Indicators will be that teachers express feeling supported by the campus coach, increased number of PLCs are being dedicated to the discussions and plans for PL classrooms and an increase in the number of new instructional strategies being tried by teachers.

Fourth, data cycles will need to be shortened and reflected on by teachers through the use of weekly Data Driven Instruction (DDI) forms in order to make adjustments to the instruction on a daily to weekly basis. (It's important to note that short data cycles are linked to indicators previously discussed such as flexible groupings, targeted instruction, student data tracking, and goal setting. The four components really are a cycle, each linking to one another.) Longer data cycles need to be in place as well which will include the SPMS monitoring. Indicators will be increased functional levels by students in IXL, increased number of students tracking to meet progress from district checkpoints and benchmarks. Long term indicators for each of the above mentioned categories will be students functional level of mathematics increased from third to fourth grade, increased number of students meeting progress on the STAAR, and increased number of teachers expressing their students made gains in the classroom overall.

Figure 2

Process Map: Math Personalized Learning Program



Theoretical Framework

BL does not have a pedagogy of its own, but it draws its strength from the three basic theoretical perspectives on learning: behaviorism, cognitivism, and constructivism (Thomas, 2010). Combining these theories into one model, The Community of Inquiry Framework (CoI) is the work of Garrison et al. (2000) in order to better structure the learning process that occurs in blended classrooms. This model of inquiry is largely based on the work of John Dewey and his constructivist theories of experimental learning (Cleveland-Innes & Wilton, 2018). Inquiry based learning is an instructional strategy in which students follow a process of discovery through experiments or observations, while asking questions. It often includes approaches to solving problems while emphasizing active participation which places the students at the heart of the responsibility for discovering new knowledge (Pedaste et al., 2015). BL is the convergence of constructivism, cognitivism, and social learning. The technology itself can be described as a cognitive tool as it is used as an engager, facilitator of thought and knowledge creation (Jonassen et al., 1999). Technology as a cognitive tool leads to the construction of knowledge in which the thinking on the content fosters the learning (Jonassen et al., 2000). Behaviorism theory is based on the idea that all behaviors are learned through interactions in the learning environment. Social interactions in the classroom, including the ones needed for BL environments such as peer interactions and self-reflections play a fundamental role in the cognitive development and thus the constructivism of knowledge (Thomas, 2010).

Impact and Significance

This study is of great importance to the organization for a number of reasons. First, the charter was set to be a lab school from the very beginning. A lab school is a place where new ideas, methods, and practices are implemented in an effort to gain insight on the effectiveness and the impact on student achievement. The lab school is in partnership with a university and therefore, undergraduate pre-service teachers are embedded in the schools as well to learn up to date teaching methods as well as see new implementation methods firsthand. The PL model under evaluation is a new one. It is imperative that the model be studied for its effectiveness not only on student achievement but the ability for teachers to be able to implement to fidelity. A lab school is based on PDSA cycles and this PL model is the newest PDSA cycle present in the district, therefore needs to be studied, and most likely changed to some degree based on this study for the next implementation phase.

Secondly, the charter was written and established to be a model not only for surrounding school districts, but for the entire nation. In fact, the vision statement for the district is, “The district seeks to be a national model for STEM education innovation as a STEM Academy and University Laboratory School." If this PL model proves to be effective and increase mathematical student achievement, not only will this benefit the district but could possibly benefit other districts and teaching practices as well. As mentioned before, and research largely supports, math is an underperforming area in most schools and the U.S. at large. If this model proves to be successful in this setting, then perhaps there is room for scalability to other districts in need of increasing the math achievement in their students.

Third, the charter was written to be a STEM school, therefore, mathematics is at the heart of who the district is and who it represents itself to be to the public and prospective students. In theory, a STEM school is one that focuses predominantly on producing students who are able to enter college with a chosen STEM major and be successful in their endeavor. The mission of this charter is to "develop students who leave school STEM College and Career Ready. STEM College Ready indicates students are prepared to enroll in a STEM Major at a university. Typically, this means they are calculus ready upon graduation or have completed calculus in high school." The evaluation of the PL model in math is imperative the goals of the charter. A strong math curriculum and instructional model directly impacts the overall mission statement of the schools. If students are not successful in math in the younger grades, it is highly unlikely that they will not be in later grades as well.

Limitations

Through this evaluation and research study, there will be data collected that could in turn present and highlight areas that have been used to personalize math content in an effort to meet students' individual needs. While some of the data that is collected may prove to reveal gains in student achievement, it is wise to assume that there very well could be other areas that led to movement in student achievement in a positive or negative direction. Factors such as teacher retention, student attrition, instructional models due to the COVID-19 pandemic, and administration turnover could all result in the data being effected inadvertently. It is also important to note that due to the small sample size of three classrooms with three teachers, perceptions may effect qualitative data as well.

In addition, it is important to note the position of the researcher in regard to the organization as a whole as well as the background of the researcher. The researcher is a white female who has been in education for sixteen years. The experiences of the researcher range from teaching high school math, professional development for teachers in math and PBL, and teaching undergraduate pre-service math and science teachers for the university connected in the study. Additionally, the researcher serves as the Director of Curriculum for the charter in which the study takes place, therefore, the positionality of the researcher is an "insider collaborating with other insiders" (Herr & Anderson, 2005). Collected data is data that is already collected regardless of this study such as teacher observations and student assessment data. Teacher observations are done through the three instructional coaches and teachers administer the assessments. Member checking is a key piece to all data collected throughout the evaluations.

Chapter 2: Literature Review

Student achievement and success in mathematics in the United States has continued to be a concern over the past several decades when data of U.S. students is compared to data of students in other countries. Educators have relied on best practices in teaching mathematics as well as emerging promising instructional methods to increase the mathematics achievement of students. Of these, BL has become a "buzz" word in education as one instructional strategy that could be a possible solution to the gaps seen in classrooms. BL, defined as any time a student learns at least in part at a supervised brick and-mortar location away from home and at least in part through online delivery with some element of student control over time, place, path, and/or pace (Staker, 2011). A newer approach, which falls under the BL umbrella, is PL. A consensus on the definition of PL is not reached in the literature, but most research points to some aspect of customization, student groupings, and flexibility of instruction (Berry, 2018). The implementation of PL practices in American schools has increased significantly over the past several years (Pane et al., 2015) however, the variances in implementation range drastically from organization to organization (Staker, 2012) resulting in little research on the effect PL has on mathematics achievement.

The following literature review will first look to discuss the impact low achievement in math has on students in elementary grades as well as the potential impact of low achievement in early grades for later years in students' lives. Instructional strategies used as well as research on a few linkages to math achievement will also be presented. Secondly, this literature review will focus on research in relation to PL models, specifically, 1) the varying approaches in PL models, 2) the impacts of one or a combination of two of the following components in a PL model has on mathematical achievement; data driven decisions, student goal setting and reflection, targeted instruction and integrated technology, and 3) the barriers to implementation of PL models. Based on the research presented early in the literature review from a student lens and an adult lens, the remainder of the literature review will focus on a working theory of improvement. Research from several approaches to possible change ideas for improvement will be outlined by 1) PL lessons learned, 2) best practices in mathematics, 3) content knowledge and experience of teachers, and 4) targeted professional development for teachers of mathematics.

Review of the Scholarly Knowledge – Student Lens

Mathematics Achievement in U.S. Students

The achievement of students in mathematics in the United States (U.S.) has fallen short to other countries for many decades. According to some studies, the achievement of U.S. students in mathematics ranks around average (Desilver, 2017). Mathematics scores seem to have leveled off to stay relatively the same since 2006 with no significant improvement in results. Other studies which include results from the TIMSS show U.S. students, particularly fourth-grade students, to be ranked eleventh out of 45 countries (Provasnik et al., 2016). A report from the Organisation for Economic Co-operation and Development (OECD), found that the average U.S. mathematics score in 2018 was lower than the average OECD mathematics score and had not measurably changed since 2003. An examination of the results of the 2019 National Assessment of Education Progress (NAEP) results in mathematics shows that the average mathematics score for students in Texas was not statistically different from their average score in 2017. Only 30% of students in Texas performed at or above the NAEP proficient level (NAEP, 2019).

Mathematics Trajectories

Various factors could influence student performance in math such as the teaching style, homelife, recent loss or trauma, but of the varying factors, researchers have reported that students' previous mathematics achievement is strongly linked to their current mathematics achievement (Yeo et al., 2022), and is called a cumulative pattern (Salaschek et al., 2014). This type of cumulative pattern assumes that the high achievers get better and the low achievers get lower, or "the rich get richer and the poor get poorer" (Stanovich, 1986). This trend held true in the most recent report by NAEP (2022) for mathematics achievement in fourth grade students. The magnitude of score declines for lower performing students were greater than the declines for higher performing students (Nation's Report Card, 2022). Once students fall behind in mathematics, catching back up to their peers on grade level has proven difficult to do, and rarely done. Research in mathematics from early grades has revealed that proficiency in math as early as kindergarten and first grade has been linked to later achievement in grades eight (Claessens et al., 2009). In addition, math proficiency and courses taken by students in grade eight have been largely linked to students taking robust math courses in high school and later in college (Wang & Goldschmidt, 2003). Research supports the linkage between early math success in elementary grades and future ability to enroll and be successful in more rigorous math courses in college.

Mathematics Achievement and STEM Fields

Math achievement and math growth trajectory plays a significant role in not only predicting future math outcomes but also in future career choices as well (Yeo et al., 2022). Creating educational programs that result in STEM oriented outcomes for all students is an important federal policy initiative and educational objective in the United States (Shanley et al., 2019). Students' ability to demonstrate proficiency in mathematics is largely related to the STEM outcomes for students and their future degrees and careers. Elementary education in mathematics and science are the foundation for entry into postsecondary STEM majors and occupations (NSF, 2022). According to a report by the National Center for Education Statistics, of all the bachelor's degrees awarded in 2015-2016, only 18% were in STEM fields (NCES, 2016). Success in mathematics has been associated positively with occupational success, leadership roles, and educational attainment (Lubinski et al., 2014). Reports by the OECD showed that the United States ranked higher in science, coming in at 7th overall out of 37 OECD countries than it did in math, placing 25th out of 37 (NSF, 2022) and as mentioned earlier, math scores remained basically stagnant for the past twenty or so years.

Mathematics Instructional Strategies and the Impact on Student Achievement

When considering mathematics instruction, two main instructional practices have been distinguished over the past several decades; direct (or traditional) instruction and dialogic (ore reformed) instruction (Campbell & Yeo, 2022). Direct instruction relies mainly on the teacher, and therefore is considered teacher centered instruction. In direct instruction, the teacher is responsible for relaying all information, demonstrating mathematical strategies and students are responsible for following along and ultimately applying such strategies in new situations or problems. The discourse among students is not of high concern in teacher centered classrooms and the main dialogue occurs between teacher and student. While some studies have found that direct instruction lowers the difficulty of the tasks (Chandler & Sewer, 1991), other studies have shown that due to the nature of scaffolding that occurs in direct instruction it can promote long term memory in students (Kirschner, 2006).

Dialogic instruction, sometimes referred to as discovery learning or inquiry-based learning is more student centered in nature. This type of instruction was first advocated for by the National Council of Teacher of Mathematics (NCTM) as early as the late 1980's. In this type of instruction, collaboration, communication, the use of manipulatives, and problem solving is a high focus. Students are given autonomy to a degree and are engaged in strategy invention and the use of multiple representations (Yeo et al., 2022). While these two approaches are drastically different, it has been known that many teachers implement both strategies in their classroom. The research has reported mixed findings from the students' perceptions about which instructional method is more beneficial to them. While some research has found that dialogic methods have promoted higher achievement and perseverance in math, particularly in problem solving, (Boaler, 2008) other studies have concluded that non-significant differences between the teaching styles and even positive results when direct instruction is the focus. In one study in grade eight mathematics in Sweden, it was found that classrooms that focused on teacher explanations, student listening, and student memorization were positively linked to secondary achievement in mathematics (Eriksson et al., 2019).

Links to Mathematics Achievement in Elementary Students

To further expand on dialogic instructional strategies, in 2014 NCTM affirmed that the mathematics connection ability is the most important factor in understanding mathematics concepts. This idea of mathematical connectedness means that students are able to connect math ideas to other previously learned math ideas thus strengthening their understanding, and their ability to connect math concepts to real world concepts (NCTM, 2014). Mathematics connections must first be recognized by the teacher and realized by students under dialogic instructional strategies provided that the teacher. In a research study done in fifth grade math students in Indonesia, it was found that the mathematics connection ability determines students' mathematics learning achievement (Ndiung & Nendi, 2017). This study also discussed that students' ability to connect math ideas not only relates to their math achievement but also to other subject areas and everyday life. Mathematics instruction needs interdependence to other disciplines as well as real world concepts so the learning can be meaningful. The perceived usefulness of mathematics, its value, has also largely been linked to the success of students in math (Adelson & McCoach, 2011).

It has been found that success in mathematics in early grades is linked to the students' enjoyment of math (Villaviicencio & Bernado, 2013). Lambic and Lipkovski (2012) argue that the enjoyment of mathematics seems to have a greater influence on math achievement than any other factor. In a study conducted by Garcia et al. (2016) it was found that the enjoyment of mathematics, defined as the degree to which a person takes pleasure in doing and learning the subject (Adelson & McCoach, 2011), was the single greatest indicator of student mathematics achievement. The study included 524 students from 12 primary schools in Spain and the ages of the students ranged from 10 to 13 years old. In the same study, it was also found that surface level teaching of mathematics the greatest negative impact on the students' mathematics achievement (Garcia et al., 2016). Surface level approaches to learning involve more traditional teaching methods described earlier such as rote memorization. This type of teaching rarely entails deep elaboration into the content and is characterized by the repetitious rehearsal of information. As a result of this type of learning, knowledge acquired fades quickly (McInerney et al., 2012).

Review of the Scholarly Knowledge – Adult Lens

Variations in Blended Learning Models

Once BL started to take hold in many schools, Watson (2008) set out to observe and study the implementation methods across nine different organizations. He concluded that there is no single type of blended education. It is unique and requires new methods of instruction, content development and professional development. In the early years of implementation, many terms started to emerge among BL implementations. After studying more than 80 programs among American schools, Staker categorized BL into four main categories; rotation model (includes station-rotation, lab-rotation, flipped-classroom and individual-rotation), flex model, self-blend model, and enriched-virtual model (Staker, 2012). While research in any of the four areas mentioned above exists, the remainder of this literature review will focus on the rotation model, in particular station-rotation and individualized-rotation under the umbrella of PL.

The Shift to Personalized Learning

As society becomes more diverse the need for customization over standardization in education has created a paradigm shift (Aleven et al., 2017). PL is considered a way to address the customization needs since it acknowledges students' differences in background knowledge, interests, and abilities (Holmes et al., 2018). The International Association for K-12 Online Learning (iNACOL) claimed that "K-12 education is at the beginning of what many hope will be a systematic transformation toward personalized learning". The Every Student Succeeds Act (ESSA), which is the federal K-12 education law in the United States, was signed into effect in 2015 and replaced the former education law, No Child Left Behind (NCLB). ESSA supports PL models to close achievement gaps of diverse K-12 students (Lee et al., 2021). Many programs such as The Race to the Top Program funded states to transform their systems to develop local PL instructional models in their schools unique to their districts and needs (U.S. Department of Education, 2012). The promotion of PL has also been endorsed by the U.S. Department of Education's National Education Technology Plan (2017). In the state of Texas, many grant funded programs like Math Innovation Zone (MIZ) and Blended Learning Grant (BLG) were awarded to districts to create and adopt their own local PL instructional models.

Personalized Learning Models

While there are claims that PL models are positively related to gains in achievement for students, there is very little research on which PL components or models yield the most desirable achievement results (Lee et al., 2021). Perhaps this is largely due to the wide varieties in definition, components and design models. A consensus on the definition of PL is mainly non-existent, but most research points to some sort of customization, student groupings, and flexibility of instruction (Berry, 2018). In addition, terms such as personalized learning, adaptive learning, individualized learning, and customized learning are all used when referring to PL. Differences between these terms are often unclear and used interchangeable which can lead to further confusion around PL (Xie et al., 2019).

When PL models utilize online content delivery, PL is one strategy to blended learning. In a review by Park and Lee (2004) of the theoretical perspective of customized learning that have appeared in history, three main types came to the forefront. First, macro-adaptive instruction which can be traced to the 1970's, focuses on adaptation from a macro level. The adaptation is mostly on the group level on some sort of premeasured level before the instruction begins and the adaptations are minimal. Students were provided with a new task after mastery was shown (Park & Lee, 2004). Second, aptitude treatment interaction is based on learner characteristics that are often measured in advance (Van Schoors et al., 2021). Third, and the type referred to in this literature review most often, micro-adaptive instruction which occurs not only by a premeasure but also during moments of the student interacting with the system (Parker & Lee, 2004). This type of customization typically comes from a diagnostic assessment at the beginning of the year in the chosen software program that then tailors lessons and assignments to the functional level of the students and/or the needs of the student within the grade level curriculum. Micro-adaptive instruction also customizes lessons based on real-time interaction between the student and the online software platform. Additionally, in some PL models, customization can also come from the teacher who is able to use the data from the online content to create unique instructional experiences for their students. These types of models focus on student data to create learner profiles which are then the personalized recommendations for each unique student (Bulger, 2016).

While PL can be thought of as an instructional approach that focuses on the individual needs of students, oftentimes it has been implemented in such ways that either focus solely on the academic achievement of students or the interests and social impacts on students. When PL classrooms are inclusive of students' interests and their personal goals academically, such programs have the potential to increase student learning and engagement (Pane et al., 2017). In *The Truth About Personalized Learning*, Pane et al. (2016) highlighted that PL has significant effects on students' mathematical performance but also noted that PL doesn't have any implementation standards or official methodology. Out of the 32 schools studied, there were 32 unique implementations (Pane et al., 2015). There is also little research on how different PL instruction are associated with academic performance on standardized tests (Lee et al., 2021). When considering PL models and design, especially in subjects with high stakes tests for accountability, it is important to incorporate research based best practices in the content area.

Well Balanced Math Classrooms and Components of Personalized Learning Models

The National Council of Teachers of Mathematics (NCTM) promotes eight practices that are considered to be best practices in mathematics which were recapped in their *Personalized Learning and Mathematics Teaching and Learning* publication:

> Establish mathematics goals to focus learning, implement tasks that promote reasoning and problem solving, use and connect mathematical representations, facilitate meaningful mathematical discourse, pose purposeful questions, build procedural fluency from conceptual understanding, support productive struggle in learning mathematics, and elicit and use evidence of student thinking. (Berry, 2018, para. 5)

When reviewing and implementing PL practices into the classroom, NCTM encourages educators to ask questions around each of the eight practices and how they are addressed in the PL model (Berry, 2018). While there is research on different implementation models, the majority of the research lends itself to studies addressing only one or two components of a PL model. Four components of PL will be reviewed in the following paragraphs; data driven decisions, personal goal setting and reflection, targeted instruction, and use of technology.

First, data driven decisions utilized through PL models have been reported to have the single greatest impact on student achievement when used for instructional grouping, and meeting students at their functional level (Zdeb, 2018). Data-driven decision making is the process of collecting, analyzing, and applying many forms of data from a variety of sources in order to change instruction. The goal is to enhance student performance while addressing the learning needs (Marsh et al., 2006). Earl and Katz (2002) noted, that the use of data to change instruction is no longer an option for school improvement reasons but the issues of using data to do so still remains an issue. Of these issues, the timely availability of data was a top concern (Earl & Katz, 2002). Longer data cycles can cause problems in addressing student needs in a timely manner and thus the need for shorter data cycles is prevalent. The data collected in these shorter data cycles are typically called formative assessments. Formative assessments is defined as "a systematic process to continuously gather evidence about learning" (Heritage, 2007, p. 2). This type of assessment is used to gather real-time information about a student's current level of understanding and thus allow teachers to meet students where they are and address gaps. Data from assessments not only help a teacher make data driven decisions, but also have the potential to provide early interventions to students to address issues before they become major learning shortcomings later on (Wilson, 2017). Such use of real-time data is a major component of PL models. Since multiple forms of data are key to making data driven decisions, the data obtained from online learning platforms is not only a piece of data that can be beneficial for teachers, but for students as well. Students having access to his or her own data coupled with discussions of their data with their teacher, has been reported from schools with the greatest achievement gains in mathematics (Pane et al., 2015).

Second, personal goal setting based on academics such as past grades and current level of understanding has been found to be a strong motivator for increasing student performance on standardized tests (Smithson, 2012). In an action research study done by Smithson (2012), it was found that teacher assisted student reflection and personal goal setting positively impacts the academic performance on assessments in elementary students in all subjects. Goal setting can be defined as specifying requirements for personal success by initiating self-monitoring and self-judgments of performance and the attainment towards the goal (Bandura, 1991). When comparing the goals of underachievers to achievers, underachievers had no particular goals whereas achievers not only set goals, but realistic, attainable ones that were related to their academics. For students to achieve success in the classroom, teachers should engage in the process of teaching students how to set goals, reflect on them and reach them (Simthson, 2012). It is important for students to have a role in setting their own goals which promotes autonomy and ownership in their academic achievements. Guidelines for setting goals should include stating the goal in written form, making the goal as concrete as possible, conceptualizing what it looks like to accomplish the goal, identifying the steps to do so, receiving feedback from the teacher on the progress towards the goal, and communicating what did and did not work (Nunez, 2011). Zimmerman et al. (1992) found that when students are taught to obtain distant goals in a process of achieving smaller goals connected to the larger one, they can make progress in learning skills and content faster. Additionally, it has been found that a student's desire to become proficient is a predictor of academic achievement, particular with students in grades first through third. Studies showed that higher math and reading grades were linked to higher levels of achieving desired goals (Broussard & Garrison, 2004). Since research tends to point to the positive effects goal setting can have on student academic achievement it is interesting that research to support PL models including this component is scarce. After studying practices in PL models across 62 schools, Pane et al. (2015) reported few models that used personal goal setting among students.

Third, targeted instruction in regard to flexible grouping in PL classrooms is shown to have the greatest impact on student achievement in math when using data to drive the groupings as mentioned above. Differentiated instruction by grouping has been a widely accepted practice in math classrooms, however, some studies have shown it to have little to no effect on elementary students (Maxey, 2013). In another study conducted with sixth grade math students, it was found that flexible instructional groups can help teachers meet the varied needs of their students which in turn allows for students to be more engaged and therefore successful in the classroom (Mainini & Banes, 2017). It is almost certain that a single classroom will present itself with mixed levels of ability, learning styles, strengths and needs of students. It is imperative that teachers are equipped with tools to meet the needs of all of their students (Mainini & Banes, 2017). Through flexible grouping, teachers are able to perform small group instruction which is defined as "situations in which three or more students work on a common mathematical task" (Jansen, 2012). In these small group situations, skill grouping is a top priority. Skills grouping has been found to be effective when done fluidly and for a short period of time (Gibbons, 1991). In addition, Webb et al. (2009) found that teacher probing of student ideas and answers in small group instruction may be more effective than their probing of such ideas in whole group instruction. This probing led to higher instances of correct and complete mathematical explanations. In other research on flexible groupings, it was found that teachers rarely use it due to time constraints and pressure they feel from state testing (Sanders, 2015) and that often times student groupings are seen as either one-on-one or whole class but rarely as a combination of the two (Berry, 2018).

Fourth, integrated technology is included in all PL programs and most reported using the technology to be the personalization as well as the place data is mainly drawn from in making decisions (Pane et al., 2015). A study on fourth-grade math students in Taiwan, reported that personalized computer assisted math programs improved student performance and attitude (Chen, 2007). A contrasting study on fourth graders in the U.S. showed that after seven weeks students placed on a computer based program in math showed no significant difference in student achievement than students receiving hands-on instruction (Ravenel et al., 2014). In another study observing integrated technology in elementary classrooms the use of virtual versus concrete manipulatives were compared. The results showed no significant difference in achievement on a post-test amongst the two groups (Burns, 2011).

PL Challenges and Barriers to Implementation

PL is often used solely to improve test scores and leaves out the humanizing aspect such as discussions and groupings, even though those are known as best practices in mathematics (Berry, 2018) as well as having the greatest impact when implemented in PL models (Pane at al., 2015). PL is most commonly seen as an extension of a previously established instructional model in a school rather than a brand-new innovative program. More often than not, it takes an average of two full years of implementation before gains are seen in mathematics (Pane et al., 2017) and far too often the practice of PL is abandoned before results in achievement are seen.

PL can be a tough transition for schools and educators who are used to more traditional instructional practices. PL requires teachers to change their instructional practices and become a coach or facilitator rather than a lecturer, requires them to use new tools (online software, new assessments and diagnostics), and possibly even restructure their classrooms into more flexible grouping set-ups (Staker & Horn, 2012). PL is often challenging for teachers to put into daily practice since it requires a great amount of commitment, real-time correct judgment, and thorough preparation which can be time consuming (Van Schoors et al., 2021). Each of these factors present challenges for schools in terms of training and supporting their teachers with such endeavors. A strong barrier to the implementation of PL models has been the teacher preparation, development strategies, and teacher practices have not yet caught up with the demands and needs of PL teachers (Bingham et al., 2018).

Schools are also hesitant in some cases to implement PL models due to the pressures of high stakes testing in the U.S. The insert of high stakes testing under NCLB placed much pressure on teachers and administrators and even caused some teachers who used student centered instruction to abandon these practices and shift to teaching to a test (Lee et al., 2021).

Even though ESSA encourages PL models for instruction, the high stakes testing are still present as well as the accountability measures that are presently in place in U.S. schools. Perhaps another hesitation comes from the wide varieties of models and the very little research on which PL components or models yield the most desirable achievement results (Lee et al., 2021). It has been found that a leading implementation barrier for PL models is the expectations from the users (teachers, students, school leaders) to outside expectations such as standardized testing. The way in which school and student success was measured in PL models was not aligned with how outside stakeholders measure student success (Bingham et al., 2018).

One major barrier in the implementation of PL models are often the school systems and policies in place that inhibit innovative practices (Zedb, 2018). Of these could be assessment schedules, scope and sequences, and even the technology available. The amount in which technology is used in the classroom as well as the way in which it is used is largely based on the comfort level of the teacher (Ertmer, 2005), their belief about teaching with technology (Zhao et al., 2002), and the quality of the technology itself (Groff & Mouza, 2008). Each of these factors can have a correlation to teachers using the technology in alignment to PL practices or more traditional teaching practices. In a collective study by Bingham et al. (2018), school infrastructure and available technology not yet aligning with PL teachers' needs was one of the top implementation barriers.

Further Research Studies Needed to Define Best Strategies for PL Models

The study conducted by Pane et al., (2015) is the largest PL study conducted and shows PL has positive impacts on mathematical achievement but also noted the results needed to be confirmed against more rigorous experimental study designs (Pane et al., 2017). The field lacks evidence on which instructional strategies, or combinations of instructional strategies, within a

PL model are most effective. There is also little evidence on what policies need to be in place to maximize student benefits (Pane et al., 2017). In a study on the differences in PL models and technology use, it was found that teachers in high-performing schools implemented PL more thoroughly and utilized the functions of the technology more than those in low-performing schools (Lee et al., 2021). In other studies, a positive trend was observed on student learning outcomes although this same studied noted that methodological differences need to be considered and more research was needed to determine which models lead to stronger trends (Van Schoors et al., 2021).

When PL models utilize online content delivery, typically used to customize the learning for the student at their functional level, PL is one way to implement blended learning. While PL can be thought of as an instructional approach that focuses on the individual needs of students, it has often been implemented in such ways that either focus solely on the academic achievement of students or the interests and social impacts on students. When PL classrooms are inclusive of students' interests and their personal needs and goals academically, such programs have the potential to increase student learning and engagement (Pane et al., 2017).

Early evidence suggests that PL can improve achievement of students in math, regardless of their starting level of achievement. At the present time, there are no common PL models among implementation sites and not even one common definition. While research does show that data driven decisions used to target instruction is a common strategy used in PL schools that reported the largest gains (Pane et al, 2017), there is little research outside of this strategy as to what other strategies may prove to have positive impacts. Some schools in the studies described above also showed large negative results when implementing PL (Pane et al., 2017), while other studies showed no significant differences at all between PL control groups and groups of students not receiving PL (Ravenel et al., 2014). While it is believed that PL is not going away, and in contrast will only continue to grow in practice, it may take some time to fully see the impacts on mathematical achievement (Zdeb, 2018). It is also unclear what the long-term effects of PL will be as more and more schools are shifting to this instructional model in today's schooling. It is likely that even more educational tech companies, as well as curriculum companies, will produce PL materials, but at this time many questions still remain as to what they should contain.

Models for didactic design for PL that are based on theoretical concepts are missing (Kerres & Witt, 2003), and no research was found that combines all four components described above into one PL model. After looking at a variety of approaches to PL, the models can be grouped into four broad strategies; learner profiles, personal learning paths, competency-based progression and flexible learning environments (Pane et al., 2015). The majority of the schools studied by Pane et al. (2017), used one to two of these strategies but not all four. It is still unclear what combination of PL practices has the greatest impact (Pane et al., 2017).

Potentially Impactful Change Efforts

Previously in this literature review, impacts of low performing students in math in the early grades was discussed as well as strategies to address this concern such as PL models. Concerns about the achievement level of elementary U.S. students in math as well as challenges in implementing PL models were present prior to Covid-19. Post Covid-19 student achievement gaps and the need for sound instructional strategies only grew. Technology is being used more in recent years than prior to Covid-19 and students are experiencing larger deficits than ever. The remainder of this literature review will focus on working theories of improvement for these challenges faced by educators post Covid-19.

The learning loss from the last day of school to the beginning of the next school year is typical in any given year in mathematics, however, after the COVID-19 related shutdowns, the learning losses were expected to be larger than ever. In fact, studies predicted that students returning in the 2020-2021 school year would have only on average 37-50% of the learning gains when compared to a typical school year (Kuhfeld et al., 2020). In one study which compared more than 8,000 schools and the students' scores in reading and math from the fall of 2019 to the fall of 2020 showed that while reading scores remained relatively the same, math scores in 2020 were five to ten percentage points lower than in 2019 (Bailey et al., 2021). Educational researchers predicted that the gap will only grow in the 2020-2021 school year causing students to start the fall of 2021 even further behind than in the fall of 2020. In a recent study conducted by NAEP, this prediction rang true. In 2022, fourth-grade mathematics scores declined at all five selected percentiles for the first time since the initial mathematics assessment in 1990 (NAEP, 2022).

Kraft and Falken (2020) provide an evidence-based blueprint for scaling effective tutoring strategies in order to help close the learning gaps. The Texas Education Agency has changed laws recently to require schools to provide accelerated instruction to students who did not approach grade level standards on the 2021 STAAR assessment. As a requirement of the accelerated instruction, tutoring is the strategy presented by the state. The Effective Schools Framework by The Texas Education Agency (2020) says that effective classroom routines and instruction are foundational essential actions present in schools that are effective. Further, it states that campus instructional leaders provide training and on-going supports so that teachers can effectively use research-based teaching practices. More and more schools will find themselves needing to refer to the Effective Schools Framework as a result of the learning losses from the pandemic. In order to do so, best practices in math will have to become more known to many schools and teachers.

Best Practices in Mathematics

In 1989, The National Council of Teachers of Mathematics launched a standards-based movement in America with their *Curriculum and Evaluation Standards for School Mathematics*. This initiative was based on best teaching practices in mathematics and was revised in 2000 in the launch of Principles to Actions which includes eight guiding practices; establish mathematics goals to focus learning, promote reasoning and problem solving, connect mathematical representations, facilitate mathematical discourse, purposeful questioning, procedural fluency and conceptual understanding, productive struggle, and evidence of student thinking (NCTM, 2000). Inquiry oriented instruction and concepts-based teaching in math is positively related to student achievement (Blazar, 2015). Teacher practices such as these, as well as student participation and mathematical discussions when students are able to explain their own ideas, have also been linked to positively increasing student achievement (Ing et al., 2015). While personalized learning is on the rise in schools across the nation, there is still a place for best practices in math. The key might just be figuring out how to best align personalized learning and mathematical best practices. Personalized learning strategies must blend with mathematical best practices, they cannot be stand-alone ideas but must work together seamlessly in the classroom to be successful (Berry, 2018).

Content Knowledge and Experience

Beyond best practices in math, there is also a strong relationship between the teacher's content knowledge and how the mathematics is enacted in the classroom which ultimately leads to students being successful or not (Blazar, 2015). Best practices have been outlined by NCTM for decades, and now there is research to support their theories. Most teachers will even, when asked, say that they believe that best practices and innovative teaching strategies are the best way to teach math in the elementary grades. However, it is often found that their perceptions do not align to their practices and they rely more on how they were taught than their own beliefs about math (Harbon & Newton, 2013). More inexperienced teachers have an even harder time implementing best practices into their classrooms in the beginning of their careers (Ing et al., 2015). The question then becomes, how do we get more teachers who not only believe in best practices, understand how to teach them and then actually implement them into the mathematics classrooms. This can particularly be a problem in elementary classrooms as teachers in Texas who teach elementary grades are not math content experts. Math is only one portion of the EC-6 certification exam. Therefore, unless elementary teachers come to schools with strong content knowledge and a profound understanding of teaching mathematics, schools might just have to take on the role of growing their teachers in this regard. Teachers today are younger, less experienced and more diverse in their preparation for the field than even two decades ago. This is a problem because inexperience and lower levels of content knowledge prove to be less effective in the classroom (Henry et al., 2014). Teacher content knowledge of concepts and connections is directly related to student achievement (Tchoshanov, 2011). Harbison and Hanushek (1992) stated, “At fourth grade, a ten-point improvement in the mean teacher’s command of her mathematics subject matter…would engender a five-point increase in student achievement; this is equivalent to a 10% improvement over the mean scores of fourth graders” (p. 114).

Personalized Learning Lessons Learned

In addition to the content knowledge of teachers being an issue in relation to student achievement, with the rise in personalized learning programs, there have been many pitfalls that have come along with it. COVID-19 related environments only increased the use of technology in classrooms but with little to no training or support for teachers on how to best implement the technology into their classrooms. Pitfalls of PL models include the overuse of technology. Learners can become mere consumers of the technology and the technology can divide, disconnect, and alienate students if used too heavily (Sulceio de Alvarez, 2018). Students must be given the time to learn and engage with challenging problems and to think. Schools need to return to previously successful experiences again used prior to COVID-19 and the reliance on technology (Sullivan et al., 2020). Teaching mathematics requires consistent guidance from teachers through collaborative opportunities (Khirwadkar et al., 2020) that can only come through discussions and group work in the classroom.

Working Theory for Improvement in Math Achievement and PL Models

Previously reviewed research suggests that one major barrier to implementing PL is the support and training for teachers to successfully change their practices. Other research surrounding low achievement in math suggests that the content knowledge of the teacher as well as the implementation of best practices in math play a large role in math achievement of students. A possible working theory for improvement is job embedded, targeted professional development for teachers that involves developing the content expertise of teachers as well as support in PL models.

Job-Embedded Professional Development

Job-embedded learning has been defined as learning that occurs as teachers engage in their daily work activities (Wood & Killian, 1998) and also as learning by doing, reflecting on the experience, and creating and sharing new ideas and insights while learning with oneself and others (Wood & McQuarrie, 1999). Job-embedded professional development (JEPD) refers to teacher learning that occurs in day-to-day teaching practice. It is designed to enhance teachers' content-specific instructional practices with the ultimate goal of improving student outcomes (Croft et al., 2010). Job-embedded professional development typically has the following characteristics; (1) occurs regularly, aligned with academic standards, curriculum, and school improvement goals, (3) involves educators working together collaboratively, often facilitated by school instructional leaders, coaches, or mentors, (4) requires active engagement rather than passive learning, and (5) focuses on understanding what and how students are learning and how to address student needs (Zepeda, 2019). JEPD should be centered on finding solutions to immediate problems of practice as part of cycles of continuous improvement as it makes a direct connection between learning and application to daily practice (Croft et al., 2010).

In order for JEPD to be effective, teachers must be given multiple opportunities to learn. JEPD can take many forms but can vary between departmental, cross-departmental and vertical teams of teachers. Activities for JEPS can include mentoring, instructional coaching, peer observations, lesson studies, book studies, action research, student work protocols, professional learning communities, critical friends, and portfolios (Croft et al., 2010; Zepeda, 2019). The school's professional culture must foster continuous learning, openness to feedback, and open door policies for peer observations and other visitors. Research based knowledge about how adults learn should be at the core of the design for JEPD. According to Knowles (1980) and

Knowles et al. (1998), adults learn best when they are self-directed, can use their life experiences, linked to real life situations, applied immediately, motivated internally related to their job, and know why they need to learn.

Targeted Professional Development

In order to address teacher content needs and best practices in math and/or PL, there needs to be a targeted emphasis on the development of the teachers' mathematical conceptual knowledge through content-focused professional development geared toward student achievement (Tchoshanov, 2011). A large study done by Blank and de las Alas (2009), provided scientifically based evidence for the positive effects of content-focused teacher professional development on student learning in math. Teachers who received math based professional development had students who performed better than teachers who did not. In a different study by Cave and Brown (2010), it was concluded that the professional development elementary math teachers received resulted in students making greater gains in math than expected. Providing more professional development opportunities to enhance teacher content knowledge is one way schools can support their teachers and help students achieve.

Another way for teachers to gain a better understanding of how to teach elementary mathematics is through frequent observation cycles and intensive coaching and support programs (Blazar, 2015). There is research to support that teacher behaviors will change with targeted professional development and reflections from the teachers (Thomas, 2008). Job embedded professional development has also been shown to increase teachers' self-efficacy in teaching mathematics (Aulthauser, 2010). Other studies have shown that students in schools whose teachers receive professional development in math perform higher than schools whose teachers do not (Brendefur, 2020).

Post Covid-19, the need for teachers to receive support in technology platforms has increased as well. Abaci et al. (2020) provided a framework as a result of a study done to understand pitfalls with technology and usage by teachers during the pandemic shutdown. The framework consists of six findings that are considered important components to supporting teachers, improving the professional development for teachers in online platforms. These six elements consist of (1) designing and developing a supportive professional development environment, (2) acknowledge the existing context regarding the online platforms, (3) address the teacher change that needs to take place with the transition to more digital learning, (4) determine the overall goals for the professional development, (5) acknowledge the professional development strategies, and (6) disseminate the knowledge, skills, and attitude about digital learning and evaluate the professional development.

While there is a great deal of research on best practices in math and some research on the effect of personalized learning in some grade levels and subjects, there is less research or case studies in relation to the lessons learned from implementing personalized learning in math classrooms. As the research above suggests, best practices, content knowledge, and targeted professional development for teachers are linked to an increase in student achievement in math. The intervention of job-embedded professional learning for teachers in math and PL is an attempt to not only enhance the content knowledge of the teachers, but to also promote best practices in math and PL through targeted professional development sessions for elementary math teachers provided by an instructional coach who also provided regular observations and feedback.

Chapter 3: Evaluation Study

Introduction

The public university charter school was established within the College of Education and Psychology in 2012. The schools serve two purposes: to design and deliver innovative STEM education to K-12 students, and to serve as a research and demonstration platform in teaching, learning and assessment for faculty in the School of Education, providing training of future teachers, educational experimentation, educational research, and professional development. Administrators and teachers within the academy serve as adjunct faculty to the School of Education and partner to support educator preparation and research. The laboratory schools allow for longitudinal research and the testing of interventions to fidelity not always possible in traditional partner schools not managed by the university. A Network Improvement Community (NIC) has been established to implement improvement science practices and leverage the expertise of professors, content experts, school administrators, teachers, and students when appropriate. The school curriculum is co-managed by university faculty and district personnel. The NIC helps identify and guide school improvement priorities and research.

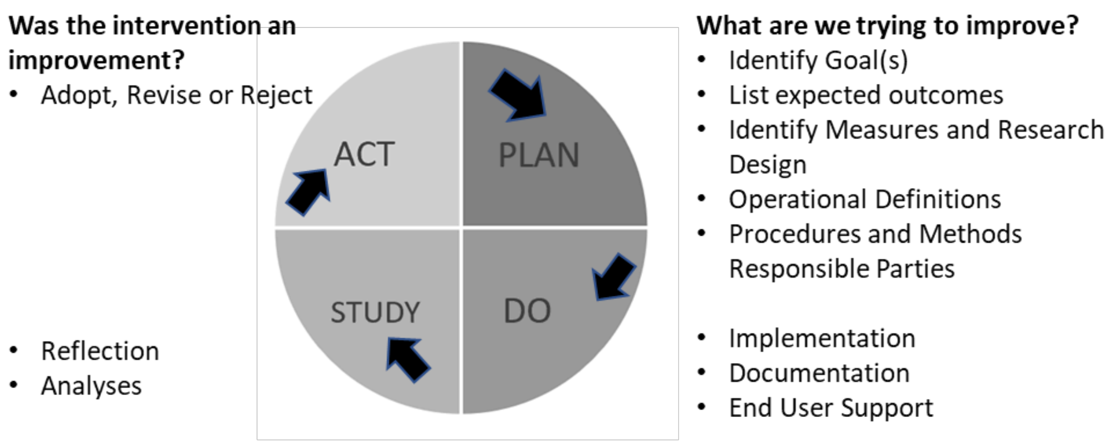
This chapter focuses on improvement science cycles implemented in mathematics classrooms at one laboratory school over the course of nine years, with an emphasis on the third PDSA cycle for improvement and the evaluation of a personalized learning model. Each of the cycles presented are grounded in the use of disciplined inquiry to drive changes that lead to improvement. In *Learning to Improve,* Bryk et al. (2016) describes that:

> All activity in improvement science is disciplined by three deceptively simple questions: 1. What specifically are we trying to accomplish? 2. What change might we introduce and why? 3. How will we know that a change is actually an improvement? (p. 114).

Three distinctive improvement science cycles will be described and each one will focus on the previous questions by stating a problem of practice, describing the interventions and their primary drivers which took place in order to address the problem, and reviewing data to measure the effectiveness. The Plan-Do-Study-Act (PDSA) Inquiry Cycle was applied to each of the cycles, and often, multiple PDSA cycles were used in one larger improvement cycle. Figure 1 highlights the PDSA model used by the NIC and staff to drive improvement.

Figure 3

The Plan, Do, Study, Act (PDSA) Model Approach



Literature Review

Student achievement and success in mathematics in the United States has continued to be a concern over the past several decades when data of U.S. students is compared to data of students in other countries. Various instructional strategies, as well as resources, have emerged along the years as possible solutions to this lagging problem. Of these, PBL, PrBL, Phenomenon-Based Learning (PhBL), and BL are practices that are garnering attention in schools and the research literature. Each of these instructional methods are addressed separately for ease of presentation for their individual PDSA cycle.

PBL and PrBL are both inquiry types of instructional methods, however they are not the same. These terms are sometimes used interchangeably but there are clear distinctions in the methodologies. The foundational concept behind PBL and PrBL is to develop students who can manage their own learning (Odell & Pedersen, 2020). Students learn by designing, applying, and problem-solving while collaborating with other students and presenting their ideas and findings. PhBL shares similarities with both PBL and PrBL, however PhBL extends learning into a global context through both topical and thematic instruction while focusing on real-world issues or phenomena (Drew, 2020; Finnish National Board of Education, 2016; Prakash Naik, 2019).

In the past several decades, there has been much support for the use of both PBL, PhBL, and PRBL in STEM classrooms. Today's educators face the challenge of preparing students for jobs that have are yet to be created and problems that are yet to arise (Bybee & Fuchs, 2006; National Science Teachers Association, 2011). Inquiry methods such as these could possibly be a solution to this problem in that both methods focus on 21st Century Skills in addition to the standard content.

BL can be defined as any time a student learns at least in part at a supervised brick and-mortar location away from home and at least in part through online delivery with some element of student control over time, place, path, and/or pace (Staker, 2011). A newer approach, which falls under the blended learning umbrella, is PL. A consensus on the definition of PL is not reached in the literature, but most research points to some aspect of customization, student groupings, and flexibility of instruction (Berry, 2018). The implementation of PL practices in American schools has increased significantly over the past several years (Pane et al., 2015) however, the variances in implementation range drastically from organization to organization (Staker, 2012) resulting in little research on the effect PL has on mathematics achievement.

When PL models utilize online content delivery, typically used to customize the learning for the student at their functional level, PL is one way to implement blended learning. While PL can be thought of as an instructional approach that focuses on the individual needs of students, it has often been implemented in such ways that either focus solely on the academic achievement of students or the interests and social impacts on students. When PL classrooms are inclusive of students' interests and their personal needs and goals academically, such programs have the potential to increase student learning and engagement (Pane et al., 2017).

The National Council of Teachers of Mathematics (NCTM) promotes eight practices that are considered to be best practices in mathematics which were summarized in the *Personalized Learning and Mathematics Teaching and Learning* publication:

> "Establish mathematics goals to focus learning, implement tasks that promote reasoning and problem solving, use and connect mathematics representations, facilitate meaningful mathematics discourse, pose purposeful questions, build procedural fluency from conceptual understanding, support productive struggle in learning mathematics, and elicit and use evidence of student thinking. (Berry, 2018, para. 5)"

When reviewing and implementing PL practices into the classroom, NCTM encourages educators to ask questions around each of the eight practices and how they are addressed in the PL model (Berry, 2018).

Background

The setting is an open-enrollment public charter school in Texas made up of two K-12 campuses and one 1-12 campus. The charter was written as a funding mechanism to support a laboratory school for a university and it is modeled after the 2015 T-STEM Academy blueprint (Texas High School Project, 2015). The district implements PBL, PrBL and BL as the primary methods of instruction, with occasional opportunities for PhBL scenarios. PBL has been identified as an instructional model used to improve the achievement of students in STEM classes (Odell et al., 2019). Students choose either the PLTW Engineering or Biomedical pathway for their chosen designation in grades 9-12. Students also take dual credit classes beginning their ninth-grade year and are able to graduate with 42+ hours of university credit.

Even though the district is considered high performing currently, it has not always been the case. There have been three major improvement cycles in the area of mathematics that have led to the current performance rating. The remainder of this chapter will focus on three distinct problems of practice; the need for overall improvement in mathematics achievement, the need to close the gaps, and the need for individual student progress. Each problem of practice will be addressed with specific components of PDSA cycles.

PDSA Improvement Cycle 1

Prior to the opening of the laboratory school, little was done curriculum-wise in terms of foundational systems to align the curriculum, instruction, and assessment. Instead, teachers completed six full weeks of professional learning that focused on best practices of PBL, the main instructional approach of the laboratory school. PBL was the only method of instruction, and the only foundational system in place. Teachers created their own PBLs, with resources they found. Standards were taught in any order teachers deemed instructionally appropriate. There were minimal checks and balances for monitoring that all standards were taught. Practice assessments were not given prior to the state assessments. At the time, it was the belief that teachers could plan the projects without a mandated scope and sequences in place and without any standardized assessments to test for mastery other than authentic assessment products such as PBL presentations or other products.

The theory in the early years was that students could be successful without testing and that projects could drive learning even if they were designed without a sequence of standards in mind. One of the goals was to provide teachers flexibility to collaborate and integrate content from different disciplines. A standardized scope and sequence would minimize opportunities for collaboration and interdisciplinary projects. Projects were fun and engaging for students and teachers were quoted as saying, “it was the most fun year I’ve had in my career.” However, projects at this time could be described more as interest projects and upon reflection, they were not tightly aligned to standards.

Mathematics is tested each year in Texas starting in grade 3. There is also a high school end of course exam in Algebra required for graduation. As one might imagine, the academic results on state assessments in year one were not only poor but, landed the charter in the bottom five percent of the state.

The problem of practice for PDSA Cycle 1 was clear, the need for overall improvement in mathematics achievement according to state assessment results. Academic achievement is Domain 1 of the state accountability system. Once the results came back and were analyzed internally, we realized that our students were well below the state average in every grade level in mathematics. Even worse, students who had transferred to us with previous test scores from the prior year, had dramatically declined. The proposed intervention by the NIC was better curriculum alignment to the state standards and the state assessments. Planning and implementing aligned instruction can be difficult when using inquiry instructional strategies.

The first improvement science cycle spanned over a period of two years. Its important to note that this is longer than a typical improvement cycle. However, its technically made up of two major change ideas, each one needing its own time and space to be planned, studied, implemented and tweaked. We believed it was necessary to start with one change idea that would eventually roll into a larger one in order to make one overall PDSA cycle rather than focusing on both right away. This was done intentionally considering how much needed to be done but trying to avoid overwhelming the teachers in the process which could ultimately lead to set backs instead of gain. At this point in the school's development, structures from the college of education and the laboratory schools were not closely aligned.

The primary drivers of the low-test scores in mathematics were the lack of a scope and sequence, the lack of assessments, standards not being tightly aligned to PBLs and overall alignment of curriculum, instruction and assessment. Along with identifying the problem, and root causes, users (teachers described in the case studies reviewed in this chapter) needed to be engaged in instructional decisions made thereafter. With the data in mathematics being drastically lower than the state average, overall systems had to be put into place for year two. Teachers were consulted and came together to make a plan for year two. The intervention would be simple; alignment of curriculum, instruction, and assessment, however, our questions was what would that need to entail since nothing was in place? The co-founding faculty members of the laboratory schools and school leadership developed an advisory structure to better support the school. This advisory structure would eventually become the foundation for the NIC that exists presently.

Change ideas were identified by a team of stakeholders which resulted in a scope and sequences for each grade level, revision of PBLs to align the standards, creation of standards-based classroom assessments, the use of a district PBL coach, and PBL content rubrics based on standards.

First, a team of teachers created scope and sequences and then revised PBLs from year one as a team to tightly align them to standards. This same team of teachers created post-tests for the PBLs (remember no tests were given in year one). Second, a coach was identified to support mathematics for all schools districtwide that would travel to each of the three campuses, observe teachers, give feedback to teachers, co-teach lessons, and lead PLCs in the afternoons, embedded into the workday. Lastly, content rubrics were implemented for each PBL lesson which were based solely on the standards included in the PBL. Thus, PBLs were now directly aligned to standards.

The changes were studied over the course of year, with many revisions to the PBL model based on new scope and sequences, instruction of the PBLs based on feedback from observations, and different versions of the content rubrics to adequately assess student knowledge through PBL products. By the end of year two, the laboratory schools gained eleven percentage points in the area of mathematics on the state assessment. We knew we were on the right track after analyzing the data since we saw improvements, but we knew more improvements were needed. Even though the district had seen gains in mathematics, scores were still below the state average, which told us our focus still, needed to be alignment of state standards to state assessments.

Year two of PDSA Cycle 1 continued to focus on the primary intervention of alignment but as the year progressed, additional drivers were identified through the PDSA process such as the need for teachers to receive timely and relevant feedback directly tied to the content. In response to primary driver, a mathematics content coach was hired for the district with the main responsibility of supporting teachers by writing model PBL lessons with and for teachers, coaching mathematics specific strategies in the classroom and modeling mathematics instruction for teachers while giving feedback and helping pull resources. Two extra interventions in year three included the addition of district benchmark tests aligned to the state assessments and the purchase of a mathematics textbook (the first district adopted resource in mathematics).

By the end of the third year of the school (second year of PDSA Cycle 1), the laboratory schools again saw improvements in overall performance in mathematics by increasing 36 percentage points. For the first time, the district mathematics score had exceeded the state average.

The main takeaway from the early years is that there is an incredible need for alignment between curriculum, instruction, and assessment. This may seem obvious but keep in mind there were not readily available PBL curricula available for implementation. PBL and PrBL Inquiry-based lessons are primarily developed and implemented by teachers to this day. To refine the model, in the context of PBL instruction, it is necessary to provide:

1. a written scope and sequence and mathematics resources for teachers;
2. instructional feedback that is aligned to the curriculum, and
3. formative assessments based on the curriculum and state requirements.

Table 1 provides annual data by PDSA Cycle. It should be noted that the trend in scores has been increasing annually and through each PDSA cycle. State averages have remained flat over time.

Table 3

Mathematics Achievement by Intervention Cycle

	PDSA 1				PDSA 2		PDSA 3		
	2013	2014	2015	2016	2017	2018	2019	2020	2021
State	79	78	81	76	79	81	81	82	N/A
District	48	59	95**	83	86	88	88	91	N/A

**New TEKS were implemented and tested this year. Fewer TEKS were tested and Special Education tests were not included in scores.

PDSA Improvement Cycle 2

At the conclusion of Cycle 1, mathematics scores were trending in the positive direction. Cycle 2 is characterized by the NIC and school personnel focusing on a related but finer grained problem of practice. The identified aim for Cycle 2 was focused on equity in mathematics achievement and the improvement of mathematics achievement for all students. The problem of practice focused on closing the achievement gaps, Domain 3 of the state accountability system, between sub-populations of students including race, socio-economic-status, special populations, and language. The proposed interventions included:

- Intervention 1: Data Tracking
- Intervention 2: PrBL and Mathematics Best Practices

Once improvements had been achieved in overall student performance, a new challenge became increasing equity and thus ensuring success in the student sub-populations of the district. Clearly, gains had been made in mathematics achievement overall, but the state assessment data revealed significant achievement gaps by sub-groups when compared to the overall score. For example, data for the end of year three showed there was a ten-percentage gap or higher when comparing the progress that Hispanic students made from the year before to the overall students, and in some cases a twenty-percentage gap when comparing African American students to overall. In this second PDSA iteration cycle, two interventions were implemented simultaneously. This improvement cycle, which centered on closing the gaps, required a four-year span with numerous adjustments to complete.

The first intervention was grounded in assessments and data tracking. As previously described, district benchmarks were implemented at the end of cycle one. However, data at that point was simply being tracked by overall student performance (percentage of each grade level who was having success on the state exams). During cycle two, a plan was made to track data by subgroups to measure the gaps by demographics. Stakeholders agreed that there needed to be a common system for tracking this data and thus a primary driver became district spreadsheets. For the first two years in this cycle, spreadsheets were used to track benchmark data and compare it to the state assessment data. It is important to note that this data was held at the leadership level. Occasionally, teachers were asked to review their data with the instructional content coach, but this was not a practice that was utilized often.

During this two-year span, some gaps were starting to narrow but not significantly. Based on the "study" component of the PDSA plan, a revision was required. The plan needed to be revised to keep narrowing the gaps. The district that started almost six years ago looking at no data at all, had come to the realization that they didn't have enough data or a robust system to manage data.

A second driver was the introduction of common district assessments, which would be administered at the end of each nine weeks. These assessments were developed by the Director of Curriculum, in conjunction with the mathematics instructional coach. The assessments would be modeled on the state assessments but would include more open-ended questions to assess deeper understanding. The purpose behind these types of assessments was to increase the rigor of the instruction in the classroom.

In her President's Message for NCTM, Linda Gojak shared, "Rigorous teaching and learning require rigorous formative assessment throughout a unit so the teacher knows what the student has learned and can plan additional activities, or adjust them, to address student needs" (Gojak, 2013, para. 7). By increasing the rigor of assessments at the end of each quarter, teachers would need to increase the rigor in the classroom in order for students to be successful. The data at this point had variations by classrooms, where some classrooms were proving to close gaps more quickly, others were not having the same success. One major observation and takeaway was that the classrooms experiencing greater rates of change in their data were classrooms where the level of instruction was higher in rigor and teachers were paying close attention to their data.

As a result, a final primary driver that was addressed under this category was more frequent data meetings and teacher empowerment. Teachers were now asked to keep their own data spreadsheets, which would include their classroom assessments, common district assessments, benchmarks, and state assessment data. Data meetings were called once per quarter, where members of the curriculum team would work with teachers to help them analyze their data while paying attention to their subgroups. As Bryk et al. (2017, p. 87) describes as one of the main principles of improvement science in the fourth chapter of *Learning to Improve*, “we cannot improve at scale what we cannot measure.”

The second intervention during this cycle, was a significant change to the mathematics instructional model and curriculum. Even though there were improvements to the overall mathematics scores, when examining the equity gaps related to progress and keeping the observational data in mind mentioned before, stakeholders attributed some of the gaps to the rigor and variances in mathematics instruction.

Two major drivers under this intervention were (1) mathematics instruction switched from PBL to PrBL, due to better alignment with mathematics inquiry, and (2) a larger range of mathematics resources were implemented to create a more well-rounded mathematics classroom.

After three years of PBL in the mathematics classrooms, district personnel recognized that mathematics was always an afterthought when it came to the planning of the PBLs. All PBL lessons up to that point had been interdisciplinary and upon analysis, mathematics was rarely the driving discipline for the PBL.

There were many elements of best practices in mathematics that were missing from the curriculum and from instruction in general. With inquiry-based learning still being the foundational model of the charter, the switch to problem-based learning was made. Problems were introduced at the beginning of units and used to drive the learning of the standards throughout. PrBL also doesn’t take as long to implement as PBL, therefore, opportunities emerged for supporting best practices in mathematics to be included in the classrooms. Supporting instructional practices were included in the mathematics instructional model such as spiral reviews and skills and drills practice, while other best practices like mathematics discussions and questioning were left in place. New resources were acquired, and teachers were provided intensive professional development on how to use each one in relation to each best practice. It is important to note that the role of the content coach in the organization of materials and training of teachers was instrumental in these interventions leading to eventual success of closing the gaps.

By the end of improvement cycle two, the laboratory school saw the gaps narrow for subgroups as seen in the state assessment data for Domain 3, closing the gaps. The gap for Hispanic students when compared to all students was narrowed to two percentage points and the gap for African American students narrowed to seven percentage points in terms of students progressing. Even though both interventions were done simultaneously, each year there were adjustments to the existing drivers and/or additional drivers added. The data used to inform these decisions were based on studying the intervention through observations, testing data both in-house and at the state level, as well as feedback from the users (teachers).

Table 2 provides data from 2013 to 2019 in terms of student progress as determined by the state. Student progress is measured from year to year with predetermined growth rates. In the table, Lim% is the percent of students who showed limited growth from the prior year and results in the state classifying the student as having zero growth. Exp% is the percent of students who had the set expected growth from the prior year and Acc% is the percent of students who showed accelerated growth (more than the standard growth rate that was expected) from the prior year. The table is also broken down into sub-populations in order to see how the gaps were closed in student progress from 2013 to 2019.

Table 4

PDSA Cycle 2 Closing the Gaps in Student Progress

	Limited		Expected		Accelerated	
Subgroup	2013	2019	2013	2019	2013	2019
All Students	62	35	5	38	2	15
Hispanic	83	38	6	36	0	14
Asian	20	0	0	31	20	31
African American	83	49	0	29	0	14

Table 3 depicts the change in mathematical student progress by sub-populations from 2013 to 2019. The students' progress by sub-groups grew at close to the same rate over the course of the seven years and the change in the percent of students making accelerated progress was almost the same for each sub-group. This table, however, is a preview to Improvement

Cycle 3 which will be centered on student progress alone. As the data below shows, there is much need for improvement in overall student progress for all sub-populations.

Table 5

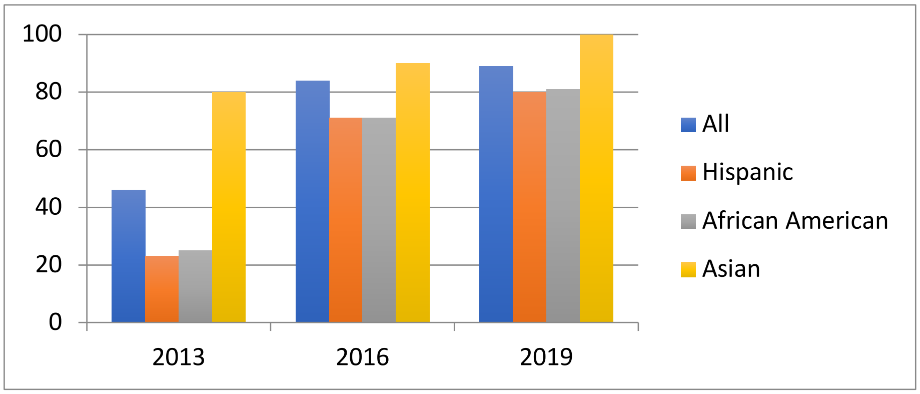
PDSA Cycle 2 Change in Percent in Student Progress from 2013 to 2019

Subgroup	Limited	Expected	Accelerated
All Students	-27	33	13
Hispanic	-45	30	14
Asian	-20	31	11
African American	-34	29	14

Figure 1 is a comparison chart of the academic achievement in math by sub-populations from the end of year one to the most recent accountability data. The chart also displays the end of the first year of Cycle 2 (2016) to the end of Cycle 2 (2019). It is important to consider how these data relate to data from the state. While the district began with a twenty-one point gap from all students to lowest sub-pop, narrowed it to a thirteen point gap in 2016 (which is the current average gap size in the state according to 2019 data reports), and closed it even tighter to an eight point gap by the end of 2019.

Figure 4

PDSA Cycle 2 Academic Achievement by Sub-Groups



PDSA Improvement Cycle 3

Purpose

As noted, the third improvement science cycle focused on the problem of practice of individual students making progress from year to year. As mentioned in the beginning of the chapter, students not meeting progress in mathematics was a problem for the laboratory school as evaluated by the state assessment in 2019.

The purpose of this study is to evaluate the effects of four components of a BL model (data driven decisions, student reflection, targeted instruction and integrated technology) integrated into one PL model on the mathematical achievement of fourth-grade students. The study will seek to answer the following questions: (1) How have students' mathematical functional levels changed?, (2) How have students progressed as defined by a state assessment, State of Texas Assessment of Academic Readiness (STAAR)?, and (3) How have teachers' views of their success in the classroom changed?

Setting

The district had an enrollment of approximately 827 students and employed 57 teachers. The student population was 66.5% White, 17% Hispanic, 6.3% African American, 4.7% Asian, 49.3% female, and 50.7% male. The district was also 38% economically disadvantaged across the three campuses, with 6.4% special education population and 10.3% Section 504 students. The teachers have various backgrounds ranging from novice probationary teachers to veteran teachers with thirty plus years of experience. Each campus has one Director (much like a principal) and one instructional coach whose main role is to help assist in the implementation of the instructional model as outlined in the charter.

The district is considered a high performing district in the state, earning an overall rating of an “A” on the 2019 accountability ratings, a rating only given to ten percent of districts. However, the district received a “B” rating in the domain related to student progress. The rating in student progress was largely related to math. To address the lack of academic growth in math, the Math Innovation Zones (MIZ) program was put into place for the 2019-2020 school year. The MIZ program is a BL grant awarded by the state with the purpose of improving mathematics achievement in students. The district chose to implement the MIZ program as a PL model. The math PL model was designed by stakeholders, which included the Director of Curriculum, three campus instructional coaches, and teacher representatives from each campus. The design team created the PL model to include four main components: data driven decisions, student reflection, targeted instruction and integrated technology.

Participants/Demographics

For the purpose of this evaluation, criterion sampling was used in order to evaluate the effects the PL model had on all fourth-grade students in terms of growth. The district’s three fourth-grade classrooms, one from each campus, will be studied. In total, these classrooms include 63 students and three teachers. Each teacher is in his or her second year of teaching under the PL model and total years of experience range from three to ten years. Students are demographically represented 25% Hispanic, 6% African American, 55% White, and 35% economically disadvantaged. The students are almost an even split between male and female, with 52% being male and 48% being female. The evaluation will be the study of the implementation of the PL model by the three classroom teachers and the effect it has on student outcomes.

Personalized Learning Model

The mathematics PL model was designed by stakeholders, which included the Director of Curriculum, three campus instructional coaches, and teacher representatives from each campus. The design team created the PL model to include four primary drivers:

1. data driven decisions,
2. student reflection,
3. targeted instruction, and
4. integrated technology.

Stakeholders theorize that if students are provided a PL model in mathematics designed to meet students on their functional level and address their individual mathematics needs, then they will achieve academic growth from one grade level to the next and ultimately meet progress on the state assessment.

To support this intervention, an instructional coach is present on each campus to support the teachers to implement PL, a new instructional approach. Each teacher was placed at least on a tier two in terms of support level for year one. Teachers are placed in tiers by the Director of

Curriculum and the Campus Instructional Coaches and reviewed once a quarter throughout the year. Teachers are tiered in levels one through three where teachers in tier one are teachers who can implement the instructional model with fidelity and are experts in their respective contents. Teachers in tier two are placed in this tier because they typically have one or areas to work more in-depth on which could include either the instructional model or their content. Teachers in tier three typically have areas of growth in more than one area such as the instructional model, content, or classroom management.

It is important to note that due to Covid-19, there was a disruption for a full year of implementation of the PL model in the 2019-2020 school year. There was also limited data available beginning the 2020-2021 school year since 2020 state assessments were cancelled. The laboratory schools, like many other schools, were forced to implement remote learning in the spring of 2020. However, students and teachers were able to seamlessly transition from face-to-face blended learning to fully remote learning which can be credited to the limited initial experience with the PL model.

For the 2020-2021 school year, teachers began the implementation of the PL model for the second time. Implementation began in August and students were taught under this model for the length of the school year. Coaches observed and met with teachers over the course of the school year. Data was collected periodically throughout the year and the evaluation of results were determined after the school year ended.

Theory of Change

The software programs and learning management systems were already in place to support students remotely. Students and teachers already had enough experience with the platforms to transition completely to an online environment. The theory behind the PL model is based on three guiding principles:

1. Students met on their functional level will fill gaps more quickly.
2. Teachers making data driven decisions (in short data cycles) to inform their instruction will help each student progress.
3. Students taking ownership in their own learning leads to academic growth.

Research Methodology

A mixed methods research design was identified as the best way to evaluate the mathematics PL program. Mixing methods can offer insight, complement one another, and offer further questions for future research opportunities (Caruth, 2013). A mixed method study was necessary in this evaluation since different evaluation questions require different methods. The two types of data will serve to answer different questions in order to evaluate the math PL program over the course of a school year which consists of a nine-month period. Quantitative data will include the functional level scores from the software program used by students, previous STAAR scores and final STAAR scores for the current year. The quantitative data will be used to track the students' functional level as well as the students' progress over the course of the year. Qualitative data will include teacher surveys and classroom observations. This data will assist in understanding the degree to which the components of the PL model are implemented and the views from the classroom teacher in regard to the model, as well as their perceived level of success.

The improvement science framework in education, although a relatively new methodology, is a natural fit for practitioners who are often engaging in improvement science efforts on their own daily by gathering data and distributing new ideas (Carnegie Foundation for the Advancement of Teaching, 2015). Improvement science is rooted in the work of Deming (1993) and his idea of Profound Knowledge which helped organizations organize their thoughts and approaches to improve (Perry et al., 2020). The methodology of PDSA cycles, based on the work of Langley et al. (2009), were used in this evaluation to test the theory of the PL model in an effort to understand what worked (and for who), under what circumstances, and why (Bryk, 2018).

Research Design

The embedded mixed method design was chosen for numerous reasons but the main one being that the quantitative data is the driving force behind the research and the root of the problem of practice. However, the quantitative data is not sufficient alone and therefore needs the support of qualitative data to see the entire picture. For this reason, the embedded experimental design was chosen to fully evaluate the PL model. A one phase embedded experimental design was chosen which is also referred to as a concurrent nested mixed methods design (Creswell et al., 2003). The significance of the one phase embedded experimental design is that the qualitative data is embedded within the intervention period along with the quantitative data. The qualitative data is needed in addition to the quantitative data to examine the intervention as a whole. A quantitative pre and post measure are given before and after the intervention. Interpretations of the results of the evaluation are then made using primarily the quantitative data with the support and storytelling of the qualitative data (Creswell & Plano Clark, 2010).

For the quantitative design of this study, a quasi-experimental design was used to attempt to establish a cause and effect relationship between the PL model implemented and later with an intervention on the outcome of student achievement and progress scores on the state assessment as well as their individual functional levels. Conclusions about any such cause and effect relationship will be carefully considered due to any other variables outside of the research that could possibly account for any of the dependent variable outcomes.

The qualitative design used for this study is phenomenological since existing observational data as well as perception surveys were given to understand the experiences teachers had during the study and intervention. The meaning that teachers tied to their feeling of success during the study as well as the intervention was of particular importance to the study to determine what changes, if any, need to be made to the PL model as the "users" are the center of any improvement science research. The insight gained from the teacher's attitudes about the PL model is significant information for the overall study and can be used to continue the improvement science in this setting.

Data Collection

All quantitative data will be stored in the teachers' data spreadsheets and collected throughout the year. Teachers are responsible for inputting the data into the spreadsheets and the teacher and campus instructional coach will meet twice quarterly to review the data. IXL functional levels are placed in the spreadsheets as soon as the diagnostic is completed within the first few weeks of the school year. IXL is a personalized software program that calculates the students' functional levels. A student's functional level can be different from the student's actual grade level. For example, a student can be in the fifth grade, third month of the school year, which is denoted as 530. For the student to be considered on grade level, the student's functional level score would need to match the student's grade level and corresponding month of the school year when the functional level as taken. A higher functional level would indicate the student is above grade level and a lower functional level indicates that the student is below grade level. The

IXL functional level is updated at the beginning of each quarter for the remainder of the year and tracked to monitor student progression. Data from the state assessment will be collected to assess the long-term goals for the program once the scores are released.

Qualitative data will be collected throughout the school year as well. Teacher surveys will be administered twice during the year. Each of the surveys will contain identical questions and format. The surveys will include closed and open-ended questions related to the use of student goal setting and student data tracking. The surveys will also include questions related to the affective factors impacting how teachers structure support and empowerment in PL with relation to the success of their students.

Teacher observations will be conducted throughout the year by campus instructional coaches and will be reviewed with the Director of Curriculum. Coaches use a standardized district created PL observation form. The purpose of the observations is to collect data on the use of stations, the implementation of the PL model, the self-directedness and engagement of students, flexible groupings of students, teachers helping students set goals and students tracking their data. Observations will be done on a bi-weekly basis throughout the year.

Data Analysis

IRB approval was obtained prior to the collection of data for this study even though this study is one that would have been done anyway, as it is tied to a grant given by the state of Texas. Teacher surveys are anonymous and collected using a Qualtrics account owned by the Director of Curriculum. Reports of the survey are distributed to the instructional coaches and reviewed in curriculum meetings following each administration. Results are analyzed for themes in the reflections of the teachers as well as the frequency of the responses. The surveys provide useful information for each of the components of the model as well as the perceptions of the teacher.

Teacher observations made by instructional coaches are shared with the teacher and with the Director of Curriculum only and are common practice in the district regardless of this study. Instructional coaches do not see the observations made by other coaches. Observations are analyzed and coded into themes by the Director of Curriculum without any identifiers noted. Quotes and anecdotal notes are abstracted from observation forms and field notes into separate documents. Once all notes are organized by themes in separate documents, it is unidentifiable as to which teacher it belongs to.

Member checking is involved in each of the data collection methods. As the Director of Curriculum, this is an important part of the process in order to make sure my own preconceived notions do not interfere with analysis. In my role, it is imperative to not make generalizations about all classrooms if I am only seeing evidence of a certain theme in one. By each member of the curriculum team individually analyzing data first, then discussing in curriculum meetings to safeguard accuracy and consistency, the team is able to ensure data is being portrayed accurately.

Data Analysis Limitations

As previously mentioned, the study had many issues going into the 2020-2021 school year due to the lasting impacts of Covid-19. The study was designed to be able to fully evaluate the PL model used in the fourth-grade classrooms across the district. However, shortly after the evaluation plan was written, it was determined that the full study would not be able to take place due to many factors impacting teachers' ability to be able to implement the model to any form of fidelity.

First, students across the state were able to choose to stay home and attend school remotely or come to school in person attendance in response to the still unknowns of the Covid-19 virus and the possible rise in cases once schools started again. Due to the size of our school, teachers were responsible for teaching all of their classes face to face as well as remotely to students who were at home and attending virtually through a Zoom platform or some other remote meeting software. This type of instruction proved to be very challenging for teachers and it was soon noted through observations as well as meetings with instructional coaches that teachers were not able to implement the PL model as designed. Teachers at this time were in "survival mode" and relied on whatever methods they felt best to just get through the content. Qualitative analysis and discussion will be extracted from observation forms, DDI and MLE instead of surveys. Due to the nature of the study and the effects of Covid-19, surveys were not given during the 2020-2021 school year.

Second, since the quantitative data was largely dependent on the progress of fourth grade students on the STAAR test, this proved to be faulty as well. Shortly after the school year began in 2020-2021, the state of Texas decided that progress would not be measured for the state accountability. Students, however, did utilize the IXL software in each of the fourth-grade classrooms which will serve to answer the first research question in regard to functional level. The second research question is not able to be evaluated, however, students did take a pre-assessment for the 2020-2021 school year which was provided by the state. This assessment was designed to mirror the STAAR test students would have taken the prior year. In other words, students in the fourth grade took a beginning of the year assessment which mirrored the STAAR test they would have taken in the previous spring semester (2020) in the third grade if schools had not been shut down due to COVID-19. The results of this pre-assessment and the results of the 2021 STAAR will be used as an alternative data set for this study.

Results

Qualitative

Observations from coaches, DDI forms, and MLE forms were sorted by semester one and semester two. Semester one data included all observation forms, DDI forms, and MLE forms available. At the end of semester one, these forms were extracted into one major document and analyzed by open coding. This method was performed a second time at the conclusion of semester two with the same available documents as before. Once open coding had been done on all the forms at the conclusion of each semester, the open codes were then compared for similarities between the two. The codes were then categorized into major themes. Early analysis of the qualitative data sources leads to three major themes emerging from the observations, DDI, and MLE.

The first scenario that is worthy of discussion is the switch from larger data cycles to shorter ones which enables the ability to use targeted instruction more frequently. According to observational data, teachers are using data from a variety of sources (software, formative assessments and summative assessments) in order to pull small groups and target students' immediate needs. One instructional coach noted in an observation, "teacher had a small group station where they were working individually and with student groups to target instruction." Another direct quote from an observation was, "some students were on the computer, some were working with the teacher in a small group based on recent data." These quotes are evidence that teachers, regardless of their current level of capacity of the concept, are embracing small group instruction based on student needs. The DDI forms also pointed to the switch of targeted instruction from the beginning of the year to the end. Early forms were largely based on data cycles referencing benchmark data which are long cycles of data. However, more recent DDI forms frequently referenced weekly to even daily classroom data which indicates that the data cycles have indeed been shortened. Teachers expressed in their weekly reflection comments such as, “Students are working at their functional level, so they are able to be successful”, and “students are eager to participate in the small groups I lead”.

The second major theme to materialize, which is tied to the research question of student progress, was the goal setting methods used by the teachers. Goal setting was present in each of the classrooms according to data, which is a positive factor considering it is one of the core components of the PL model. Goal setting as a standard component of the PL model was to be based on students’ academic data from the standards and visited frequently in order to reflect and update as goals were mastered. However, the observations indicate that this wasn’t necessarily the case. First of all, it seems that students were writing goals around personal behaviors instead of content driven goals. Comments such as, “I see students writing goals around personal goals” were found in observations made by instructional coaches.

A third theme found is the goal setting frequency. Teachers reflected in ways such as, “An area of growth would be tracking the students’ data regularly. We are moving so much slower that I cannot see growth from Pre- to Post assessments in a timely manner. Students do not see the connection throughout the unit. There is a disconnect.” Another teacher reflected that her class had “discussed goal setting and set an overall goal for the year.”

Quantitative

The first data set is the functional level IXL data for the fourth-grade students. Although there are a total of 63 students among the district in fourth grade, the sample size for the IXL data was cut to 54 to only include students who had a beginning of the year functional level score as well as a corresponding end of the year functional level score. The functional level scores were captured at the beginning of quarter one (August 2020) and placed in the student progress monitoring spreadsheets, then captured again close to the end of quarter 4 (May 2021). All scores were abstracted from the spreadsheets and placed into a master copy of IXL data for the 2020-2021 school year. BOY was used to name the quarter one functional level score and EOY was used to name the quarter four functional level score.

A paired-samples t-test was used to compare students' mathematical functional level score before and after exposure to the IXL program through the personalized learning model. Prior to data analysis, the primary assumptions of the analytic procedure were checked. The assumption of normality was not violated as assessed by Shapiro-Wilk's test ($W = 0.96$, $p = .11$). The results indicated students' functional level increased from quarter one ($M = 392.78$, $SD = 56.32$) to quarter four ($M = 451.67$, $SD = 57.55$). This improvement was statistically significant with $t(53) = -9.47$, $p < .001$. The effect size for this analysis ($d = - 1.29$) fell above Cohen's (1992) convention for a large effect. See Table 4.

Table 6

Results of IXL Data from Quarter 1 to Quarter 4

BOY		EOY		t(53)	*p*	Cohen's *d*
M	SD	M	SD			
392.78	56.32	451.67	57.55	-9.47	<.001	-1.29

A one-way ANOVA was done to determine if IXL growth differed among based on the location; campus one, campus two, and campus three. A check for the assumptions of the analytic procedure were done and found all assumptions were met. Results indicated there was not a significant difference in IXL scores based on the campus, with F (2, 51) = 2120.33 p = .55.

Effect size estimates suggest approximately 2% of the variance in the functional level scores was explained by the campus ($n^2 = .02$). See Table 5 for the descriptive.

Table 7

Difference of Growth Scores by Campus

Campus	Mean	SD	N
1	58.24	49.15	17
2	67.78	59.37	18
3	51.05	23.07	19

The second data set is the STAAR pre-assessment data for the fourth-grade students and the actual STAAR data. Although there are a total of 63 students among the district in fourth grade, the sample size for the pre-assessment to STAAR was cut to 40 to only include students who had a score for both assessments. The STAAR pre-assessment data were captured at the beginning of quarter one (August 2020) and placed in the student progress monitoring spreadsheets and the STAAR data was captured and placed in the spreadsheets at the end of the 2020-2021 school year (June 2021). All scores were abstracted from the spreadsheets and placed into a master copy of STAAR data for the 2020-2021 school year. BOY STAAR was used to name the pre-assessment and STAAR '21 was used to name the STAAR scores for the 2020-2021 school year.

A paired-samples t-test was used to compare students' STAAR scores before and after exposure to the personalized learning model. Prior to data analysis, the primary assumptions of the analytic procedure were checked. The assumption of normality was not violated as assessed by Shapiro-Wilk's test ($W = 0.98$, $p = .51$). The results indicated STAAR scores increased from the pre-assessment ($M = 60.8$, $SD = 519.19$) to STAAR ($M = 69.75$, $SD = 16.76$). This improvement was statistically significant with $t(39) = -3.32$, $p = .002$. The effect size for this analysis (d = - 5.24) fell above Cohen's (1992) convention for a large effect.

Table 8

Results of BOY STAAR to STAAR '21

BOY STAAR		STAAR '21		t(39)	*p*	Cohen's *d*
M	SD	M	SD			
60.8	19.19	69.75	16.76	-3.32	.002	-.52

Discussion

There are perhaps a few conclusions this data, both quantitative and qualitative can tell us. First, while the IXL data did prove to increase the students' functional level scores from the BOY to the EOY, the extent to which they increased is perhaps less than desirable. To answer the second research question, "to what extent" would be to state it in terms of grade level and months. The students' functional level scores increased on average from a third-grade month nine to ta fourth-grade month five. This is on average a six-month gain. I would like to think with the effects of Covid-19 and the varying settings for students to learn this school year that this six-month gain is an adequate gain, but still falls short of long-term goals. Another key takeaway in terms of the IXL data, is that regardless of campus, the gains seemed to be about the same with no real variances among the campuses. This is an ideal situation seeing that the same model was expected on each campus to be implemented and evaluated. It is important to note that the qualitative data reveals that the functional level being made known to the teacher could have also played a role in the increases we see in the quantitative results. Throughout the observations as well as the DDI forms, teachers were noted to be using students' functional levels to make decisions in regard to targeted instruction and small group pull-outs, a practice that had not been in place before the PL model.

As data driven decisions and targeted instruction are two of the four components of the PL model, it is a good sign that early indicators prove to show that they are being implemented more frequently and teachers are having early success with these instructional strategies. The future impacts of this strategy are extremely beneficial in regard to the overall goal of the PL model, which is to have students progressing from year to year. As pointed to in the literature earlier, some PL models choose to have the software programs take care of not only the personalization for the student but the targeted instruction as well. I found it surprising, in a good way, that the teachers in this model were going a step further and tailoring their instruction as well. Even though targeted instruction was a core component, it could have been very easy for the teachers to teach as they had before with the only adjustment being the software expectations. The fact that the teachers wanted to be just as much a part of the targeting of their student needs is a good sign of the future impacts of this model. The impact of targeted instruction on the future of the campus is a large one in that if the belief and assumptions of this model prove to be correct, students will progress and ultimately be more successful in mathematics. This strategy could also have potential impact on other subject areas in the future.

Secondly, the BOY STAAR to the STAAR ’21 results revealed that on average students increased from 60% to 69%. In 2021, an average of 69% means that the students fell just short of meeting grade level expectations. In order to approach grade level, a student needed to make 50%, and in order for a student to meet grade level, a score of 71% was needed. As mentioned before, we are unable to accurately answer research question three due to the no progress measure in 2021 and the lack of STAAR scores in 2020, however, using the BOY STAAR to STAAR ’21 does give us some insight to the students’ progress over all. If the BOY STAAR is used to give some indicator of what students would have made in 2020 on STAAR (minus the

Covid-19 and summer slide) then the students would have been considered in the medium range of approaches at 60%. Approaches ranged from 50% to 70%. By the STAAR ’21, students ended in the high approaches at 69%. This leads me to believe most students would have made progress had there been measures in place to gather this data.

It’s interesting to consider what the STAAR quantitative data tells us about the goal setting qualitative data. As stated earlier, we know from the qualitative data that students for the first time were setting goals in their math classrooms, but most goals were centered around personal initiatives rather than academic ones. While goal setting was present in each of the classrooms, surprisingly how teachers had students set goals and the content in which they were set varied dramatically among the classrooms and from the intention of the component. It does beg the question, if goals had been set consistently on how students were performing on assessments in respect to progress on STAAR, how the quantitative data might differ, if at all. Early analysis perhaps indicates that goal setting, which is a new initiative to the district, might take a longer time period in order for students to make the switch to reflecting on academic goals. It could be that personal goals come more naturally than goals based on the standards.

Conclusion

The laboratory schools have institutionalized improvement science principles and tools as part of their structure and operations. The NIC that helps guide the work of the laboratory schools leverages resources and expertise of the College of Education and Psychology, the School of Education, and its related research centers focused on STEM Education and School Improvement.

The iterative improvement cycles as described in this chapter provide evidence that implementing improvement science principles and tools to develop a learning culture can lead to continuous improvement and closing achievement gaps among historically underserved populations as measured by state assessments in the area of mathematics. As a laboratory school, our charge from day one was to be a model school for our surrounding districts and eventually a larger region. Closing the gaps was imperative for us in order to prove that our model works for all students, not just a select few. Our focus on continuous improvement using principles of Improvement Science has resulted in the district moving from one of the lowest achieving districts to one the top achieving districts in a relatively short time-period.

Recommendations

Due to the limitations of Covid-19, the third iterative cycle (PL model), needs further evaluation but perhaps with some narrowed focus. Due to the results of the IXL, STAAR, observations, and DDI, some themes emerged that could help narrow this focus. First, IXL needs to be continued into the next school year as the software used to diagnose and reach students at their functional level. Teachers can continue to use it as they did in the 2020-2021 school year such as a station or set block of time in the class period, but with all students face to face, it will be interesting to see if further gains are made in relation to the overall mean of students' functional level scores.

Secondly, teachers need to continue to focus on students' functional level score and use this score to target student needs as well as provide small group instruction. It seems that this strategy proved to be beneficial in the 2020-2021 school year and with intentional focus, as well as all students face to face, this one strategy could further increase math scores in the district. IXL data can be used to pull small groups in levels as seen in reading level groups in reading classrooms. IXL data can also be used to re-teach concepts that the class might be struggling with as a whole or to reach students individually in a tangible way through conferences and tutorials.

Third, goal setting is an area that needs a great deal of work in order to truly evaluate whether this method, as research suggests, can have a big impact on student success. From the 2020-2021 school year, all we really know is that goals were set but the nature of the goals and the frequency varied widely amongst classrooms and even individual students. Students need to understand their areas of growth academically and reflect on them with their teacher frequently as to update, revise or set new ones. I believe this one area alone could have a huge impact on the individual progress of students. However, in order to gather data to see if that is indeed the case, academic goal setting and the frequency of it, needs to be done with fidelity amongst the three campuses.

Chapter 4: Evaluation of Intervention Plan

Problem of Practice

A problem of practice was determined as a response to a PDS cycle from the prior year. The previous PDSA cycle studied the effects of a PL model on the mathematical achievement of fourth grade students. An initial review of the quantitative data from the previous PDSA cycle, the 2021 STAAR results along with comparison trends over the past four STAAR assessments by grade level and subject, exposed many strengths and weaknesses for the district. Among the weaknesses, the math scores quickly emerged as a central area of focus once again. It appears that only 24% of third grade students met state standards in math, which causes them to be behind when entering the fourth grade. In addition, the overall average of math achievement for third grade was only 36%. The average comes from the percent of students who approached grade level (64%), the percent of students who met grade level (24%) and the percent of students who mastered grade level (19%). To gain some perspective here, the district has a goal of the overall average for every grade level and subject to be 60% which classifies the district with an "A" rating in Domain 1 for accountability purposes.

In addition, only 52% of fourth graders met state standards. However, it's important to note that the 52% is an increase of 13 percentage points from the percent of students who met grade level standards in fourth grade math in 2019. The overall average of math achievement for third grade was only 53%. The average comes from the percent of students who approached grade level (81%), the percent of students who met grade level (52%) and the percent of students who mastered grade level (25%). Since there was not a progress growth measure tied to the 2021 STAAR data, meeting state standards is the measure that will be used as well as the overall average of the percent of students who approached grade level, met grade level and mastered grade level. There is a need as a district to improve the math scores of elementary students, not only in regard to Domain 1 of the accountability system (overall achievement) but also of individual student progress from year to year.

Historically, fourth and sixth grade are the lowest percentages for math in the state each year. In comparison to the state, as well as Region 7, (the region for all three campuses), it's important to note that the district's math scores have been higher than both in every grade level in math for the past five years, generally on average by 15 to 20 percentage points. However, the third-grade scores for the district for 2021 only outperformed the state scores by one percentage point and fell three percentage points below the region. To put this in perspective, the district's average math scores for third grade in 2019 were 18 percentage points higher than the state's according to the Texas Academic Performance Reports (TAPR) (TEA, 2020). After reviewing the district data for the 2021 STAAR, it seems that while the overall average for fourth grade math did rise by 11 percentage points, third grade math fell by a dramatic 33 percentage points for the district. This decline was not campus specific, and in fact materialized on all three campuses.

Intervention

A NIC, was formed in the Spring of 2021 before STAAR results were in. The purpose of the NIC at this time was to review qualitative data from the previous PDSA cycle in the 2020-2021 school year which revealed the district's math instructional model, PL, not being implemented to fidelity across the district. Possible reasons for the PL model not being implemented fully the past school year are outlined in the methods section but one reason is due to the online learning environments brought on by COVID-19. The district leadership team decided that a hard "re-set" of the instructional model will be needed in the Fall of 2021 with all teachers in the district to get back to the best practices outlined in the model. With this “re-set” in mind, the NIC will proposed a support program specifically for math teachers.

After a careful review of the 2021 math STAAR data, the NIC met again and the intervention plan was re-confirmed, this time even stronger, with the analyzation of the quantitative data. In particular, the greatest need for an intervention is in the elementary grade levels. The specified intervention is ongoing, job embedded professional learning for third and fourth grade math teachers; specifically content support for the PL model.

The purpose of this professional learning is to provide teachers with the guidance and support needed to implement best practices of math instruction in all math classrooms, according to the district’s math PL model. The plan for content support includes two full days of math professional development in August prior to the start of the school year and bi-weekly content support sessions in horizontal teams which will be called “Math Mondays”. All sessions will be led by the district math specialist. In addition, third and fourth grade math teachers will not only receive coaching support from their campus instructional coach but will also receive one-on-one monthly visits with the math specialist for the district.

The two days prior to school starting will be to review the district PL math model, and to plan for the first unit of the school year. The bi-weekly content support sessions (Math Mondays) will be centered on the upcoming unit in the scope and sequence. During each session, standards will be deconstructed and best practices in math will be reviewed. These sessions will also be centered on reviewing and analyzing data, helping teachers plan for upcoming lessons and tailored to meet their needs at the given time based on qualitative and quantitative data explained below. Finally, a book study will be suggested for the summer of 2022. The book study is based on the book *Making Sense of Mathematics for Teaching Grades 3-5* by Dixon et al. (2016) and used for the purpose of developing the teachers' content knowledge further. The book was chosen carefully by the NIC as it is meant to help teachers in these specific grade levels develop a deeper understanding of mathematics by the representation of mathematical modeling, exploring, and applying the best practices in mathematics simultaneously. A course will be developed in the district's learning management system which will be broken into modules based on chapters of the book and allow teachers to discuss and collaborate as they move through the book.

Purpose

The purpose of this study is to evaluate the effects of a job-embedded professional learning for the PL math model for third and fourth grade math teachers on the mathematical achievement of third and fourth-grade students. The study will seek to answer the following questions: (1) How have third and fourth grade students met or mastered state standards as defined by STAAR?, (2) How have fourth grade students progressed as defined by STAAR)?, (3) How have students' mathematical functional levels changed?, and (4) How has job-embedded professional learning impacted teachers' ability to implement the PL math model?

Literature Review

The learning loss from the last day of school to the beginning of the next school year is typical in any given year in mathematics, however, after the COVID-19 related shutdowns, the learning losses were expected to be larger than ever. In fact, studies predicted that students returning in the 2020-2021 school year would have only on average 37-50% of the learning gains when compared to a typical school year (Kuhfeld et al., 2020). In one study which compared more than 8,000 schools and the students' scores in reading and math from the fall of 2019 to the fall of 2020 showed that while reading scores remained relatively the same, math scores in 2020 were five to ten percentage points lower than in 2019 (Bailey et al., 2021). Educational researchers predict that this gap will only grow in the 2020-2021 school year causing students to start the fall of 2021 even further behind than in the fall of 2020. Kraft and Falken (2020) provide an evidence-based blueprint for scaling effective tutoring strategies in order to help close the learning gaps. The Texas Education Agency has changed laws recently to require schools to provide accelerated instruction to students who did not approach grade level standards on the 2021 STAAR assessment. As a requirement of the accelerated instruction, tutoring is the strategy presented by the state. The Effective Schools Framework by The Texas Education Agency (2020) says that effective classroom routines and instruction are foundational essential actions present in schools that are effective. Further, it states that campus instructional leaders provide training and on-going supports so that teachers can effectively use research-based teaching practices. More and more schools will find themselves needing to refer to the Effective Schools Framework as a result of the learning losses from the pandemic. In order to do so, best practices in math will have to become more known to many schools and teachers.

Best Practices in Mathematics

In 1989, The National Council of Teachers of Mathematics launched a standards-based movement in America with their *Curriculum and Evaluation Standards for School Mathematics*. This initiative was based on best teaching practices in mathematics and was revised in 2000 in the launch of Principles to Actions which includes eight guiding practices; establish mathematics goals to focus learning, promote reasoning and problem solving, connect mathematical representations, facilitate mathematical discourse, purposeful questioning, procedural fluency and conceptual understanding, productive struggle, and evidence of student thinking (NCTM, 2000). Inquiry oriented instruction and concepts-based teaching in math is positively related to student achievement (Blazar, 2015). Teacher practices such as these, as well as student participation and mathematical discussions when students are able to explain their own ideas, have also been linked to positively increasing student achievement (Ing et al., 2015). While personalized learning is on the rise in schools across the nation, there is still a place for best practices in math. The key might just be figuring out how to best align personalized learning and mathematical best practices. Personalized learning strategies must blend with mathematical best practices, they cannot be stand-alone ideas but must work together seamlessly in the classroom to be successful (Berry, 2018).

Content Knowledge and Experience

Beyond best practices in math, there is also a strong relationship between the teacher's content knowledge and how the mathematics is enacted in the classroom which ultimately leads to students being successful or not (Blazar, 2015). Best practices have been outlined by NCTM for decades, and now there is research to support their theories. Most teachers will even, when asked, say that they believe that best practices and innovative teaching strategies are the best way to teach math in the elementary grades. However, it is often found that their perceptions do not align to their practices and they rely more on how they were taught than their own beliefs about math (Harbin & Newton, 2013). More inexperienced teachers have an even harder time implementing best practices into their classrooms in the beginning of their careers (Ing et al., 2015). The question then becomes, how do we get more teachers who not only believe in best practices, understand how to teach them and then actually implement them into the mathematics classrooms. This can particularly be a problem in elementary classrooms as teachers in Texas who teach elementary grades are not math content experts. Math is only one portion of the EC-6 certification exam. Therefore, unless elementary teachers come to schools with strong content knowledge and a profound understanding of teaching mathematics, schools might just have to take on the role of growing their teachers in this regard. Teachers today are younger, less experienced and more diverse in their preparation for the field than even two decades ago. This is a problem because inexperience and lower levels of content knowledge prove to be less effective in the classroom (Henry et al., 2014). Teacher content knowledge of concepts and connections is directly related to student achievement (Tchoshanov, 2011). Harbison and Hanushek (1992) stated, "At fourth grade, a ten-point improvement in the mean teacher's command of her mathematics subject matter…would engender a five-point increase in student achievement; this is equivalent to a 10% improvement over the mean scores of fourth graders" (p. 114).

Personalized Learning Lessons Learned

In addition to the content knowledge of teachers being an issue in relation to student achievement, with the rise in personalized learning programs, there have been many pitfalls that have come along with it. COVID-19 related environments only increased the use of technology in classrooms. Learners can become mere consumers of the technology and the technology can divide, disconnect, and alienate students if used too heavily (Sulceio de Alvarez, 2018). Students must be given the time to learn and engage with challenging problems and to think. Schools need to return to previously successful experiences again used prior to COVID-19 and the reliance on technology (Sullivan et al., 2020). Teaching mathematics requires consistent guidance from teachers through collaborative opportunities (Khirwadkar et al., 2020) that can only come through discussions and group work in the classroom.

Targeted Professional Development

In order to address teacher content needs and best practices in math and/or PL, there needs to be a targeted emphasis on the development of the teachers' mathematical conceptual knowledge through content-focused professional development geared toward student achievement (Tchoshanov, 2011). Providing more professional development opportunities to enhance teacher content knowledge is one way schools can support their teachers. Another way for teachers to gain a better understanding of how to teach elementary mathematics is through frequent observation cycles and intensive coaching and support programs (Blazar, 2015). There is research to support that teacher behaviors will change with targeted professional development and reflections from the teachers (Thomas, 2008). Job embedded professional development has also been shown to increase teachers' self-efficacy in teaching mathematics (Aulthauser, 2010). Other studies have shown that students in schools whose teachers receive professional development in math perform higher than schools whose teachers do not (Brendefur et al., 2020).

While there is a great deal of research on best practices in math and some research on the effect of personalized learning in some grade levels and subjects, there is less research or case studies in relation to the lessons learned from implementing personalized learning in math classrooms. As the research above suggests, best practices, content knowledge, experience of teachers and targeted professional development are linked to an increase in student achievement in math. The intervention proposed below is an attempt to not only enhance the content knowledge of the teachers, but to also promote best practices in math and personalized learning through targeted professional development sessions for elementary math teachers.

Theory of Change

The district's math model was already in place but as discussed was not implemented to fidelity as written. Students and teachers have varying levels of experience with the math model thus far as some pieces were implemented more in depth than others depending on classroom, grade level and some students being new the district for the current year. Students and teachers proved to implement the software component of the model to fidelity the previous year so it will stay intact and data will be collected as in the past from the IXL platform. In order to support the remaining pieces of the model, the professional learning by way of content support will be implemented. The following is a brief description of the data being collected and the theory behind this intervention is that if teachers are provided job embedded professional learning support for the UA Math Model and the standards, then

(1) Teachers will be able to implement the math model to fidelity

(2) Students will be able to meet state standards on 3rd and 4th grade math

(3) Students will meet progress on 4th grade math STAAR

Setting

The district is an open-enrollment public university charter school in Texas made up of two K-12 campuses and one 1-12 campus spread across three distinct communities in east Texas. The charter was written to be a lab school for the university, and it is modeled after the T-STEM blueprint (Texas Education Agency T-STEM, 2020). The district seeks to be a national model for STEM education innovation as a STEM Academy and University Laboratory School. The district implements project based, problem based and blended learning as the primary methods of instruction while also implementing PLTW Engineering and Biomedical pathways for students.

Students take dual credit classes beginning their ninth-grade year and are able to graduate with 42+ hours of university credit.

The district had an enrollment of approximately 853 students and employed 57 teachers. The student population was 64% White, 19.5% Hispanic, 6.7% African American, 4.1% Asian, 49.6% female, and 50.4% male. The district was also 33.5% economically disadvantaged across the three campuses, with 9.1% special education population and 10.6% Section 504 students. The teachers have various backgrounds ranging from novice probationary teachers to veteran teachers with thirty plus years of experience. Each campus has one Director (much like a principal) and one instructional coach whose main role is to help assist in the implementation of the instructional model as outlined in the charter.

The district's organizational structure is slightly different than a traditional independent school district. The district is led by a superintendent, however, since it is housed under The University of Texas at Tyler's School of Education and Psychology, the superintendent reports directly to the Dean of the college who reports to the president of the university. Under the superintendent there is a Director of Curriculum, Director of Administration and Director of Special Programs. Each of the campuses has a Campus Director and Instructional Coach. Each of the Campus Directors are largely responsible for the organizational aspects of their campuses, while the curriculum team which is made up of the Director of Curriculum and the instructional coaches is largely responsible for the curriculum, instruction and assessment. The directors work closely with the curriculum team to ensure quality instruction is taking place in each classroom but also that the instructional model is implemented with fidelity to the district's as outlined in the charter.

The partnership between the curriculum team and the campus directors is extremely important and vital to the success of the district. As outlined in the Effective Schools Framework, the importance of strong leadership, culture and staffing along with strong instructional materials and instruction is what ultimately makes an effective school (Texas Education Agency ESF, 2020). While the curriculum team and the campus directors share equal responsibility in each of these levers, Lever 5: Effective Instruction, is worth a deeper discussion. The system currently in place in regard to ensuring effective instruction in all classrooms, is that teacher capacity is built mainly by the instructional coaches through observations, tailored professional development by teacher need, frequent meetings, modeling, and feedback cycles. The Campus Directors then work closely with the instructional coaches to provide additional walkthrough feedback and formal evaluations. The directors evaluate while the coaches provide support to the teachers to build their capacity. It’s important to note that the support teachers receive is classified mainly in three areas; instructional support to the model, classroom management, and technology.

Methodology

A mixed methods research design was identified as the best way to evaluation the intervention plan for content support through professional learning. The embedded mixed method design was chosen for numerous reasons but the main one being that the quantitative data is the driving force behind the research and the root of the problem of practice. However, the quantitative data is not sufficient alone and therefore needs the support of qualitative data in order to see the entire picture. For this reason, the embedded experimental design was chosen to fully evaluate the professional learning intervention. More specifically, the two-phase embedded experimental design (Creswell, 2010) was chosen since the intervention is the second phase of a

PDSA cycle and the second iteration of the evaluation. Therefore, the study of the intervention began with the qualitative data results from the first PDSA cycle. The qualitative data playing a supportive role is gathered and a quantitative pre-measure is given before evaluation. Qualitative and quantitative data will be collected during the evaluation and then a quantitative post measure will be given. Quantitative data will include functional level scores of each student from the IXL database (chosen software) and state assessment data. Qualitative data will include teacher surveys and observations. Results of each type of data will be used to make recommendations for the next course of action for possible future iterations of PDSA cycles.

Further, this study is a design-based implementation research study (DBIR) which is iterative in design. DBIR was largely influenced by the term “teaching experiment” which refers to testing approaches to support student learning, observing how students respond, and articulating potential learning trajectories based on what was learned (Campanella & Penuel, 2021). DBR is a good choice in when the research is intended to evaluate an existing program and then a design team plans to intervene based on the results of the evaluation. The intended outcome of the intervention is to hopefully improve student results (Campanella & Penuel, 2021). That is the case in this research evaluation. Iterative cycles will be planned to refine the design and the theory (Svihla, 2014). Furthermore, in this research study, Design Based Implementation Research (DBIR) will be used. DBIR includes a focus on a problem of practice from a team of stakeholders, a commitment to iterative, collaborative design, systematic inquiry, and a concern to developing and sustaining change in schools (Svihla, 2014). The start of this study, was with a team of stakeholders, which is a key component of DBIR. This team of stakeholders is referred to as a Network Improvement Committee (NIC).

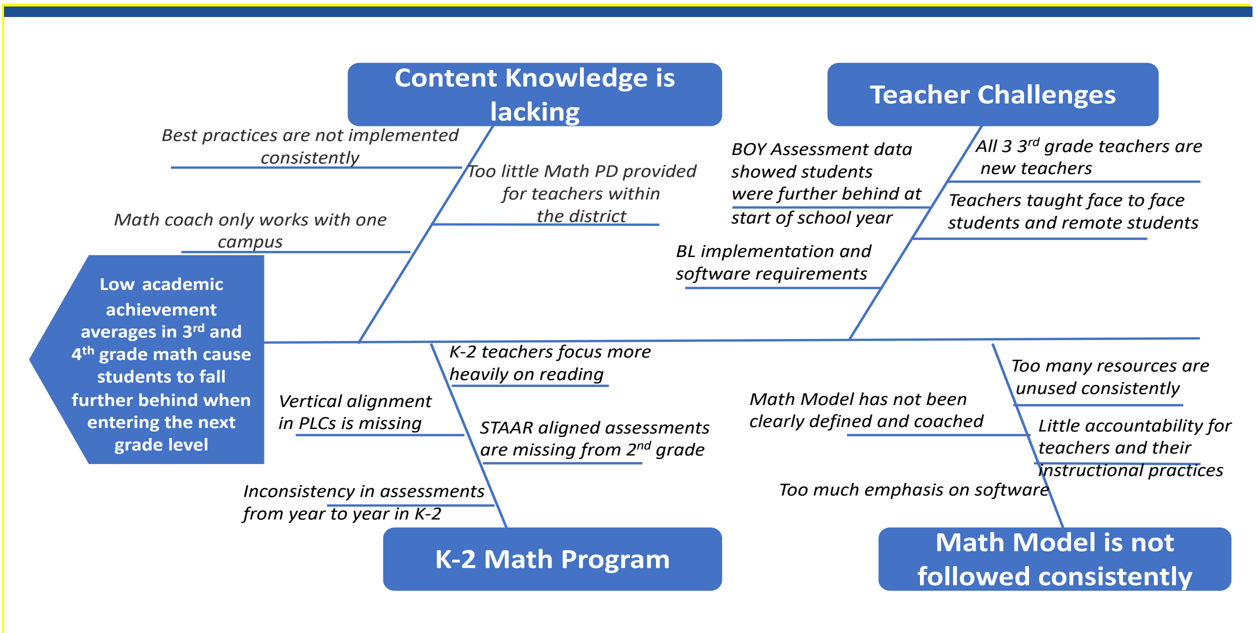
The NIC and DBIR Design

The NIC is made up of various individuals within the organization in order to challenge the participants to think differently about the work to be done and how they relate to and help one another see different perspectives of the problem (Bryk et al., 2017, p. 150). The NIC included the following members: Director of Curriculum, three Campus Instructional Coaches (one from each campus), three Campus Directors (one from each campus), three elementary math teachers (one from each campus), and the District Finance Manager. In order to develop a clear system improvement map, stakeholders from each department need to be included in the NIC. Teachers will be included because they will be the ones actually doing the work to be done. It’s crucial to engage teachers in the design changes that will be needed to address the problem (Bryk et al., 2017, p. 32)

After the NIC met and reviewed the available quantitative (STAAR) and qualitative (observations) data from the previous school year, it was decided that third and fourth grade math would be the focus. In order to dive deeper into the primary contributing factors, the team created a fishbone diagram. The primary factors that contributed to this problem are overall math content knowledge, the district math model not being followed consistently, teacher challenges, and vertical alignment issues from the K-2 math program in test grade levels above. See Figure 1 below for the fishbone diagram.

Figure 5

Fishbone Diagram



Once these primary factors were identified, the NIC developed a system improvement map in order to better understand the distinct pieces of the system that were responsible for the contributing factors. The following paragraphs will outline each of the primary contributing factors in more detail as well as the system improvement map.

First, the overall content knowledge of the teachers in the third-grade classrooms is lacking. When reviewing observational data from the past year, it is clear that best practices are not being implemented consistently. Math professional development has also not been implemented in-house over the past several years. Each of these factors listed are lodged within the curriculum department which consists of the Director of Curriculum and three Campus Instructional Coaches. In addition, the instructional coach who specializes in math no longer travels from campus to campus which is tied to institutional governances related to the budget.

This leaves the other two instructional coaches trying to fill the math gaps on their campus and provide feedback through a math lens when math is not their area of expertise.

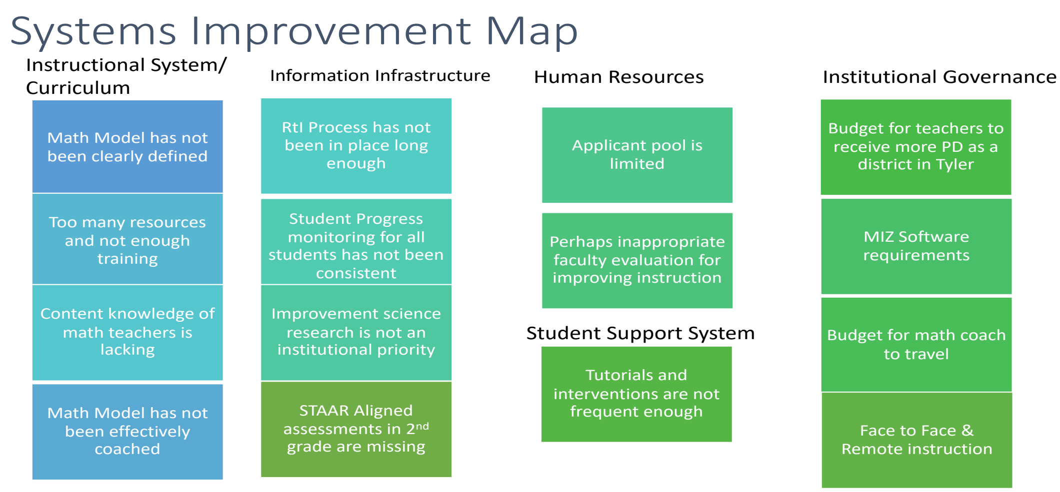
Second, the district math model has not been followed consistently. Several factors contribute to this such as the math model has not been clearly defined to teachers nor been coached throughout the school year. In the past, it was a belief that teachers should have ultimate autonomy in their daily instruction, which has led to drastic variances in the implementation of standards in the district. Additionally, there are too many resources for teachers to choose from and most have not been thoroughly tied to specialized training, which is problematic if the resources are not implemented effectively. Both of these factors are also a direct tie to the curriculum department as well as human resources in term of the budget and time constraints for teachers. Another factor is that teachers were not held accountable to best instructional practices in math, which in some ways was a result of the challenging year presented by COVID-19 and the added stresses on the teachers which will be discussed in the next section. This factor is tied to the leadership of the campuses as well as human resources due to the lack of a quality faculty evaluation system.

Third, overall teacher challenges contributed to this problem in the past year in ways that were more remarkable than years' past. To start, all third-grade teachers were inexperienced. One was a first-year teacher, one was a first-year teacher still in the certification program and one was a first-year teacher to third grade mathematics. This factor falls under human resources and the limited applicant pool for the past hiring year. Additionally, students came in this year lower than previous years due to school shutting down in the Spring of 2020. Beginning of the year assessments revealed that students were lower than they were when they left in the spring. According to data released nationwide by NWEA MAPS assessments, students lost significant ground in math but generally started school where they should be in reading (Fensterwald, 2020). Lastly, there was too much emphasis on software programs this year and not enough emphasis on quality instruction. Some of this can be tied to institutional governances due to software requirements set by the state under the requirements of the MIZ grant, but it can also be linked to district institutional governances and teachers being required to teach face to face learners as well as remote ones. In an article from Ed Weekly, educators were warned that schools may see more loss in math because it is difficult to teach conceptual math and have rich mathematical discussions in an online setting. It also discussed that teachers may rely too heavily on apps and online worksheets which only help in mathematical procedures and is missing the higher rigor of mathematics which will lead to learning loss (Sawchuk & Sparks, 2020). Our teachers faced both of these challenges with the added pressure of teaching students in both settings.

Fourth, the K-2 math program is not vertically aligned to third grade accountability standards. Currently the K-2 classrooms are self-contained, and teachers tend to put more emphasis on reading than math. There are other factors, such as assessments, that are problematic as well; STAAR aligned assessments are missing from second-grade classrooms and there has been an inconsistency in the past several years with the K-2 math assessments. Both of these factors can be linked to the curriculum department as well as the information infrastructure for the district. Lastly, vertically aligned PLCs from K-2 to third and fourth are missing. When reviewing the PLC schedule, it became evident that these teachers do not ever truly get the chance to meet which is problematic in terms of aligning the instruction and expectations from the primary grades to the elementary grades. This factor is linked to institutional governances for the district. See Figure 2 below for the system improvement map created by the NIC.

Figure 6

System Improvement Map



Participants/Demographics

For the purpose of this intervention evaluation, criterion sampling was used in order to evaluate the effects the professional learning model had on all third and fourth-grade students in terms of meeting academic standards and their attributed growth. The district's four third-grade classrooms (two from one campus and one each on the other two campuses) and three fourth-grade classrooms, one from each campus, will be studied. In total, 145 students and six teachers will be participants in the professional learning study. Each teacher is in his or her second year of teaching under the PL model and total years of experience range from one to thirteen years. Students are demographically represented 25% Hispanic, 6% African American, 55% White, and 35% economically disadvantaged. The students are almost an even split between male and female, with 52% being male and 48% being female.

Data Collection

Before the data collection procedures were determined, the NIC met to decide the needs of the program as well as the best way to measure its effectiveness. A program evaluation tool which is used frequently in continuous school improvement was created (Bernhardt, 2018). The tool includes the needs assessment similar to the ones mentioned above, the purpose of the program, intended participants, implementation to fidelity, and how the results will be measured. Table 9 below describes the program evaluation tool in detail.

Table 9

Program Evaluation Tool

Needs Assessment	**Purpose**		**Participants**	**Implementation**		**Results**
What are your data telling you about the need for the program?	***What is the purpose of the program?***	***What are the intended outcomes?***	***Who is the program intended for?***	***How should the program be implemented with integrity and fidelity to ensure attainment of intended outcomes?***	***How is the implementation being monitored?***	***How will results be measured?*** UA Math Model Survey will be administered twice a year to better understand the perceptions of the teachers. Walkthrough data is reviewed once a quarter as a curriculum team. Teacher self-reflection will be done at the end of the year and beginning of
• Only 25% of third grade students met standard on STAAR in 2021 • Only 52% of fourth grade students met state	The purpose of the UA Math Model is to… • Increase the number of students who meet and master STAAR in each grade level	After implementation, the following will result… • Every student will grow in math each year • Every teacher will feel confident and successful in their content	The UA math model is intended to serve all students in K-Algebra II classes (DC comes after) ***Who is being served? Who is not being served?***	• Every math teacher was trained on the UA math model for a week in August prior to school • Semester 1, teachers were	• Instructional coaching • Walkthrough forms reviewed by Director of Curriculum ***How should implementation be monitored?***	

standards on STAAR in 2021 • Only 62% of students met progress in math on STAAR in 2019 • Only 27% of fourth grade students met progress in math on STAAR in 2019 • Not very teacher implements the math model with fidelity • Not every teacher has post-tests aligned to rigor of STAAR or district created tests	• Increase the number of students who make progress on STAAR • Ensure an alignment in math K-12 • Ensure a strong inquiry based, problem solving pattern K-12 in math • Ensure students are university STEM ready upon graduation • Provide continuous job-embedded PD in math content as well as the UA math model • Support teachers through a	knowledge and implementation of the model each year • Every teacher will understand how to target instruction • Every teacher will work with students in small groups • Every student will be exposed to personalized plans via technology • Overall mathematics performance will increase on STAAR • Over time, overall TSI scores will increase in math • Over time, numbers of students who are university STEM ready will increase • Every student will be engaged	Students K-10 are being served through the UA Math Model Math teachers K-10 are being supported to implement the UA math model and participate in Math PD embedded in day	engaged in afternoon PD twice a month • Teachers engage in PLCs vertically as well as horizontally each month • Coaches observe, meet with and provide feedback to teachers • Semester 2, teachers will attend 2 full day pull-outs to plan with the math instructional coach • Teachers are trained in new assessments and on student progress monitoring	• Teachers have the UA math model guide and should use that to gauge where they are in the process. Teachers and coaches should review the walkthrough forms together after each observation. Teachers should be on a scaffolded degree of the model depending on their preps and their experience level as well as content knowledge. ***To what degree is the program being implemented with fidelity?***	the next to determine where teachers believe they are. Coaches will meet with teachers to help teachers set goals. IXL data is monitored each month. Common district assessments are given 3 times a year. District benchmark is given once a year. Common post-tests are given at the end of each unit. MAPS assessments are given BOY/MOY/EOY Student progress monitoring sheet are monitored, and teachers are met with 4 times a year. STAAR data will be used

• K-2 program focuses heavily on reading • Lack of vertical alignment • 3rd grade teaching team is all new • Very little math content PD is provided • Too much emphasis on software following Covid-19	coaching model • Ensure there is a balance of blended learning and best practices in mathematics in every classroom • Ensure consistent assessments are given and aligned • Increase student goal setting and ownership in math	in learning mathematics • No student will need college remedial math • Aligned K-12 curriculum in math		• Post-tests were created by the district for the first time and reviewed with teachers	• All math classrooms are implementing the UA math model. Some classrooms have been scaffolded for teacher needs according to the number of preps and the experience level of the teacher.	

All quantitative data will be stored in the Student Progress Monitoring (SPM) spreadsheets used across the district and collected throughout the year. The teachers are responsible for inputting the data into the spreadsheets and the teacher and campus instructional coach meet twice quarterly to review the data. The data review meetings are used to calculate the percent of students currently on track to meet grade level standards on STAAR as well as the number of students on track to meet progress in fourth-grade. The SPM spreadsheets are designed to be one place where all relevant test data is stored and tracked by grade level and subject. The data in the SPM spreadsheets consists of common district assessments, benchmark data and classroom post-assessments. Classroom post-assessments currently vary by teacher and therefore are reviewed and monitored but not calculated in the percentages for the quantitative in-process measure. Common district assessments are given at the end of each quarter and are based on the standards from the scope and sequence taught within the respective quarter. Common district assessments are aligned to the content and rigor of STAAR. Teachers will place the scores in the SPM spreadsheets at the end of each quarter. A district benchmark, which is a released STAAR test, will be given in March and the scores from the benchmark will be placed in the SPM spreadsheet once it is given. Data from the STAAR will also be collected in order to assess the percent of students who met grade level standards once the scores are released in June.

IXL functional levels are placed in the spreadsheets as soon as the diagnostic is completed within the first few weeks of the school year. IXL is a personalized software program that calculates the students' functional levels. A student's functional level can be different from the student's actual grade level. For example, a student can be in the fifth grade, third month of the school year, which is denoted as 530. For the student to be considered on grade level, the student's functional level score would need to match the student's grade level and corresponding month of the school year when the functional level as taken. A higher functional level would indicate the student is above grade level and a lower functional level indicates that the student is below grade level. The IXL functional level is updated at the beginning of each quarter for the remainder of the year and tracked to monitor student progression.

Qualitative data will be collected throughout the school year as well. Teacher surveys will be given in December and May. The surveys will be identical in questions and format in order to be able to see changes over time, if any. The surveys will include closed and open-ended questions related to the bi-weekly content support sessions and the support received in relation to their ability to effectively implement the district's math model. The teachers will be asked to rate and/or choose the frequency of each question and then asked to reflect as to why they chose the closed answer response. Closed and open responses are required for each question. The surveys are intended for teachers to be able to openly express their satisfaction with the content support and if they feel they are better able to implement the model as a result of the extra support. The goal of the surveys is also to be able to evaluate a correlation, if any, from the content support to the implementation of the model, to the overall success of the students. Teachers will be encouraged to be as open and honest as possible and assured that the survey is for feedback purposes only and not an evaluative measure of their teaching performance.

Teacher observations will also be conducted bi-weekly by the campus instructional coach and monthly by the math specialist. The curriculum team will meet to analyze the teacher observation data in October, February and May. All observers will use a standardized walkthrough form designed to reflect math best practices and the district PL math model. In order to ensure there is interrater reliability in the use of the forms, the curriculum team will meet and perform walkthroughs using the form on each campus with the goal of calibration. The purpose of the walkthroughs will be to collect data in order to evaluate the fidelity of implementation of the PL math model across the third and fourth-grade classrooms in the district.

Data Analysis

IRB approval was obtained prior to the collection of data for this study even though this study is one that would have been done anyway, as it is tied to a grant given by the state of

Texas. The state requires data collection from a quantitative standpoint but when the district wrote the implementation plans, the surveys were written into the program and data collection methods. It is important for the instructional coaches to have access to the data (1) because they are responsible for helping with data collection for their respective campus and (2) they are involved in the data analysis in order to make decisions for future cycles of this work, and (3) the data helps inform them of how to best support their teachers. While ultimately the quantitative data is reviewed at the end of the school year once all STAAR data is in, the curriculum team monitors the input of district assessment data as well as IXL data quarterly in curriculum meetings.

Teacher surveys are anonymous and collected using a Qualtrics account owned by the Director of Curriculum. Reports of the survey are distributed to the instructional coaches and reviewed in curriculum meetings following each administration. Results are analyzed for themes in the reflections of the teachers as well as the frequency of the responses. The surveys provide useful information for each of the components of the model as well as the perceptions of the teacher.

Teacher observations made by instructional coaches are shared with the teacher and with the Director of Curriculum only and are common practice in the district regardless of this study. Instructional coaches do not see the observations made by other coaches. Observations are analyzed and coded into themes by the Director of Curriculum without any identifiers noted. Quotes and anecdotal notes are abstracted from observation forms and field notes into separate documents. Once all notes are organized by themes in separate documents, it is unidentifiable as to which teacher it belongs to.

Member checking is involved in each of the data collection methods. As the Director of Curriculum, this is an important part of the process in order to make sure my own preconceived notions do not interfere with analysis. In my role, it is imperative to not make generalizations about all classrooms if I am only seeing evidence of a certain theme in one. By each member of the curriculum team individually analyzing data first, then discussing in curriculum meetings to safeguard accuracy and consistency, the team is able to ensure data is being portrayed accurately.

Limitations

It's important to note that there are a few limitations and barriers with the evaluation of this intervention. For example, all bi-weekly content support sessions will be conducted remotely via the Zoom platform. One limitation to the district is that the three campuses are extremely spread apart, anywhere from 45 to 91 miles in distance between the respective campuses. Therefore, bi-weekly meetings are remote only options. Another limitation is that the math specialist for the district is also the campus instructional coach for the Longview campus which gives the third and fourth-grade classroom in Longview an advantage perhaps over the other two campuses in terms of the types of feedback from the walkthrough forms and coaching sessions. The other two campus instructional coaches are not math specialists. An additional barrier could be the collection of the data coming from three different sources in terms of the classroom observations. This is the reason that specific dates are set for the analyzation of the data as a curriculum team as well as the collective walkthroughs as a team prior to any individual walkthroughs at the beginning of the year.

Another limitation to the data is that the data in terms of the first research question is that the results are from different groups of students. Since the first research question pertains the percent of students who meet or exceed progress on STAAR in third and fourth grade, this data will be compared to the previous year's data from where the problem of practice originated, even though the groups of students are not the same. The students' data reflected in the 2020-2021 STAAR results were fourth and fifth grade students at the time of this study, therefore, the data is not following a cohort for question number one.

Results

Quantitative

As mentioned in the problem of practice, only 24% of third graders met state standards in 2021 and 52% of fourth graders met standard. The first research question was centered on the extent that would meet standard in 2022 as a result of the professional learning intervention and teachers' ability to implement the math model. After a review of the 2022 STAAR results, it appears that 47% of third graders and 60% of fourth graders met grade level standards. This is a percent increase of 23 in third grade and a percent increase of 8 in fourth-grade. In addition, in 2021 only 18% of third grade students mastered grade level expectations but in 2022, 23% mastered. Not only is this an increase of 5 but its worth noting that the percent of students who mastered standards in 2022 is about equal to the number of students who met standard in third grade in 2021. Finally, in 2021, 25% of fourth grade students mastered standards and in 2022, 35% mastered standards; an increase of 10. See results presented in Table 10 below.

Table 10

Percent of Meets and Masters of STAAR '21 and STAAR '22 by Grade Level

Grade	Meets21	Meets22	Growth	Masters21	Masters22	Growth
3	24	47	23	18	23	5
4	52	60	8	25	35	10

A Mann-Whitney U test was conducted to students' STAAR scores before and after teacher exposure to the content support. The Mann-Whitney U test, a nonparametric statistical procedure, was used because our data violated the assumptions of the independent samples t-test. Specifically, a test for normality of for each level of the independent variable indicated the data were not normally distributed. We found no significant differences in average STAAR scores from 2021 to 2022. (*M* = 66.91, *SD* = 20.51) and intervention conditions (*M* = 67.57, *SD* = 20.28), with W = 8012.00, *p* < .856. See table 11 below.

Table 11

Results of STAAR '21 to STAAR '22

STAAR '21			STAAR '22			t(255)	*p*	Effect Size
N	M	SD	N	M	SD			
112	66.91	20.51	145	67.57	20.28	8012.00	.856	-.01

Fourth-Grade Progress

The second research question, 'To what extent did fourth grade students meet or exceed progress as defined by STAAR?,' was looked at next. It's important to note that the personalized learning model used in math was created in response to the low number of fourth grade students meeting progress on STAAR in 2019. The personalized learning model was implemented for the first time in 2019-2020 school year and then the pandemic hit in the spring of 2020 which resulted in no STAAR test that school year. STAAR progress was not measured in 2021 since there was not a test in 2020. Thus, the 2022 STAAR progress scores were the first ones since 2019 and therefore the first chance the math model has been able to be evaluated for the true reason it began. In 2019, 27% of fourth grade students met progress as deemed by STAAR. After a review of the 2022 STAAR results, it appears that 86% of students met or exceeded progress on STAAR. This is a percent increase of 59 in this grade level since the implementation of the math model. See table 12 below.

Table 12

Percent of Students Meeting Progress in 2019 and 2022

N	Meets Progress '19	N	Meets Progress '22	Difference
56	27	61	86	59

Mathematical Functional Levels

To address research question three, 'to what extent did students' mathematical functional level improve?', the chosen software of IXL will be used for this data set. The district has used the IXL software since the 2019-2020 school year and fourth-grade data sets are available for BOY to EOY for the past three years. Due to the change in software for third graders over the course of the three years, only fourth grade data can be longitudinally studied. However, third grade data in IXL will be discussed for the most recent year 2021-2022. The number of students in each data set varies from the reported overall students in the respected grade levels and may also differ from STAAR data sets. This is due to students who did not have a score for both BOY and EOY being removed from the IXL data sets.

A paired-samples t-test was used to compare third grade students' mathematical functional level score before and after exposure to the IXL program in the 2021-2022 school year under the intervention of teacher content support for the program. Prior to data analysis, the primary assumptions of the analytic procedure were checked. The assumption of normality was not violated as assessed by Shapiro-Wilk's test ($W = 0.99$, $p = .48$). The results indicated students' functional level increased from quarter one (*M* = 272.50, *SD* = 80.63) to quarter four (*M* = 390.88, *SD* = 75.71). This improvement was statistically significant with $t(79) = -20.56$, $p < .001$. The effect size for this analysis ($d = -2.30$) fell above Cohen's (1992) convention for a large effect. See Table 13.

Table 13

Results of 3rd Grade IXL Data from Quarter 1 to Quarter 4

BOY		EOY		t(79)	*p*	Cohen's *d*
M	SD	M	SD			
272.50	80.63	390.88	75.71	-20.56	<.001	-2.30

A paired-samples t-test was used to compare fourth grade students' mathematical functional level score before and after exposure to the IXL program in the 2021-2022 school year under the intervention of teacher content support for the program. Prior to data analysis, the primary assumptions of the analytic procedure were checked. The assumption of normality was not violated as assessed by Shapiro-Wilk's test ($W = 0.98$, $p = .67$). The results indicated students' functional level increased from quarter one (*M* = 388.70, *SD* = 76.83) to quarter four (*M* = 470.37, *SD* = 82.67). This improvement was statistically significant with $t(53) = -10.75$, $p < .001$. The effect size for this analysis ($d = -1.46$) fell above Cohen's (1992) convention for a large effect. See Table 14.

Table 14

Results of 4th Grade IXL Data from Quarter 1 to Quarter 4

BOY		EOY		t(79)	*p*	Cohen's *d*
M	SD	M	SD			
388.70	76.83	470.37	82.67	-10.75	<.001	-1.46

Lastly, a one-way ANOVA was done to determine if IXL growth differed amongst the year of implementation and interventions: Year1 = personalized learning model (2019-2020), Year 2 = modified personalized learning model under Covid-19 protocols (2020-2021), and Year 3 = personalized learning model with content support for teachers (2021-2022). It's important to note that in year one, one fourth grade classroom was not included in the data set due to the campus using an alternative software program for functional levels than IXL. In year two and three, all campuses were used in the data and only students who had scores for BOY and EOY were used in the data set. A check for the assumptions of the analytic procedure were done and found all assumptions were met. Results indicated there was a significant difference in IXL scores based on year of implementation, with $F(2, 1149) = 5473.91$, $p < .001$. Effect size estimates suggest approximately 21% of the variance in functional level scores was explained by the year of implementation ($n^2 = .21$). See Table 15 for the descriptive.

Table 15

Difference of Growth Scores by Year

Year	Mean	SD	N
1	151.48	112.04	44
2	58.89	45.71	54
3	81.67	55.82	54

Qualitative

Observations forms completed by coaches were sorted by semester one and semester two. Semester one data included all observation forms from each third and fourth grade classroom in the district. At the end of semester one, these forms were extracted into one major document and analyzed by open coding. This method was performed a second time at the conclusion of semester two with the same available observation forms as before. In addition, a middle and end of year math content support survey was also extracted into one major document and analyzed using open coding. Once open coding had been done on all the forms as well as the two surveys, the open codes were then compared for similarities between the two; observations and survey results. The codes were then categorized into major themes. Analysis of the qualitative data sources leads to three major themes emerging from the observations and survey results.

Semester One Results

The observation results from semester one and the middle of the year survey were found to somewhat contradict one another. It is evident from the observation forms that teachers were having success implementing the math model more times than not when instructional coaches were in their rooms observing, however, the feedback left in the survey overwhelmingly indicated that teachers were not satisfied with the Math Mondays.

From the observations, there were a few comments left by coaches on how to better spend class time such as, “You had students all get on IXL after the whole group lesson. How can you try IXL as a station and not have all your students on it at one time?” This comment reveals that some teachers might not be comfortable enough with the math model at that point in the year and need help breaking up the class time into the components of the model. Another comment made by a coach which is very similar, “How could you maximize your time to be able to pull small groups? Could you have had the 15 minutes of IXL or Reflex in a station, then the TEKS task cards in a station and you as a group? This would have enabled you to review or reteach items needed. That small group time is so important and crucial to filling the gaps.” Again, reveals that this teacher needs more support in being able to implement the model. Comments like this indicate that the teacher isn’t necessarily following the math model and is struggling in the areas of small-group pull-outs and stations. Other than these two comments in the observations, all other feedback left by coaches indicate that teachers were having success in the model.

Once the observations were sorted, analyzed with codes and sorted into themes, the teachers implementing the model and seeing success with it was the majority of the notes left by instructional coaches. Coaches consistently noted the use of whole group instruction followed by stations, small-group pull-outs based on data, students tracking data, integration of technology as stations, fluid stations, and students self-monitoring their progress. “Students are self-monitoring their progress throughout the classroom assignments. This is a huge improvement in their ability to prioritize and stay on task” was one comment left by a coach. Another one, “Other students are moving around the room answering questions placed on cards stationed around the room. Students have a choice where they go. When they finish they get on IXL independently. Students know where to put their papers when they are done and what to do when they are done. The teacher is able to facilitate small group with few distractions.” One comment that sums up the success coaches were seeing from the implementation of the model was, “The teacher is implementing the UA Math Model with fidelity while providing the extra support of whole group time for this group of students. The teacher has adjusted the instruction to provide higher level direct instruction to model metacognition for all students.”

Even with the overwhelming evidence from the observations made by coaches of the success of the implementation of the math model, the middle of the year survey results indicated that teachers were not contributing this to the implementation of the Math Mondays. When asked if students were making progress, out of the six teachers surveyed, four agreed and two strongly agreed. However, when asked since starting the Math Mondays if they had been able to implement the math model in their classrooms, only one teacher agreed, four were neutral and one strongly disagreed. Lastly, and most likely the most important, when asked if the Math Mondays directly supported them in the understanding of their content and resources, three disagreed, one strongly disagreed and two were neutral. The survey contained comments such as, “I personally do not see the benefit of Monday PD Sessions”, “I do not find the Monday PD Days beneficial.”, and “I do not feel the sessions are beneficial.” While these comments were important because they let the team know that teachers were not valuing the days as intended, other comments proved to be more helpful such as, “There is no focus on grade level content in the meetings. Having all grades meet together does not allow us to focus on particular content. Even when we break into groups, we only briefly discuss content” which alluded to a specific problem. A few other comments such as, “I think to be more beneficial, the time could be utilized making and preparing actual engaging activities to increase learning”, “I do not feel like Math Monday is impacting my students in a positive way because I think the approach is wrong. It is also not happening at a convenient time for teachers, which makes it difficult for teachers to feel receptive and engaged” and “I do think it would be great to have a way to share ideas and discuss things as math teachers, but I am not sure this is the most effective way” revealed that the problem was in the approach, timing, and setting and not the extra support. A final comment in the survey led the team to change the approach entirely, “I would like to add one of the most helpful things I have received is the planning day with the math specialist. That was valuable in so many ways.”

Due to the observational data providing evidence that the model was being implemented and being implemented well, but the survey data showing that the teachers were not happy with the delivery of the support (Math Mondays) but still valued the time with the other teachers and the math specialist, a switch in the job embedded professional development was made at semester. Math Mondays were eliminated based on the teacher feedback and in place of them came Math pull-out days for teachers with the math specialist. Three days were embedded for each grade level group of teachers for the spring semester. Substitutes were placed in the classroom for the three days and teachers were pulled out to plan for upcoming quarters together and with the guidance from the math specialist.

Semester Two Results

The observational data remained relatively consistent with semester one, proving that teachers were still implementing the model and having success with it. "Small group instruction. Groups pulled on data. Students working on different items based on their playlist" was a comment left in one observation. Another one similar, "Next, the students moved into their flexible stations. You chose a group to pull for small groups based on their unit assessment the day prior". It seems from comments like this one, "I love the consistency that you have built into your Math classroom! It is evident that students know the routine and it gives them a sense of safety in the procedures to being open to learning the content. In math the content is difficult, so giving them the sense of peace in the environment is wonderful!!!" that teachers had found a rhythm and students knew what was expected of them.

One major shift in the observational data from semester one to semester two was in the feedback left for teachers who might not be implementing the model all the way. In semester one, the feedback was more related to the use of class time and not spending so much time on whole group activities but providing space for stations and small-group pull-outs. These comments were not present in semester two which may be due to the fact that it had been fixed. However, there were comments in the observational data from semester two around the use of technology being over utilized which is also a fidelity issue. One coach commented, "After pulling some IXL data, I noticed that there seems to be an increase in the number of questions that 4th grade is completing compared to other campuses/grade levels, so I wanted to see if there was a reason. This month your 4th grade students have completed an average of 140 questions per week. Other classes are just below the 65/week mark for this month, so that is a little concerning. Technology is a wonderful tool, but it should never replace your content knowledge. The technology is to support and reiterate your classroom instruction."

The survey data from semester two shifted in tone from semester one. Three themes emerged from the comments; expression of success in the model, growth in students and the preferring the pull-out days to the Math Mondays. Comments such as, "I am very successful with the math model" and "the station expectations make it easier to narrow down what to do in class" were present in the second round of the survey. One comment in the survey in regard to the growth in their students was, "According to MAP and the benchmark, and the increase in mastery on unit tests, my students have grown. Even in class, when we are having math discussions or working together, their understanding and their ability to use math vocabulary has grown exponentially." And finally comments such as, "I definitely prefer the pull-out days to the Math Monday PD. We were not having to give up planning time to hear things that weren't always specifically relevant. These pull-out days were more targeted. It was like the difference in whole group versus small group instruction" were made in support of the change of the delivery in job embedded professional development. "I honestly loved the pullout days better than the Monday PD sessions. Not saying the Monday sessions were bad, I just got more benefit and information from the pull-out days than I did from the Monday sessions" was another comment in support for the new method used in the spring semester.

Discussion

One important factor to consider when discussing the quantitative and qualitative results is that the qualitative results showed that most teachers implemented the model to fidelity across the three campuses and it was consistently supported and coached throughout the year. The quantitative data revealed that growth was made in each area studied. The percent of students who met and exceeded state standards on STAAR increased from the previous year in both third and fourth grade and the percent of students who made progress in fourth grade increased significantly from the last time it was measured. We know from the previous year that the model was not implemented with fidelity. The intervention of embedding support for the implementation of the model did prove to lead to an increase in the fidelity of the model and we know that there was also an improvement in student achievement data.

When asked about their current understanding of the math model, the survey results shifted from the way this question was answered after semester one to the way it was answered after semester two. The middle of the year survey revealed answers such as, “PrBL was at the core of our model when I started, then blended learning, and now we are trying to find a way to combine the two. Best practices have remained the same, but the classroom model has been constantly evolving” and “we started with launching a PrBL and based workshops throughout the unit to teach the TEKs, but also to give the understandings needed to complete the PrBL. Now I feel there are some that still follow this, but not all. Also we now have worked with personalized learning and blended learning. It is hard to integrate all of this successfully in my opinion.” Comments such as these are somewhat vague, they don’t speak to specific pieces of the model. It seems that when teachers wrote these, they were still unsure of the exact model and how it all fit together. Confusion seems to be an undertone in these remarks.

However, when the same question was asked in the end of the year survey, the answers became more specific and clear in reference to the pieces of the model. “PRBL, Small group, differentiated instruction, stations” was one comment left by a teacher. Another teacher wrote, “We are supposed to start with a PrBL, and then use blended learning throughout the unit to learn and master concepts. There should be a pre-test and a post-test. We should use stations that include fluency, hands-on, independent, teacher-led, and technology” which encompasses the model with each piece and how it blends together.

Students’ functional levels also improved from beginning below grade level averages in both third and fourth grade to ending averages falling just under the desired ending levels. For instance, IXL results for third grade went from an average of 272 (second grade, seventh month) to 391 (third grade, ninth month), an increase of an entire year plus two months. Fourth grade rose from 388 (third grade, month eight) to 470 (fourth grade, month seven), a gain of almost a year. When reviewing the observational data, we know that teachers were implementing IXL consistently rather it be in stations, flexible stations or as a whole class (even though that is the least preferred method). The quantitative and qualitative data from the technology shows that when used consistently, the chosen software is able to improve the functional level of the students.

Conclusion

Student growth was at the heart of this study and whether it be growth from previous years’ data to the growth shown from beginning to end of year, there was an improvement in all areas measured quantitatively. When comparing the PL math model from the previous year to the current study, the model was the same, however the mathematical achievement improved the second year. It would seem that the difference in the two years could be attributed to the continuous job embedded professional development on the PL math model for the teachers. Teachers commented that they are very successful with the math model and “it was nice to talk to other teachers to see what was working or not working in the classroom”. Teachers want to be successful, but they have to be given the support and the time to implement new strategies or expectations to fidelity. While the current study was not necessarily a study of how well the math model works for this setting, it does appear that when implemented consistently, it leads to gains. The most important takeaway from this study is the intervention led to more successful implementation of the model than previously and thus led to improving student achievement. Unlike prior year’s implementations of training and then coaching, it appears that the intervention of on-going training, support and intentional development of the teachers is the key to true change in what was seen in the classroom. Teachers commented that they felt very supported and knew that they could readily ask for help and it be provided.

Recommendations

“Student success directly correlates with teacher success, as teachers are the number one contributing factor to student success” (Cleary, 2018). In order for teachers to be successful, they need to be supported. They need on-going job embedded professional development that affords them the time and space to learn and grow. Each year for a teacher is different because each year brings a new group of students with new strengths and struggles. Teachers need the time within their workday to communicate with one another, to learn from specialists and each other, to ask questions, plan together, and ultimately impact student achievement in a positive way.

While the intervention for this study shifted mid-year from Math Mondays to pull-out days, it was an important change because it was an immediate response to teacher feedback and teacher needs. Even though observational data was showing that the model was being implemented in semester one, teachers were not pleased in the way they were being supported in order to do so. A change mid-year most likely also showed the teachers that their voices matter and the district does listen. A recommendation is that is what needs to continue to happen.

The math pull-out days should continue for the 2022-2023 school year and it might be important to consider these days for teachers in other subjects. While this study was done on math teachers, the intervention of job embedded professional development should be considered for all teachers moving forward.

Additionally, data should continue to be collected both quantitatively and qualitatively. Lastly, and most perhaps most importantly, a response to the data needs to happen. In order for teachers to be most successful, they have to know they are cared for and their concerns are heard.

Early Chapter 5: Discussion of the Results

The final chapter of this dissertation in practice, chapter five, will provide an overview of the results the consecutive year-long studies. In addition to a discussion of the results of the initial evaluation study as well as the intervention study, implications for future practice and research will be important to note. The improvement science framework is a “learning by doing” model that is based on disciplined inquiries (Bryk et al., 2017). The study for the purpose of this dissertation may have come to an end but the future practice in the school setting as well as future study will not, otherwise, the improvement science framework would have failed. Improvement science frameworks are designed to achieve high quality results and therefore demand sustained attention to refining, redesigning, and re-evaluating on a continuous basis. Improvement science is not something that is done for the purpose of stand-alone studies by outside researchers but rather actively engaged participants who have the same goal and that is to improve the school. Plan-Do-Study-Act (PDSA) cycles are meant to be iterative in design. What is learned from one is acted upon and planned in the next only to be studied again. Everyone involved is now an “improver” and seek to generate evidence about how to achieve better outcomes on a more regular basis (Bryk et al., 2017). Subsequently, this chapter serves to discuss what is next for future PDSA cycles in practice for the district as well as future research that can be studied. Lastly, discussions of how this research can impact other districts or education at large will be considered.

Discussion

Discussions for both years of study, year one evaluation and year two intervention will be described separately below. A review of the quantitative and qualitative for each year will be discussed separately but also analyzed for themes and possible correlations between the two. The reasons for decisions made after the year one evaluation into the year two intervention, as well as a change to the delivery of the intervention midyear will also be discussed.

Year One Evaluation Study

The evaluation study from year one was an evaluation of the existing, yet new at the time, personalized learning model for math particularly in grade four. The study was intended to evaluate if the personalized learning model had any impact on student growth in the personalized software (IXL) as well as if the model had any impact on student achievement on the STAAR test. Additionally, the study sought to understand the teachers' level of satisfaction with the personalized model and if they believed their students were being successful under it as well. Year one proved to have many limitations beyond the obvious ones of small sample size of only three classrooms, therefore only three teachers. Additional limitations or barriers to the study year one was that the model ultimately wasn't implemented to fidelity due to instructional challenges as a result of Covid-19. The study was done in the 2020-2021 school year, coming off an early end to the previous school year in the spring of 2020 due to Covid-19. The 2020-2021 school year began with the option for students to attend face to face or virtually which resulted in instructional changes and challenges for many teachers. Findings from year one will be discussed in this chapter but if we were to discuss what did not go well from any aspect of the study, year one fidelity issues would be a top concern.

Research question two was designed to see if the progress of fourth-grade students in 2021 increased from the progress of the fourth-grade students in 2019. The progress results in

2019 drove the why behind the personalized learning model in the first place. The model was designed to address the lack of progress in all math grades after trends were seen for several years in math data but, in particular the lowest group which was fourth grade. However, since the STAAR test was not given in 2020 due to the shutdowns caused by Covid-19, there was no progress measure for 2021. As a result, the Texas Education Agency (TEA) released optional beginning of year assessments (BOY) that were designed to assess students' achievement in the grade level before and given in the fall of 2020. The results of this BOY were compared to the actual STAAR test given in the spring of 2021. Students scored an average of 60.8% on the BOY STAAR test and average of 69.75% on the actual STAAR test that spring in 2021. At first glance, a growth of only 9% from the beginning of the year to end of year seems surprising.

To gain further insight into these numbers, it might be important to bring in some wider context from state and local data. The BOY STAAR given by TEA was a third grade STAAR designed to see where the students would have been had they taken in at the end of third grade had Covid-19 not shut schools down from face-to-face learning. According to state data, of all the schools who took the BOY, only 15% of students met standard in fourth grade math according to data released on the Texas 2036 website (Texas 2036, 2021). Internal data shows that on the same exact BOY the district had 23% of students meeting standard. By the actual STAAR test that spring, state data indicates that 35% of students in the state of Texas met standard and 38% of students in Region 7 (the education service center in which the district lies) met standard. The district, however had 52% of students meeting state standards for fourth grade math. So thus, while the average overall scores for the district only gained 9% from BOY to STAAR, the percent of students meeting state standards increased by 29% while the state had an increase of 20%. One might want to assume that the greater gains were seen in the district as a result of the personalized learning model, but in order to gain greater insight or be able to relate the gains to the model, its important to look at the functional levels within the personalized software as well as research around the Covid-19 learning loss.

Research question one, the functional level growth in the personalized software was the second area of focus quantitatively for the evaluation study. The software is designed to reach students at their functional level and in theory help them catch up quicker to grade level or exceed grade level. In a study produced by IXL (the personalized software used in the district) it was found that Texas schools using IXL outperformed schools without IXL in math grades 3-8 and schools that used IXL for two or three years outperformed schools who had only used for one year (IXL, 2019). These findings are consistent with the state data for the district mentioned above. The evaluation study in year one found that fourth grade students' functional level in IXL grew on average from a third grade, ninth month score (392) at the beginning of the 2020-2021 school year to a fourth grade, fifth month score (451) at the end of the year. This is an average growth from all three campuses of six months. Like the state assessment data, this growth seems surprisingly low at first glance.

In a study produced by TEA, it was estimated that students returning in August of 2020 had a 5.7 months of summer loss as opposed to the typical 2.5 months reported in typical years prior (TEA, 2020). In addition, historically only 4% of students who are below grade level catch up to grade level over a two-year span in Texas. According to the Texas 2036 report mentioned above, by the end of 2021 students in Texas were on average five months behind grade level in math. To put this in perspective, that would mean that the average IXL scores for the state, if taken, would be somewhere in the low 400's or lower. An average end of the year score for the district of 451 is on grade level, even though it is less than desirable. Covid-19 had many impacts on student learning and the learning loss was severe. It appears that the personalized learning model in the district was able to offset some of that learning loss.

The qualitative results from year one can be summed up with three major takeaways. As a result of the personalized learning model, data cycles had moved from large ones to shorter ones, goal setting was present but more so on personal goals than academic ones, and the model was not implemented with fidelity across the district. The first, data cycles shortening which impacts data driven decisions, is in line with prior findings that data driven decisions utilized in short cycles in personalized learning models have been reported to have single greatest impact on student achievement (Zdeb, 2018). The use of shorter cycles in conjunction with data driven decisions in the district under the personalized learning model could have been a contributing factor the increase in state assessment data as compared to the state or region. In addition to shorter data cycles, the qualitative data reported from year one showed that teachers were using the shorter data cycles to make grouping decisions as well as pull small groups to work with. In previous research, it was noted that when teachers use data to drive groupings in their classroom, it has a significant impact on student achievement (Zdeb, 2018). Shorter data cycles, using daily formative assessment data, has the potential impact to address students who are struggling and address their needs before they become larger issues (Wilson, 2017).

Secondly, the qualitative data revealed that goal setting among students was present but not based mainly on academic goals as intended in the writing of the PL model. According to an action research study done by Smithson (2012), teacher assisted goal setting positively impacts academic performance of elementary students. In addition, when students have access to his or her own data and are setting goals for themselves in relation to their own data, have been reported from schools who have seen the greatest achievement gains in math (Pane et al., 2015).

However, in conjunction with this evaluation study it is important to note that Pane et al. (2015) reported that very few models use, or are able to implement goal setting effectively with students which is in line with the findings from this year one study. In a follow-up study, Pane et al. (2017) reported that it takes on average two full years of PL implementation before a school starts to see achievement gains in math. Early findings from the first year of implementation of the PL model for the district point towards small gains even in year one under Covid-19 challenges.

As mentioned before, the personalized learning model was not implemented in year one to fidelity and Covid-19 was considered a major contributing factor for that. However, it's important to note that this study is in line with many previous studies that have found that personalized learning models are rarely implemented the same. One study even found that out of 32 schools that implemented PL models, there were 32 unique models. It does make one question that had Covid-19 not been an issue, would there still have been three variances of the proposed model across the three classrooms. The PL model implemented by the district was based on four components: data driven decisions, goal setting and reflection, targeted instruction, and the use of personalized technology. Prior research has revealed that there are very few if any models that utilize all four components. The most that was found of any model have been two. The field lacks evidence on which instructional strategies will yield in achievement in math. According to the year one study done in this district, the two components of the model that were consistently used were the personalized software and the use of shorter data cycles to target instruction. This study would assume that these two components can yield achievement results even in the first year of implementation.

Year Two Intervention Study

The intervention study was designed similarly to the evaluation study. The same personalized model was to be used but this time, the teachers would receive embedded professional development along with the traditional instructional coaching, delivered by the district math specialist. The intervention began as a refresh of the model prior to school starting along with follow-ups bi-weekly in joint professional development sessions on Mondays. Research question one sought to answer if the intervention had any effect on student achievement on STAAR in terms of students meeting state standard or mastering state standard. The year two intervention study also included third grade classrooms in addition to the fourth-grade classrooms due to the results of the third-grade state assessment data for the district in 2021. After a review of the 2022 STAAR results, it appears that 47% of third graders and 60% of fourth graders met grade level standards. This is a percent increase of 23 in third grade and a percent increase of eight in fourth-grade. In addition, in 2021 only 18% of third grade students mastered grade level expectations but in 2022, 23% mastered. Not only is this an increase of 5 but its worth noting that the percent of students who mastered standards in 2022 is about equal to the number of students who met standard in third grade in 2021. Finally, in 2021, 25% of fourth grade students mastered standards and in 2022, 35% mastered standards; an increase of 10. One surprising takeaway from this data is that the percent of students meeting grade level in third grade in 2021 almost doubled in 2022. It’s important to note that while third grade was expected to do the personalized learning model in 2021, it was not studied or as much of a focus as the fourth grade was. For year two intervention study, the third-grade results far exceeded the fourth-grade ones. This is likely contributed to the fact that it became focus for the 2021-2022 school year. Fourth grade likely didn’t grow as much due to the simple fact that it had less room to grow. To put this data in perspective with state data, the third-grade data is somewhat surprising.

The state had 43% of students meeting standard and 21% mastering standard. The district had percentages only slightly higher than the state in third grade unlike in fourth grade where the state had 43% meeting standard and 23% mastering standard in which the district was much higher than the state. This data could mean that the personalized learning model with the intervention support for teachers was higher for fourth grade since it was year two and therefore needs to continue for third grade for third grade to see greater gains in year two.

Research question two sought to answer if the intervention of job embedded professional development for the personalized learning model had any impact on the growth of fourth grade students. This question is of particular importance because the extreme low growth of only 27% in 2019 is a major contributing factor for the implementation of the PL model in the first place. It's also the first data set due to Covid-19 that reveals if the model has had any effect on the growth scores in this grade level. While these are two completely different cohorts of students and that has been discussed in the limitations of this study, it is still interesting to discuss the data. In 2019, only 27% of students met progress on STAAR and in 2022, 86% met progress on STAAR. The PL model by 2022 had been in place for the third year with a full year of intervention support for teachers. This data likely reveals that a major contributing factor the progress of fourth grade would be the PL model.

Research question three sought to answer if the intervention support for teachers contributed to gains in IXL software data. In 2020-2021 as discussed above, fourth grade saw an average of six months growth. In the 2021-2022 school year, fourth grade grew on average by eight months (from 388 to 470) and third grade grew on average by twelve months (from 272 to 390). When compared to the prior year, it would seem that the intervention support was a contributing factor the growth in IXL for both grade levels. While the results for fourth grade growth were a desirable result for the district, the third-grade ones were surprising in a positive way. The greater gains in third grade for IXL are also in line with the greater gains in third grade for the STAAR test as well. The gains in achievement in both the functional levels of the students as well as the overall grade level achievement can likely be contributed to the PL model and the intervention.

Research question four sought to answer if teachers felt the intervention for the PL model contributed to success in students and their ability to implement the PL model with fidelity. The qualitative results were a bit surprising in the fact that the observational data from semester one conducted by instructional coaches revealed that teachers were implementing the PL model as intended. The was an improvement from year one evaluation study where only a few components were being implemented which was also in alignment with prior research studies. However, the survey data after semester one revealed that teachers were not happy nor contributing their success in the model to the intervention. In particular teachers noted they were unhappy in the way it was being delivered (on Mondays after school). Improvement science frameworks are designed to be directly centered on the users. In this case, the user is the teacher. The first principle in improvement science is “make the work problem-specific and user-centered” (Bryk et al. 2017). While the work here in this dissertation in practice is problem-specific due to using data from assessments to pinpoint the work, it must also be user-centered. If the teachers who are carrying out the work are not satisfied with the “solution” then there must be a change. Therefore, in the middle of the intervention year the delivery of the intervention was changed to meet the needs of the user. The teachers had to be engaged in the process and the focus had to be shaped on how they carry out their work (Bryk et al., 2017).

Semester two intervention was switched from Mondays after school to pull-out days for teachers to plan with their grade level teams as well as the district math specialist twice. The subject and focus of the pull-out days were the same as the after-school sessions, however, they were extended periods of time due to them being all day as opposed to spread out in shorter chunks. As a result of this change for semester two, observations by coaches continue to show that teachers were having success in implementation of the model but the surveys from teachers also revealed that teachers enjoyed the pull-out days and contributed their ability to implement the model, as well as their students' success to these days. Job embedded professional development has been shown to increase teachers' ability to teach mathematics (Aulthauser, 2010) especially when the professional development is targeted to meet the teachers' needs (Thomas, 2008). The achievement data of the students in year two also aligns with other research around this intervention in that students in schools who support teachers with job embedded professional development in math perform higher than schools whose teachers do not (Brendefur, 2020). It seems that the intervention support for teachers led to an increase in their ability to implement the PL model which we saw from year one had a positive impact on student achievement, thus causing year two data to have even more significant gains as a result. In a study done to investigate the challenges of implementing PL models, one suggestion from the study was on-going professional development if the model was to be effective (Bingham et al., 2018). While that study was not done to see if it made an impact, this intervention study reveals that it most likely did.

Recommendations for Practice and Further Study

The district in which this two-year study took place is a STEM charter school who has a mission of students being able to enter college university ready to enter a STEM field. This typically means that students are able and ready to enter Calculus I. For students to be prepared to take Calculus I as dual credit their senior year of high school or as a college freshman, they must be prepared mathematically, and this begins at a young age. Many studies have examined the trajectory of math achievement over time. Such studies appear to confirm that early mathematics achievement is positively related to later mathematics achievement (Yeo et al., 2022). More so, mathematics trajectories of students in early grades may have important ramifications for STEM field entries later on (Shanley, 2019). It is imperative for the organization that math is a top priority and students are successful, so the charter adheres to its mission and purpose. Supports students with a strong foundation in the early grades is a crucial aspect of the long-term goal for students.

Based on the gains seen from year one evaluation to year two intervention in the math achievement of the third and fourth grade students, the most important next step should be to sustain what is working. The recommendation would be that these classrooms not only continue with the PL model as is but also to continue with the intervention of professional development in the format of pull-out days since it was well received by teachers. The recommendation would be to have built in days for the math teachers by grade level to work with the district math specialist before school starts as well as continue these days twice a semester throughout the school year. The days should be strategically aligned to the school calendar to be organized at times when planning can be centered on recent district assessment data.

The second recommendation would be to spread the intervention to additional grade levels in math, specifically grades K – 2 and 5 – Algebra 1. Seeing that the PL math model is already an expectation for the other grade levels in math, it makes sense to support these teachers in the same way as the third and fourth grade teachers based on the results of this two-year study.

In addition to expanding the grade levels in mathematics, a third recommendation would be to expand this intervention to the English Language Arts (ELAR) classrooms as well. The ELAR classrooms began their own PL models two years after the math classrooms did based on STAAR data and a new grant awarded by the state. Seeing that the intervention for teachers proved to be effective for the math teachers, it would be wise to offer the same support for the ELAR teachers in terms of job embedded targeted professional development.

As noted in prior research, PL models are likely not only not going away, but are likely to expand in the future (Zdeb, 2018). In the state of Texas, there continue to be grants promoted by TEA such as the Math Innovation Zones (MIZ) grant that this study was based on as well as other such ad Blended Learning Grants (BLG), Blended Learning Sustainability Grants and Raise Your Hand Texas that are all geared towards incentivizing schools to start blended learning or PL models of their own. As a result of the early findings in this two-year study, it would be recommended for other schools to do so. Schools who are interested, should start with what their data is telling them just as in this research design. Focus on the problem as well as the users involved. It would be encouraged to start small in a few grades spans and focus on one subject alone. The intervention of job embedded professional development should also be a component of any PL model that is started so teachers are supported throughout the process of implementation and should be supported by content specialist who are involved in the design of the PL model for the school.

Further research is also a recommendation for this study. While notable improvements were made over the course of the two years, the study is still “young” in terms of research and findings. Further research is needed to determine the effects after three years of implementation. The data could be collected over a five-year period, especially considering year one of implementation wasn't done to the full fidelity of the model. Third year data could be centered not only the intervention supports again of professional development but other aspects of the model that were lacking in year one and two such as students setting academic goals. Longitudinal data that tracks not only grade level data but students by cohort would be another area of interest for future studies. It would be important to know the trends of growth over time after students have been in the PL model for consecutive years as well as the teachers' comfortability in teaching under such a model after three plus years. Statistical regression studies would be another area of focus for possibly predicting trends between IXL and STAAR progress and eventually other forms of measurement outcomes such as students entering college STEM ready.

Conclusion

The two-year improvement science dissertation in practice was originally designed to evaluate a newly implemented grant program in the district, MIZ. The MIZ grant was awarded to the district in the summer of 2019 leading into the 2019-2020 school year. The district chose to apply for the grant for several reasons; 1) the grant aligned with the overall mission and vision of the district and 2) the progress from year to year on STAAR in math for the district was below desired results. After the 2019 data was released and only 27% of fourth grade students met progress on STAAR, fourth grade became the focus of the year one study evaluation. The PL model for the district was written to align to improvement science framework principles. A Network Improvement Committee (NIC) came together to analyze the data and this team was made up of stakeholders such as leadership but also centered on the users (the teachers). The PL model was written to contain four components mentioned earlier. Data was collected quantitatively in the form of IXL and STAAR scores and qualitative data was based on observations made by campus instructional coaches. Year one revealed that quantitative data did increase as seen on 2021 STAAR data and end of year IXL data, although perhaps not to the desired amount. The qualitative data revealed this may have been due to the PL model not being implemented to fidelity with all four components. Some of the contributing factors were deemed to be the after math of Covid-19 and the instructional challenges present in the 2020-2021 school year. It was determined at the end of the school year in year one that since the quantitative data improved, it would be wise to continue with the PL model for another year.

During the summer of 2021, the NIC met again to review the latest STAAR data as well as the IXL and qualitative data from year one. The NIC determined that more time was needed to see additional improvements in the student achievement data. Root causa analysis were performed as well as fishbone diagrams to address where the issues lied. It was after these improvement science tools were conducted that the NIC determined the best course of action for year two was to sustain the PL model but to provide an intervention to support teachers through it. Research was conducted and on-going job embedded professional development came to the forefront as the intervention of choice. Year two data showed improvements in fidelity to the model as seen in observations, but teachers were not completely satisfied in the chosen delivery as seen on semester one survey results. A tweak to the intervention was made for semester two but the intervention of job embedded professional development remained. At the end of year two gains were seen in both third and fourth grade classrooms across the district and teachers concluded they not only liked the newest delivery of professional development but wanted to see it continue for the following school year. More research is needed in order to determine the long term effects of this type of PL model, if any, on the mathematical achievement of students who were exposed to it in early grades and continuously exposed year after year.

References

Abaci, S., Robertson, J., Linklater, H., & McNeill, F. (2020). Supporting school teachers' rapid engagement with online education. *Education Tech Research, 69*(1), 29-34.

Adelson, J. L., & McCoach, D. B. (2011). Development and psychometric properties of the math and me survey: Measuring third through sixth graders' attitudes toward math- ematics. *Measurement and Evaluation in Counseling and Development, 44*(4), 225–247. http://dx.doi.org/10.1177/0748175611418522

Aleven, V., McLaughlin, E. A., Glenn, R. A., & Koedinger, K. R. (2017). Instruction based on adaptive learning technologies. In R. E. Mayer & P. Alexander (Eds.), *Handbook of research on learning and instruction* (2nd ed., pp. 522–560). Routledge.

Althauser, K. (2010). *The effects of a sustained, job-embedded professional development on elementary teachers' math teaching self-efficacy and the resulting effects on their students' achievement*. Eastern Kentucky University, Online Theses and Dissertations.

Bailey, D. H., Duncan, G. J., Murnane, R. J., & Au Yeung, N. (2021). Achievement gaps in the wake of COVID-19. *Educational Researcher*, *50*(5), 266–275. https://doi.org/10.3102/0013189X211011237

Bandura, A. (1991). Self-regulation of motivation through anticipatory and self-reactive mechanisms. Nebraska Symposium on Motivation. http://des.emory.edu/mfp/Bandura1991Nebraska.pdf

Bernhardt, V. L. (2018). *Data analysis for continuous school improvement (4th ed.).* Routledge.

Berry, R. Q. (2018). *Personalized Learning and mathematics teaching and learning*. National Council of Teachers of Mathematics (NCTM). https://www.nctm.org/News-and-Calendar/Messages-from-the-President/Archive/Robert-Q_-Berry-III/Personalized-Learning-and-Mathematics-Teaching-and-Learning/

Bingham, A. J., Pane, J. F., Steiner, E. D., & Hamilton, L. S. (2018). Ahead of the curve: Implementation challenges in personalized learning school models. *Educational Policy, 32*(3), 454–489. https://doi.org/10.1177/0895904816637688

Blank, R.K., & de las Alas, N. (2009). *Effects of teacher professional development on gains in student achievement: How meta analysis provides scientific evidence useful to education leaders*. Washington, DC: Council of Chief State School Officers.

Blazar, D. (2015). Effective teaching in elementary mathematics: Identifying classroom practices that support student achievement. *Economics of Education Review, 48*, 16-29.

Boaler, J. (2008). Promoting 'relational equity' and high mathematics achievement through an innovative mixed-ability approach. *British Educational Research Journal, 34*(2), 167–194.

Brendefur, J., Champion, J., Strother, S., Thiede, K. W., & Osguthorpe, R. D. (2021). The effects of mathematics professional development on elementary student achievement. *International Journal of Science and Mathematics Education*, *20*(1), 1079-1097. https://doi.org/10.1007/s10763-021-10184-z

Broussard, S.C., & Garrison, M. E. B. (2004). The relationship between classroom motivation and academic achievement in elementary school-aged children. *Family and Consumer Research Journal, 33*(2), 106-120.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2017). *Learning to improve: How America's schools can get better at getting better.* Cambridge, Massachusetts: Harvard Education Press.

Bulger, M. (2016). Personalized learning: The conversations we're not having. *Data and Society*, *22*(1). https:// www.datasociety.net/pubs/ecl/PersonalizedLearning_primer_2016.pdf

Bybee, R. W., & Fuchs, B. (2006). Preparing the 21st century workforce: A new reform in science and technology education. *Journal of Research in Science Teaching, 43*(4), 349–352. http://onlinelibrary.wiley.com/doi/10.1002/tea.20147/epdf

Campanella, M., & Penuel, W. (2021). Design-based research in educational settings. In Z. A. Philippakos, E. Howell, & A. Pellegrino (Eds.), *Design-based research in education: Thoery and applications* (1st ed., pp. 3-22). The Guilford Press.

Campbell, T., & Yeo, S. (2022). In defense of dichotomies in mathematics education research. *For the Learning of Mathematics, 42*(2), 51–52.

Carnegie Foundation for the Advancement of Teaching. (2015). *Our ideas.* https://www.carnegiefoundation.org/our-ideas/

Caruth, G. (2013). Demystifying mixed methods research design: A review of the literature. *Mevlana International Journal of Education, 3*(2), 112-122. http://dx.doi.org/10.13054/mije.13.35.3.2

Cave, A., & Brown, C. W. (2010). When learning is at stake: Exploration of the role of teacher Training and professional development schools on elementary students' math achievement. *National Forum of Teacher Education Journal, 20*(3), 1-21.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293–332.

Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth-grade achievement: Evidence from the ECLS-K. *Economics of Education Review, 28(*4), 415–427. https://doi.org/10.1016/j.econedurev.2008.09.003

Cleary, J. (2018, October 24). *The number one factor in student success? Relationships with teachers.* The Learning Sciences International. https://www.learningsciences.com/blog/number-1-factor/

Cleveland-Innes, M., & Wilton, D. (2018). *Guide to blended learning*. Commonwealth of Learning.

Cohen, J. (1992). Statistical Power Analysis. *Current Directions in Psychological Science, 1*(3), 98–101. https://doi.org/10.1111/1467-8721.ep10768783

Creswell, J. W., Plano Clark, V. L., Gutmann, M., & Hanson, W. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.),*Handbook of mixed methods in social and behavioral research* (pp. 209–240). Sage.

Croft, A., Coggshall, J. G., Powers, E., & Killion, J. (2010). *Job-embedded professional development: What it is, who is responsible, and how to get it done well*. National Comprehensive Center for Teacher Quality.

Desilver, D. (2017, February 15). *U.S. academic achievement lags that of many other countries*. Pew Research Center. https://www.pewresearch.org/fact-tank/2017/02/15/u-sstudents-internationally-math-science/

Dixon, J. K., Nolan, E. C., Adams, T. L., Tobias, J. M., & Barmoha, G. (2016). *Mathematics for teaching grades 3-5*. Solution Tree.

Drew, C. (2020, March 3). What is Finland's Phenomenon-based Learning approach? *Teacher*. Australian Council for Educational Research (ACER).

https://www.teachermagazine.com/au_en/articles/what-is-finlands-phenomenon-based-learning-approach

Earl, L., & Katz, S. (2002). Leading schools in a data-rich world. In K. Leithwood & P. Hallinger. (Eds.), *Second international handbook of educational leadership and administration.* (pp. 1003-1022). Kluwer Academics.

Eriksson, K., Helenius, O., & Ryve, A. (2019). Using TIMSS items to evaluate the effectiveness of different instructional practices. *Instructional Science, 47*(1), 1–18.

Ertmer, P. A. (2005). Teacher pedagogical beliefs: The final frontier in our quest for technology integration? *Educational Technology Research & Development*, *53*(4), 25-39.

Fensterwald, J. (2020). *Early data on learning loss show big drop in math, but not reading skills*. EdSource. https://edsource.org/2020/early-data-on-learning-loss-show-big-drop-in-math but-not-reading-skills/644416

Finnish National Board of Education. (2016). *New national core curriculum for basic education: Focus on school culture and integrative approach*. https://www.oph.fi/sites/default/files/documents/new-national-core-curriculum-for-basic-education.pdf

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education model. *The Internet and Higher Education, 2*(3), 87-105.

Gibbons, P. (1991). *Learning to learn in a second language*. Heinemann.

Gojak, L. M. (2013, February 5). *What's all this talk about rigor?* National Council of Teachers of Mathematics. https://www.nctm.org/News-and-Calendar/Messages-from-the-President/Archive/Linda-M_-Gojak/What_s-All-This-Talk-about-Rigor_/

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

Harbin, J., & Newton, J. (2013). Do perceptions and practices align? Case studies in intermediate elementary mathematics. *Education, 133*(4), 538-543. http://www.projectinnovation.com/education.html

Henry, G. T., Purtell, K. M., Bastian, K. C., Fortner, C. K., Thompson, C. L., Campbell, S. L., & Patterson, K. M. (2014). The effects of teacher entry portals on student achievement. *Journal of Teacher Education*, *65*(1), 7–23. https://doi.org/10.1177/0022487113503871

Herr, K., & Anderson, G. L. (2005). The continuum of positionality in action research. In Herr, K., & Anderson, G. L. *The action research dissertation : A guide for students and faculty* (pp. 29 - 48). Sage Publications, Inc.

Holmes, W., Anastopoulou, S., Schaumburg, H., & Mavrikis, M. (2018). *Technology-enhanced personalized learning: Untangling the evidence*. Robert Bosch Stiftung GmbH. http://www.studie-personalisiertes-lernen.de/en/

Ing, M., Webb, N. M., Franke, M. L., Turrou, A. C., Wong, J., Shin, N., & Fernandez, C. H. (2015). *Student participation in elementary mathematics classrooms: The missing link between teacher practices and student achievement?* Springer Science + Business Media Dordrecht.

IXL. (2019). Measuring the impact of IXL math and IXL language arts in Texas schools. *The IXL Effect.* https://www.ixl.com/research/Impact-of-IXL-in-Texas.pdf

Jansen, A. J. (2012). Developing productive dispositions during small-group work in sixth-grade mathematics classrooms: Teachers' facilitation efforts and students' self-reported benefits. *Middle Grades Research Journal*, *7*(1), 37-56.

Jonassen, D., Peck, K., & Wilson, B. (1999). *Learning with technology: A constructive perspective.* Prentice Hall.

Khirwadkar, A., Khan, S. I., Mgombelo, J., Obradovic-Ratkovic, S., & Forbes, W. A. (2020). Reimagining mathematics education during the COVID-19 pandemic. *Brock Education Journal, 29*(2), 42-46.

Kirschner, P., Sweller, J., & Clark, R. E. (2006). Why unguided learning does not work: An analysis of the failure of discovery learning, problem-based learning, experiential learning and inquiry-based learning. *Educational Psychologist, 41*(2), 75–86.

Kraft, M., & Falken, G. (2020). *A blueprint for scaling tutoring across public schools*. Annenberg Institute at Brown University. https://doi.org/10.26300/dkjh-s987

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the Potential Impact of COVID-19 School Closures on Academic Achievement. *Educational Researcher*, *49*(8), 549–565. https://doi.org/10.3102/0013189X20965918

Langley, G., Moen, R., Nolan, K., Nolan, T., Norman, C., & Provost, L. (2009). *The improvement guide: A practical approach to enhancing organizational performance* (2nd ed.). Jossey-Bass.

Lee, D., Huh, Y., Lin, C. Y., Reigeluth, C. M., & Lee, E. (2021). Differences in personalized learning practice and technology use in high-and low-performing learner-centered schools in the United States. *Educational Technology Research and Development*, *69*(1), 1221-1245.

Lubinski, D., Benbow, C. P., & Kell, H. J. (2014). Life paths and accomplishments of mathematically precocious males and females four decades later. *Psychological Science, 25*(12), 2217–2232. https://doi.org/10.1177/0956797614551371

Mainini, M. J.,& Banes, L. C. (2017). Differentiating instruction to increase conceptual understanding and engagement in mathematics. *Journal of Teacher Action Research, 4(*1), 80-100.

Marsh, J.A., Pane, J.F., & Hamilton, L.S. (2006). *Making sense of data driven decision making in education.* RAND Corporation.

Maxey, K. S. (2013). *Differentiated instruction: Effects on primary students' mathematics achievement* (Publication No. 3573708) [Doctoral dissertation, Northcentral University]. ProQuest Dissertations Publishing.

McInerney, D. M., Cheng, R. W., Mok, M. M. C., & Lam, A. K. H. (2012). Academic self concept and learning strategies: Direction of effect on student academic achievement. *Journal of Advanced Academics, 23*(3), 249–269. http://dx.doi.org/10.1177/1932202X12451020

National Center for Education Statistics. (2019). *Nation's report card: 2019 NAEP Mathematics Assessment*. National Assessment of Educational Progress. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education.

National Science Foundation (NSF). (2022). *The state of U.S. science and engineering 2022: U.S. and global STEM education and labor force*. https://ncses.nsf.gov/pubs/nsb20221/u-s-and-global-stem-education-and-labor-force

Nation's Report Card .(2022). *NAEP mathematics assessment: Largest score declines in NAEP mathematics at grades 4 and 8 since initial assessments in 1990*. National Assessment of Educational Progress. https://www.nationsreportcard.gov/highlights/mathematics/2022/

Ndiung, S., & Nendi, F. (2018). Mathematics connection ability and students mathematics learning achievement at elementary. *School SHS Web Conference, 42*(1), 1-5.

Nunez, L. (2011). *Setting goals in the classroom*. CE-Learning. http://celearning.ph/index.php/community/teachers/teaching-ideas/30- setting-goals-in-the-classroom

OECD. (2019). *PISA 2018 Results (Volume I): What Students Know and Can Do.* PISA, OECD Publishing, Paris. https://doi.org/10.1787/5f07c754-en

Odell, M. R. L., Kennedy, T. J., & Stocks, E. (2019). The Impact of PBL as a STEM School Reform Model. *Interdisciplinary Journal of Problem-Based Learning,13*(2). https://docs.lib.purdue.edu/ijpbl/vol13/iss2/4/

Odell, M. R. L., & Pedersen, J. L. (2020). Project and Problem-Based Teaching and Learning. In B. Akpan & T. J. Kennedy (Eds.), *Science Education in Theory and Practice* (Chapter 23, pp. 343-357). Switzerland: Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-030-43620-9_23

Pane, J.F., Steiner, E. D., Baird, M.D. & Hamilton, L.S. (2015). *Continued progress: Promising evidence on personalized learning,* RAND Corporation. https://doi.org/10.7249/RR1365

Pane, J., Steiner, E., Baird, M., Hamilton, L., & Pane, J. (2017). *How does personalized learning affect student achievement?* RAND Corporation. https://doi.org/10.7249/RB9994

Park, O., & Lee, J. (2004). Adaptive instructional systems. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 651–684). Lawrence Erlbaum Associates Publishers.

Pedaste, M., Maeots, M., Siiman, L., Jong, T., van Riesen, S., Kamp, E., Manoli, C., Zacharia, Z., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14*(2), 47-61. https://doi.org/10.1016/j.edurev.2015.02.003

Perry, J. A., Zambo, D., & Crow, R. (2020). *The improvement science dissertation in practice: A guide for faculty, committee members, and their students*. Myers Education Press.

Provasnik, S., Malley, L., Stephens, M., Landeros, K., Perkins, R., & Tang, J. H. (2016). *Highlights from TIMSS and TIMSS advanced 2015: Mathematics and science achievement of U.S. students in grades 4 and 8 and in advanced courses at the end of high school in an international context.* U.S. Department of Education, National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2017002

Salaschek, M., Zeuch, N., & Souvignier, E. (2014). Mathematics growth trajectories in first grade: Cumulative vs. compensatory patterns and the role of number sense. *Learning and Individual Differences, 35*(1), 103–112.

Sawchuk, S., & Sparks, S. (2020). *Kids are behind in math because of COVID-19: Here's what research says could help.* Education Week. https://www.edweek.org/teaching learning/kids-are-behind-in-math-because-of-covid-19-heres-what-research-says-could help/2020/12

Shanley, L., Biancarosa, G., Clarke, B., & Goode, J. (2019). Relations between mathematics achievement growth and the development of mathematics self-concept in elementary grades. *Contemporary Educational Psychology, 59*(1). doi:https://www.sciencedirect .com/science/article/abs/pii/S0361476X19304096?via%3Dihub

Shojania, K. G., & Grimshaw, J. M. (2005). Evidence-based quality improvement: The state of the science. *Health Aff (Milwood), 24(*1), 38-50. doi: 10.1377/hlthaff.24.1.138

Staker, H. (2011). *The rise of K-12 Blended learning: Profiles of emerging models.* Innosight Institute. https://www.christenseninstitute.org/wp-content/uploads/2013/04/The-rise-of-K-12-blended-learning.emerging-models.pdf

Staker, H., & Horn, M. B. (2012). *Classifying K-12 blended learning*. Innosight Institute. https://www.christenseninstitute.org/wp-content/uploads/2013/04/Classifying-K- 12-blended-learning.pdf

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360–407.

Sulecio de Alvarez, M., & Dickson-Deane, C. (2018). Avoiding educational technology pitfalls for inclusion and equity. *TechTrends*, *62*(4), 345–353. https://doi.org/10.1007/s11528018-0270-0

Sullivan, P., Bobis, J., Downton, A., Feng, M., Hughes, S., Livy, S., McCormick, M., & Russo, J. (2020). Threats and opportunities in remote learning of mathematics: Implication for the return to the classroom. *Mathematics Education Research Journal*, *32*(3), 551–559. https://doi.org/10.1007/s13394-020-00339-6

Svihla, V. (2014). Advances in design-based research. *Frontline Learning Research*, *2*(4), 35-45. https://journals.sfu.ca/flr/index.php/journal/article/view/114

Tchoshanov, M. A. (2011). Relationship between teacher knowledge of concepts and connections, teaching practice, and student achievement in middle grades mathematics. *Educational Studies in Mathematics*, *76*(2), 141–164. https://doi.org/10.1007/s10649

Texas 2036. (2021, January 14). *Texas education agency: Optional beginning-of-year assessments data help quantify initial impact of COVID-19 on Texas students*. https://texas2036.org/posts/tracking-covid-learning-loss/

Texas Education Agency. (2020). *Effective schools framework: A clear vision for what districts and schools across the state do to ensure an excellent education for all Texas students*. https://texasesf.org

Texas Education Agency. (2020). *State of Texas assessments of academic readiness end-of-year and beginning-of-year results*. https://tea.texas.gov/sites/default/files/covid/Supt-Call-Learning-Acceleration%20Slides.pdf

Texas Education Agency. (2020) *T-STEM Blueprint*. Blueprint. https://tea.texas.gov/sites/default/files/2020_T-STEM_Blueprint_2-13-20.pdf

Texas High School Project. (2015). *Texas science technology engineering and mathematics academies design blueprint, rubric, and glossary*. Texas Education Agency. https://www.texasccrsm.org/sites/default/files/1_-_2015_Blueprint_Final.pdf

Thomas, L. B. (2008). *Reflecting on practice: An exploration of the impact of targeted professional development on teacher action.* ProQuest: Dissertations & Theses Global. (304492523). https://ezproxy.uttyler.edu/login?url=https://www.proquest.com/dissertationstheses/reflecting-on-practice-exploration-impact/docview/304492523/se 2?accountid=7123

Thomas, P. (2010). *Towards developing a web-based blended learning environment*. University of Botswana, University of South Africa, Pretoria, http://hdl.handle.net/10500/4245

Van Schoors, R., Elen, J., & Raes, A. (2021). An overview of 25 years of research on digital personalized learning in primary and secondary education: A systematic review of conceptual and methodological trends. *British Journal of Educational Technology, 52*(1), 1798-1822. doi:10.111/bjet.13148

Villavicencio, F. T., & Bernardo, A. B. I. (2013). Positive academic emotions moderate the relationship between self-regulation and academic achievement. *British Journal of Educational Psychology, 83*(2), 329–340. http://dx.doi.org/10.1111/j.2044-8279.2012.02064.x.

Wang, J., & Goldschmidt, P. (2003). Importance of middle school mathematics on high school students' mathematics achievement. *The Journal of Educational Research, 97*(1), 3–17. https://doi.org/10.1080/00220670309596624

Webb, N. M., Franke, M. L., De, T., Chan, A. G., Freund, D., Shein, P., & Melkonian, D. K. (2009). 'Explain to your partner': teachers' instructional practices and students' dialogue in small groups. *Cambridge Journal of Education*, *39*(1), 49-70.

Wilson, S. (2017). Exploring the importance of using formative assessment in informing instruction to improve student learning. *Northwestern College, Iowa: Master's Theses & Capstone Projects.* doi:https://nwcommons.nwciowa.edu/cgi/viewcontent.cgi?article=1073&context=education_masters

Wood, F. H., & Killian, J. E. (1998). Job-embedded learning makes the difference in school improvement. *Journal of Staff Development, 19*(2), 52-54.

Wood, F. H., & McQuarrie, F., Jr. (1999). On-the-job learning. *Journal of Staff Development, 20*(3), 10-13.

Xie, H., Chu, H. C., Hwang, G. J., & Wang, C. C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education*, *140*(1), 103599. https://doi.org/10.1016/j.compedu.2019.103599

Yeo, S., Shim, H., Hwang, S., & Campbell, T. (2022). Relations between elementary mathematics growth trajectories and middle school mathematics achievement:

Moderating effects of student perception of teaching styles. *International Journal of Science and Mathematics Education*. doi: https://link-springer com.ezproxy.uttyler.edu/content/pdf/10.1007/s10763-022-10332-z.pdf

Zdeb, N. (2018, March 22). The truth about personalized learning. *Teach. Learn. Grow. The Education Blog.* https://www.nwea.org/blog/2018/the-truth-about-personalized-learning/

Zepeda, S. J. (2019*). Professional development: What works (3rd ed.)*. Routledge.

Zimmerman, B., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Research Journal, 29*(3), 663-676.