2015

A Program Evaluation on Implementing Investigations in Number, Data, and Space® in Three Elementary Schools

Alexandra Leigh Smith
Gardner-Webb University

A Program Evaluation on Implementing Investigations in Number, Data, and Space® in Three Elementary Schools

By
Leigh Smith

An Applied Dissertation Submitted to the
Gardner-Webb University School of Education
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Education

Gardner Webb University
2015

Approval Page

This applied dissertation was submitted by Leigh Smith under the direction of the persons listed below. It was submitted to the Gardner-Webb School of Education and approved in partial fulfillment of the requirements for the degree of Doctor of Education at Gardner-Webb University.

E. Ray Dockery, Ed.D. Committee Chair	Date
Steven Bingham, Ed.D. Committee Member	Date
Dixie Abernathy, Ed.D. Committee Member	Date
Jeffrey Rogers, Ph.D. Dean of the Gayle Bolt Price School of Graduate Studies	Date

Abstract

A Program Evaluation on Implementing Investigations in Number, Data, and Space® in Three Title I Elementary Schools. Smith, Leigh, 2015: Applied Dissertation, Gardner-Webb University, Investigations/Elementary School/Title I/Mathematics Programs

This applied dissertation was designed to provide perceptual teacher data as well as summative testing data to educational leaders concerning the effects of implementing Investigations in Number, Data, and Space® (Investigations) in three Title I elementary school settings, two Title I schools, and one non-Title I school. Data collected during this dissertation will be of use to educational stakeholders in selecting mathematics programs for elementary age students.

The purpose of this applied dissertation was to assess the effects of the Investigations program utilizing Stufflebeam's CIPP program evaluation model. End-of-grade math test data for third, fourth, and fifth grade from the 2010-2011 to 2013-2014 school years in a southeastern school district were analyzed along with teacher perceptual data.

Teacher perceptions of the effectiveness of Investigations were measured by a survey developed by the researcher. Specific process and product research questions asked, "What are the teacher perceptions about the impact of Investigations on student achievement," "What were the unanticipated effects of the Investigations program on student academic development," and "What are the teacher perceptions about any unanticipated effects of Investigations on student academic development?"

The survey data indicate that more than half of the teachers in the researched school district believed their opinions were not used in the selection of materials to implement Balanced Active Math strategies and the trainings offered did not adequately prepare them to deliver the Investigations program. All three schools dropped in proficiency following Investigations implementation in the 2012-2013 school year and increases in proficiency rates in the second year of implementation.

Table of Contents

Figure

Chapter 1: Introduction

Statement of the Problem

A Nation at Risk was first published in 1983 by the United States Department of Education (National Commission on Educational Excellence, 1983). This report called American educators to awaken from a slumber of mediocrity concerning the quality of education provided to students from kindergarten to twelfth grade. Educators were not preparing students who successfully graduated high school, were literate, and able to perform adequately in the American workforce. Math data collected in the report showed that in 1983, only four of every 20 students were proficient in math. More recently, a 25-year review of the progress of American education was published in 2008. In *A Nation Accountable: Twenty-five Years after a Nation at Risk,* eight of 20 students were considered proficient in math (United States Department of Education, 1997).

According to data analyzed by the Trends in International Mathematics and Science Study (TIMSS) collected by the National Center for Education Statistics (2011), student achievement data on standardized tests in fourth and eighth grade had increased from previous data collection years of 1995 and 2007. Even with an increase from previous years in proficiency, the United States Department of Education Secretary Arne Duncan stated that “While student achievement is up since 2009 in mathematics it’s clear that achievement is not accelerating fast enough for our nation’s children to compete in the knowledge economy of the 21st Century” (National Center for Education Statistics, 2011, p. 42).

National Assessment of Educational Progress (NAEP) data are collected periodically as an assessment tool to measure American student progress in various academic areas such as reading, mathematics, writing, science, and the arts. After significant NAEP gains in the 1990s, particularly in mathematics, the 2011 results continue a pattern of modest progress (NBC News Services, 2011). In examining the test scores assessed by NAEP, only 40% of fourth graders and 35% of eighth graders reached proficiency according to NAEP standards.

The landmark publication, *A Nation at Risk*, first collected education data in 1983 (National Commission on Educational Excellence, 1983). Sponsored by the U.S. Department of Education, the report found disturbing data concerning the performance of American schools. Reassessing school performance in 1997, the commission gathered and examined data in the following areas: curriculum content, standards and expectations, time, teacher quality, and leadership and financial support (United States Department of Education, 1997). Results in student proficiencies are reported to have only made slight gains in proficiencies since 1983. Of children born in 1983, only 20% would be proficient in reading and 4% would be proficient in math. Compared to students born in 2007, 7% of students were proficient readers and 8% of students were considered proficient in mathematics.

Comparison of American student proficiencies to their international counterparts was also discussed at length in this report. According to the study, international students are outperforming American students in both reading and mathematics. When data were collected and reexamined in 1997, the United States was found to have slipped to tenth place in the number of high school graduates it produces each year (United States Department of Education, 1997).

One important point to be made is that the United States Department of Education, publisher of *A Nation at Risk*, included the total student population that is testing in American schools. Students in American schools are tested regardless of poverty level, special needs status, or limited English proficient. At best, it is unclear whether international schools that are compared to American schools include such student groups in their testing. In one southeastern state, the number of students who participated in end-of-grade (EOG) or end-of-course (EOC) tests who had a disability totaled over 14% of students tested for the 2009-2010 school year (North Carolina Department of Public Instruction, Accountability Services, 2011).

On the local level, the researched school district's mathematics data reported through the EOG tests reflect the following student proficiencies from the 2010-2011 school year.

Table 1

Math Proficiency—Overall

Researched School District	Grade 3	Grade 4	Grade 5
Math proficiency	80.0%	81.8%	80.1%

Student performance broken down by gender and ethnicity who passed math and standardized tests reveal that student performance in minority subgroups is lower than that of White students. Students with Disabilities group has the lowest proficiency scores of any group listed.

Table 2

Math Proficiency—Subgroups

Researched School District	Male	Female	White	Black	Hispanic	Economically Disadvantaged	Students with Disabilities
Passed Math	62.3%	66.3%	71.3%	44.7%	55.2%	53.1%	22.8%

The overall purpose of the research school district is to provide rigor, relevance, and relationships leading to a student who is globally competitive. The mission of the district is to provide quality educational opportunities to ensure student success and a lifetime of learning. The researched school district put forth strategic goals for the years 2011-2014. First on the list of goals was to have students demonstrate competency in reading, math, and science. In 2012, district-wide data indicated that 80% of students were proficient on EOG testing in mathematics for Grades 3-8. The goal by 2014 was 86%. Proficiency goals are of special interest to the researcher as they provide a context of school system priorities and implementation of math initiatives such as Investigations in Number, Data, and Space® (Investigations). While the combined mathematics proficiency stated above may not raise red flags to those outside of education, it is important to note that individual schools have varying proficiency scores. This is especially true of schools that receive federal funding through the Title I program. In the district, schools that have an overall average of 65% of its students who receive free or reduced lunch qualify for additional funding. In the researcher's school, 2011-2012 EOG data report great disparity from the overall district proficiency. Research school proficiency numbers are listed below.

Table 3

Math Proficiency—Grade Level

Grade 3	Grade 4	Grade 5	Overall Proficiency
62.6%	70.3%	64.4%	65.6%

Note. Information obtained from 2011-2012 North Carolina School Report Card.

Student demographics are also of interest to the researcher. The district is among the top 10 school districts in the state according to student population. Of these students, more than 3,990 students receive special education services and 1,235 receive second language programs.

The researched school district employs over 3,500 people to educate and serve more than 30,000 students in Grades Prekindergarten through 12. In total, 55 schools serve the needs of students; 30 are elementary and primary, two are intermediate, 11 are middle, 10 are high, one is a special needs school, and one is an alternative behavior school. Elementary schools serve Grades Kindergarten through 5. Primary schools serve K-2 students, while intermediate schools serve Grades 3-5. Middle schools serve students in Grades 6-8, while high school finishes the spectrum from Grades 9-12.

Student ethnicity is varied in this school district with 64.7% of students being Caucasian, 20.3% being African-American, and 9.5% being Hispanic. The researched district also provides free or reduced meal benefits to 59% of its students as well as transports 7,360 elementary students via school buses.

Students had an attendance rate of 95.2% during the 2011-2012 school year as well as a graduation rate of 78.8% in 2012. Special education services are provided to 9.1% of the student population. Students identified as Academically Gifted and Talented make up 8.2% of the population. Additionally, 27 Advanced Placement courses are offered in the researched school district with a total Advanced Placement enrollment of more than 3,200 students.

Increases and decreases in math achievement scores can be related to a myriad of possible causes. One such reason could be the inclusion of more students with disabilities in the testing program. From the testing year 2010-2011, in North Carolina alone, more than 33,000 students with disabilities participated in the testing program in Grades 3-10 (North Carolina Department of Public Instruction, Accountability Services 2011). Students who have diagnosed differences in learning are held to the same standard as their nondisabled peers. Student test scores may not be as high as their peers therefore reflecting poorer performance when looking at the total number of students who are considered proficient.

The Topic

The topic of this dissertation was a program evaluation that examined aspects of the implementation of Investigations into the researched school district as the stand-alone math instructional tool. The effects of this implementation on EOG or EOG scores given in May 2014 were examined. Staff of three elementary schools in the researched school district were given a survey to ascertain their perceptions of the implementation of Investigations. Two of the research schools are considered Title I schools, and one is considered a non-Title I school.

The Research Problem

Math achievement scores have been decreasing in the researched school district for several years. This trend mirrors slow increases in student proficiency as well as state and national decline in overall math scores.

To remedy this effect, instructional specialists in the researched school district have mandated implementation of Investigations in all of its elementary schools. This study conducted a program evaluation to determine the impact of the implementation of Investigations on EOG test scores as well as teacher perceptual data of the effects of implementation.

Background and Justification

Information gathered from the 2011-2012 North Carolina Report Card indicated students in the researched school district are being outperformed by their peers across the state of North Carolina. The tables below indicate that in Grades 3-8, the researched district is below the state average in math performance on EOG tests as well as in reading. When examining data based on subgroups of students at the district level, the research district underperformed in the following subgroups: All, Male, Female, White, Black, Asian, Economically Disadvantaged, Not Economically Disadvantaged, and Students with Disabilities.

Table 4

Percent Proficient on ABC EOG Tests—Mathematics2011-2012

Grade	District	State
Third	80.0%	82.8%
Fourth	81.8%	85.1%
Fifth	80.4%	82.1%

Table 5

Percent Proficient on ABC EOG Tests—Subgroups—Mathematics 2011-2012

	All	Male	Female	White	Black	Hispanic
District	64.3%	62.3%	66.3%	71.3%	44.7%	55.2%
State	67.5%	65.0%	70.1%	79.3%	49.4%	55.1%

Theoretical Framework

"Evaluation is a very young discipline—although a very old practice," said Scriven (1996) in describing the history of program evaluation (p. 393). Organizational stakeholders and decision makers want and need to verify that programs are accomplishing their stated purposes. To that end, questions must be asked from various contexts of the implementation process to evaluate effectiveness. Processes, procedures, and outcomes must all be inspected by examiners. Program evaluation can also judge the merit or worth of something (Scriven, 1991).

Program evaluations can involve ongoing monitoring of programs or one-time evaluations of students of processes, outcomes, and program impact (Stufflebeam & Shinkfield, 2007). Formative evaluations can help to strengthen or improve a program and help to determine what works best in an organization. Additionally, formative evaluations can help provide feedback for improvement while shedding light on any negative results of an implemented program.

On the other hand, summative evaluation examines the overall quality and outcome of a program. Summative evaluations are designed for decision-making purposes to ascertain if a program has met its planned outcomes. Both formative and summative evaluations are needed in the development of a product or service

(Stufflebeam & Shinkfield, 2007).

Once program evaluations have been conducted, assessors can also use results to plan effective staff development for areas of programming that need to be strengthened (Centers for Disease Control, 2010). Furthermore, results from program evaluations can also help to celebrate successes within the program itself as well as strengthen the program design through rigorous examination (Alleghany Evaluation Specialists, 2014).

Deficiencies in the Evidence

To date, no assessment of the Investigations mathematics program has been conducted as it impacts the North Carolina EOG test scores. This study examined the effect as well as assessed teacher perceptions of the planning and implementation process of Investigations into teacher classrooms.

Audience

Practitioners in education in the elementary school setting as well as curriculum leaders in the district will benefit from the results of the study. The results of this study could be helpful to schools and districts in implementing future mathematics programs. It is also important to consider teacher viewpoints of the effectiveness of mandated educational programs such as Investigations before, during, and after implementation.

Definition of Terms

Investigations. The instructional program used by the researched school district to instruct students in Grades K-5 in mathematics instruction.

EOG tests. The assessments students in North Carolina take in May of every academic year in order to assess their proficiency in reading and mathematics in Grades 3-8.

Balanced Active Math (BAM). An approach to teaching mathematics that follows a constructivist approach whereby students build upon their own level of understanding with mathematical foundations often in cooperative groups where more than one method can be used to derive answers. Students often work in small cooperative groups during this time.

Title I schools. Schools that receive additional federal money through the Department of Education based on the number of students who qualify in each school district for free and reduced lunch. Each school district can set its own threshold for schools qualifying for these funds.

Context, Input, Process, and Product (CIPP). The program evaluation model developed by Stufflebeam (2011) to help the public and private sector evaluate the effectiveness of programs based on four quadrants.

Purpose of the Study

The purpose of this study was to determine the impact of Investigations on academic achievement, teacher perceptions of the implementation, and effectiveness upon EOG tests scores at the researched site. The implementation of the Investigations program has been mandated as the stand-alone math strategy and resource to be used by elementary school teachers within the school district where the study was located. This study also examined teacher perceptions of the implementation of Investigations using the CIPP model of program evaluation.

Research Questions

This program evaluation utilized Stufflebeam's CIPP model to determine the impact of implementation of Investigations on student achievement at three elementary schools in the southeastern United States as measured by achievement data in the third grade. Teacher perceptions were measured by a survey developed by the researcher.

Research questions addressed in this program evaluation center on the Process and Product components of the CIPP program evaluation model.

Process Evaluation Questions

Were the various components of Investigations implemented as they were originally intended?

a. What are the teachers' perceptions about the implementation of strategies and activities within Investigations?

b. How did teachers have an opportunity to ask questions and voice concerns during the implementation stage?

c. How were any program adjustments made by teachers during implementation?

Product Evaluation Questions

1. Based on EOG data from 2010-2011 through 2013-2014, what impact did Investigations have on student achievement through proficiency scores in Grades 3-5?

    a. What are the teacher perceptions about the impact of Investigations on student achievement?

2. What were any unanticipated effects of the Investigations program on student academic development?

3. What are teacher perceptions about any unanticipated effects of Investigations on student academic development?

To answer these questions, the researcher conducted a quantitative study using EOG assessment scores from the third grade from the 2012-2013 school year. The usage of Investigations in teacher classrooms is mandated for all elementary teachers who teach

Grades K-5.

Chapter 2: Literature Review

Importance of Periodic Assessment and Evaluation

In the world of education, periodic assessments can often look like teachers giving students common assessments which cover curriculum that has been previously taught in the classroom. Periodic assessments are important to teachers and managers as they can provide regularly scheduled feedback to improve performance (Thompson, 2011). In a similar manner, the evaluation of programs can also provide crucial feedback to program managers. In both the academic and nonacademic worlds, program evaluation has proven to be a valuable tool in strengthening the quality of existing programs (Behavioral and Social Science Volunteer Program, 2012). Benefits to conducting periodic assessments and evaluations include evidence of effectiveness and justify the need for more support of the program, increasing the program's ability to contribute to the knowledge of the field, improving upon skills and quality of the program, streamlining services, and promoting the effectiveness of the program (Behavioral and Social Science Volunteer Program, 2012).

Effective program evaluation also collects, provides, and analyzes data that can be used to learn about the program itself and any strengths and limitations (Centers for Disease Control, 2010). Routine program evaluation can offer improved documentation, learning opportunities, and common understanding about functions of a program that are successful and those that are not. Evaluations can also promote accountability within institutions, solidify requests for increased funding, identify ineffective practices, and produce credibility to outside agencies (Pell Institute, 2014).

Within the educational setting, conducting frequent assessments and evaluations is one of the methods in which educators can keep up with the ever-changing and multifaceted endeavor of teaching. Leaders in the field must be able to confront and challenge current methods and be prepared to change various parts of a program to develop more effective implementation (Ross, 2010). Leaders must also be versed in multiple approaches to evaluation in order to accommodate varied sources of data and purposes of evaluation.

Effective assessments and evaluations can be designed as formative or summative. Formative evaluation provides information that forms and refines the program. A review of practices and their interpretation is one crucial step in formative evaluation (CDC, 2010). Data collection and research on staffing, training, materials, and implementation processes can provide managers and educators with information to improve upon the program. If provided in a timely manner, the data can facilitate making corrections and adjustments that can refocus a program.

Summative evaluation can be most helpful in clearly defining the benefits created by a program as well as the costs and conditions necessary for maximum effectiveness. When using summative evaluation, educators and managers must look at their measurement and data collection tools to ensure the correct tools are used to report the most effective data.

Program Evaluation

The field of program evaluation began in the mid-1900s as a way to "judge the worth or merit of something or the product of the process" (Scriven, 1991, p. 139). Defined by seven time periods of evolution, program evaluation has been used by businesses, government agencies, and the private sector to assess the effects of programs.

In determining effectiveness of new ideas in the workplace, program evaluations have been used to obtain reliable, valid, and credible data (Scriven, 1991) in order to judge the performance of programs. Hogan (2007) described five evaluation approaches to program evaluation that are in current use by practitioners. Objectives-orientated approaches focus on setting clear goals and objectives of the given program and describe the degree to which the goals have been attained. The recognized pioneer in this approach is Ralph Tyler. Tyler (1949) stated that the goals and objectives of any program must be defined in order for evaluation to take place. While considered to be the pioneer, Tyler was not without his critics. Critics claimed that the selection of objectives for evaluation was faulty as not all objectives could be evaluated. In addition, selecting objectives to be evaluated was also open to bias (Stufflebeam & Shinkfield, 1985).

Expertise-oriented program evaluation is the most widely used and oldest method to judge an institution, activity, or program. A panel of judges or experts evaluates a program and makes recommendations based on their perceptions and opinions. A formal and informal review of internal systems can be used in this approach. Critics of this approach claim that judgments made by experts are biased and not based on program objectives (Worthen, Sanders, & Fitzpatrick, 2004).

Judicial processes are used when utilizing the adversary-oriented approach. Pros and cons of any issue are debated by two teams who then defend their positions in a public debate until an agreement can be made on a common position. Within this evaluation system, hearings, prosecutions, juries, charges, and rebuttals are integral components. By using this approach, positive and negative viewpoints are brought into the open; however, the truth is believed to emerge from a hard but fair fight (Worthen et al., 2004).

Experience plays a huge part in the participant-oriented approach which values firsthand experiences with activities as well as the importance of the participant in the process. The least powerful stakeholders are used in this process from start to finish wherein the evaluator and stakeholder work alongside each other as partners to solve problems.

Empowerment evaluation is considered a subclassification within participative-oriented evaluation. Using the empowerment approach, participants develop a clear purpose, identify program strengths and weaknesses to assess where the program currently stands, and plan for the future by establishing goals.

Based on the four aspects of training—context, input, reaction, and outcome—the CIRO model was proposed in 1970. Context evaluation identifies an organization's training needs and the setting of goals and objectives. Input evaluation is focused on the design and delivery of the training activity. The CIRO model also takes into account the objectives and training materials that are to be used in the evaluation. Reaction evaluation examines the quality of the trainees' experiences, while outcome evaluation highlights achievements gained from the activity that is assessed at three levels: immediate, intermediate, and ultimate evaluation.

The management-orientated approach was intended to be utilized by organizational leaders in providing information to decision makers on the managerial level. Stufflebeam (2011) created the widely used management-orientated tool, CIPP model, to evaluate programs. This model is intended for service providers ranging from university administrators, physicians, and military leaders when conducting internal evaluations for examination of the social acceptability, cultural relativity, and technical adequacy of programs (Stufflebeam & Shinkfield, 2007). Context evaluation investigates the program objectives to determine if they are acceptable given social, cultural, and technical characteristics. Input evaluation examines the intended content of the program. Process evaluation relates to what degree the program was delivered as it was planned. Finally, product evaluation assesses program outcomes.

CIPP evaluations are conducted to complement rather than supplant other reviews of existing programs (Stufflebeam, 2011). Throughout the evaluation process a meta-evaluation, or evaluation of an evaluation, is completed. The model's main theme is that the evaluation's most important purpose is not to prove but to improve.

History of Mathematics Reform

As the 20th century dawned, American culture saw a change in its character. It was during this time that the works of Thorndike (1923) called upon school psychologists to make schools more efficient and effective in educating large populations of children (Ellis & Berry, 2005). Through his research, Thorndike believed that drill and repeated practice were the best mathematical practices to instruct children in mathematics. Furthermore, Thorndike called mathematics a "hierarchy of mental habits or connection" (p. 52). As such, mathematics should be clearly taught through carefully planned sequences with much repetition so as to enable learning.

The Progressive Education Association or (PEA) entered the American educational forum in the early 1920s as a counter measure to Thorndike's (1923) call for rote memorization of skills and the use of repetition. Influenced by the works of Dewey (1961), PEA favored providing direction to children learning activities. Learning then, according to the progressives, occurs when it is connected to student experiences and interests (Przychodzin, Marchand, Martella, & Azim, 2004). Student interests should also be a factor in developing instructional practices in the classroom. Furthermore, progressives also called for the role of the teacher to change from a taskmaster to one of facilitator or guide.

Fast forward to the mid-20th century where the New Math phenomenon was developed out of concerns of the Russians launching the satellite, Sputnik, into space prior to the Americans. In response, Congress created the National Science Foundation (NSF) in 1950 to help promote science education in the United States. From the funding provided through the NSF, numerous projects set their sights on overhauling mathematics education. Developed programs contained strategies such as usage of student manipulatives, intensive in-service workshops for teachers, the development of textbooks heavily influenced by early constructivist thoughts, and the creation of Advanced Placement testing by the College Entrance Examination Board for advanced students with mathematical aptitudes. These new approaches failed to gain pervasive success in American classrooms; however, they were beneficial to the next generation of educators as they laid the groundwork for future reform (Klein, 2003).

As a pushback against the New Math movement, Back-to-Basics was launched in the early 1970s. Proponents called for the simplification and orderly development of mathematical skills. This movement was connected closely with the competency test movement in American education in the 1970s and 1980s. Modest improvements were seen in test scores; however, critics espoused that the Thorndike-like math textbooks did little to prepare students for higher levels of cognition and understanding (Wilson, 2003).

In review, many of the revisions of mathematics education formulated over the past century have been created within the procedural-formalist paradigm (Ellis & Berry, 2005). The procedural-formalist paradigm asserts that mathematics is a set of organized facts, skills, and procedures that exist apart from human experience which in turn make it difficult to learn. In stark contrast, the cognitive-cultural paradigm believes that all students can learn interconnected concepts that come from human experience as long as they are presented in a culturally relevant way.

Sensing a need to influence change in American mathematics, the National Council of Teachers of Mathematics (NCTM, 1989) published updated standards. With the implementation of these new goals, students would be able to apply knowledge to new situations, explain mathematical arguments, and make sense of conceptual connections.

Burrill (1998) explored implications of the NCTM standards on mathematics curriculum reform by reviewing the changes that have occurred in mathematics education, myths about mathematics, and the mathematics that children like to do, and where we are headed given the tremendous changes in technological advances of the 20th century (Jackson, 1998). Burrill saw the need to create a curriculum that flows from various grade levels into one coherent whole in which students are expected to have a shared common knowledge base by a given grade level and in which teachers act on this expectation. A curriculum designed in this vein, Burrill argued, will reduce the emphasis on the repeat and remediation cycle in which today's mathematics education is often embedded, allow for a broader and more useful base of mathematics to be explored in the classroom, and make mathematics consistent across grade levels nationally.

With the increase in usage of technology such as virtual classrooms and online tutoring, the questions of which of the programs will best serve students becomes more difficult. The technology of today provides numerous avenues for differentiating instruction that can both teach the logically sequenced set of mathematical skills as well as provide for engaging group discussions and project-based learning opportunities for students (Kuhn & Dempsey, 2011). The push for an increase in technology integration will no doubt influence future reforms in mathematics.

Cooperative and Constructivist Instructional Design

People construct their own reality based on their experiences and knowledge. Constructivism provides an understanding of learning where individuals create new understandings or knowledge sets based on interaction with ideas, activities, and events in their daily lives. Teachers then serve as guides, co-explorers, and facilitators who encourage students to question, challenge, and formulate their own ideas, opinions, and conclusions (Cannela & Reif, 1994). It is important to note however that "Constructivism is not a theory about teaching . . . it is a theory about knowledge and learning . . . the theory defines knowledge as temporary, developmental, socially and culturally mediated, and thus non-objective" (Brooks & Brooks, 1993, p. 8).

Dewey (1961) was a major force in progressive education in the United States in the early to mid-20th century. Dewey's work led the way for other researchers such as Abraham Maslow, Carl Rogers, Lev Vigotsky, and Jean Piaget. All of these thinkers had their unique perspective of human development; however, they all shared Dewey's belief that education naturally facilitates the developing tendencies and potential of each child (Matthews, 2003). From Dewey's perspective, knowledge is not a representation of reality. The relationship between knowledge and reality is the result of individual and social experiences. Enriched experiences change people's perception of right. Classroom teachers understand this theory well as plan field trips for students to zoos, courthouses, capitals of states, and industries so they can experience and internalize life situations that may not be feasible under ordinary circumstances.

Self-direction in both adult and child learning was also of importance to Dewey (1961). He believed that active participation and self-direction were crucial to student success. Dewey believed the "contents of the child's experience" are more important than the "subject-matter of the curriculum" (p. 342).

Piaget (1953), another pioneer of constructivist theory, centered his main focus on constructivism around how the individual builds knowledge. According to Piaget, the nature of knowledge should be studied empirically through experimentation of learners in their natural environments such as schools and homes. Humans cannot be given information they immediately understand and use; instead humans must construct their own knowledge (Piaget, 1953). According to Piaget, three kinds of knowledge exist that are used to structure and build knowledge: physical, social, and logico-mathematical. Physical knowledge is knowledge of objects in external reality. Social knowledge includes knowledge of certain social norms such as Father's Day or saying "good morning" under specific circumstances. Logico-mathematical knowledge contains relationships formed by each person.

In terms of mathematical theory and instruction, Kamii (1996) believed the traditional goal of memorizing facts in order to internalize sums is incorrect. Kamii asserted that mathematical sums must be internalized by each child on the inside. Methods of classroom practices should not then include superficial mastery of concepts through repetition and reinforcement from external sources. Instead, students should be exposed to numerical reasoning through daily life experiences, group games, and problem-solving discussions. Repetition is important; however, it should be accomplished through games where students are motivated to learn arithmetic. Group work is also a foundational cornerstone of effective mathematics instruction as Piaget

(1971) pointed out that exchange of ideas and points of view are essential for intellectual and socio-moral development.

Leading Piaget student, Kamii (1994) believed so strongly in cooperative instruction that she called for the end of teaching carrying and borrowing in first- through fourth-grade classrooms. In examination of schools through Hoover, Alabama, City Schools, Kamii (1996) compared the processes in which students answered algorithms in classrooms that did and did not teach the direct instruction (DI) approach of carrying and borrowing to procure sums. Students in each classroom were heterogeneously placed for ability. Two hundred and twenty students and their algorithms were examined during the study.

In summation of her research, Kamii (1994) asserted that students who used traditional algorithms to answer questions were more likely to answer the question incorrectly but could also not articulate how the numbers were related to each other and why they had to borrow and carry. Teaching algorithms is harmful, asserted Kamii (1996), because they do not allow for children to develop their own thinking. Algorithms remove the knowledge of place value children have already constructed, which in turn prohibits them from developing number sense (Kamii, 1996).

The concept of adaptation is also pivotal to the constructivist learning theory. Piaget (1971) believed that assimilation occurs when children bring new knowledge to their own schemas or experiences and accommodation occurs when children have to change their schema to accommodate the new knowledge or information. As the learner processes how to fit new information into existing memory files, this adjustment process occurs (Powell & Kalina, 2009).

Fellow Piaget student, Duckworth (1995) also believed in the constructivist approach to teaching and learning through student exploration of developing meaning of instructional materials. In her text, Duckworth asserted that students should be made to feel wonderful about their ideas and the process of developing them to their understanding. Children should be exposed to a number of new ideas and theories to gain their attention whereby they can create their own understanding and meaning. A predetermined pace of intellectual development is not found in children, declared Duckworth, as students develop understanding based upon their experiences, actions, and connections.

Teachers, who are facilitators, present information to students; however, they do not assign meaning to the material. Students must be placed in situations where they develop their own understanding. As facilitators, teachers must present broad ideas to students, not just narrow goals and objectives (Meek, 1991). Teachers should also occasionally disagree with student viewpoints in order for a deeper level of thinking to occur by the student. These disagreements must be made respectfully; however, they can lead a student to accept another's viewpoint as interesting and thereby increase understanding of a topic.

To delve deeper, teachers must also refrain from providing hidden meanings or signals that could interfere with a student's development of meaning and context of curriculum (Fusaro, 2014). As a result, teacher knowledge of the subject matter is increased as they try to take student thoughts and deepen them based on their own understanding.

Montessori schools bear the name of the first woman admitted to practice medicine in Italy, Maria Montessori. Trained as a physician, Montessori originally developed her program to assist children with various health disorders. These programs challenged the traditional classroom model of students sitting at desks and memorizing facts. In Montessori models, students were given the opportunity for movement and interaction in a structured manner that supported students' natural curiosity. Social skills, academic lessons, exercises in daily living, and concern for health, hygiene, and self-discipline were all fundamental components of the Montessori experience (Hedeen, 2005).

The cooperative aspect of the constructivist classroom is essential. Duff (2012) found that students prefer cooperative and constructivist components of instruction. Such components can include working in teams on an assignment or project and making team members accountable for the content and degree to which the project is completed. Duff conducted a survey of middle school students in which she asked the students if they felt their achievement scores would improve with more or less group work time. A total number of 15 students were in the sixth-grade classroom. Seven boys and eight girls ages 11-12 made up the study. According to questionnaire results, students believed their achievement would increase after group learning techniques in part due to their belief that the approaches used allowed them to understand the presented material better. Duff extrapolated that students benefited from cooperative learning strategies because they allowed students to make real world connections using learning styles and group work.

Student perception of effective math instruction techniques can also play a critical role in motivating students to participate fully in mathematics applications. The National Research Council—Mathematics Learning Study Committee released a report in 2001 that recommended teachers of mathematics use a mixed-methods approach to engage students in the five integrated competencies of conceptual understanding, strategic competency, adaptive reasoning, productive dispositions, and procedural fluency. By considering learning styles, work completion rates, modes of expression, and student-centered techniques, students are predicted to be able to attend to tasks at a higher rate as well as process information to a higher extent. The council further recommended that a mixture of both DI and cooperative approaches be used to instruct students. Strategies such as collaborative group work, open-ended tasks, games, and student presentations are also predicted by the council to be effective for learners.

Researchers in parts of lower socioeconomic levels of Melbourne, Australia, have utilized student input to ferret out their preference of instructional activities in mathematics classrooms. As part of the Task Types in Mathematics Learning (TTML) research project, researchers used student responses to help refine research questions to student surveys. Of the 12 students surveyed in this lower socioeconomic part of Melbourne, students created math stories in which their ideal classrooms were designed so that children would work together, play games, and move around. Of particular importance to students in the primary classrooms was the need to be outside and move while they were learning. Cooperative structures were also clearly preferred by students during the study. Seven students preferred working in groups or pairs, five students wanted to share their work with the rest of the class, and one student wanted to help younger students in their acquisition of mathematical concepts as well as to be able to sit and talk during math class (O'Shea, 2009). In creating their own learning experiences, students embarked on shared and cooperative strategies that as a whole group combine to form a "community of practice" (Lave & Wenger, 1991; Rogoff, Turkanis, & Bartlett, 2001).

Furthermore, students in this study stated they disliked conventional types of lessons that included sample problems posted on the board with instructions for students to copy down in their notebooks, rote memorization of simple math facts, and an overabundance of worksheets.

Proponents of DI

Influenced by traditional models of mathematics education, the procedural-formalist paradigm asserts that mathematics is a group of logically organized skills, procedures, and facts that have been refined over centuries. Human experience does not factor into this mental model and therefore increases the difficulty of learning this material as students would not be influenced by life experiences. This viewpoint guided the works of back-to-basic advocates as well as those of psychologist Edward Thorndike in the early 20th century. Thorndike's (1923) Stimulus-Response Bond theory had a penetrating influence on mathematics instruction. Thorndike and his believers felt that mathematics instruction was best internalized through drill and practice that is deliberately sequenced and explicitly taught. Repetition and frequent practice are also hallmarks of the Stimulus-Response Bond Theory as is a student's non-ability to reflect on mathematical concepts. Much of Thorndike's data and philosophy shaped an entire generation of math students and teachers alike. Thorndike believed the learner is seemingly unable to formulate any thoughts of his or her own, let alone develop a new model of meaning.

To deliver this traditional view of delivering instruction, DI models have been developed and used in the classroom. DI models are highly segmented and sequenced and consist of design and effective presentation techniques (Stein, Silbert & Carnine, 1997). Four such DI programs—DISTAR Arithmetic I and II, Corrective Mathematics, and Connecting Math Concepts—were examined by researchers from Eastern Washington University. A meta-analysis was conducted of 12 studies that examined the effectiveness of DI programs being used in United States schools from 1990 to 2004.

All four programs examined were organized through the usage of formats, tasks, and tracks. As defined by Engelmann and Carnine (1975) identified programs consisted of major skills or strategies. The purpose of each track is to teach students to solve simple, written story problems independently. In all four programs, students must have prerequisite skills for success in completion of each track. Student success can be seen through formatted and successive lessons. Furthermore, tasks are created by inserting a new set of numbers into a pattern in which the wording is unchanged.

The DI approach, according to Przychodzin et al. (2004), aligns to the principles for improving math instruction as set forth by NCTM (1989). The first principle validated by this meta-analysis is the Equity Principle that sets high expectations for all students through strong support systems. DI math programs utilize flexible skill groups which are based on current levels of performance using frequent progress monitoring. The Curriculum Principle asserts mathematical tasks must be a collection of activities that are coherent and articulated across grade levels. Advocates of DI advance this principle in designing strands of instruction that are organized around big ideas. Such math programs are also designed to guide students in the acquisition of basic mathematical concepts and operations.

Third, NCTM's (2006) Teaching Principle believes the needs of learners should be well understood by teachers as well as how to challenge and support these needs. Intensive preservice training is provided for teachers prior to the implementation of the four researched models. The Learning Principle claimed students must learn math skills with understanding, actively building new knowledge from prior knowledge and experiences. NCTM also believes that students should study mathematics for 1 hour per day. Teachers using the DI approach are encouraged to spend 30-35 minutes on daily group instruction along with guided practice and modeling. Students then spend 20-30 minutes completing independent seatwork. DI math programs also offer teachers predesigned instructional formats to use in teacher preparation of lessons so as to minimize the amount of time the teacher needs to spend in lesson plan development. Fifth, NCTM's Assessment Principle states that assessment should furnish useful information to both teachers and students. DI programs provide frequent in-program mastery tests to allow teachers to tailor instructional planning to student needs. Various forms of assessments are used in DI programs such as mastery tests, fact games, and take-home assignments.

The final principle from NCTM, Technology Principle, states that technology should influence and enhance skills that are taught to students. Technology, according to NCTM, should be embedded in the math program rather than being a supplemental element. Calculator usage is the predominant method of technology delivery in the four models of DI examined in the meta-analysis.

Managing DI initiatives can be time-consuming and radical in its change on mental models for teachers as well as administrators. Hill and MacMillan (2004) professed that the implementation of DI models is essential to school success in the wake of federal and state mandates such as No Child Left Behind. In order to guide and successfully implement DI instruction, teachers and administrators must understand the essential components of the approach. Hill and Macmillan defined DI has having been based on the theory that instruction erases student misinterpretations and can improve and accelerate learning. Correct and immediate feedback is also a feature of this model where students are not allowed to learn concepts, facts, and skills in an incorrect manner.

Repeated presentation of tasks throughout and at the culmination of the lesson, called the firming cycle, is needed to ensure that students have a solid mastery of the taught skills (Kameenui & Simmons, 2008).

Within this approach, positive reinforcement and immediate feedback are considered to be hallmarks of the program and teacher toolkit. Feedback given to students is essential for success and specifically targeted to the student and task. In short, students are not expected to guide their own learning and feedback. Instead, it is presented in an explicit manner and directly reinforced. Subsequent activities such as concrete learning activities as well as shaping and scaffolding take the place of student-driven learning.

Because of its specific and sequential delivery of instruction education, professionals who work with students with learning struggles and disabilities often tout the DI approach, however Hill and MacMillan (2004) sited that the DI approach can be used with diverse levels of student learning abilities such as those who speak English as a second language. Since its branding in the 1960s, compelling research supports that if students are taught basic skills in a clear and direct format that is strategic in design, they will learn to read and process mathematically (Martells, MacMillan, & Slocum, 2004). It is important to note that researchers point out that DI is one of the processes that promote academic achievement in students in reading and mathematics.

DI formats can be applied to any age student and in numerous instructional contexts. One such context is the application of DI to students with learning difficulties. Researchers at Al-Balqa Applied University in Jordan have examined the effect of DI on math achievement in fourth- and fifth -grade students with learning disabilities (Abdulhameed & Al-Makahleh, 2011).

Sixty students in fourth- and fifth-grade mathematics classes who attended special education classes in the resource setting were selected via a random sample through learning centers within the city of Amman, Jordan. Students were randomly assigned to experimental and control groups. Two tests were administered to students that measured student mathematical achievement, usage of mathematical skills in everyday life, and student attitudes towards mathematics.

Results indicated that a statistically significant difference existed among achievement scores of the experimental and control subjects on mathematics posttests. Experimental students received training on basic math skills using DI strategy, whereas the control groups were taught using more cooperative measures of student grouping. Pretest and posttest mean scores of students in the experimental group were higher than those of the control group. The experimental group pretest scores were M=16.80, and posttest scores were M=40.73. Control group pretest scores were M=15.93, and posttest scores M=22.70.

Abdulhameed and Al-Makahleh (2011) pointed to the format of DI to its success across international borders and cognitive ability groups. Subsequently, the researchers claimed these steps benefited students in the experimental DI group: measuring student performance directly, accurate goals are set for students, tasks are arranged sequentially and systematically, adequate time is set for each task, feedback is provided for students, DI is provided to all students, student performance is displayed in suitable graphical forms, and appropriate problem-solving forms are matched to specific skills.

Saxon Math

Saxon Math is a DI mathematics program that has been developed for students in Grades K-12. Saxon Math programs are based on Gagne's (1965) theory of cumulative learning that stated that skills can be broken down into simple skills and then broken down further into even simpler tasks. Intellectual skills, research states, are organized into an arrangement that reveals prerequisite relationships among them (Gagne & Briggs, 1974). Anderson's (2007) ACT theory explains development of experience in three stages: cognitive, associative, and autonomous. In the cognitive stage, learners memorize and rehearse facts related to a particular skill. During the associative stage, learners can detect errors and misconceptions through practice and feedback. Finally, during the autonomous stage, learners have practiced a skill so that it becomes routine therefore decreasing the amount of working memory needed to perform the skill.

Saxon publishers have created their math programs around incremental instruction, continual practice, and cumulative assessment that are dispersed across the span of a school year. Based on the studies from Brophy and Everston (1976), incremental steps must be taken in order to teach new information to students. Likewise, effective skill development requires additive skill distribution throughout the school year. Research studies praised by Saxon publishers have seen that students who are taught with a curriculum that uses consistent and continual math review have greater math achievement than those taught with a mass approach (Good & Grouws, 1979).

Cumulative assessment is also a critical component of the Saxon Math program. According to Fuchs (1995), periodic assessments can enhance instruction by monitoring student learning, evaluating instructional programs, and revealing student remediation needs. Routine classroom assessments that are designed to be a part of the instructional program instead of an interruption are also recommended from NCTM (1989).

Research studies conducted by PRES Associates in South Carolina, California, Georgia, and Texas from 2005-2007, all examined longitudinal state achievement test data to document the effectiveness of Saxon's elementary and middle schools over time (Resendez & Azin, 2007). Performance was measured using student-level achievement data that compared users of Saxon Math to students who used other math curricula during the same years. In South Carolina, using the Palmetto Achievement Challenge Test, researchers found that students who used Saxon Math from third through fifth grade had an increase in scale scores from 306.8 to 502.9 by the end of the fifth-grade year. Scale scores of middle school students also increased from 590.8 at the end of the sixth-grade year to 782.0 at the end of the eighth-grade year.

Oklahoma City Public Schools Planning, Research, and Evaluation Department also conducted quasi-experimental studies on the effectiveness of Saxon Math. Using achievement data from the Iowa Tests of Basic Skills, students were compared against a matched sample of students who were in the control group that used other mathematics curricula (Nguyen & Elam, 1993). In the study, students were matched by grade level, previous test scores on the Iowa Basic Skills Test, socioeconomic status, race, and gender. Composite performance of students in the Saxon Math classrooms had scale scores of 55.0, while the students with other math curricula had a composite scale score of 52.59.

Researchers from the National Center for Education Evaluation and Regional Assistance (2013), which is a part of the Institute of Education Sciences, examined how four math curricula affected student's achievement scores from first to second grades (Agodini & Harris, 2010). This study began in 2009 and concluded in 2010 while examining four math curricula that were balanced between constructivism and DI formats. Two constructivist approaches, Investigations and Math Expressions, were compared against the DI formats of Saxon Math and Scott Foresman-Addison Wesley

Mathematics, which was later renamed enVision.

Of importance to researchers in this study were the effects of switching math curricula for students between different theoretical types of curricula as well as student achievement. Publishers who were selected to be a part of the design had to apply and submit proposals. Programs were selected according to appropriateness for usage in a Title I school, capacity to train the number of teachers to be used for the study, quality of training and materials, empirical evidence of effectiveness, and research support for the conceptual foundations of the curriculum. The four programs listed above differ in mathematical emphasis in areas such as cognitive demand on the student, frequency and length of repeated practice and routine as well as the teacher's role, pathway for learning, and support for teachers. While all curricula have varied amounts of each activity, stark differences do exist. Saxon Math for example contains 0% of students doing math or actively engaged in math problems, versus Investigations which has 40% of students using active engagement each day. Conversely, Saxon Math is heavy in repetition with procedures for completing math problems each day, 95% of tasks each day, while Investigations only spends 60% of its daily allotted math activities to procedures.

In this study, 111 schools from 12 districts across the United States enrolled in the study and agreed to participate for at least 1 year. In the second year of the study, only 58 schools participated. Random assignment was used to assign a curriculum to each school. Following assignment, publishers from each curriculum made presentations and delivered needed materials to the school staff.

To assess the outcomes of the curricula, researchers administered the Early Childhood Longitudinal Study-Kindergarten (ECLS-K) Class of 1998-1999 assessment. The ECLS-K assessment is given to students on an individual basis and is norm tested nationally. Both open-ended and multiple-choice questions which measure number sense, operations and measurement, geometry and spatial sense, data analysis, statistics, probability, and algebraic patterns were used in the assessment.

Examination of data from the study indicated that after 2 years, Math Expressions, Saxon, and SFAW/envision all outperform Investigations in both first and second grades. Students using Investigations for 2 years had an average score of 65.5, whereas Math Expressions, Saxon, and SFAW/enVision represented scores of 69.8, 69.2, and 69.2, respectively.

Reform-Based Mathematical Programs

Instructional math programs have been caught between the so-called Math Wars waging in the United States since the 1989 publication of *Curriculum and Evaluation Standards for School Mathematics* by NCTM (1989). This publication called for new methods in mathematics instruction, often called reform mathematics, to be used in American classrooms. Reform mathematics differs from traditional mathematical practices of highly sequenced instruction with attention to algorithms and rote memorization of facts.

Exposing students to carefully paced instruction is essential for an adequate foundation to lifelong learning. Foundational skills are often taught in the home and then transferred to the kindergarten classroom in the K-12 spectrum of learning. Students with mathematical difficulties are typically deficient in three areas of mathematics: long-term memory retrieval; ability to solve word problems; and the ability to organize, monitor, and evaluate information (Mercer & Miller, 1997). In kindergarten, student acquisition of the ability to gain knowledge of number sense is a crucial step in order to provide students with the necessary requisite skills for future learning.

One study examined the use of Investigations to research its effects on the acquisition of number sense among kindergarten students (Agodini & Harris, 2010). The Investigations unit is organized into six units that provide units of study that focus on the math strands of number sense, data analysis, and geometry. Units of study can vary in length from 3-8 weeks in duration. Sample topics of instruction include collecting, counting for ourselves and others, counting and measuring, collecting, pattern trains, and hopscotch path, among others.

Twenty-three students in the study were given various forms of assessment including the Stanford Achievement Test-10 (SAT10) as well as a set of Early Numeracy-Curriculum Based Measures (EN-CBM). The SAT is a norm-references achievement test given in group settings. The subsets of tests from the EN-CBMs performed on students in this study included Oral Counting Fluency, Counting From, and Number Identification. Research results were examined using a Pearson Chi-Square analysis that indicated there were no statistically significant differences in gender and ethnicity. Based on these findings, researchers support the usage of carefully sequenced activities with the use of explicit instruction combined with student practice in order to affect student achievement.

In order to determine adequate mathematics instructional resources and practices, Slavin and Lake (2009) completed a meta-analysis of reviews produced at Johns Hopkins University and the University of York that are part of the Best Evidence Encyclopedia (BEE). The BEE uses research methods similar to those of the What Works Clearinghouse yet are broader in focus and not contained to measures that are inherent in research treatments. Five researchers studied thousands of studies and found a total of more than 400 that met the inclusion standards for publication. Reviews suggest that strategies likely to improve student performance are those that increase student active participation and focus in the classroom, help students to learn about their mathematical thinking, and improve the quality of daily mathematical instruction.

Best evidence synthesis (Slavin, 1986) seeks to apply consistent, well-justified standards to identify meaningful information from experimental studies that are unbiased and provide meaningful information. From this analysis and review of literature and research studies, the following conclusions were found:

1. Cooperative learning programs are consistently affiliated with the most positive learning outcomes. Programs that teach metacognitive strategies also have noted positive effect sizes.
2. Traditional computer-assisted instruction programs produced positive effects in math.

In review of mathematics data and traditional mathematics instruction models, Kohn (1999) continued to be an outspoken critic of traditional DI techniques. Kohn asserted that there exists no differences between first-grade classrooms and high school algebra classrooms where the teacher demonstrates the correct way to complete the algorithm and assigns a plethora of examples of the same problem whereby students should imitate the method in which they were shown. The transition model of information leads this type of classroom where students are given facts and procedures by the textbook and the teacher. Students have little, if any, decision about diverse strategies they may be able to formulate to produce the correct answer. Kohn proclaimed one consequence of this drill and kill method is a society where students casually explain that they hate math and lack any skills in the area.

Due to drill and kill repetition model of teaching, students are not able to take the methods taught and transfer them to other areas of mathematics (Brownell, 1944). Drill does not develop meanings. Repetition does not lead to understandings (Resnick, 1980). Kohn (1999) recommended that avoiding the scripted mathematics curriculum is most important for students who are below grade level or experiencing difficulty. The more students are given algorithms to solve and told how to solve them, the farther they fall in understanding.

Math Anxiety

A person with math anxiety feels negative emotions when engaging in an activity that requires numerical or math skills (Sparks, 2011). A stress response is caused when students with math anxiety are faced with numerical problems. This stress and anxiety can cut off the brain's working memory that is needed to learn and solve problems. During stress, more activity is seen in the amygdale than the prefrontal cortex which is responsible for the brain's working memory and critical thinking. The amygdale is a cluster of nerve cells. People suffering from math anxiety are thought to use up their brainpower to cope with anxiety rather than solve the problem at hand. In studies conducted at the Numerical Cognition Laboratory, one researcher found adults with high math anxiety have a lower than typical ability to recognize differences in numerical magnitude, or the total number of items in a set, which is considered a form of dyscalculia. Dyscalculia is the severe inability to complete mathematical computations. Under normal development, children grow in ability to identify which of two numbers is larger; however, in those students with high math anxiety, this ability is less accurate and slower with their tasks, which could explain why students who have math anxiety are often lower mathematical performers when it comes to standardized testing.

Children in the earliest of stages of formal acquisition of math skills in elementary school are often the persons who exhibit the first signs of math anxiety. In a study of 54 children from the San Francisco area, researchers examined the link between the patterns in the amygdale, or nerve cell clusters, believed to be responsible for processing emotions and memory and between students with math anxiety compared to those without. Students in the study were given an MRI as well as the Wechsler Abbreviated Scale of Intelligence to measure IQ. Students completed subtraction and addition activities. Each group of activities contained problems such as number identification, simple arithmetic, and complex arithmetic problems. Examples of these problems would include 5+1=7 and 7-1=5. Subjects were given the full equation and asked to select whether the equation was correct or incorrect.

Results from this study reported by Young, Wu, and Menon (2012) indicated that in children aged seven to nine, math anxiety is associated with hyperactivity and abnormal effective amygdale, which is the region of the brain associated with processing negative emotions and stimuli. Simply put, children with abnormal effective amygdale are oversensitive to outside stimuli and become more anxious in stressful learning situations than their peers under normal circumstances.

Current brain research is important in determining the underlying causes of poor math performance. While brain scans and IQ tests are important, researchers are beginning to understand the human element and how it can contribute to the student's development of math anxiety. Elementary schools consist of 90% female teachers. Math anxiety is more often found among women than men. Researchers from Columbia University studied the impact of female teachers' math anxiety on girls' math achievement at the beginning and end of the school year. Fourteen first- and second-grade female teachers were given measures of math anxiety to complete. Student math achievement was tested in the first 3 months of school and again during the last 2 months of school. Beilock, Gunderson, Ramirez, and Levine (2010) found that no relation existed between a teacher's math anxiety and her student's math achievement at the beginning of the year. The higher the level of anxiety, the more likely girls were to believe stereotypes that boys are often better at math than girls. In summation, girls who held fast to these stereotypes had significantly poorer math achievement scores than girls who did not; however, this study did not report specific achievement test scores for participants in the study at the end of the year (Beilock et al., 2010).

A longitudinal study conducted by a team of researchers in 2013 examined the effects of math anxiety on 113 students between the second- and third-grade years of elementary school. These students attended Title I schools in an urban area of the Northeastern United States. The study examined various aspects of mathematical anxiety but focused on the role of working memory (Vukovic, Kieffer, Bailey & Harari, 2013). Specifically, researchers studied if mathematics anxiety in second-grade students affected their mathematical performance in both second- and third-grade school years.

Students were interviewed by researchers using a 20-item questionnaire adapted from the MARS-Elementary Scale (Suinn, 1988). This questionnaire was given to students at the end of the second and third grades. Additionally, areas of student mathematical performance were also measured during the same timeframe. These areas included working memory, calculation skills, mathematical applications, and geometry.

Results from this study concluded that calculation skills and mathematical applications were negatively correlated to mathematical anxiety. These findings were present even when researchers controlled for student reading ability, level of working memory, and early number sense ability. Researchers felt these three variables may affect student mathematical performance during the study. Furthermore, researchers concluded that math anxiety may have a greater impact on those areas in math where students must understand and manipulate numbers. Support for this claim came from research findings where little evidence existed that math anxiety existed to geometric reasoning with second- and third-grade students. In conclusion, researchers point to the importance of educators pinpointing the specific area of mathematical weakness in students as a way to decrease levels of mathematical anxiety.

Math anxiety cannot only be seen on using brain imaging scans but can also include physical symptoms such as an upset stomach, clammy hands, and increased heart rate. Psychological symptoms can include an inability to concentrate and feelings of helplessness. Furthermore, students often exhibit behavioral symptoms of avoidance of math classes, not studying regularly, and delaying math homework until the last minute. Teachers can help students of any age overcome and reduce math anxiety. Researchers have found that students who develop strong skills and positive attitudes towards math and relate math to real life are less likely to suffer from math anxiety. In addition, teachers should encourage critical thinking and not accept surface-level memorization as understanding of foundational concepts. Active learning with manipulatives as well as active learning strategies can turn passive learners into active learners. Teachers seeking to reduce math anxiety in students should also place less importance on computational speed and correct answers. Finally, teachers can also utilize cooperative groups where competition among students is not emphasized but rather students are allowed to ask questions freely, verbalize their thoughts, and justify their answers.

Research suggests that parental involvement reduces math anxiety and that math anxiety reduces mathematical achievement (Geist, 2010). To support this claim, a team of researchers in New York City examined parents and students at two Title I elementary schools in 2009. Seventy-eight second-grade children and their parents participated in this study. Researchers hypothesized that math anxiety would mediate the relation between parental involvement and mathematics achievement (Roberts & Vukovic, 2011).

Investigators visited classrooms to assess students as well as send home parent questionnaires and information packets. Student mathematical ability was measured using the Key Math-Third Edition test which measured algebraic reasoning. To assess story problems, students were asked to solve 15 problems involving three story types: changing problems, comparing problems, and equalizing problems. Procedural skills were assessed through the Stanford Diagnostics Mathematics Test-Fourth Edition. Parent involvement was measured by a researcher-developed survey.

Results were analyzed through multiple regression analysis. Findings supported the researcher's earlier hypothesis that parental involvement was positively correlated with mathematics achievement. Furthermore, parental involvement was negatively associated with math anxiety. Finally, researchers concluded that schools can help decrease math anxiety in students by enhancing their parental outreach efforts.

Chapter 3: Methodology

The purpose of this program evaluation was to examine the effects of the implementation of Investigations on EOG test scores as well as teacher attitudes toward the preparation of and implementation of Investigations. The researcher used a quantitative approach. The study was grounded in the CIPP model of program evaluation (Stufflebeam, 2011). This chapter is organized around research questions associated with the model in an effort to focus the reader on how each research question was evaluated.

Participants

The sample size consisted of 65 teachers ranging from kindergarten through fifth grades. The EOG data were collected for students in the third through fifth grades, and teacher perceptual data were collected among teachers in all elementary grade levels as this program has been mandated for their use. Demographic information gathered also included grade level taught, total years of teaching experience, and years taught at their current school. Research schools are located in a suburban setting of the southeastern United States. Schools used for perceptual research data include two Title I schools where schools receive additional federal funding due to a free and reduced lunch population of over 60%. One additional school was also used in this study that does not qualify for Title I funds due to the researcher's desire to broaden the validity of results to all elementary school practitioners.

EOG test data were collected from 300 third-grade students from the 2012-2013 and 2013-2014 school years. Investigations was first mandated as an instructional program at the beginning of the 2012-2013 school year. Of the students tested, 63.8% are White and 41.2% are Black, while 36.6% are economically disadvantaged, and 30.4% are limited English proficient.

Instruments

The instrument used in this program evaluation was a survey given to certified teachers in kindergarten through fifth grades at the research schools. The teacher survey focused on the Process and Product portion of evaluation. Statements contained in the survey utilized a Likert rating scale of 1-5 with "1" responses indicating with a Strongly Disagree, "2" representing Disagree, "3" and "4" indicating Agree, and "5" signaling Strongly Agree.

These surveys were first field-tested using 31 teachers at the researcher's school utilizing Google Forms. Through the emailed Google Forms, the researcher reviewed the purpose of collecting the data and content of the questions with the faculty prior to distribution. A 1-week timeline was given for teachers to complete the survey and return to the school's Instructional Facilitator who then turned them into the researcher. Anonymity was assured for all teachers by not asking for teacher names or requiring staff to turn in surveys to the principal. The surveys included demographic questions such as total years taught, years taught in current school, and current grade level. These questions were asked to determine if responses are impacted by experience, school, or grade level.

Field testing at the researcher's school was completed to refine data collection procedures among teachers prior to sending out the survey to other elementary schools. In addition, field testing allowed the researcher to determine if more appropriate data collection procedures should be put into place. Once field testing was completed, three additional elementary schools in the same county were invited to participate in gathering teacher perception data.

Schools selected for the study closely mirrored the field-test research school site in student body size and demographics. To broaden the appeal of this dissertation to varying levels of practitioners, data collection also included one non-Title I status school. The researcher followed the same process in contacting willing participants and sending consent forms and the survey.

EOG tests in mathematics were used to determine the impact of Investigations on standardized testing. The EOG is the most appropriate instrument for usage as it is valid, reliable, and matches state standards to test questions. The framework of the types of questions asked on the EOG tests is located in the appendix of this program evaluation.

Procedures

Research design. This program evaluation utilized Stufflebeam's (2011) CIPP model to determine the impact of implementation of Investigations on student achievement at a suburban elementary school in the southeastern United States as measured by achievement data in third grade. Teacher perceptions of the effectiveness of Investigations were measured by a survey developed by the researcher. Research questions to be addressed in this program evaluation centered on the Process and Product components of the CIPP program evaluation model.

The researcher asked the following questions: "Is it being done?" and "Did it succeed?" Each question correlates with that of the CIPP model. Because the researcher's school district mandated the usage of Investigations to all of its elementary schools, information regarding the context and input in the selection of Investigations as the reform mathematics program to be used was not available for review. As such, research only focused around the process and product portions of the CIPP model.

Numbered responses of "1" correlating with a Strongly Disagree, "2" representing

Disagree, "3" and "4" indicating Agree, and "5" signaling Strongly Agree were measured by Google Forms analysis tools. Open-ended response comments on the teacher survey were analyzed and organized thematically. EOG test data were available to the researcher in various formats from the state accountability division as the tabulation of school, district, and state-wide trends, proficiencies, and subgroup information is generated prior to its public release. Descriptive statistics were used to include measures of central tendency including mean and proficiency percentages.

Process Evaluation Questions

Were the various components of Investigations implemented as they were originally developed?

a. What are the teachers' perceptions about the implementation strategies and activities of the Investigations program?
b. How did teachers have an opportunity to ask questions and voice concerns during the implementation stage?
c. How were any program adjustments made by teachers during implementation?

To answer the process evaluation questions, the researcher gathered empirical data through a teacher survey that was sent to teachers in three elementary schools who have implemented Investigations in their classrooms during the 2012-2013 and 2013-2014 school years.

Product Evaluation Questions

1. Based on EOG data from 2010-2011 through 2013-2014, what impact did Investigations have on student achievement through proficiency scores in Grades 3-5?

   a. What are the teacher perceptions about the impact of Investigations on student achievement?

2. What were the unanticipated effects of the Investigations program on student academic development?
3. What are the teacher perceptions about any unanticipated effects of Investigations on student academic development?

To answer the product evaluation questions, the researcher examined EOG math test data of students in third through fifth grades from the 2010-2011 through 2013-2014 school years as well as teacher perceptual data gathered from the teacher survey to determine the effects of Investigations on student achievement.

In 1973, educational researchers Worthen and Sanders described evaluation as the process of information gathering in order to determine the worth or value of a program (Worthen et al., 2004). Furthermore, it is also crucial to compare possible strategies or approaches that are valuable in order to reach specific objectives or needs (Stufflebeam & Shinkfield, 1985). The overarching concept of evaluation through the CIPP model is not to prove but rather improve so that an organization can be as efficient and effective as possible for future successes (Stufflebeam, 2011).

The research program, Investigations, was evaluated using EOG test scores in Grades 3-5 from the school years of 2010-2011 through 2013-2014. In addition, teacher survey data were collected to determine perceptions of implementation, professional development, adequate materials and resources, and opportunity to ask questions and/or express concerns regarding the program. Scholarly literature was also used to facilitate this program evaluation.

As suggested by the Association for Supervision and Curriculum Development

(2003), using various sources of independent data is crucial to establish truth and accuracy of a claim. This process was aligned with each research question so triangulation of data was possible.

EOG tests in mathematics were used to determine the impact of Investigations on standardized testing. The EOG is the most appropriate instrument for usage as it is valid, reliable, and matches state standards to test questions. The framework of the types of questions asked on the EOG tests is located in the appendix of this program evaluation.

Data analysis. Numbered responses of "1" correlating with a Strongly Disagree, "2" representing Disagree, "3" and "4" indicating Agree, and "5" signaling Strongly Agree were measured by Google Forms analysis tools. Open-ended response comments on the teacher survey were analyzed and organized thematically. EOG test data were available to the researcher in various formats from the state accountability division as the tabulation of school, district, and state-wide trends, proficiencies, and subgroup information is generated prior to its public release. Descriptive statistics were used to include measures of central tendency including mean and proficiency percentages.

Limitations

Elementary teachers in the research schools are required to use Investigations as their sole source of DI in mathematics education for students. This mandate was brought upon county officials to insure fidelity of implementation throughout all elementary classrooms in the school district. The researcher cannot verify that all teachers who completed the survey have implemented Investigations as prescribed by the curriculum leaders in the district due to the possibility of teachers supplementing other instructional materials and strategies in addition to the Investigations materials. Furthermore, the researcher cannot guarantee that teachers have not supplemented any outside resources for mathematics instruction that may vary in approach from Investigations. These two limitations could have an impact on the validity of this dissertation's findings.

It is also important to note that during the 2012-2013 school year, the researched school district underwent a major change in curriculum delivery as the Common Core State Standards were expected to be taught in all classrooms. In alignment with the Common Core Standards, portions of the test questions on the EOG test were analyzed and rewritten which may have an effect on standardized test scores.

Delimitations

Teachers used in the survey portion of this program evaluation were a sample of convenience for the researcher as the researcher was also a principal in the school district. Access to other teachers in schools as well as confidence in other teachers in returning surveys may have influenced the size of the sample. Empirical data were gathered from EOG test scores as well as teacher perceptual surveys in order to examine possible impacts of implementation.

Chapter 4: Results

Introduction

The results of this study are presented in two sections. The first section includes the descriptive statistics including the means, standard deviations, and 95% confidence intervals for each of the survey questions that reflect teacher perceptions of the Investigation mathematics program. The means for Title I and non-Title I schools are also presented. The interpretation of the means with respect to degree of agreement is also provided. The interpretation followed the guidelines and is outlined in Table 6. Each survey item is presented separately.

The second section presents the changes in student proficiency rates on the EOG math tests from the 2 years prior to the implementation of the Investigations program to the 2 years after its implementation. Since the EOG math tests and the criteria for proficiency changed over the 4 years under consideration, the district's proficiency rates were compared to the state proficiency rates to assess if the district proficiency rates improved in comparison to the state's proficiency rates. If the EOG math proficiency rates improved in comparison to the state, then it would lend evidence that the Investigations math program had a positive effect on the math performance of the students.

Table 6

Lower and Upper Mean Limits for the Interpretation of Degree of Agreement with the 11 Statements

Mean Lower Limit	Mean Upper Limit	Interpretation
1.0	1.25	Strongly Disagree
1.26	1.75	Disagree-Strongly Disagree
1.76	2.24	Disagree
2.25	2.75	Slightly Disagree
2.76	3.25	Neutral
3.26	3.75	Slightly Agree
3.76	4.25	Agree
4.26	4.75	Agree-Strongly Agree
4.76	5.00	Strongly Agree

Teacher Perception of Investigations Math Program

The means for the 11 items of the survey are presented in Table 7. The means, standard deviations, and percent of teachers agreeing, disagreeing, and having a neutral position broken down by Title I/non-Title I schools are presented for each survey item separately in Tables 8 through 18. The discussion of these descriptive statistics takes place for each survey item separately.

Process Evaluation Survey Statement Findings

Survey statement: "The training I received prior to implementing Investigations program was sufficient for effectively using the program." An inspection of Table 8 reveals that the mean for the entire sample fell in the slightly disagree range (M=2.73) as did the means for both Title I (M=2.66) and non-Title I schools (M=2.71). An inspection of the percentage of teachers disagreeing, agreeing, and being neutral were similar to each other for Title I and non-Title I schools, with approximately 60% of teachers in both schools disagreeing with the statement that they received sufficient training prior to the implementation of the Investigations program. The mean for Title I schools was very close to the mean for the non-Title I school, again reflecting the similar responses for the two types of schools.

Survey statement: "The materials I received to teach the Investigations program were appropriate and adequate." An inspection of Table 9 reveals that the mean for the entire sample fell in the agree range (M=3.79), as did the means for both Title I (M=3.86) and non-Title I schools (M=3.80). An inspection of the percentage of teachers disagreeing, agreeing, and being neutral were similar to each other for Title I and non-Title I schools, with approximately 70% in both types agreeing with the statement that materials they received for the Investigations math program were appropriate and adequate. Almost 16% of teachers in Title I schools disagreed with this statement in comparison to teachers in the non-Title I school (0.00%). The mean for Title I schools was very close to the mean for the non-Title I school, reflecting the similar perception regarding the materials for the two types of schools.

Table 7

Means and Interpretation for Each of the Items on the Survey for Entire Sample

Survey Statement	Mean	Interpretation
The training I received prior to implementing investigations program was sufficient for effectively using the program.	2.73	Slight Disagreement
The materials I received to teach the Investigations program were appropriate and adequate.	3.79	Agreement
Classroom teachers were given an opportunity to view components of the Investigations program prior to implementation.	2.70	Slight Disagreement
Classroom teachers were given an opportunity to express their preference for the math program to use for implementation of BAM strategies.	2.13	Disagreement
Opportunity was given to ask questions and express concerns before, during, and after implementation.	2.97	Neutral
Opportunity was given to express concerns, make suggestions, or ask questions during the implementation process.	2.98	Neutral
Materials and activities used in implementation of Investigations are appropriate for the age of students I teach.	3.48	Slight Agreement
Adjustments to the delivery of Investigations were made in my classroom after district-wide implementation.	3.30	Slight Agreement
Implementation of Investigations has made a positive impact on student achievement in my classroom.	2.92	Neutral
Implementation of Investigations has made a negative impact on student achievement in my classroom.	3.08	Neutral
Students are comfortable using strategies taught in Investigations in my classroom.	3.41	Slight Agreement

Table 8

Teacher Responses to Statement, "The training I received prior to implementing Investigations program was sufficient for effectively using the program."

		Response Category					
		Disagree	Neutral	Agree	Total	Mean SD	Interpretation
Title I Schools	Count	26	6	12	44	2.66	Slight Disagreement
	Percent	59.1%	13.6%	27.3%		1.06	
Non-Title I School	Count	12	1	8	21	2.71	Slight Disagreement
	Percent	57.1%	4.8%	38.1%		1.35	
Total	Count	38	7	20	65		
	Percent	58.5%	10.8%	30.8%			

Table 9

Teacher Responses to Statement, "The materials I received to teach the Investigations program were appropriate and adequate."

		Response Category					
		Disagree	Neutral	Agree	Total	Mean SD	Interpretation
Title I Schools	Count	7	5	32	44	3.86	Agreement
	Percent	15.9%	11.4%	72.7%		.655	
Non-Title I School	Count	0	6	15	21	3.80	Agreement
	Percent	0.0%	28.6%	71.4%		.972	
Total	Count	7	11	47	65		
	Percent	16.9%	10.8%	58.5%			

Survey statement: "Classroom teachers were given an opportunity to view components of the Investigations program prior to implementation." An inspection of Table 10 reveals that the mean for the entire sample fell in the slightly disagree range (M=2.70), as did the mean for Title I schools (M=2.64). The mean non-Title I schools (M=2.86) fell in the neutral range. An inspection of the percentage of teachers disagreeing, agreeing, and being neutral were somewhat dissimilar with a higher percentage of teachers in Title I schools (45.5%) disagreeing that they were given an opportunity to view the components of Investigation prior to its implementation when compared to teachers in non-Title I schools (33.3%). A greater percent of teachers (38.1%) in non-Title I schools agreed with this statement in comparison to teachers in Title I schools (29.5%). The mean for the Title I schools was somewhat disparate from the mean for non-Title I schools, reflecting a slight difference in perception regarding being given an opportunity to view the components of the Investigations program.

Survey statement: "Classroom teachers were given an opportunity to express their preference for the math program to use for implementation of Balanced Active Math strategies." An inspection of Table 11 reveals that the mean for the entire sample fell in the disagree range (M=2.13), as did the means for both Title I (M=1.98) and non-Title I schools (M=2.33). The percentage of teachers disagreeing, agreeing, and being neutral were somewhat different from each other for Title I and non-Title I schools, with a higher percent of teachers in Title I schools (72.7%) disagreeing that they were able to express their preference when compared to teachers in the non-Title I school (57.1%). While both types of schools had means that fell in the disagree range, the means were somewhat disparate from each other with those teachers in Title I schools showing a relatively greater level of disagreement when compared to teachers in the non-Title I school.

Table 10

Teacher Responses to Statement, "Classroom teachers were given an opportunity to view components of the Investigations program prior to implementation."

		Disagree	Neutral	Agree	Total	Mean SD	Interpretation
Title I Schools	Count	20	11	13	44	2.64	Slight Disagreement
	Percent	45.5%	25.0%	29.5%	100.0%	1.12	
Non-Title I School	Count	7	6	8	21	2.86	Neutral
	Percent	33.3%	28.6%	38.1%	100.0%	1.15	
Total	Count	27	17	21	65		
	Percent	41.5%	26.2%	32.3%	100.0%		

Table 11

Teacher Responses to Statement, "Classroom teachers were given an opportunity to express their preference for the math program to use for implementation of Balanced Active Math strategies."

		Disagree	Neutral	Agree	Total	Mean SD	Interpretation
Title I Schools	Count	32	7	5	44	1.98	Disagreement
	Percent	72.7%	15.9%	11.4%		1.02	
Non-Title I School	Count	12	6	3	21	2.33	Disagreement
	Percent	57.1%	28.6%	14.3%		1.02	
Total	Count	44	13	8	65		
	Percent	67.7%	20.0%	12.3%			

Survey statement: "Opportunity was given to ask questions and express concerns before, during, and after implementation." An inspection of Table 12 reveals that the mean for the entire sample fell in the neutral range (M=2.97), as did the means for both Title I (M=2.98) and non-Title I schools (M=2.81). The percentage of teachers disagreeing, agreeing and being neutral were similar to each other for Title I and non-Title I schools, with approximately and equal number agreeing, disagreeing, and having a neutral position regarding being able to ask questions and express concerns before, during, and after the implementation of the Investigations program. The mean for the Title I schools was very close to the mean for the non-Title I school, again reflecting the similar responses for the two types of schools regarding being able to ask questions and express concerns.

Survey statement: "Opportunity was given to express concerns, make suggestions, or ask questions during the implementation process." An inspection of Table 13 reveals the mean for the entire sample fell in the neutral range (M=2.98), as did the means for both Title I (M=3.11) and non-Title I schools (M=2.76). The percentage of teachers disagreeing, agreeing, and being neutral were disparate from each other with teachers in non-Title I schools disagreeing more (47.6%) that they were given the opportunity to express concerns, make suggestions, or ask questions during the implementation process when compared to teachers in the Title I schools (31.8%). While both types of schools' means fell in the neutral range, the mean for the Title I schools was somewhat higher than the mean for the non-Title I school, reflecting a greater agreement for teachers in Title I schools with this statement.

Table 12

Teacher Responses to Statement, "Opportunity was given to ask questions and express concerns before, during, and after implementation."

		Disagree	Neutral	Agree	Total	Mean SD	Interpretation
Title I Schools	Count	16	12	16	44	2.98	Neutral
	Percent	36.4%	27.3%	36.4%	100.0%	1.11	
Non-Title I School	Count	9	5	7	21	2.81	Neutral
	Percent	42.9%	23.8%	33.3%	100.0%	1.30	
Total	Count	25	17	23	65		
	Percent	38.5%	26.2%	35.4%	100.0%		

Table 13

Teacher Responses to Statement, "Opportunity was given to express concerns, make suggestions, or ask questions during the implementation process."

		Disagree	Neutral	Agree	Total	Mean SD	Interpretation
Title I Schools	Count	14	11	19	44	3.11	Neutral
	Percent	31.8%	25.0%	43.2%		1.01	
Non-Title I School	Count	10	5	6	21	2.76	Neutral
	Percent	47.6%	23.8%	28.6%		1.22	
Total	Count	24	16	25	65		
	Percent	36.9%	24.6%	38.5%			

Table 14

Teacher Responses to Statement, "Materials and activities used in implementation of Investigations are appropriate for the age of students I teach."

		Disagree	Neutral	Agree	Total	Mean SD	Interpretation
Title I Schools	Count	8	11	25	44	3.52	Agree
	Percent	18.2%	25.0%	56.8%	100.0%	1.09	
Non-Title I School	Count	6	4	11	21	3.38	Slightly Agree
	Percent	28.6%	19.0%	52.4%	100.0%	1.20	
Total	Count	14	15	36	65		
	Percent	21.5%	23.1%	55.4%	100.0%		

Survey statement: "Materials and activities used in implementation of Investigations are appropriate for the age of students I teach." An inspection of Table 14 reveals that the mean for the entire sample fell in the slight agreement range (M=3.48), as did the means for teachers in a non-Title I school (M=3.38). The mean for teachers in the Title I schools fell in the agree range (M=3.52). Although there was a difference in the level of agreement, the difference in means was not substantial. The percentage of teachers disagreeing, agreeing, and being neutral were similar to each other for Title I and non-Title I schools, with approximately 55% of teachers in each type agreeing that the activities in the Investigations program is age appropriate for their students. The mean for the Title I schools was very close to the mean for non-Title I schools, reflecting the similar responses for the two types of schools regarding the age appropriateness of the materials for their students.

Survey statement: "Adjustments to the delivery of Investigations were made

in my classroom after district-wide implementation." An inspection of Table 15 reveals that the mean for the entire sample fell in the slight agreement range (M=3.30), as did the means for teachers in a Title I school (M=3.34). The mean for teachers in a non-Title I schools fell in the neutral range (M=3.20). Although there was a difference in the level of agreement, the difference in means was not substantial. The percentage of teachers disagreeing, agreeing, and being neutral were similar to each other for Title I and non-Title I schools, with approximately 25% of teachers in each school disagreeing that there were adjustments made to the delivery of Investigations after its implementation. The mean for the Title I schools was very close to the mean for the non-Title I school, even though there was a difference in interpretation of level of agreement. This reflects the similar teachers' perceptions in the two types of schools regarding adjustments being made after implementation.

Survey statement: "Implementation of Investigations has made a positive impact on student achievement in my classroom." An inspection of Table 16 reveals that the mean for the entire sample fell in the neutral range (M=2.92), as did the mean for teachers in the Title I schools (M=3.05). The mean for teachers in the non-Title I school fell in the slightly disagree range (M=2.62). This difference in level of agreement indicates that teachers in Title I schools were in more agreement with the statement that Investigations has made a positive impact on student achievement. The percentage of teachers disagreeing, agreeing, and being neutral in Title I and non-Title I schools were dissimilar, with 57.1% and 19% of teachers in the non-Title I school agreeing and disagreeing with this statement, respectively, while 29.5% and 36.4% of teachers in Title I schools expressing agreement and disagreement, respectively.

Table 15

Teacher Responses to Statement, "Adjustments to the delivery of Investigations were made in my classroom after district-wide implementation."

		Disagree	Neutral	Agree	Total	Mean SD	Interpretation
Title I Schools	Count	9	13	22	44	3.34	Slightly Agree
	Percent	20.5%	29.5%	50.0%	100.0%	1.58	
Non-Title I School	Count	5	7	8	20	3.20	Neutral
	Percent	25.0%	35.0%	40.0%	100.0%	1.06	
Total	Count	14	20	30	64		
	Percent	21.9%	31.3%	46.9%	100.0%		

Table 16

Teacher Responses to Statement, "Implementation of Investigations has made a positive impact on student achievement in my classroom."

		Disagree	Neutral	Agree	Total	Mean SD	Interpretation
Title I Schools	Count	13	15	16	44	3.05	Neutral
	Percent	29.5%	34.1%	36.4%		1.08	
Non-Title I School	Count	12	5	4	21	2.62	Slightly Disagree
	Percent	57.1%	23.8%	19.0%		.974	
Total	Count	25	20	20	65		
	Percent	38.5%	30.8%	30.8%			

Survey statement: "Implementation of Investigations has made a negative impact on student achievement in my classroom." An inspection of Table 17 reveals that the mean for the entire sample fell in the neutral range (M=3.08), as did the means for both Title I (M=3.07) and non-Title I schools (M=3.10). The percentage of teachers disagreeing, agreeing, and being neutral in non-Title I and Title I schools were similar to each other with approximately one-quarter of teachers agreeing that Investigations has made a negative impact on student achievement. The mean for the Title I schools was very close to the mean for the non-Title I school, reflecting the similar responses for the two types of schools regarding the belief that Investigations had a negative impact on student achievement.

Survey statement: "Students are comfortable using strategies taught in Investigations in my classroom." An inspection of Table 18 reveals that the mean for the entire sample fell in the slightly agree range (M=3.41), as did the means for both Title I (M=3.37) and non-Title I schools (M=3.48). An inspection of the percentage of teachers disagreeing, agreeing, and being neutral in Title I and non-Title I schools were dissimilar to each other with 23% of teachers in Title I schools disagreeing with this statement when compared to only 9.5% in the non-Title I school. A higher percentage of teachers in the non-Title I school (42.9%) were neutral toward this statement while only 23.3% of teachers in Title I schools were neutral. Despite these differences, the means were not different from each other, which reflects overall that there was a similar perception about students being comfortable using Investigations strategies in Title I and non-Title I schools.

Table 17

Teacher Responses to Statement, "Implementation of Investigations has made a negative impact on student achievement in my classroom."

		Disagree	Neutral	Agree	Total	Mean SD	Interpretation
Title I Schools	Count	14	17	13	44	3.07	Neutral
	Percent	31.8%	38.6%	29.5%		.707	
Non-Title I School	Count	5	8	8	21	3.10	Neutral
	Percent	23.8%	38.1%	38.1%		.889	
Total	Count	19	25	21	65		
	Percent	29.2%	38.5%	32.3%			

Table 18

Teacher Responses to Statement, "Students are comfortable using strategies taught in Investigations in my classroom."

		Disagree	Neutral	Agree	Total	Mean SD	Interpretation
Title I Schools	Count	10	10	23	43	3.37	Slightly Agree
	Percent	23.3%	23.3%	53.5%		1.07	
Non-Title I School	Count	2	9	10	21	3.48	Slightly Agree
	Percent	9.5%	42.9%	47.6%		.814	
Total	Count	12	19	33	64		
	Percent	18.8%	29.7%	51.6%			

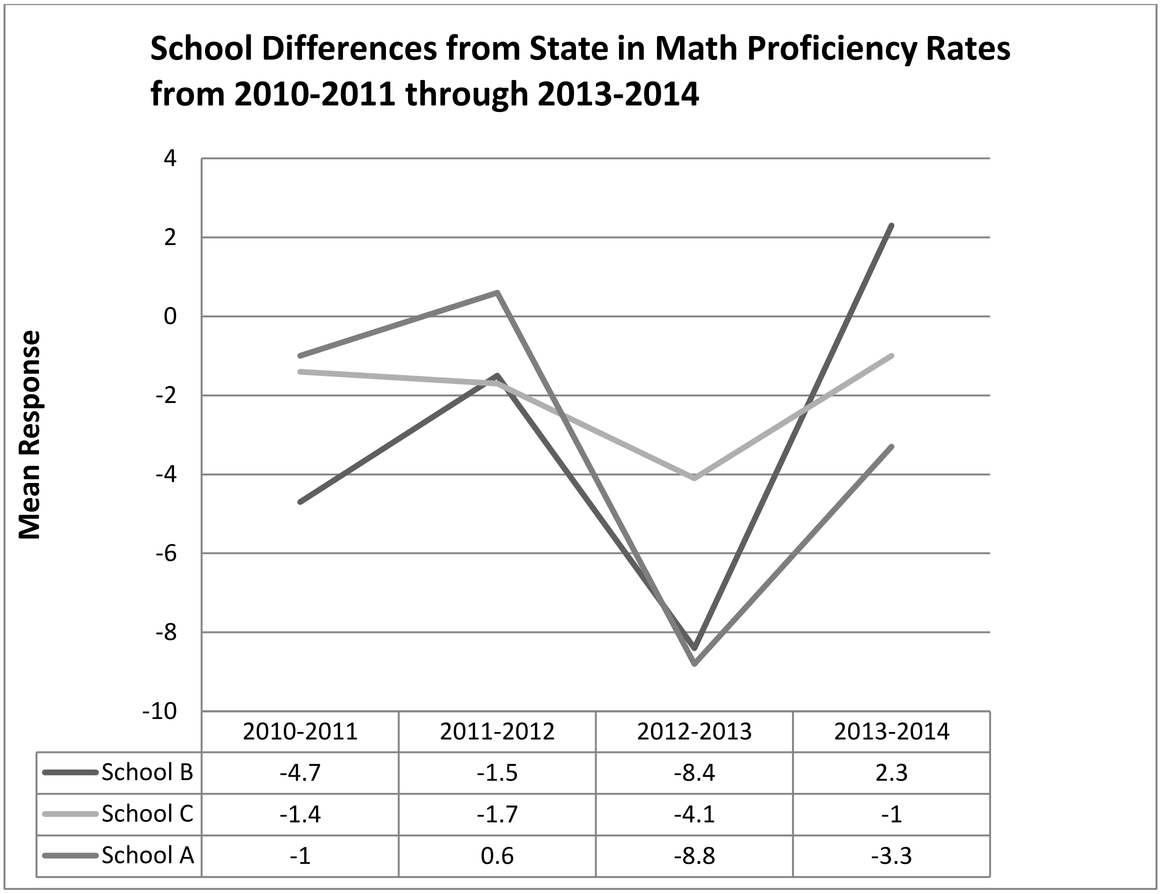

Figure. School Differences from State in Math Proficiency Rates from 2010-2011 through 2013-2014.

Math Proficiency Rate Differences from State for 2010-2011 through 2013-2014

Table 19 contains the math proficiency rates for the three elementary schools and their respective differences from the state proficiency rates for the school years 2010-2011, 2011-2012, 2012-2013, and 2013-2014. The mean differences for the three schools combined are also provided. An inspection of this table reveals that the three schools combined were consistently below the state level, with the exception of School A for 2011-2012. There was also a drop in relative proficiency rates from 2011-2013 (M=-.867%) to 2012-2013 (M=-7.1%), the year after the implementation of the Investigations math program. However, there was also a relative gain in proficiency rates from 2012-2013 (M=-7.1%) to 2013-2014 (M=-.667%), the second year after the implementation of the Investigations math program.

Table 19

School Differences from State in Math Proficiency Rates from 2010-2011 through 2013-2014

		School Year		
	2010-2011	2011-2012	2012-2013	2013-2104
School B	77.7	81.3	38.4	53.4
State	82.4	82.8	46.8	51.1
Difference	-4.7	-1.5	-8.4	2.3
School A	81.4	83.4	38	47.8
State	82.4	82.8	46.8	51.1
Difference	-1	0.6	-8.8	-3.3
School C	81	81.1	42.7	50.1
State	82.4	82.8	46.8	51.1
Difference	-1.4	-1.7	-4.1	-1
Mean Difference	-2.37	-.867	-7.1	-.667

An inspection of the EOG proficiency rates reveals that the three schools were consistently below the state level, with the exception of School A for the 2011-2012 school year. There was also a drop in relative proficiency rates from 2011-2013 (M=.867%) to 2012-2013 (M=-7.1%), the year after the implementation of the Investigations math program. However, there was also a relative gain in proficiency rates from 2012-2013 (M=-7.1%) to 2013-2014 (M=-.667%), the second year after the implementation of the Investigations math program.

Chapter 5: Discussion

Introduction

The purpose of this study was to determine the impact of the Investigations mathematics program on standardized mathematics achievement of three elementary schools in the researched school district. The implementation of the Investigations program was mandated as the stand-alone math strategy to be used by elementary school teachers within the school district where the study was located. This study also examined teacher perceptions of the implementation of Investigations using the CIPP model components of Process and Product.

Stufflebeam and Shinkfield's (2007) CIPP model was used to guide the study. The acronym CIPP denotes the four evaluation types in the model: context, input, process, and product. Perceptual data were gathered from teacher survey results from three elementary schools. Two of these schools were Title I and one school was non-Title I. Quantitative data were gathered by examining mathematics EOG proficiency scores for the 2010-2011 to the 2013-2014 school years. These years represented student performance before and after implementation of the Investigations mathematics program in the researched school district.

This program evaluation focused on the Process and Product evaluation components of Stufflebeam and Shinkfield's (2007) CIPP model. Process evaluation questions were answered by a teacher survey that yielded 65 teacher responses.

Process Evaluation Questions

Process Research Question 1 asked, "What are the teachers' perceptions about the implementation of strategies and activities within Investigations?" Teacher survey statements one, two, and seven were used to answer this question. Survey statement one stated, “The training I received prior to implementing the Investigations program was sufficient for effectively using the program.” Respondents in the Title I and non-Title I schools had similar percentages of disagreement with this statement, 59.1% and 57.1% respectively. Survey statement two stated, “The materials I received to teach the Investigations program were appropriate and adequate.” Respondents in the Title I schools and the non-Title I school had similar percentages with agreed percentile scores of 72.7% and 71.4% respectively. Finally, survey statement seven stated, “Materials and activities used in implementation of Investigations are appropriate for the age of students I teach.” In Title I schools, 56.8% of respondents agreed with this statement, while 55.4% of respondents in the non-Title I school agreed. Overall, teachers in both Title I and non-Title I schools were within 1.4% agreement of the appropriateness of materials provided to deliver the program, the appropriateness of the materials for the age of students taught by teachers, and that training was sufficient for teacher needs.

The second process evaluation question asked, “Did teachers have an opportunity to ask questions and voice concerns during the implementation stage?” This question was addressed through survey statements three, four, five, and six. Survey statement three stated, “Classroom teachers were given an opportunity to view components of the Investigations program prior to implementation.” In Title I schools, 45.5% of teachers disagreed with this statement. In non-Title I schools, 33.3% of respondents also disagreed with this statement. Twenty-six percent of teachers rated this statement in the neutral category of the survey.

Survey statement four stated, “Classroom teachers were given an opportunity to express their preference for the math program to use for implementation of Balanced Active Math strategies.” Teachers in both types of schools overwhelmingly disagreed with this statement. Disagreement was higher in Title I schools at 72.7% than in the non-Title I school at 57.1%. This statement provided the highest level of disagreement of the 11 statements in the teacher survey.

Statement five was, “Opportunity was given to ask questions and express concerns before, during, and after implementation.” Results were similar in types of schools with a total of 38.5% of respondents indicating disagreement, 26.2% indicating agreement, and 35.4% indicating a neutral response.

The final process evaluation question asked, “Were any program adjustments made by teachers during implementation?” Survey statement eight was, “Adjustments to the delivery of Investigations were made in my classroom after district-wide implementation.” Overall, 46.9% of teachers stated that adjustments had been made to the delivery of Investigations since implementation in both Title I and non-Title I schools. In Title I schools, 50% of the respondents indicated that changes had been made to instructional delivery of Investigations since implementation.

Product Evaluation Questions

Product evaluation questions were analyzed through empirical data that were gathered through standardized achievement tests given in the spring of school years 2010-2011 through 2013-2014. In addition, teacher perceptual data were gathered from survey statements nine, 10, and 11 to address the product evaluation questions.

Third-grade end-of-year standardized test data from three schools in the researched school district were examined in order to determine shifts in mathematics proficiency ratings. Two of these schools were Title I schools and one was a non-Title I school. All three schools dropped in proficiency following Investigations implementation in the 2012-2013 school year. In all, the mean difference of the three schools in comparison to their state scores was -.07 during the implementation year; however, a relative gain in proficiency was noted in all three schools after the second year of implementation in 2013-2014.

Product Research Question 1 asked, “What are the teacher perceptions about the impact of Investigations on student achievement?” Survey statement nine stated, “Implementation of Investigations has made a positive impact on student achievement in my classroom.” Over 38% of teachers agreed that Investigations had made a positive impact on student achievement in the classroom. Survey statement 10 examined teacher perceptions of the Investigations program and whether they had a negative impact on student achievement. Of the 65 respondents, 32.3% agreed, 29.2% disagreed, and 38.5% were neutral with the statement, “Implementation of Investigations has made a negative impact on student achievement in my classroom.” Three times the number of teachers in Title I schools believed Investigations had made a positive impact on student achievement. Ogolla (2003) investigated the struggles and successes of elementary teachers implementing Investigations into their classrooms. Results indicated that teachers believed in the student-centered problem-solving approach to teaching mathematics that is an essential component of the Investigations pedagogical framework.

Product Research Question 2 asked, “What are the teacher perceptions about any unanticipated effects of Investigations on student academic development?” Survey statement 11 stated, “Students are comfortable using strategies taught in my classroom.” Overall, 51.6% of teachers agreed, 18.8% disagreed, and 29.7% provided neutral responses to this statement.

Froyd and Simpson (2010) indicated that students are comfortable using student-centered learning activities such as collaborative groups, peer tutoring and editing, and question-directed learning. The Investigations mathematics protocol utilizes both collaborative groups and question-directed learning.

Conclusions

Process Evaluation Questions

The evaluation of the Investigations mathematics program indicated concerns about teacher stakeholder participation. Teachers in this survey had the highest percentages of disagreement with survey statements one, three, and four. These survey statements dealt with teacher perceptions of the training received prior to implementation, the opportunity to view materials prior to implementation, and the opportunity to express their preference for the mathematics program to be used to deliver BAM strategies. The survey data indicated that teachers in the researched school district believed their opinions were not used in the selection of materials to implement BAM strategies and the trainings offered did not adequately prepare them to deliver the Investigations program. Teachers wanted and should be given a voice in selecting materials that will be used in their classroom. As recommended by Confrey and Krupa (2010) of the Center for the Study of Mathematics Curriculum, all stakeholders should have adequate opportunities to learn about new instructional materials presented in conjunction with the Common Core Essential Standards. Teachers want to provide their input and feedback in selecting instructional programs. Seifert and Seifert (1999) recommended involving members of every portion of the organization before models of instructional change can be made in the classroom.

Teacher support is needed for any instructional program to work. To obtain buy-in, teachers need to feel their opinions are valid and meaningful in the selection process of materials that will be used in their classrooms. It is also important to seek other stakeholder input from students and parents in the materials selection process. Fullan (2007) implored governing agencies to utilize the school, community, and district/state to build capacity in instructional programs that can be linked to results such as standardized tests.

Once new instructional programs or activities are selected for usage, professional development needs to be sufficient, meaningful, and ongoing (Garet, Porter, Desimone, Birman, & Yoon, 2001). Professional training sessions could be revised to step away from traditional models of one time, sit and get workshops to more reform-based sessions such as study groups, coaching, and mentoring sessions.

Product Evaluation Questions

The implementation of Investigations cannot be solely linked to decreases in student proficiency scores in third-grade mathematics during the implementation school year of 2012-2013 as content for third-grade mathematics testing was changed in the research state due to alignment of the Common Core Essential Standards along with changes to the state's accountability model. Standardized test data from the 2013-2014 school year yielded a 32.2% increase in student mathematics proficiency in the third grade when compared to the 2012-2013 school year. Moreover, 3-5 years of consistent implementation of an instructional program are needed before correlations can be drawn between standardized test scores and the program's effectiveness (Fielding, Kerr, & Rosier, 2007).

Teachers in the non-Title I school had nearly twice the percentage of disagreement responses (57.1%) to survey statement nine, "Implementation of Investigations has made a positive impact on student achievement in my classroom," compared to the percentage of disagreement responses in Title I schools. More non-Title

I teachers agreed with statement number 10, "Implementation of Investigations has made a negative impact on student achievement in my classroom." The demands of meeting the academic needs of students in Title I schools can often overwhelm teachers and influence their viewpoints on academic or social reform initiatives (Long, 2011).

Implications

Over 67% of teachers surveyed felt they did not have an opportunity to express their preference for the mathematics program to implement BAM strategies. Dufour and Eaker (1998) stated reform changes often do not succeed because of the absence of strong leadership along with a lack of support from faculty and staff. In the rush to adopt top-down mandates and searching for the student proficiency silver bullet, school officials can often fail to develop a sufficient level of support prior to initiating change. In addition, Olivier, Hipp, and Huffman (2003) noted that in order for change to be effective, school administrators should establish an environment in which teachers share in the power, authority, and decision-making process. A lack of teacher buy-in with new initiatives could affect the extent to which teachers make implementation changes to the program.

Almost half of teachers surveyed, 46.9%, reported making changes to the Investigations program when delivering instruction to students. Cook (2005) noted that difficulty in creating education change stems from poor implementation efforts. Called educational change agents, teachers can often underestimate the time, energy, and sustained efforts that are required in implementing new programs in the classroom. As a result, educators can be swayed to make changes to program resources, materials, and timelines. These changes can often cause an instructional program to fall short of its intended goals.

Recommendations

Additional research studies should examine the impact of teaching in a Title I school and what effects it has on implementation of instructional programming. Teachers in Title I schools often face additional stressors such as parental involvement, lower academic student achievement, and higher teacher turnover (Association for Supervision and Curriculum Development, 2005). These factors could impact the implementation process of programs as well as influence teacher perception.

Future research studies could explore teacher perceptions of newly adopted instructional programs through the lens of teacher demographics such as number of years taught, number of years taught in the current grade level, and grade level taught. Collected data could help practitioners discern if the success of an instructional program is based on one of those factors. Studies conducted in the future could examine the differing demographics of teacher and student makeup that might contribute to differences in teacher perception to newly implemented instructional programs.

Finally, the researcher would recommend additional studies in 3-5 years to assess the long-term impact of Investigations upon standardized mathematics testing. This will allow sufficient length of time to see trends in data from implementation (Fielding et al., 2007).

Limitations

The sample size of teacher respondents was a limitation of this study. Sixty-five teacher surveys were collected in three elementary schools. A larger sample size would have resulted in more generalizable results. In addition, the research state's accountability model along with end-of-year achievement test content and questions were changed during the 2012-2013 school year in order to align with Common Core Essential

Standards implementation. As this realignment occurred during the same year as implementation of the Investigations mathematics program, it is expected that a drop in student proficiency scores was seen due to lack of continuity of test content. It is therefore unrealistic to correlate decreased mathematics proficiencies to only the implementation of Investigations methods in classrooms. Furthermore, the researcher cannot guarantee that classroom teachers implemented instructional strategies as prescribed by the publishers of Investigations.

References

Abdulhameed, A., & Almakahleh, A. (2011). The effect of direct instruction strategy on math achievement of primary 4th and 5th grade students with learning difficulties. *International Education Studies, 4*(4), 199.

Agodini, R., & Harris, B. (2010). An experimental evaluation of four elementary school math curricula. *Journal of Research on Educational Effectiveness, 3*(32), 199.

Alleghany Evaluation Specialists. (2014). *Program evaluation*. Alleghany, PA: Alleghany Intermediate Unit.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY: Oxford University Press.

Association for Supervision and Curriculum Development. (2003). *Differentiated instruction stage 2: An ASCD professional development planner*. Alexandria, VA: Author.

Association for Supervision and Curriculum Development. (2005). A *school reclaims itself.* Alexandria, VA: Author.

Behavioral and Social Science Volunteer Program. (2012). *The role and importance of program evaluation and evaluation.* Atlanta, GA: Centers for Disease Control.

Beilock, S., Gunderson, E., Ramirez, G., & Levine, S. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences, Washington, DC, 107*(5), 1860.

Brooks, J. G., & Brooks, M. G. (1993). *In search of understanding: The case for constructivist classrooms.* Alexandria, VA: Association of Supervision and Curriculum Development.

Brophy, J., & Everston, C. (1976). *Learning from teaching: A developmental perspective.* Boston, MA: Allyn & Allyn.

Brownell, W. (1944). The progressive nature of learning mathematics. *Mathematics Teacher, 1*(37), 147.

Burrill, G. (1998). Notices of the AMC. *American Mathematical Society, 1*(45), 87.

Cannela, G. S., & Reif, J. C. (1994). Individual constructivist teacher education. *Teacher Education Quarterly, 2*(66), 27.

Centers for Disease Control. (2010). *Program evaluation*. Atlanta, GA: Centers for Disease Control.

Confrey, J., & Krupa, E. (2010). Curriculum design, development, and implementation in an era of common core state standards. *Center for the Study of Mathematics Curriculum, Arlington, VA, 1*(1).

Cook, C. J. (1995). *Critical issue: Implementing curriculum, instruction, and assessment standards in mathematics*. Retrieved May 9, 2005, from http://www.ncrel.org/sdrs/areas/issues/content/cntareas/math/ma600.htm

Dewey, J. (1961). *John Dewey on education (selected writings).* London: Macmillian Publishers.

Duckworth, E. (1995). *The having of wonderful ideas* (2d ed.). New York: Teachers College Press.

Duff, J. (2012). Cooperative learning vs. direct instruction: Using two instructional models to determine their impact on student learning in a middle school math classroom. (Master's Thesis, Valparaiso University). *Education Senior Action Research Project,* (Paper 5), 1.

DuFour, R., & Eaker, R. (1998). *Professional learning communities at work: Best practices for enhancing student achievement.* Bloomington, IN: Solution Tree.

Ellis, M., & Berry, R. (2005). The paradigm shift in mathematics education: Explanations and implications of reforming conceptions of teaching and learning. *The Mathematics Educator, 15*(1), 7.

Englemann, S., & Carnine, D. (1975). *DISTAR arithmetic: Teacher presentation book A.* Columbus, OH: Science Research Associates.

Fielding, L., Kerr, N., & Rosier, P. (2007). *Annual growth for all students, catch-up growth for those who are behind* (1st ed.). Kennewick, WA: The New Foundation Press.

Froyd, J., & Simpson, N. (2010). *Student-centered learning addressing faculty questions about student-centered learning*. College Station, Texas: Texas A & M University.

Fuchs, L. S. (1995). Connecting performance assessment to instruction: A comparison of behavorial assessment, mastery learning, curriculum-based measurement, and performance assessment. *ERIC Digest E530, ED 381984*

Fullan, M. (2007). *The new meaning of educational change* (4th ed.). New York, New York: Teachers College Press.

Fusaro, M. (2014). Critical exploration in the classroom. *Harvard Graduate School of Education Usable Knowledge, 4*(3), 68.

Gagne, R. M. (1965). *The conditions of learning*. Austin, TX: Holt, Rinehart, and Winston.

Gagne, R. M., & Briggs, L. J. (1974). *Principles of instructional design.* Fort Worth, TX: HBJ College Publisher.

Garet, M., Porter, A., Desimone, L., Birman, B., & Yoon, K. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915.

Geist, E. (2010). The anti-anxiety curriculum: Combating math anxiety in the classroom. *Journal of Instructional Psychology, 37*(1), 8.

Good, T. L., & Grouws, D. A. (1979). The Missouri mathematics effectiveness project. *Journal of Educational Psychology, 3*(71), 355.

Hedeen, T. (2005). Dialogue and democracy, community and capacity: Lessons for conflict resolution education from Montessori, Dewey and Freire. *Conflict Resolution Quarterly, 23*(2), 185.

Hill, J., & MacMillan, B. (2004). *An effective, research-based instructional approach to meet the needs of all students: Direct instruction* (Electronic). Bridgewater, MA: Bridgewater State College.

Hogan, L. (2007). The historical development of program evaluation: Exploring the past and present. *Online Journal of Workforce Education and Development, 4*(2), 2.

Jackson, A. (1998). Interview with Gail Burrill. *Notices of the American Mathematical Society, 3(11*), 87.

Kameenui, E. J., & Simmons, D. C. (2008). *Designing instructional strategies: The prevention of academic learning problems*. New York, NY: Macmillan Publishing Company.

Kamii, C. (1994). *Young children continue to reinvent arithmetic, 3rd grade.* New York: Teachers College Press.

Kamii, C. (1996). Piaget's theory and the teaching of arithmetic. *Prospects, 25*(1), 5.

Klein, D. (2003). A brief history of American K-12 mathematics education in the 20th century. In J. Royer (Ed.), *Mathematics cognition* (pp. 1). Charlotte, NC: Information Age Publishing.

Kohn, A. (1999). *Getting the 3 R's right: The schools our children deserve.* Boston, MA: Houghton Mifflin.

Kuhn, M., & Dempsey, K. (2011). End the math wars. *Learning & Leading with Technology,* November 18, 2011. Retrieved March 18, 2014, from http://www.cfegrants.org/wp-content/gallery/files/2015/08/Jennifer-Schultz-Final-Paper.pdf

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation* (1st ed.). Cambridge: University of Cambridge Press.

Long, C. (2011). How do we increase teacher quality in low income schools. *NEA Today,* May 24, 2011. Retrieved June 15, 2015, from http://neatoday.org/2011/05/24/how-do-we-increase-teacher-quality-at-low-income-schools/

Martells, N., MacMillan, R., & Slocum, T. (2004). *Introduction to direct instruction.* Boston, MA: Pearson.

Matthews, W. J. (2003). Constructivism in the classroom: epistemology, history, and empirical evidence. *Teacher Education Quarterly*, Summer.

Meek, A. (1991, March). On thinking about teaching: A conversation with Eleanor Duckworth. *Educational Leadership, 3*(8), 30.

Mercer, C., & Miller, S. (1997). Educational aspects of mathematical disabilities. *Journal of Learning Disabilities, 30*(1), 37.

National Center for Education Evaluation and Regional Assistance. (2013). *After two years, three elementary math curricula outperform a fourth*. Washington, DC: Institute of Education Sciences.

National Center for Education Statistics. (2011). United States Department of Education. *Trends in international mathematics and science study (TIMSS)-mathematics achievement of fourth and eighth graders in 2011.* Retrieved November 27, 2013, from http://nces.ed.gov/timss/

National Commission on Educational Excellence. (1983). *A nation at risk: The imperative for educational reform* (No. 1). Washington, DC: United States Department of Education.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Retrieved July 23, 2015, from http://www.nctm.org/Standards-and-Positions/Principles-and-Standards/

National Council of Teachers of Mathematics (2006). *Principles and standards for school mathematics: An overview of principles and standards.* Retrieved June 19, 2013, from http://standards.nctm.org

National Research Council. (2001). Adding it up: Helping children learn mathematics. J. Kilpatrick, J. Swafford, and B. Findell (Eds.). Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

NBC News Services. (2011, November 1, 2011). Best-ever US math scores-but still behind? Retrieved May 4, 2014, from http://www.nbcnews.com/id/43556727/ns/us_news-math?q=math

Nguyen, K., & Elam, P. (1993). *The 1992-1993 Saxon mathematics program evaluation report.* Oklahoma City, OK: Oklahoma City Public Schools.

North Carolina Department of Public Instruction, Accountability Services. (2011). *Number of participating students with disabilities in math assessments with and without accommodations at the state level school year 2009-2010*. Raleigh, NC: North Carolina Department of Public Instruction, Accountability Services.

Ogolla, P. (2003). *Practicing elementary teachers' perspectives of "investigations" curriculum*. (January 24-27, 2003). Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education, New Orleans, LA.

Olivier, D. F., Hipp, K. K., & Huffman, J. B. (2003). Professional learning community assessment. In J. B. Huffman & K. K. Hipp (Eds.). *Reculturing schools as professional learning communities.* Lanham, MD: The Scarecrow Press.

O'Shea, H. (2009). The ideal mathematics class for grades 5 and 6: What do the students think. *Australian Mathematics Primary Classroom, 14*(2), 18.

Pell Institute. (2014). *Evaluation myths and benefits*. Washington, DC: The Pell Institute.

Piaget, J. (1953). *To understand an event*. New York: Grossman.

Piaget, J. (1971). *Science of education and the psychology of the child.* New York: Viking Press.

Powell, K. C., & Kalina, C. J. (2009). Cognitive and social constructivism: Developing tools for an effective classroom. *Education, 130*(2), 241.

Przychodzin, A., Marchand, N., Martella, R., & Azim, D. (2004). Direct instruction mathematics programs: An overview and research summary. *Journal of Direct Instruction, 4*(1), 53.

Resendez, M., & Azin, M. (2007). *The relationship between using Saxon elementary school math and student performance on California statewide assessments.* Austin, TX: Harcourt.

Resnick, D. P. (1980). Minimum competency testing historically considered. *Review of Research in Education, 14*(8), 3.

Roberts, S., & Vukovic, R. (2011). The relation between parental involvement and math anxiety: Implications for mathematics achievement. *Society for Research on Educational Effectiveness.* Washington, DC: Journal of Research for Educational Effectiveness.

Rogoff, B., Turkanis, C. G., & Bartlett, L. (2001). *Learning together: Children and adults in a school community.* New York: Oxford University Press.

Ross, M. (2010). Designing and using program evaluation as a tool for reform. *Journal of Research on Leadership Education, 5*(12.7), 481.

Scriven, M. (1991). *Evaluation thesaurus*. Newbury Park, CA: Sage Publications.

Scriven, M. (1996). The theory behind practical evaluation. *Evaluation, 2*(4), 393.

Seifert, E., & Seifert, J. (1999). From the teacher to the top. *National Staff Development's Annual Conference,* Dallas, TX.

Slavin, R. (1986). Best evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher, 15*(9), 5.

Slavin, R., & Lake, C. (2009). Effective educational programs: Meta-findings from the best evidence encyclopedia [Abstract]. *Society for Research in Educational Effectiveness, 1*(19), 26.

Sparks, S. (2011). "Math anxiety" explored in studies. *Education Week, 30*, 31.

Stein, M., Silbert, J., & Carnine, D. (1997). *Designing effective mathematics instruction: A direct instruction approach.* Upper Saddle River, NJ: Prentice-Hall.

Stufflebeam, D. (2011, June). *The CIPP model for evaluation.* Retrieved March 15, 2013, from http://www.scribd.com/doc/58435354/The-Cipp-Model-for-Evaluation-by-Daniel-l-Stufflebeam

Stufflebeam, D., & Shinkfield, A. J. (1985). *Systematic evaluation*. New York: Kluwer-Nijhoff Publishing.

Stufflebeam, D., & Shinkfield, A. J. (2007). In Jossey-Bass (Ed.), *Evaluation theory, models, and applications.* ([Chapter 19] ed.). San Francisco, CA: Jossey-Bass.

Suinn, R. (1988). *Mathematics anxiety rating scale, MARS.* Fort Collins, CO: Rocky Mountain Behavioral Sciences Institute.

Thompson, S. (2011). *Positives of negatives of informal assessments*. Santa Monica, CA: Demand Media.

Thorndike, E. (1923). *The psychology of arithmetic*. New York: The Macmillan Company.

Tyler, R. (1949). *Basic principles of curriculum and instruction.* Chicago, IL: The University of Chicago Press Books.

United States Department of Education. (1997). *A nation accountable: Twenty-five years after A nation at risk*. Washington, DC: United States Department of Education.

Vukovic, R., Kieffer, M., Bailey, S., & Harari, R. (2013). Mathematics anxiety in young children: Concurrent and longitudinal associations with mathematical performance. *Contemporary Educational Psychology, 8*(38), 1.

Wilson, S. M. (2003). *California dreaming: Reform mathematics education.* New Haven, CT: Yale University Press.

Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (2004). *Educational evaluation: Alternative approaches and practical guidelines*. Boston: Allyn & Bacon.

Young, C., Wu, S., & Menon, V. (2012). The neurodevelopmental basis of math anxiety. *Psychological Science, 23*(5), 492.

Appendix

Investigations Implementation Survey

Grade Level Currently Taught:______________
Number of Years Teaching in Current Grade Level: ____________
Number of Years Teaching Total:____________

Using a Scale of 1 to 5, with one being the strongly disagree to five being strongly agree, please answer the following questions.

1. The training I received prior to implementing Investigations program was sufficient for effectively using the program.

Strongly Disagree 1	Disagree 2	3	Agree 4	Strongly Agree 5

2. The materials I received to teach the Investigations program were appropriate and adequate.

Strongly Disagree 1	Disagree 2	3	Agree 4	Strongly Agree 5

3. Classroom teachers were given an opportunity to view components of the Investigations program prior to implementation.

Strongly Disagree 1	Disagree 2	3	Agree 4	Strongly Agree 5

4. Classroom teachers were given an opportunity to express their preference for the math program to use for implementation of Balanced Active Math strategies.

Strongly Disagree 1	Disagree 2	3	Agree 4	Strongly Agree 5

5. Opportunity was given to ask questions and express concerns before, during, and after implementation.

Strongly Disagree 1	Disagree 2	3	Agree 4	Strongly Agree 5

6. Opportunity was given to express concerns, make suggestions, or ask questions during the implementation process.

Strongly Disagree 1	Disagree 2	3	Agree 4	Strongly Agree 5

7. Materials and activities used in implementation of Investigations are appropriate for the age of students I teach.

Strongly Disagree 1	Disagree 2	3	Agree 4	Strongly Agree 5

8. Adjustments to the delivery of Investigations were made in my classroom after district-wide implementation.

Strongly Disagree 1	Disagree 2	3	Agree 4	Strongly Agree 5

9. Implementation of Investigations has made a positive impact on student achievement in my classroom.

Strongly Disagree 1	Disagree 2	3	Agree 4	Strongly Agree 5

10. Students are comfortable using strategies taught in Investigations in my classroom.

Strongly Disagree 1	Disagree 2	3	Agree 4	Strongly Agree 5

11. Additional comments:

According to the Accountability Department of the North Carolina Department of Public Instruction, the following table represents the proficiency levels and descriptions for students taking the 2012-2013 READY EOG assessments.

Mathematics Interpretive Achievement Level I (Limited Performance).
Typically, a student:
_ Exhibits minimal performance.
_ Shows very limited evidence of conceptual
understanding and use of strategies.
_ Frequently responds with inappropriate answers and/or
procedures.
_ Very often displays misunderstandings.
_ Infrequently completes tasks appropriately and
accurately.
_ Needs assistance, guidance, and modified instruction.
NCDPI Division of Accountability Services/Testing Section Page 4 Grades 3–5_Revised
March 2003
Mathematics Interpretive Achievement Level II (Not Yet Proficient).
Typically, a student:
_ Exhibits inconsistent performance and
misunderstandings at times.
_ Shows some evidence of conceptual understanding.
_ Has difficulty applying strategies or completing tasks
in unfamiliar situations.
_ Sometimes responds with appropriate answers or
procedures.
_ Frequently requires teacher guidance.
_ Needs additional time and opportunities.
_ Demonstrates some Level III competencies but is
inconsistent.

Mathematics Interpretive Achievement Level III (Proficient).
Typically, a student:
_ Exhibits consistent performance.
_ Shows conceptual understanding.
_ Applies strategies in most situations.
_ Responds with appropriate answers or procedures.
_ Accurately completes tasks.
_ Needs minimal assistance.
_ Exhibits fluency and applies learning.
_ Shows some flexibility in thinking.
_ Works with confidence.
_ Recognizes cause and effect relationships.
_ Applies, models, and explains concepts.

Mathematics Interpretive Achievement Level IV (Exceeds Expectations).
Typically, a student:
_ Consistently performs beyond grade level.
_ Works independently.
_ Understands advanced concepts.
_ Creatively applies strategies.
_ Analyzes and synthesizes.
_ Shows confidence and initiative.
_ Justifies and elaborates responses.
_ Makes critical judgments.
_ Makes applications and extensions beyond grade level.
_ Applies Level III competencies in more challenging situations.

Students taking the READY End of Grade Assessments are also assessed using scale scores that correlate with each level of proficiency. The following table provides the mathematics scale scores for students ranging from Level I and II which are considered not proficient to Level III and Level IV which are considered proficient. Survey data from teachers will also be examined.

Achievement Level Ranges for the North Carolina EOG Tests

Mathematics at Grades 3–8 Subject/Grade		Level I	Level II	Level III	Level IV
Mathematics					
	3	311-328	329-338	339-351	352-370
(Starting with the 2005-06 school year)	4	319-335	336-344	345-357	358-374
	5	326-340	341-350	351-362	363-378
	6	328-341	342-351	352-363	364-381
	7	332-345	346-354	355-366	367-383
	8	332-348	349-356	357-367	368-384