A DESCRIPTIVE STUDY COMPARING READING ACHIEVEMENT SCORES

OF

FORMATIVE ASSESSMENTS USED IN PREPARATION FOR

SUMMATIVE ASSESSMENTS

by

Nerrissia T. Kemp-Murray

SUSAN SLATER, EdD, Faculty Mentor and Chair

CHRIS STABILE, EdD, Committee Member

AMY KUO-NEWHOUSE, PhD, Committee Member

James Wold, PhD, Interim Dean, School of Education

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

Capella University

August 2015

Abstract

In the years prior to 2011, a North Carolina school system utilized a paper-pencil delivery format for formative assessments in reading for third grade students. In 2012, the school system changed to a computerized-online delivery format for formative assessments. With the mandated change in delivery formats for formative assessments, it was not known if these changes would impact student achievement on the summative end of grade assessments, as measured by the North Carolina End of Grade Assessment. This study utilized descriptive statistics to interpret summative assessment scores in comparison to the change in formative assessment delivery formats. This study provided data that rejected the null hypotheses, and found there was a significant difference in the End of Grade reading achievement scores between the third grade students who received paper-pencil, formative assessment preparation (2010-2011) and third grade students who received computerized-online, formative assessment preparation (2011-2012). The comparison of the independent two tailed t-Test results showed that there were significant differences. Individual schools' student reading percentile scores were higher for the paper group at School A (t [182] = 2.14, p = .03) and at School B (t [634] = 1.32, p = .19) but lower at School C (t [120] = 2.77, p = .007), and similar scores at School D (t [173]= 0.52, p= .61). Although there were not large differences in students' summative scores, the educational relevance of the study encourages schools to investigate other factors at their sites. It would be valuable for educators to monitor instruction and technology to discover why certain schools had little or no difference in student summative reading assessment scores.

Dedication

I dedicate this dissertation unto my family; my husband Manon, my son Milan, my parents Michael and Cleandrea Kemp, spiritual mom Jackie Harris and my uncle Marlon Kyles. I also thank my pastor Dr. Ernest Jones for the prayers and setting a great example. I thank God for you all supporting me, praying for me, and encouraging me to complete this journey. I pray that this journey will allow me to help you all fulfill your greatest dreams.

Acknowledgments

I would like to thank my mentor Dr. Susan Slater, and dissertation committee for their assistance in helping me complete this educational goal. I would not be able to have completed this process without their team effort.

Table of Contents

v

List of Tables

CHAPTER 1. INTRODUCTION CHAPTER

Introduction to the Problem

During the 2010-2011 school year, third grade students in a southeastern North Carolina school district completed formative assessment benchmarks in reading, specifically *Progress Assessments*, which were utilized in preparation for the states End of Grade summative reading assessments. During the 2011-2012 school year, this North Carolina school system no longer mandated the use of *Progress Assessments* a paper and pencil testing booklet format, and adopted a different formative assessment benchmark in reading, specifically *ClassScape* in a computerized format, for the third grade students.

In the past, teachers relied on student responses in the paper and pencil testing booklet formative assessments to guide the trajectory of reading instruction, and help improve student outcomes in reading. The reliance on *Progress Assessment* data and the analysis process of the student data helped teachers to inform instruction. *Progress Assessments* were also utilized as practice for third grade students to improve skills using multiple choice answer sheets, prior to End of Grade assessments, since third grade is the first year students utilize this format for summative assessments. Starting in the 2011-2012 school year, the testing format of formative assessment benchmarks changed to the use of computers, and teachers began to also provide instruction to students on how to take online assessments. The change in the mandated formative assessments (from *Progress Assessments* to *ClassScape*) significantly changed the testing format and use of testing strategies for students, since students no longer marked in testing booklets, and could not eliminate answers on the computer (as they did previously by marking in the test booklet). As teachers participated in the newly mandated computerized formative

1

assessment (*ClassScape*), questions arose as to how this would impact (increase or decrease) students' summative assessment scores in reading as measured by the North Carolina End of Grade Assessment, since testing formats were no longer the same for both formative and summative assessments.

The purpose of this study was to examine and analyze the difference between two groups of student achievement scores; one group using the formative assessment, *Progress Assessments*, and the other group using the formative assessment, *ClassScape*, in order to compare the reading achievement scores of the third grade students using these differing testing delivery formatted formative assessments. This study compared the reading achievement of third grade students taking formative assessments in each of these testing delivery formats (pencil and paper versus computer-based), as measured by the North Carolina End of Grade Assessments during the 2010-2011, and 2011-2012 school years.

In order to improve student reading achievement, educators should be provided with analysis of student outcomes of summative assessment data when different testing delivery formats are used for formative assessments. How do achievement scores of students who took a pencil and paper, testing booklet and answer sheet, formatted formative assessment compare to the achievement scores of students who used an online computer answer response formatted formative assessment? This study described the answer to this question for third grade students in a North Carolina urban school district, revealing if the testing format effected the overall achievement of third grade students on the End of Grade reading assessment.

2

Therefore, the researchable problem is a need to know if there was a difference in student reading achievement scores between students whose preparation included *Progress Assessments*, utilizing a pencil and paper testing booklet and answer sheet format, and students whose preparation included *ClassScape*, utilizing a computerized online format. The change in the delivery format represents inconsistency of test taking strategies, for third grade is the first year that students took standardized assessments. According to the constructivist theory, a child may never reach their zone of proximal development if the scaffolding is not consistent (Powell & Kalina, 2009). When students are not utilizing common strategies and skills taught while practicing how to properly take an assessment, then students may not use the most appropriate test taking strategies on the summative End of Grade test, impacting scores and not representing accurate student knowledge.

Background, Context, and Theoretical Framework

The theoretical framework upon which this study was based is the constructivist theory as defined by Piaget and Vygotsky; the central focus is that learning is characterized by learners constructing meaning through experiences, and by processing interactions as they learn (Ozer, 2004). Constructivism also defines the learner as the one who is in control of what he or she is learning, and the learning is impacted by their activity, beliefs, attitudes and prior knowledge (Almala, 2006). In this study, students did not have a choice of the delivery format they used to take the reading benchmark (formative assessment) or End of Grade summative assessments; paper and pencil or computer based. In this study the delivery format of the test may or may not have

3

effected on the students' ability to apply skills they had learned, but the change in delivery format can be an important factor when comparing reading summative reading assessment scores for the two groups of students.

According to the constructivist theory, humans may formulate a hypothesis based on their experiences and may be prejudiced about which can affect their learning; so if students are not computer savvy, this may pose as a variable that caused a decrease in student End of Grade summative assessment scores in reading (Ozer, 2004). According to Constructivism theory, an individual has the ability to create or construct new knowledge based on prior knowledge, and to apply what was learned to new experiences (Ozer, 2004). Piaget's constructivist theory was birthed out of his cognitive development theory as "He proposed that humans cannot be given information, which they immediately understand and use; instead, humans must construct their own knowledge" (Katherine & Cody, 2009, p. 241). Piaget described constructivism as an individual's ability to construct knowledge by discovering the world around them. Piaget influenced Vygotsky's theory of Zone of Proximal Development (ZPD); the new knowledge that individuals constructed from their experiences occurred through the processes of accommodation and assimilation (Katherine & Cody, 2009).

Constructivism is a theory that describes how learning occurs, and what learners understand based on their experiences. The constructivist model requires the teacher to encourage and help students apply their own knowledge through social interaction and meaningful activities. A constructivist instructor uses teaching methods that cause students to reflect, develop, evaluate and modify their own internal conceptual frameworks of knowledge (Katherine & Cody, 2009). The constructivist teacher creates

4

an environment that leads students or supports students into creating their own learning (Lane, 2007). If students were encouraged to practice test-taking strategies using paper and pencil assessments, the prior experience that the teacher utilized to help students foster their knowledge was altered when the format was changed to computerized assessments. The change in the testing delivery format could yield a difference in student achievement scores, for the daily assessment practices consisted of paper and pencil only, but benchmarks changed to computerized assessments; and the summative assessments (End of Grade) were paper and pencil.

According to Pocaro, the constructivist theory aids in the learning theories of the current age (2011). It is imperative that teachers are able to build and help students construct knowledge in reading based on their prior knowledge, and utilize what was learned to construct new meaning. In utilizing formative assessments with a computer based delivery format, students are able to apply what was learned on assessments in collaboration with web-based learning to help improve student learning. Students are now constructing knowledge with the aid of computer online software; in classroom activities and computerized assessments. Students benefited from the immediate feedback of online formative assessments as opposed to paper and pencil assessments (CTB/McGraw-Hill, 2008). Formative online assessments for reading are beneficial for students and teachers to gain immediate access to student data. The data can be retrieved and analyzed quickly to guide instruction, and identify students' strengths and weaknesses on individual skills in reading and improve teaching practices and raise student achievement (Lai & McNaughton, 2009). It is not known if there will be an expected difference of summative scores with computer-assisted instruction and formative assessments may yield higher

5

scores. Students who took the reading End of Grade assessments during the 2011-2012 school year were exposed to the computerized delivery format formative assessments, but took paper and pencil delivery format summative EOG assessments.

Four reasons explain why students learn more through formative assessments than summative assessments;

1. Frequent, ongoing assessment allows both for fine-tuning of instruction and students focus on progress. 2. Immediate assessment helps ensure meaningful feedback. 3. Specific, rather than global, assessment allows students to see concretely how they can improve. 4. Formative assessment is consistent with recent constructivist theories of learning and motivation (Kathleen & James, 2010, p. 83).

Formative assessments aid students in constructing knowledge and assist in application of knowledge obtained as an ongoing process (Hagstrom, 2006). Formative assessments aid students in learning ongoing curriculum, and computer collaborations improve student learning (Jian-Wei & Lai, 2013). Although there is evidence in literature that computerized learning activities help improve student understanding of specific curriculum; there is a significant lack of research on the computer delivery format of formative assessments in comparison to paper and pencil testing booklet formative assessments and student success in reading. This study provides information to fill the gap in the literature related to the delivery formats of pencil and pencil formative assessments and computerized formative assessments.

The themes of this study's literature review are constructivist theory, constructivism and technology, formative and summative assessments, computer based

6

assessments, computer-assisted learning, paper-pencil assessments, and summative assessments. In classrooms today, there has been a shift from traditional to student-centered learning settings and assessment procedures; from summative assessment to formative assessments (Yurdabakan, 2011). The formative assessments help teachers facilitate student involvement as active participants in the trajectory of their learning goals. "The theoretical framework that emerged with constructivism caused learning settings to be student centered, and assessment processes to include questioning the learning process assisting students to conceptualizing new information" (Yurdabakan, 2011, p. 51). Formative assessments enhance students' understanding of what they have learned in application to real world scenarios (Hagstrom, 2006).

Statement of the Problem

It is not known if there is a significant difference in student reading achievement scores when the testing format of the formative assessments varied between the pencil and paper, testing booklet and answer sheet format (*Progress Assessments*), and the computerized online format *(ClassScape)*, as measured by achievement scores of third grade students on the North Carolina End of Grade Assessment. School systems transitioning from pencil and paper testing booklet and answer sheets delivery formats to computerized online delivery formats for formative assessments may be making this decision based on faster reporting of student scores without comparing student achievement data.

There is a gap in the literature on formative assessments pertaining to the reading achievement of third grade students when the delivery system format of formative assessments change, and specifically when the testing delivery format of formative

7

assessments changes from pencil and paper, to online computer-based answer responses. Educational research has not kept up with the use of computer technology in the testing delivery format of formative assessments.

Purpose of the Study

The purpose of this study was to examine and analyze the difference between two groups of student summative reading achievement scores; one group using *Progress Assessments* formative assessment as the (paper and pencil delivery format), and the other group using *ClassScape* as the formative assessment, (utilizing computerized delivery format), in order to compare these differing testing delivery formatted formative assessments. This study compared the summative reading achievement scores of third grade students taking formative assessments in each of these testing delivery formats, as measured by the North Carolina End of Grade Assessments from the 2010-2011 (paper and pencil) and 2011-2012 (computer-based) school years.

In order to improve student reading achievement, educators have specific information about students'summative assessment data when different testing delivery formats are used for students taking formative assessments. How do achievement scores of students who took a pencil and paper, testing booklet and answer sheet formatted formative assessment compare to the achievement scores of students who used an online computer answer response formatted formative assessment?

Studies support the use of formative assessments to benefit students' learning outcomes (Sly, 1999). The benefits of formative assessments are frequent, timely multiple-choice quizzes that give immediate constructive feedback. The feedback can aid

8

in improving teaching and learning strategies that promote student academic success, and consistently have a correlation between formative assessments given and higher summative assessments (Peterson & Siadat, 2009). Although the study by Peterson and Siadat (2009) revealed formative assessments help to impact student summative scores, the study did not utilize paper pencil formative assessments and then changed to computerized formative assessments.

Research Question

One research question was developed to study whether the difference in the formative assessments delivery format affected student scores of summative assessments in reading achievement.

R$_1$ Is there a significant difference in the End of Grade reading achievement scores between the third grade students who received *Progress Assessments*, paper and pencil, formative assessment preparation and third grade students who received *ClassScape*, computerized, formative assessment preparation for the End of Grade reading achievement testing, as measured by the North Carolina End of Grade Assessment?

H$_0$ There will be no significant difference in the End of Grade reading achievement scores between the third grade students who received *Progress Assessments*, paper and pencil, formative assessment preparation and third grade students who received *ClassScape*, computerized, formative assessment preparation for the End of Grade reading achievement testing, as measured by the North Carolina End of Grade Assessment?

9

H₁   There will be a significant difference in the End of Grade reading

achievement scores between the third grade students who received

Progress Assessments, paper and pencil, formative assessment

preparation and third grade students who received *ClassScape*,

computerized, formative assessment preparation for the End of Grade

reading achievement testing, as measured by the North Carolina End of

Grade Assessment?

Rationale, Relevance, and Significance

There is a significant need for this study to provide student achievement data that

provides concrete evidence for school systems with similar third grade student

demographics to aid in choosing formative reading assessments, and curriculum tools. If

there are significant deficits or gains reported by conducting a t-Test, comparing third

grade summative assessment reading scores utilizing computerized versus paper and

pencil formative assessments; school systems and their stakeholders may consider

utilizing or discontinuing the use of formative computerized or paper and pencil

assessments. This study provides information for stakeholders to decide if the format of

testing will affect the learning outcomes of students, and what to include as an option for

testing for future programs.

This study is needed to determine if there is a significant difference in state

mandated summative assessment (EOG) scores when changing the testing delivery

format for students taking formative assessments. This study will aid in closing the gap

10

for information on the recently implemented use of computerized formative assessments by school systems.

The motivation behind the study was to find out if pencil and paper formative assessments yielded higher reading scores on End of Grade summative assessments, compared to students who took computerized formative assessments, using data from 2010-2011 and 2011-2012. The rationale and need for this study was relevant, for formative and summative assessments are moving to computer or technologically aided formats in North Carolina. Elementary schools utilize student formative assessments to guide reading instruction to aid student success on summative (End of Grade) assessments. This study provides information on student reading achievement, comparing the use of *Progress Assessments* and *ClassScape* formative assessments as preparations for end of year summative assessments.

Nature of the Study

This study is needed to determine if there is a significant difference in reading achievement scores when changing the testing delivery format for students taking formative assessments, and their end of grade summative assessment scores. The primary theoretical basis for this study was constructivism.

The study will provide insight on how the conditions of a test may affect the outcome of the students' success, represented by achievement scores. Students in North Carolina were conditioned to take paper and pencil benchmark tests in the classroom, and then benchmark tests began to be administered using a computerized test program. But the final state summative assessment was given using the paper and pencil format. Constructivist theory states that knowledge is built by experience and a learner's ability

11

to develop processes to make connections to aid in the future recall of learned experiences (Bush, 2006).

The learner's ability to construct knowledge by repeated experiences in formative assessments can assist in success in the delivery format of state summative assessments. Since North Carolina students began experiencing a different delivery format of testing, there was a shift in the method and format in which the students were previously instructed to complete standardized state assessments. Skinner's behaviorist theory of operant conditioning states that learning could be measured by observing changed behavior (Rehfeldt & Hayes, 1998). The change in the delivery format of the formative assessment; the connection between the paper and pencil formative assessment and the computerized formative assessment, could stimulate a response and produce a change in student behavior or outcomes on the summative End of Grade Assessment (Miranda, 2009).

Definition of Terms

Benchmark

Local assessment that is created district or county-wide to measure achievement of standards taught quarterly or by semesters (Bergan, Bergan & Berham, 2009).

ClassScape

A computerized, benchmark, formative assessment program that measures a student's learning during the school year in North Carolina (ClassScape, 2011).

End of Grade Assessment

12

Standardized summative assessment that assesses what a student has mastered at the end of a school year (NCDPI, 2003).

Formative Assessment

Assessments that retrieve results of what students have mastered during a certain point of an instructional session (Ronan, 2014).

Progress Assessment-Interim Assessment

An interim formative assessment that measures a student's growth progress quarterly in a certain curriculum subject (Hicks, 2013).

Summative Assessment

A culminating assessment that measures a student's achievement at the completion of a school year or course (Ronan, 2014).

Assumptions, Limitations, and Delimitations

Assumptions

The following are considered assumptions of the study:

1. Student End of Grade scores accurately reflected reading skills and knowledge.

2. Teachers and administrators have ethically proctored and returned End of Grade reading assessments.

3. Teachers have used data from formative assessments to guide reading instruction.

Limitations

The following are considered limitations of the study:

13

1. Student scores are limited to the summative assessments only, for the formative assessments scores were not recorded.

2. There were no pre-test and post-test scores for the third grade reading formative or summative assessments.

3. There were no formative or summative assessment scores available to compare the reading scores for students as they transitioned from second grade to third grade.

Delimitations

Delimitations considered were:

1. The study is not causal, but a descriptive comparative study which studied individual formative assessments (paper and pencil scores compared to computerized).

Organization of the Remainder of the Study

Chapter 2 of this study presents the literature review of what is known about the problem. The theoretical framework is presented as the literature has been analyzed and synthesized in relation to the problem statement presented in Chapter one. Chapter three is the research methods that were conducted in order to find out what would be done with the numerical data to answer the research question posed in chapter one. Chapter four is an analysis of the data using quantitative statistical reports. Chapter five is the conclusion of the study, and is a detailed discussion of what the results mean, and what should be done with the results for future research. The final Chapter five also summarizes all of the

14

chapters 1-5, and includes a synthesis of the literature and theoretical framework to aid in

the final summary of the study.

15

CHAPTER 2. LITERATURE REVIEW

Introduction to the Literature Review

In classrooms today, there is an educational "shift from traditional to student-centered learning settings" and assessment procedures from summative assessment to formative assessments (Yurdabakan, 2011, p. 51). Formative assessments help teachers to facilitate the involvement of students as active participants in the trajectory of their learning goals. "The theoretical framework that emerged with constructivism caused learning setting to be student centered, and assessment processes to include questioning the learning process assisting students to conceptualizing new information (Yurdabakan, 2011, p. 51)." Formative assessments enhance students' understanding of what they have learned in application to real world scenarios (Hagstrom, 2006).

The themes of this study's literature review are:

- constructivist theory

- constructivism and technology

- formative assessments

- computer based assessments

- paper and pencil assessment

- summative assessments

- quantitative research

This study will advance the scientific knowledge base in the field of education, for it is not known if computerized or paper and pencil formative assessments yield higher summative assessment scores. Prior research has advanced the field of education by acknowledging the benefits of utilizing formative assessments to guide student

16

instruction (Peterson & Siadat, 2010). There is a gap in research regarding the effect of computerized assessments on student achievement outcomes, both when preparation has been pencil-paper and computer-based. This study will provide new information about the application process of administering computerized, and paper-pencil formative assessments and the effects it may have on student achievement on summative reading assessments.

The study investigated whether students' previous knowledge or construct of how to take a standardized summative assessment depended on the experience of taking said formative assessments via paper-pencil or computerized delivery formats. Students' achievement or outcome can be measured by looking at the state summative End of Grade scores of students' from a school year using paper-pencil formative assessments in comparison to a year using computerized formative assessments. Achievement scores from the two different delivery formats can provide data about whether or not the change in the delivery format of the formative assessments impacted the students' summative assessment scores.

Theoretical Framework

The theoretical framework upon which this study was based is the constructivist theory as defined by Piaget and Vygotsky, which stated that the central learning of students is affected by experiences, prejudices, and application of interactions and experiences. Constructivism is a theory that describes how learning occurs, and what learners understand based on their experiences. The constructivist model requires the teacher to encourage and help students apply their own knowledge through social

17

interaction and meaningful activities. A constructivist instructor uses teaching methods that cause students to reflect, develop, evaluate and modify their own internal conceptual frameworks of knowledge. The constructivist teacher creates an environment that leads students or supports students into creating their own learning (Lane, 2007).

In this study students did not have a choice of which format in which they took the reading benchmark assessment (paper-pencil or computer). This study investigated whether the construct of the formative test had an effect on the student's ability to apply what they have learned during the school year on end of grade summative assessments.

According to the constructivist theory humans may formulate a hypothesis based on their experiences that may affect their learning; so if students have had previous practice taking computer assessments, this may pose as a variable that causes a decrease in student summative (paper-pencil) achievement scores in reading (Ozer, 2004). Constructivism is the ability of an individual to create or construct new knowledge based on prior knowledge, and apply what was learned to new experiences. Piaget's constructivist theory was birthed out of his cognitive development theory. "He proposed that humans cannot be given information, which they immediately understand and use; instead, humans must construct their own knowledge" (Katherine & Cody, 2009, p. 242). Piaget described constructivism as an individual's ability to construct knowledge by discovering the world around them. Piaget described constructivism as an individual's ability to construct knowledge by discovering the world around them. Piaget determined the knowledge that individuals constructed from their experiences occurred through the processes of accommodation and assimilation (Piaget, 1953).

18

It is implied that if student paper-pencil summative assessment scores earned after paper-pencil formative assessment are higher versus summative scores earned after computerized formative assessment, the traditional paper-pencil format for both formative and summative assessments is beneficial for student success on the summative assessments for third grade students. Scores from third grade students was the chosen population, for this is the first grade level required to take state and local formative and summative assessments in North Carolina.

Constructivist Theory

As a developmental psychologist, Piaget's theories of cognitive development and constructivist learning theory are credited with influencing teaching methods, curriculum, and educational classroom practices. The theories have formulated the historical perspectives on how knowledge is constructed. Piaget theories have impacted the trajectory of K-12 curricula. Piaget was interested in how subjects interacted with their environments, and how those interactions built knowledge for children.

Cognitive Development

Piaget believed that knowledge is based on a biological timetable. He discovered the stages of intellectual development that paralleled physical growth to the growth of the mind. Piaget believed students were creators of knowledge, and argued that children of all ages are active participants in their ability to construct knowledge. He viewed children as little scientist who explored their environment (Eagan, 2005).

Piaget worked extensively with young children under intense observation studying how the brain of a child worked. As a result of observing the children, Piaget began to form theories based on the exact changes young children displayed during the

19

developmental process of maturity. Piaget noticed that with every stage of childhood, children also changed in cognitive abilities (Eagan, 2005).

According to White (2003), Piaget's theory of cognitive development consists of four stages to correspond to the stages of development from infancy, early childhood, later childhood, and adolescence:

> Sensorimotor stage- occurs from birth to age two; in the sensorimotor stage students' experiences are filtered through their senses. Preoperational stage- occurs from ages two to seven; children motor skills are acquired. Concrete operational stage-occurs from ages seven to 11; children think logically about concrete events. Formal Operational stage, which occurs after age 11; children develop abstract reasoning. (p. 96)

Piaget believed that as children advanced through the stages biological factors and experience equilibration occurred. As a result of the steady developing within the stages, children began to formulate and generate theories about the world. "Piaget's theory has a heavy emphasis on the reasoning ability of individuals and how individuals interpret knowledge" (Powell & Kalina, 2009, p. 241). Piaget hypothesized that infants were born with schemes called reflexes. As a result of humans possessing these reflexes, "adaptation can quickly be replaced by constructing schemes" (Huitt &Hummel, 2003, pp. 2-3) He also took interest in how organisms adapt to the environment through schemes that the individual uses to represent the world and designate action. As organisms adapt, a balance occurs between schemas through the equilibration process.

Piaget described two processes that an individual utilizes to adapt within any given environment: assimilation and accommodation.

20

Both of these processes are used throughout life as the person increasingly adapts

to the environment in a more complex manner. Assimilation is the process by

which an individual uses or transforms the environment to be placed into

preexisting cognitive structures. Accommodation is the process of changing

cognitive structures in order to accept something from the environment. (Huitt &

Hummel, 2003, pp. 2-3)

Piaget noted that the assimilation and accommodation processes could be utilized

simultaneously throughout a person's life, and as a result leads to how a person

constructs knowledge.

Piaget's theory on cognitive development has shaped the contemporary

classrooms of early childhood education. Classrooms today utilize the cognitive

development theory to inform classroom practice. Cognitive development theory has

shaped the educational revolution of cooperative learning through play. Piaget believed

young children construct knowledge through play when problems arose (Hatch, 2010).

Fawcett stated "cognitive conflict that occurs through peer interaction leads to cognitive

change" (2005, p. 157). When a child engages in conflict, new knowledge is constructed;

working with peers leads to greater benefit than working alone. In some incidences

young children adapted their thoughts based on the response from another child and took

on a new belief and the conflict was solved independently of adult interference.

Constructivist Theory Impacts Assessments

Piaget's constructivist theory "proposed that humans cannot be given information,

which they immediately understand and use; instead, humans must construct their own

knowledge" (Powell & Kalina, 2009, p. 241). Piaget termed this constructivism of an

21

individual's ability to construct knowledge by discovering the world around himself or herself, and determined the new knowledge that individuals constructed from their experiences occurred through the processes of accommodation and assimilation. Constructivism theory is valuable for education because it describes how learning occurs, and what learners understand based on their experiences, such as formative assessments.

The new construction of knowledge builds frameworks that lead to internal representations for understanding. When an individual interacts with his or her environment and new experiences occur in an existing framework, but nothing changes the individual has assimilated, but if new experiences are opposite of his or her familiar framework accommodation occurs. The new experience creates new knowledge, for the past exposure or familiarity has disproven old framework interpretations, and created new learning experiences that differed from past interactions. In the areas of testing and assessments students build new schemata based on their experiences in previous testing exposures. The formative assessment delivery formats from pencil and paper to computer-based, cause's students to create new learning experiences due to the change in the testing format.

There are two key Piagetian principles for teaching and learning: learning as an active process, and learning should be whole, authentic and real (Piaget, 1953). The following studies depicted the use of constructivism in classroom practices. The constructivist view on learning focuses on helping students construct and understand concepts themselves. The teacher's role in the constructivist classroom is to encourage students to make connections between new facts and tailor teaching strategies to help students analyze, interpret and make prediction about new information. Teachers in this

22

study facilitated learning utilizing both computerized and paper-pencil formative assessment strategies and techniques, since the formative assessment format changed from paper-pencil in (2010-2011) to computerized (2011-2012), but students took the paper-pencil summative reading assessment.

Constructivist Theory Impacts Teaching and Learning

Constructivist classrooms have adapted teaching principles to become more actively engaging, and student-centered. According to Lane, teachers in constructivist classrooms create environments where students are actively engaged with their peers, activities are designed that promote problem solving using real life examples, and solutions to problems can be practiced daily (2007).

"Reform movements within United States K-12 educational system have experienced an impact on mathematics instruction based on the constructivist learning theory" (Lane, 2007, p. 161). The constructivism shift required teachers, parents and schools to change from usual pedagogical traditions of teaching how they were taught; to a more modern approach of teaching how students learn. The constructivist learning approach focuses on the process of learning, and how students engage in discourse to solve problems.

Traditional instruction of the past focused large amounts of time lecturing, drilling of facts by memorization. The constructivist model requires the teacher to encourage students to develop their own knowledge through social interaction and meaningful activities. A constructivist instructor uses teaching methods that promote independent learning or metacognition, by encouraging students to utilize prior knowledge and social interaction to connect with new learning experiences. The

23

constructivist teacher creates an environment that leads students, or supports students, into creating their own learning. "A constructivist instructor uses teaching methods that help to mold and develop, reflect, and modify their own learning experiences or frameworks" (Lane, 2007, p. 157). The teacher creates environments, activities and assignments that promote students to take control of their learning (Brewer, 2002).

The conceptual framework the student utilizes to solve a problem gives light to how and what the child is thinking, such as reading a literary or informational reading passage. The teacher creates lessons and activities to increase reading knowledge and move the child into a new conceptual framework and to formulate solutions. In the classroom, the teacher creates environments where students are actively engaged, fosters peer learning and interaction in and outside of the classroom setting, convey a multitude of solutions and representations on problems. In addition, the teacher provides experiences where students are actively learning to learn by solving problems cooperatively with others (Fawcett & Garton, 2005).

As common practice, the teachers in this study assessed each student's current reading comprehension framework or ability, by observing how the student scored, and the test-taking strategies used on the formative assessment. Yurdabakan (2011) reported a study conducted in Turkey on primary and secondary education curricula utilized the constructivist theory to analyze the relationship between constructivism, and learning and assessment. As a result, in the late 20th century constructivism shifted from traditional to students-centered learning settings. The shift also created new innovations for assessment procedures, developing the structure for summative assessments and formative assessments. The origin of alternative assessments began as a result of constructivist

24

concern of how students learn and construct meaning of knowledge. The focus on different assessment procedures, or metacognition, laid the foundational practice for alternative formative assessments such as portfolio assessments, self-assessments, peer assessments and co-assessments.

Classrooms today that have followed the constructivist approach emphasize a student's interest and guide instruction using strategies, manipulatives, and hands on experiences. The learning is built on experiences assimilations of real life. The future of constructivist classrooms must transcend the traditional classroom setting in the following areas: (EBC, 2004).

- Move from basic skills being taught to big concepts.

- Move from top down teacher-centered curriculum, to inquiry based curriculum.

- Move from lessons in textbooks, to global, technological real life lessons.

- Move from repetitive learning, to interactive based on prior knowledge.

- Move from teacher embodying a plethora of knowledge via discussion, to students leading discussions and engaging in discourse with teachers and students.

- Move from assessing students for right answers, to assessing through observations, and portfolios. (p. 1-2)

Future Practice of Piaget's Theories

In order to improve future educational practices and instruction, subsequent studies and theories would require foundational constructivist principles to be carried out. The ability to build knowledge or frameworks from a constructivist point of view can aid

in the improvement of K-12 educational practice today. Piaget's work has been studied further to guide educational reform practices and aid in greater understanding of how students learn best (Lane, 2007). Windschitl (2002) acknowledged the philosophy of constructivist learning has grounded the curricula of K-12 content subject areas.

Modern constructivist position promotes the process of thought through the social theory. The basis for learning is infused with cooperative learning, project based learning, and discovery or inquiry based learning approaches (Lane, 2007). The concept of constructivism has catapulted new theories of social learning from theorists like Vygotsky, who believed children construct knowledge based on interactions with others, but they can be successful using the scaffolding model. Piaget's theory was that cooperative learning would contribute to cognitive learning for problem solving, but only to the extent of the developmental stage of the child.

Vygotsky believed children learned best with others who were in a higher zone of proximal development. Both Vygotsky and Piaget are notably acknowledged as core contributors in theories of early childhood education. The textbooks of future teachers are inundated with the debates of Piaget and Vygotsky's theories, but Piaget's theory has maintained position, as development is more important than teaching and learning (Hatch, 2010). Methods of teaching and learning have been impacted due to Piaget's cognitive development and constructivist learning theory. Cooperative learning or the social constructivist approaches in reading, mathematics and science foster active engagement projects, and integrate curriculum with use of technology activities, group discussions and team building. Children in classrooms years ago sat in boxed rows, but after studies on constructivist theories arose, classrooms are now organized as learning

26

communities. Students sit in cooperative groups, some are heterogeneously grouped or are taught using flexible grouping (Fawcett & Garton, 2005).

The constructivist theory uses inquiry based learning to prompt students to formulate their own questions. Marzano's questioning stems are closely linked to the inquiry-based learning theory. Marzono created a model of thinking skills to guide students thought process using research-based theories fostering deeper understanding and comprehension of text. He created question stems adapted from the theories of Bloom's taxonomy to help students reach higher comprehension levels of material. The question stems are utilized to train students to improve their thinking, and become higher-level thinkers (Marzano & Kendall, 2006). Marzano's model, New Taxonomy of Educational Objective, is made of three systems and the knowledge domain: self-system, metacognitive system, and cognitive system. The systems work together along with the knowledge domain to help the brain process how to integrate and apply new knowledge (Marzano & Kendall, 2006).

Discovery learning, along with inquiry learning is also directly related to the constructivist theory. Schools in the 21st century are practicing constructivism principles in problem solving science, and mathematics activities. Inquiry-based instruction in classrooms resembles the model below. The five E's are foundational processes that help students to practice critical thinking skills (Moore, 2009). According to Brunner (1966), the traditional process of "the inquiry model follows this process:

- brainstorm and ask questions- work toward solution.

- formulate questions, investigate, and analyze –establish solutions.

27

- interpret results, discuss, reflect, make conclusions, and present results"

    (p. 72).

The five-E inquiry model, based on the constructivist approach, can be utilized with students in various age groups. Moore lists the five E's in sequence in order for teachers to exhort students to display and engage in their natural process of critical thinking and inquiry (2009). The five steps are: engage, explore, explain, elaborate, and evaluate. In the high stakes testing age that we are currently in and continuously evolving into; schools are relying on test data to determine the success of teaching and learning practices. The five steps also elevate the level of accountability of the students, and support the teacher as facilitator. The teacher becomes a catalyst by encouraging collaborative efforts, observing, listening, and asking probing questions to further spur investigations of subject matter.

Benefits of Constructivism

Constructivism aids is an application of all the theories that originated from Piaget to impact educational instructional practices. Teachers can create an educational environment that motivates and inspires students to want to learn. In a constructivist classroom children are actively engaged, and in turn learn more by doing, and not as passive listeners. The futures of students can be impacted by the constructivist approach by developing thinking skills (comparing contrasting information), and learning to analyze and synthesize information. Students are also able to develop communication and social skills, for they are encouraged to collaborate and engage in discourse with their peers, which aids in problem solving.

28

Constructivism also encourages students to self-reflect and promotes alternative methods of assessments. Students learn by creating resources to use for further reflection; journals, research reports, building artistic and personal models, or mind maps. The skills students construct are also transferrable into their everyday world; mathematics problems can assist in cooking, or science in social studies (climate, and weather are associated to geography). When students are able to construct knowledge freely they are more apt to become inspired to learn and remain intrinsically motivated. Students that are inspired by sharing new experiences, take on more complex challenges, and become confident lifelong learners.

Review of the Research Literature and Methodological Literature
Review of Research Regarding Change of Formative Assessment Format

Is there a significant difference in student reading achievement scores (summative assessment) when the testing format of the formative assessments varied between the pencil and paper, testing booklet and answer sheet format (*Progress Assessments*), and the computerized online format *(ClassScape)*, as measured by achievement scores of third grade students on the North Carolina End of Grade Assessment? This study will advance the scientific knowledge base in the field of education, for it is not known if computerized formative assessments yield higher summative assessment scores. Prior research has advanced the knowledge base in acknowledging the benefits of utilizing formative assessments to guide student instruction (Peterson & Siadat, 2009), but there is a gap in research regarding computerized formative assessments effect on student outcomes.

29

Constructivism and Technology. The constructivist practice in a classroom setting involves student learning through active participation in a discovery-oriented process in a controlled, but teacher facilitated, environment (Overbay, Patterson, Vasu, & Gamble, 2010). According to research, Becker stated that teachers who have a strong pedagogy in the constructivist approach are more likely to incorporate technology in their instructional practices and self-use (Becker, 1992, 2001; Becker & Ravitz, 1999; Ravitz, Wong & Becker, 1999). The benefits of a constructivist technology centered classroom are the deep connection to generate meaning of content through use of simulations, and problem based learning activities (Overbay, et al, 2010). In the study conducted by the IMPACTing Leadership project, it was found that the level of constructivist teaching had a strong positive association with technology use (Overbay, et al., 2010).

Formative Assessments. There is previous research that supports the use of formative assessments to monitor the progress of students to help improve student outcomes by making adjustments to instruction based on assessment data (Cauley, 2010). The benefits of utilizing formative assessments are supported by the use of immediate feedback, ongoing observation that leads to readjusting instruction to improve student understanding, and improvements in summative assessment scores (Peterson & Siadat, 2009). This study will add to the body of knowledge on how curriculum based computerized formative assessments aid in student learning outcomes.

Immediate feedback is a benefit of formative assessment, but consistent ongoing assessments enhance student achievement and motivation while assessing (Kathleen & James, 2010). If consistency is key to building student achievement and motivation to

30

improve student outcomes on assessment, then changing the format from paper and pencil to computerized format for the formative assessment, and then utilizing paper and pencil for the summative assessment may impede the students' motivation and achievement.

> Four reasons explain why students learn more through formative assessments than summative assessments; 1. Frequent, ongoing assessment allows both for fine-tuning of instruction and students focus on progress. 2. Immediate assessment helps ensure meaningful feedback. 3. Specific, rather than global, assessment allows students to see concretely how they can improve. 4. Formative assessment is consistent with recent constructivist theories of learning and motivation. (Kathleen & James, 2010, p. 83)

Formative assessments aid in students constructing knowledge and assisting in application of knowledge obtained as an ongoing process (Hagstrom, 2006).

Although research has proven that formative assessments can improve, and yield higher academic, and summative scores, the question remains if computerized formative assessments yield the same results. Research concludes that formative assessments support the constructivist theory, by focusing on how students learn, and allow teachers to reteach, and adjust instruction to meet the student's needs to promote mastery of learning (Cauley & McMillan, 2010).

Computerized Formative Assessments. Formative assessments aid students in learning ongoing curriculum, and computer collaborations improve student learning (Jian-Wei, & Lai, 2013). Although there is evidence in literature that computerized learning activities help improve student understanding of specific curriculum, there is a

31

significant lack of research on the computer delivery format of formative assessments in comparison to paper and pencil testing booklet formative assessments and student success in reading.

There is a gap in the literature on the use of computerized formative assessment delivery formats, and the effect on student summative assessment scores. It is not known if computerized formative assessments are the final picture of what a student has mastered at the end of a school year, but the formative assessments help stakeholders to devise a plan to improve deficits to improve student mastery. There are implied benefits of computerized formative assessments (if there are no uncontrollable incidents) such as immediate feedback and disaggregation of testing data. The computer is able to group curriculum objectives by strengths, and weaknesses. It can also generate a curriculum map or activity plan to remediate students' weakest areas. Other advantages of computerized assessments are (Pomplun & Custer, 2005), (Boo & Vispoel, 2012):

- reduced testing time

- increased testing security

- increased assesse interest

- reduced cost of test production

- (elimination of books/answer sheets)

- flexible administration

- scoring in less time

- reduced measurement errors

- increased interest and motivation by assessments

32

In the study conducted by Pomplun and Custer, " The Score Comparability of Computerized and Paper-and Pencil Formats for K-3 Reading Test, it was found that the computer version was more difficult for students than the paper and pencil version of the same test" (2005, p. 154).

According to Pocaro, the constructivist theory aids in the learning theories of the current age (2011). It is imperative that teachers are able to build and help students construct knowledge in reading based on their prior knowledge, and utilize what was learned to construct new meaning. In utilizing formative assessments and computers, students are able to apply what was learned on assessments in collaboration with web based learning help to improve student learning. Students are now constructing knowledge with the aid of computer online software in classroom activities and computerized assessments. Students can benefit from the immediate feedback of online formative assessments as opposed to paper-pencil assessments (CTB/McGraw-Hill, 2008). Formative online assessments for reading are beneficial for students and teachers to gain access to individual student reading achievement data. The data can be retrieved and analyzed quickly to guide instruction, identify students' strengths and weaknesses on individual skills in reading, improve teaching practices and raise student achievement (Lai, 2009).

Computer Assisted Learning. Computer aided instruction (CAI) was the basis for the creation of Classworks computerized formative assessment program. With the shift in the 21st century, schools are relying on computers to assist in classroom instruction, and remediation. The strength of CAI is the variety of curriculum based activities, and interactive lessons (Tziro & Shamil, 2002). Since computers are an

33

integral part of educational instruction and assessment, then it is imperative that studies be conducted to identify if the changes in format affect students' achievement (Korbin & Young, 2003). According to the research of Korbrin and Young, their study compared the computer and paper delivery format, and found it affected students cognitively (2003).

Paper and Pencil Assessments. Paper and pencil assessments have traditionally been utilized to gather testing data on students for formative and summative assessments. Paper and pencil assessments allow for the learner to mark in the testing booklet, highlight important information, and also see all parts of the questions and answer choices simultaneously (Korbin & Young, 2003). A study completed by Pomplun and Custer (2005) reported that the "paper and pencil version of a state reading assessment was easier for third grade students to review, and look back over the passage to look for key words or cued phrases" (p. 154).

Summative Assessments/ High-stakes Assessments. Summative reading assessments are the equivalent of a final exam for elementary school students, and are given at the end of a school year or term. The assessments are a way to determine if students have gained growth or proficiency in the subject area assessed based on the state standards. Teachers begin preparing students for high-stakes testing by teaching according to state standards, and equipping the students with testing strategies (Stahl & Schweid, 2013). The summative assessment is a subsequent test that also activates prior knowledge based on instruction given during the school year, and covers information given on formative assessments (benchmarks, progress assessments, quizzes). Third grade is the first time elementary students are tested on standardized summative assessments. The goal is summative reading assessment scores will provide data results

34

to determine if a student can read independently, across various genres, and subject areas (Felton & Akos, 2011).

Review of Methodological Issues

    Quantitative Research. Quantitative research has had a significant impact in the field of education, as a result a multitude of educational reforms and policies that require heightened accountability of schools have been completed across the nation. The history of education can be recorded through usage of quantitative data, form studies conducted in the United States of America, Europe, and Canada. Educational researchers execute quantitative studies to retrieve data in numerical form to track student growth, and deficits. Researchers also utilize data to compare programs, interventions, curriculums, teaching strategies, teacher qualifications, and group specific data. The data retrieved is also used to fund programs for educational curriculums, programs, teacher hiring, and make policy changes in education (Smith, Desimone, & Ueno, 2005).

    Quantitative research is preferred to that of qualitative research for achievement studies, statistical data is objective, and does not lead to misinterpretations. Quantitative data is based on solid numerical data retrieved by conducting experimental studies or empirical studies. Experimental studies allow educational researchers to determine what educational programs, strategies, interventions and curriculums are supporting the evidenced-based education (Mayer, 2006). Educational reforms, such as NCLB, and Common Core Standards utilize numerical data to give monetary incentives for schools that make significant improvements in student achievement. Student achievement and high stakes testing is a driving force in reinforcing curriculum mandates, reforms, and

35

federal school budgets. This study utilized descriptive statistics to interpret summative assessment scores in comparison to the change in formative assessment formats.

Key Components. Quantitative research entails testing theories or hypothesis by examining the relationship among variables. The variables are measured using instruments, and data findings are reported using statistical data (Creswell, 2009). Quantitative research identifies measures and compares variables within a group, and analyzes numerical data of a target group, theory, or concept. This study analyzed the numerical data of third grade students' summative reading assessment scores in 2010-2012. The descriptive design also resembled relational or causal (experimental) to find out how a variable affects or relates to the group. The archival data obtained from one school district that had similar demographic information. The data collected were End of Grade summative assessment test scores reported in numerical form. The researcher utilized the data to describe a current state of a school group or population, examined relationships between variables, and hypothesized the trajectory of a group, to find out if the differences in variables invoked change in other variables (Kamil, 2004).

The key components of the quantitative methodology can be placed in two categories, experimental and non-experimental research. Experimental studies utilize assessments of the effect of different treatment conditions or independent variable, and outcomes that possibly involve multiple dependent variables. Experimental research entails two types; true experiments or randomized studies, and quasi experimental-conducted where randomization is not possible.

There are "four methods to analyze data of experimental studies:

- univariate,

36

- multivariate,

- structural equation model,

- multilevel modeling." (Creswell, 2009, p. 151)

Each method of analysis involves a process of comparing group means in a particular way. The quantitative data is analyzed and reported through statistical measurements; inferential statistics. The comparisons of data can provide averages to specific exploratory procedures in new areas of inquiry, and identify results of simple to more complex inquiries (Lomax, 2004).

Historical Context, Theory and Research Initiatives. Quantitative research has aided in the historical context on the application and theory in research in education. The evolution of quantitative research can be traced to the philosophical positivist (postpositivist) worldview, a scientific method of conducting science research. The characteristics of the postpositivism worldview included "determination, reductionism, empirical observation and measurement and theory verification" (Creswell, 2009, p.152). The rationale for this philosophy is to reflect, identify, and assess causes that create outcomes found in experiments (Creswell, 2009).

Quantitative research has been directly linked to educational initiatives and reforms. The history of socialistic structures in education has shifted due to political and social philosophies, and mandates by government officials. "Quantitative research and educational reform seems to be ubiquitous in that studies have been conducted in Australia, Canada, Israel, U.K, Germany, and United States" (Deville & Delville, 2011, p. 307). Quantitative research is linked to modern educational reform policies in

37

America; which led to the evolution of the No Child Left Behind Act (NCLB) policy. Quantitative research has been utilized in past reforms such as,

Elementary and Secondary Education Act (ESEA of 1965)-President Lyndon

Johnson's administration

National Commissions on Excellence in Education (1983)-,

America 2000 Excellence in Education Act (1991),

Goals 2000 American Education Act (1994),

NCLB (2002)

Common Core Standards (2011).

The significance of past reforms in America was the connection between the federal government and allocation of federal dollars to public institutions (Duffy, 2008). The nexus between all of the aforementioned educational initiatives is the political foundation that wanted to improve accountability and promote a stronger educational system in the United States of America. The core goal of the federal No Child Left Behind federal legislation was a call to action of greater accountability for teachers and students. The accountability had a deadline with an expectation of all students making adequate progress and proficiency by 2013-2014. The accountability also required teachers to become highly qualified, as defined by the NCLB Act; all teachers must have full certification, bachelor's degree, and knowledge of all content areas in which they were certified (Smith, Desimone, & Ueno 2005). "According to Duffy, et al. (2008), "there were four principles that were developed by NCLB: greater accountability for student performance, focus on research-based practices ("what works") empowering parents with choice options, and reduction in bureaucracy and increased flexibility" (p. 53). These standards and policies were signed

38

as a law on January 8, 2002 under President George Bush's administration. The factor that caused the No Child Left Behind Act to stand out amongst all of the other educational reform policies was the focus on accountability through mandating high-stakes assessments. The NCLB Act also placed an emphasis on utilizing quantitative measures to track student achievements, and teacher qualifications (Smith, Desimone, & Ueno, 2005). The testing focused on three curricula areas; reading, mathematics and science. The assessment data determined whether students would be promoted to the next grade level-or be able to graduate.

   The results of the standardized state assessments became the schools districts, states, and countries report card form student academic achievement in the areas of reading, literacy, and math. Quantitative data also compared schools in the state, and around the country, and provided data to determine which schools met adequate yearly progress for student achievement, and expected growth for the year. The schools who made AYP received incentives and in some schools cash bonuses for their success, other schools were put on a watch list; which could lead to state takeover, loss of Title 1 funding, and teacher jobs severed. Another difference between the NCLB act and other policies was the expansion of federal control over school systems in America. "The NCLB holds schools and teachers accountable for student performance, and as a result provided and endorsed federal grants, contracts, and provisions to promote educational practices (with heavy emphasis on quantitative research" (Deville, & Delville, 2011, p. 307).

   Quantitative methods have also molded major reforms that occurred in the United Kingdom (UK). According to Carpentier, British policy makers have relied on

39

quantitative sources for funding (2008). The quantitative method was used to monitor the construction of the national system of education for the British public schools. As a result of quantitative data collection, the 1833 parliamentary voted to report statistical data on school inspections, and use of public school funds. Three principal goals were created to report quantitative inspection data: "checking the use of public funds accountability, providing information on the success or otherwise of the educational system, and advising those responsible for the running of establishments (schools)" (Carpentier, 2008, p. 201). The result of the data retrieved also led to the "creation of the Statistics of Inspection of Annual Grants Schools; presented by Privy Council on Education " (Carpentier, 2008, p. 201). Quantitative data in the U.K created the evaluation process of the state of education, and reported all financial resources, expenditures, and disbursement of salaries, books, and student programs. The significance of quantitative data ignited the use of data for policy-makers, educational researchers, and stakeholders to document and report spending in a more systematic format. The statistical data was also utilized to build a more democratic educational system, and demand equality for all students, regardless of socioeconomic status. The educational system refined studies of the history of education by conducting quantitative research to define the disparities, identify patterns and trends to challenge traditional interpretations (Carpentier, 2008).

Quantitative Approach Preferred. The NCLB Act has required policy makers to rely on quantitative data to inform the public about the state of the current educational system, and its' students. The preference of quantitative data is chosen due to the verifiable numbers that can be measured, tracked and retested to see if what was reported is concrete. If the federal government will continue to fund programs to move students

40

from failure to success in programs like Title 1, and Reading First the assessment data must prove that students have made growth, and that determines if federal funding will continue to be released. If the program is not boosting test scores, proficiency or growth in the areas mandated by NCLB, then funding will be cut.

Quantitative data is also preferred in current education initiatives to also compare programs that have proven to be successful. As a result of NCLB, some schools have created remediation and tutoring programs to improve student proficiency as mandated by NCLB. Quantitative research or disaggregated data was used to compare the differences of performance of students who received tutoring or remediation services and those who do not receive services to determine the amount of growth, or how much growth is needed for said student(s) to be successful. The NCLB act requires schools to have a school improvement plan if adequate yearly progress is not achieved. The quantitative data from test scores is utilized to create individualized programs to meet the learning needs of the students. Data was tracked to determine how much the students improved with tutoring or new intervention in place by comparing test scores before and after tutoring. Quantitative research identifies the correlation between tutoring and test scores. The design would be relevant for it would measure if one variable tutoring would change or affect another variable, student achievement test scores. Such studies are experimental or casual in nature. The experimental method is also preferred due to the explicit definition stated within the guidelines of the NCLB Act (2002).

According to the U.S. Department of Education, "scientifically based research was defined as rigorous procedures that systematically employ systematic empirical methods and objective procedures to obtain reliable and valid knowledge relevant to

41

education activities and programs" (Horn, 2004, p. 196). The definitions eludes that true scientific based research is experimental in nature, and is clearly p-referred by policy makers.

"Quantitative research results may inform all types of purposeful sampling, by providing an overview about the existence and the distribution of certain types of social problems, structures or patterns of action prevalent to the investigated domain" (Kelle, 2006, p. 196). Schools that collect and dissect data of students' assessment scores can utilize quantitative data to inform instruction and define what problems or gaps there are within instruction, and student mastery.

Quantitative research is also used to evaluate, reconstruct or stabilize schools across the nation. Student achievement and proficiency scores are filtered by quantitative studies; for student assessments scores are documented to track annual yearly progress or growth of students elementary through high school age. The scores are also utilized to create future grant funded programs for Title one, and other intervention programs and curriculums. Quantitative research is utilized to retrieve direct data, for numbers reveal the propensity of success of a school system, and can dissect the weaknesses and strengths of schools across the nation. The comparisons can show what trends occur, and what interventions are not working, and where targeted areas of instruction should be focused. The data retrieved also create teaching and specialist positions for schools to implement remediation, and school improvement programs to raise student achievement.

The quantitative method as opposed to the qualitative method promotes objectivity. Quantitative research is not subjective to the researcher's interpretation, for numbers are concrete. The quantitative method also guides data driven lesson plans,

42

instruction, and daily reports given by the teacher on student performance. The teacher is able to report the student's progress using numbers, and stakeholders can also aid the teacher by creating workshops, programs, and new curriculum text based on the numerical data of assessments.   The disadvantage of educational reforms relying on qualitative data is open ended, as opposed to quantitative data where student data can be recorded simultaneously and yields results finite results (Creswell, 2009). As a result of the devaluing of qualitative research role in NCLB, Horn states, "quantitative researchers will only be credible participants in policy decisions" (2004, p. 196). The No Child Left Behind Act requires teachers to have data that informs instruction to create lesson plans to improve student outcomes, and interviews would not help teachers collect numerical data on students in a timely manner. Qualitative research is not preferred in schools, for the data collection process would take longer, and the data retrieved are left to the researcher's interpretation.

Curriculum content area studies have found that reading and mathematics instruction also benefit from quantitative research methods, since the NCLB Act requires students to be proficient in reading and mathematics, Common Core Standards (CCS) also mandate student growth by conducting high stakes testing and tracking of assessment data; schools must also teach students to be proficient in utilizing quantitative data. It is of grave importance students have the ability to read data on maps charts and graphs. Quantitative methods became more important in reading, mathematics and science, for students must prepare for the work force that requires, logical reasoning, literacy, social science, and fine arts (Steen, 1999). The goal of NCLB and CCS is to

43

prepare students to be globally competitive, so driving quantitative lessons would help propel this goal.

The data driven society of today requires teachers to create lesson plans that produce results that prove student growth is the center focus of teaching. Quantitative data is the cornerstone of data driven instruction. It behooves school systems to train teachers on how to utilize student scores to inform instruction; for the NCLB Act and Common Core 2010 now track student scores to evaluate teachers. Teachers are held at higher accountability as the NCLB Act focuses on the high stake assessments, growth scales of student data, and standard based assessments.

Common Core national standards supports the NCLB Act, but has shifted from state standards to national standards and assessments. Common Core also utilizes quantitative research methods, in that student scores will now be shared nationally and standards are mandated over the current 48 states that have adopted Common Core Standards (CCS). States that have bought into the challenge of CCS federal Race to the Top initiative have accepted the higher accountability of raising student achievement for all students. According to Deville, "CCS will drive instructional practices, curricula, text books, professional developments, and testing systems" (2011, p. 311). There are financial benefits for school systems that meet the demands of race to the top. Schools across the United States that meet the mandates for CCS via the Race to the Top program can compete for federal funds or grants, and national recognition. The qualification for meeting CCS mandates is based on quantitative data, of assessment scores, but also requires states to amend current state standards to national CCS, and make growth in student achievement in focused content areas of reading, mathematics, and science.

44

"Quantitative research utilizes data analysis to test hypothesis, and justify general conclusions about educational issues" (Horn, 2004, p. 196). The current trajectory for educator empowerment revolves around sound data that can be referenced, retested, and applied to answer casual relationships between the implementation efforts of all parties involved in school improvement. The task is crucial for data analysis to support education, and future policies to align to the ongoing reforms, and adjustments made by the current administration of President Obama. The United States Department of Education supported concurrent and subsequent findings of quantitative data, and research driving policies, and redirecting the funds to schools that have met the challenge of *Race to the Top*.

Synthesis of Research Findings. Research supports the use of formative assessments as beneficial to monitor the progress of students to help improve student outcomes by making adjustments to instruction based on assessment data (Cauley & McMillan, 2010). Data are utilized to improve student achievement outcomes, by providing feedback that will guide the trajectory of the curriculum and individualized lessons. The benefits of utilizing formative assessments are immediate feedback, and ongoing observations that leads to readjusting instruction to improve student understanding, academic and summative assessment scores (Peterson, & Siadat, 2009).

Assessment data drives instruction, and the delivery of the testing format can help stakeholders understand the benefits of the testing delivery and its effect on student achievement. The themes of this study are formative assessments, constructivist theory, data, high stakes testing and computer assisted learning. In classrooms today, the nontraditional shift in student-centered learning settings and assessment procedures; from

45

summative assessment to formative assessments. Formative assessments help teachers to facilitate students as active participants in the trajectory of their learning goals. "The theoretical framework that emerged with constructivism caused the learning setting to be student centered, and assessment processes to include questioning the learning process assisting students to conceptualizing new information" (Yurdabakan, 2011, p. 51). Formative assessments enhance students' understanding of what they have learned in application to real world scenarios (Hagstrom, 2006).

Constructivism is a theory that describes how learning occurs, and what learners understand based on their experiences. The constructivist model requires the teacher to encourage and help students apply "their own knowledge through social interaction and meaningful activities" (Lane, 2007, p. 157). A constructivist instructor uses teaching methods that cause students to reflect, develop, evaluate and modify their own internal conceptual frameworks of knowledge. The constructivist instructor utilizes methods of teaching that cause students to reflect, develop new schema, and modify their conceptual knowledge (Lane, 2007).

The delivery format of an assessment and the reporting form can promote positive or negative student outcomes depending on the assessment constructs, and ongoing training and preparedness. The interaction of computers with student learning and assessments can be a benefit, but also a hindrance for student assessment scores. The delivery of paper-pencil assessment can provide benefits to students as well, but the strategies and constructs are also dependent on prior knowledge, instruction and training of content.

46

Critique of Previous Research

Research displayed a significant gap in reporting details about the effects of the format of formative reading assessments, and its effect on summative reading assessments. The research similar to this study implied benefits for utilizing computerized formative assessments by listing the following strengths; immediate feedback, interactive activities, and detailed testing data (Tziroe & Shamil, 2002). It is not known if there is a correlation between computerized formative reading assessment preparation and summative reading assessment scores (especially when there is not alignment of the delivery format of both formative and summative assessments). The previous research conveys the benefits of utilizing formative assessments, and the benefits of using computerized assisted lessons and testing, but there is a lack of research on how the various forms of formative assessments effect summative assessments.

Chapter 2 Summary

High-stakes testing is utilized to retrieve data on student achievement for content areas of reading, mathematics, and science. The constructivist theory supports the use of testing experiences or structured interactive lessons to be facilitated by teachers in order to support student achievement. Teachers utilize formative assessments to inform instruction and make adjustments to daily lesson plans in order to prepare students for summative assessments. The assessment format can be computerized or paper and pencil. There are benefits of utilizing both types of test delivery formats, but the key is to scaffold students by preparing them to be successful ((Fawcett, & Garton, 2005). This study will provide insight to third grade students' summative testing results using

47

quantitative data, to analyze the effect of the independent variables of formative pencil-

paper assessments and computerized assessments.

48

CHAPTER 3. METHODOLOGY

Introduction to Chapter 3

Quantitative methodology was used to analyze summative reading assessment data for statistically significant differences in student reading achievement scores (North Carolina End of Grade Assessment) of two groups of third grade students from four suburban public schools during the 2010-2012 school years. The student groups represent two different delivery formats for formative reading assessment benchmarks (pencil and paper, and computerized). This chapter states the research question and hypothesis, describes the research design, target population, sampling method, sample size, setting, recruitment, instrumentation, and data collection. The chapter concludes with data analysis procedures, limitations of the research design, internal and external validity, expected findings, ethical issues, the researcher's position statement, and chapter summary.

Purpose of the Study

The purpose of this study was to examine and analyze the difference between two groups of student achievement scores; one group using the formative assessment, *Progress Assessments*, and the other group using the formative assessment, *ClassScape*, in order to compare the reading achievement scores of the third grade students using these differing testing delivery formatted formative assessments. This study compared the reading achievement of two different groups of third grade students taking formative assessments in each of these testing delivery formats, as measured by the North Carolina End of Grade Assessments from 2010-2012.

www.manaraa.com

Research Question and Hypothesis

R$_1$ Is there a significant difference in the End of Grade reading achievement

scores between the third grade students who received *Progress*

Assessments, paper and pencil, formative assessment preparation and third

grade students who received *ClassScape*, computerized, formative

assessment preparation for the End of Grade reading achievement testing,

as measured by the North Carolina End of Grade Assessment?

H$_0$ There will be no significant difference in the End of Grade reading

achievement scores between the third grade students who received

Progress Assessments, paper and pencil, formative assessment

preparation and third grade students who received *ClassScape*,

computerized, formative assessment preparation for the End of Grade

reading achievement testing, as measured by the North Carolina End of

Grade Assessment?

H$_1$ There will be a significant difference in the End of Grade reading

achievement scores between the third grade students who received

Progress Assessments, paper and pencil, formative assessment

preparation and third grade students who received *ClassScape*,

computerized, formative assessment preparation for the End of Grade

reading achievement testing, as measured by the North Carolina End of

Grade Assessment?

50

Research Design

The research design for this study was descriptive comparative. This design allowed for the comparative differences between summative reading scores of the paper and pencil formative assessment compared to the computerized formative assessment. Differences will not be explained, nor describe why the differences in the summative reading scores occurred (Lodico, Spaulding, & Voegtle, 2010).

Target Population, Sampling Method, and Related Procedures

Target Population

The targeted population was reported reading achievement scores of third grade students in the state of North Carolina during the 2010- 2011 and 2011-2012 school years. The reading achievement scores from all the elementary schools in one district of a North Carolina school system (four total schools) that changed formative assessments in reading from *Progress Assessments* (pencil and paper testing booklet and answering sheet format) in 2010-2011 to *ClassScape* (online computer answer responses format) in 2011-2012 was the sample.

Sampling Method

The data sample consisted of approximately 600 third grade students' summative assessment scores from school years 2010-2012. The summative reading assessment scores for 2010-2011 (pencil and paper) were compared to scores for 2011-2012 (computerized) using an independent samples design t-Test for significance analysis during the assessment years 2010-2012 (Peterson & Siadat, 2009).

The data source consisted of reading assessment archival scores for third grade students that met criteria for the 2010-2011 and 2011-2012 school years. This criterion

51

was similar demographic information retrieved from the North Carolina Department of Education report card for elementary schools, and resided in the same county and school district during the school year. The sample of student reading scores represented both paper and pencil *Progress Assessment* formative assessment scores, and the End of Grade summative assessment scores in reading (2010-2011). Also, the sample of student reading scores represented both computerized *ClassScape* formative assessment scores, and End of Grade summative assessment scores in reading (2011-2012).

Sample Size

The ideal sample would be all third grade students in North Carolina that took the formative paper and pencil formative assessment format during the 2010-2011 school year and the following year changed to a computerized formative assessment format for the 2011-2012 assessment year; but sampling every population is not feasible (Lodico, 2010). Therefore, one district of four elementary schools with similar student demographic profiles (demographic data such as, socio-economics, gender, race, and special education needs) met the assessment criteria and was selected for this study. The sample is based on scores from North Carolina 3rd grade students attending four elementary schools in the same district. The sample consisted of 636 third grade students' summative assessment scores from school years 2010-2011 and 2011-2012.

Setting

The setting of the study was a North Carolina suburban school district that included four elementary schools. The schools had similar demographics, as reported in Table 1 (school year 2010-2011) and Table 2 (school year 2011-2012).

52

Table 1
Student Reading Performance Data for 2010-2011

School	White	Black	Hispanic	Asian	Two or More Races	Economically Disadvantaged	Language English Proficient	Students With Disabilities
A	<5%	15%	<5%	N/A	14%	13%	7%	31%
B	14%	12%	21%	20%	22%	18%	57%	51%
C	9%	14%	17%	N/A	15%	15%	14%	38%
D	11%	12%	<5%	<5%	7%	12%	27%	32%

School	White	Black	Hispanic	Asian	Two or More Races	Economically Disadvantaged	Language English Proficient	Students With Disabilities
A	8%	23%	28%	<5%	N/A	14%	12%	31%
B	19%	26%	24%	20%	<5%	26%	21%	27%
C	18%	40%	<5%	N/A	8%	32%	<5%	31%
D	15%	27%	33%	11%	7%	25%	46%	43%

School	White	Black	Hispanic	Asian	Two or More Races	Economically Disadvantaged	Language English Proficient	Students With Disabilities
A	71%	52%	63%	N/A	71%	55%	54%	34%
B	53%	53%	45%	60%	48%	45%	21%	22%
C	58%	42%	50%	N/A	46%	45%	71%	28%
D	58%	53%	49%	67%	61%	50%	27%	22%

School	White	Black	Hispanic	Asian	Two or More Races	Economically Disadvantaged	Language English Proficient	Students With Disabilities
A	21%	10%	7%	N/A	<5%	8%	8%	<5%
B	10%	10%	10%	<5%	26%	11%	<5%	<5%
C	15%	<5%	29%	N/A	31%	8%	14%	<5%
D	16%	9%	14%	22%	25%	13%	<5%	<5%

Table 2

Student Reading Performance Data for 2011-2012

School	White	Black	Hispanic	Asian	Two or More Races	Economically Disadvantaged	Language English Proficient	Students With Disabilities
A	12%	14%	12%	N/A	20%	15%	60%	28%
B	7%	15%	15%	13%	23%	19%	40%	25%
C	9%	14%	17%	N/A	15%	15%	14%	38%
D	11%	12%	<5%	<5%	7%	12%	27%	32%

School	White	Black	Hispanic	Asian	Two or More Races	Economically Disadvantaged	Language English Proficient	Students With Disabilities
A	12%	29%	27%	N/A	10%	29%	20%	41%
B	21%	23%	24%	13%	19%	22%	10%	23%
C	18%	40%	<5%	N/A	8%	32%	<5%	31%
D	15%	27%	33%	11%	7%	25%	46%	43%

School	White	Black	Hispanic	Asian	Two or More Races	Economically Disadvantaged	Language English Proficient	Students With Disabilities
A	59%	47%	50%	N/A	40%	44%	20%	27%
B	61%	54%	44%	75%	39%	51%	50%	40%
C	58%	42%	50%	N/A	46%	45%	71%	28%
D	58%	53%	49%	67%	61%	50%	27%	22%

School	White	Black	Hispanic	Asian	Two or More Races	Economically Disadvantaged	Language English Proficient	Students With Disabilities
A	18%	10%	12%	N/A	30%	13%	<5%	5%
B	11%	9%	18%	<5%	19%	8%	<5%	<5'%
C	15%	<5%	29%	N/A	31%	8%	14%	<5%
D	16%	9%	14%	22%	25%	13%	<5%	<5%

Recruitment

All the elementary schools within the school district were chosen based on student demographics, grade levels served, and the formative and summative reading assessments utilized during the 2010-2012 school years. Reading achievement scores were released after the Associate Superintendent of Testing signed the study's County Approval Proposal. The proposal required the Capella approved Scientific Merit Review document,

54

and IRB approval. After the documents were signed and approved, the data were downloaded onto a secure flash drive without all identifiable school information.

Instrumentation

The summative reading assessment for third grade was measured by the End of Grade (EOG) testing instrument of the state of North Carolina.. The EOG is a required curriculum based summative achievement test that is aligned to the North Carolina standard course of study. The reading assessment included 58 multiple choice test items that focused on vocabulary, reading comprehension by assessing the students on a variety of genres (selections); fiction, nonfiction, poetry, and informational passages. Students had an estimated time of 140 minutes to complete the End of Grade reading assessment.

The purpose of the assessment was to gain insight on students' ability to build literary experience, gain information and perform reading task (NCDPI, 2012). The student data was reported utilizing scaled scores, percentile ranks, achievement levels, and lexile scores, using developmental scale scores (intervals) to measure students' academic performance on the summative reading assessment (EOG).

The North Carolina Reading Comprehension (EOG) testing program has a range of reliability Coefficients from 0.82 to 0.94. Standard error of measurement is 2-3 points for students with scores within two standard deviations from theme and 4-6 points for students with scores that fall outside of two standard deviations from the mean. The North Carolina EOG tests, evidence of validity is provided through content relevance, response processes, and relationship of scores with other external variables. In the development phases of test construction, items that

55

showed no bias due to gender or ethnicity/race were identified and subsequently included in the tests. *(*The Education Alliance at Brown University, 2008)

Data Collection

The data collection procedures were conducted upon receiving county approval for release of data and review of third grade EOG reading assessment scores for school years 2010-2012. The county received a copy of this study's Scientific Merit Review research plan, proposal letter, and permit to release data from the researcher. The data requested was stripped of all student and teacher names, all identifiable information, and a release document was signed to ensure security.

Once the summative reading assessment data was collected from third grade students in a suburban North Carolina public school district who met the requirements of this study, the results were reported in chart form. The collection was a census sampling or convenience sampling of reading scores of third grade students who took formative assessments in the school years 2010-2012 utilizing paper and pencil (2010-2011), and computerized (2011-2012) formative reading assessments, and the summative End of Grade reading assessments.

Data Analysis Procedures

The data was collected and analyzed utilizing descriptive statistics (Lodico, 2010). An Independent sample t-Test was utilized to compare the End of Grade scores of the two groups of students using the comparative descriptive model; 2010-2011 student End of Grade scores were compared to the 2011-2012 EOG scores in reading. The t-Test was used to determine the result of students' EOG reading scores in 2010-2011 who took the

paper and pencil benchmark versus 2011-2012 students who took the computerized benchmark in reading. The independent sample t-Test was run in five different parts, first as one big test comparing all End of Grade summative reading assessment scores for the four schools, and then a total of four more t-Tests to run the summative reading assessment scores against each school for comparison. There were a total of five t-Tests run to determine if the paper and pencil formative assessment (2010-2011) and or the computerized formative assessment (2011-2012) differed on the summative End of Grade reading assessment scores. The summative scores were compared using Levene's test and recorded the percentiles using the mean, median, mode, frequency, range, maximum and minimum score. The results of the data analysis allowed the researcher to reject or confirm the hypothesis; as the statistical data was compared to determine if the independent samples t-Test results and the alpha was below .05. The results were reported in tables, and reflected the outcomes of the independent t-Test ran for all four schools collectively and independently.

This study utilized some of the same tools as a causal comparative design, in that the data as a mean or frequency for the computerized and paper-pencil assessments years. Inferential statistics were used to describe if there were any significant differences between the 2010-11 and 2011-12 school years' End of Grade summative reading assessment scores where the testing format changed (Lodico, et al., 2010). The student reading scores were compared and reported by student demographic characteristics for reading levels in Table 1 (2010-2011) and Table 2 (2011-2012). The variable that changed was the format of the formative assessment, and the schools were selected

57

because of the comparative demographic information for statistical matching (Lodico, et al., 2010).

Limitations of the Research Design

This study was limited in scope, for the sample only represented third grade student reading scores from four schools in one district in a North Carolina suburban school district with similar socioeconomic statuses. Second, the End of Grade assessment scores sampled were limited only to reading summative assessment scores of third grade elementary school students, since this is the first grade level summative assessments are administered. Therefore, no earlier summative assessment scores exited to use in statistical analysis for deeper comparison of student reading achievement.

Internal Validity

Validity in research focuses on ensuring that what the measurement "claims to measure is truly what it is measuring" (Lodico, Spaulding, & Voegtle, 2010, p. 93). Creswell (2009) referred "validity in quantitative research to whether one can draw meaningful and useful inferences from scores on particular instruments" (p. 235).

Internal validity refers to the collection of accurate information. An instrument is valid if it accurately and appropriately measures what it intends to measure, in this case, academic knowledge (Lodico, et al., 2010). The North Carolina End of Grade reading assessments provide a benchmark of measurement to evaluate student skills and progress toward state standards for reading, as measured by the requirements under the NCLB federal act. The End of Grade testing instrument in reading provided data that was valid, for it accurately measured the intended to measure, reading percentile scores (Lodico, et al., 2010). The research assumed all the retrieval data via district and county approval

58

were accurate, private and secure. The increase security measure taken by the researcher to only have the county designee access the identifiable information and recode student names with identification numbers as it pertained to their school and testing scores per year for paper and pencil 2010-2011, and computerized 2011-2012 summative reading assessment years.

External Validity

The external validity was confirmed by the use of the independent sample t-Test, for the results of the End of Grade summative assessment scores for the 2010-2011 and 2011-2012 school years. The summative assessment scores were compared to ascertain if the format of the formative assessment training impacted the End of Grade reading scores for third grade students in North Carolina.

Expected Findings

The expected finding of the data analysis was a significant difference in End of Grade reading assessment developmental scaled scores for students who took the computerized formative assessment compared to students who took the paper pencil formative assessments. According to formative assessment feedback studies, students who receive immediate feedback can receive instruction swiftly based on current data (Cauley, 2010).

Ethical Issues

Researcher's Position Statement

Conflict of interest assessment. The researcher did not have a conflict of interest with regard to the research and study being conducted. At the time of this study, the researcher did not work in the district where the research was conducted.

59

Position statement. It is my position that the study was void of any conflict of interest, and was conducted without ethical issues. The researcher followed every guideline and protocol according to the Capella Dissertation Manual. All research was handled in an ethical manner according to the approval and mandates of IRB.

Ethical Issues in the Study

There were no ethical issues or conflict of interest in this research study. No identifiable information of the sample was collected, and IRB approval was received. At the time of this study, I was employed by the county in another setting, as an instructional coach, and not as a teacher during the data collection phase of this study.

Chapter 3 Summary

This purpose of this proposed study was to examine if there were statistical differences in reading summative achievement scores of third grade students that took formative assessments in two different delivery formats in school years 2010-2011 (paper and pencil) and 2011-2012 (computerized). The study's research question targeted a population based on third grade student reading scores in North Carolina, from one school district (four schools) that had similar elementary school demographic profiles and the sample consisted of student reading achievement scores in the district during third grade.

The summative assessments from the End of Grade summative reading assessment scores for the 2010-2011 and 2011-2012 archival data were collected by release of county district officials per the IRB approval of Capella. The archival End of Grade reading scores were uploaded onto a secure, and preapproved flash drive. The

60

archival End of Grade scores were utilized to run independent sampled t-Test to determine if there was a significant difference in the summative assessment reading scores of third grade students who took formative assessments in the two different delivery formats. This study included research that supported the use of formative assessments to prepare for summative assessments, and theoretical supports of the constructivist theories.

61

CHAPTER 4. DATA ANALYSIS AND RESULTS

Introduction

This chapter explains data collection and analysis utilizing the descriptive statistics method (Lodico, Spaulding, & Voegtle, 2010). An independent samples t-Test was utilized to compare the End of Grade reading achievement scores of the two groups of students using the comparative descriptive model; 2010-2011 student End of Grade reading achievement scores were compared to the 2011-2012 student End of Grade scores in reading. The independent sample t-Test was employed to determine the result of students' End of Grade reading scores in 2010-2011 who took the paper and pencil benchmark versus the 2011-2012 End of Grade reading scores for students who took the computerized benchmark in reading. There were a total of five t-Test run to compare schools collectively and independently for each tested school year utilizing the two delivery formats of the formative reading assessment.

The purpose of this study was to examine and analyze the difference between reading achievement scores of two groups of third grade students; one group using the formative assessment, *Progress Assessments*, and the other group using the formative assessment, *ClassScape*, in order to compare the reading achievement scores of the third grade students using either a pencil and paper or computerized testing delivery formatted formative assessment. Archival data for 636 students were used.

The following research question guided this study:

R_1 Is there a significant difference in the End of Grade reading achievement scores between the third grade students who received *Progress Assessments*, paper and pencil, formative assessment preparation and third

62

grade students who received *ClassScape*, computerized, formative

assessment preparation for the End of Grade reading achievement testing,

as measured by the North Carolina End of Grade Assessment?

H0   There will be no significant difference in the End of Grade reading

achievement scores between the third grade students who received

Progress Assessments, paper and pencil, formative assessment

preparation and third grade students who received *ClassScape*,

computerized, formative assessment preparation for the End of Grade

reading achievement testing, as measured by the North Carolina End of

Grade Assessment?

H1   There will be a significant difference in the End of Grade reading

achievement scores between the third grade students who received

Progress Assessments, paper and pencil, formative assessment

preparation and third grade students who received *ClassScape*,

computerized, formative assessment preparation for the End of Grade

reading achievement testing, as measured by the North Carolina End of

Grade Assessment?

This research study investigated whether the change in the delivery format of the

formative assessment affected the outcome of the summative reading assessment. During

the 2010-2011 school year, third grade students took paper and pencil formative reading

assessments, and in the 2011-2012 school year, the second group of third grade students

took computerized formative assessments. Both groups of students completed the state's

summative assessment for the End of Grade reading assessment, which was in a paper and pencil delivery format.

Description of the Sample

The sample consisted of 636 third grade students' whose summative reading assessment scores from the 2010-2011 and 2011-2012 school years were used as a data source for testing the research hypothesis. Therefore, one district of four elementary schools with similar student demographic profiles (demographic data such as, socio-economic, gender, race, and special education needs) was selected for this study. Only reading achievement scores that met the criteria of this study were used for analysis; third grade students attending the 2010-2011 school year with an EOG reading achievement score, and third grade students attending during the 2011-2012 school year with an EOG reading achievement score.

Table 3 displays the frequency counts for selected demographic variables. Of the 636 third grade students in the study, 316 scores represented the Paper group (49.7%) and 320 scores represented the Computer group (50.3%). The largest group of students' scores was from School A (28.9%), while the smallest group was from School C (19.2%). Of the total number of achievement scores, 316 were female students (49.7%), and 320 were male students (50.3%). The majority of achievement scores were from students who represent Black (56.4%), White (16.7%), or Hispanic (14.9%) ethnic groups. Comparing reading levels, 312 student scores were rated as "Far Below" grade level standards (27.7%) or "Below" (21.4%), but the largest individual group was "At Grade" with 255 students (40.1%). Sixty-nine students were rated as "Superior" to grade level in reading (10.8%).

64

Table 3

Frequency Counts for Selected Variables (N = 636)

Variable	Category	*n*	%
Groups by Formative Assessments			
	Paper	316	49.7
	Computer	320	50.3
Scores by School			
	A	184	28.9
	B	155	24.4
	C	122	19.2
	D	175	27.5
Scores by Gender			
	Female	316	49.7
	Male	320	50.3

Table 3 *Continued*

Variable	Category	*n*	%
Scores by Ethnicity			
	Asian	13	2.0
	Black	359	56.4
	Hispanic	95	14.9
	Indian	11	1.7
	Multicultural	49	7.7
	Pacific Islander	3	0.5
	White	106	16.7
EOG Scores by Level			
	Far Below	176	27.7
	Below	136	21.4
	At Grade	255	40.1
	Superior	69	10.8

Table 4 displays the frequency counts for reading percentile groups (total sample $M = 45.48$, $SD = 25.81$, $Mdn = 46.00$, Mode = 46). The percentile scores ranged in size from 1st to 98th, based on the lowest score received by a student on the End of Grade reading summative assessment was 1, and the highest score received was 98. As for statistical assumptions, the primary dependent variable (reading percentile score) was

66

normally distributed based on inspection of a frequency histogram and a box plot. The

scores were normally distributed based on inspection of a frequency histogram and box

plot (Appendix A).

Table 4

Frequency Counts for Reading Percentile Score (N = 636)

Variable	Category	n	%
Reading Percentile [a]			
	1 – 24	176	27.7
	25 – 43	136	21.4
	44 – 65	172	27.0
	66 – 98	152	23.9

[a] Score: $M = 45.48$, $SD = 25.81$, $Mdn. = 46.00$, Mode = 46.

By design of the study, all 636 reading scores were independent from scores of

other children and not considered to be repeated measurements of the same student. The

Levene's Test for Equality of Variance for each of the five t tests was not significant

(Appendix B). Given these preliminary analyses, the statistical assumptions for the t tests

were adequately met (George & Mallery 2011).

67

Summary of the Results

The primary hypothesis found no significant overall differences ($p = .19$) between those scores of students that were in the paper sample and those in the computer sample, t(634) =1.32, p=.19 (Table 5). When the t-Tests were repeated for each of the four individual schools, significant differences were found between the groups of students at three of the four schools. However, the differences, though statistically significant, but were not educationally relevant. There were weak correlations, for the p values were not less than .05 (George & Mallery, 2011). Statistical differences were found in schools A and D, with student scores higher on the End of Grade summative reading assessment during the year formative assessments were completed using a computerized format. School C scores were higher on summative End of Grade reading assessments the year that paper and pencil formative assessments were given. School B showed no significant difference in the summative End of Grade reading assessment scores, whether students took the computerized or paper and pencil reading formative assessments delivery formats.

Table 5
Comparison of Reading Percentiles based on Group t-Tests for Independent Means
(N = 636)

Sample	Group	n	M	SD	r_{pb}	t	p
Total Sample (*N*= 636)					.05	1.32	.19
	Paper	316	46.84	24.88			
	Computer	320	44.14	26.67			
School A (*n* = 184)					.16	2.14	.03
	Paper	84	46.30	24.22			
	Computer	100	38.03	27.50			
School B (*n* = 155)					.18	2.29	.02
	Paper	76	51.83	24.28			
	Computer	79	42.80	24.84			
School C (*n* = 122)					.25	2.77	.007
	Paper	62	41.50	27.41			
	Computer	60	55.10	26.82			
School D (*n* = 175)					.04	0.52	.61
	Paper	94	46.80	23.78			
	Computer	81	44.89	25.09			

69

Detailed Analysis

Table 5 displays the results of the five t-Tests for independent means for the entire sample ($N = 636$) and the four individual schools. These tests compared the two groups of students (paper versus computer). Inspection of the table found all 10 standard deviations to be similar ranging in size from $SD = 23.78$ to $SD = 27.50$ suggesting similar distributions among the subset scores. In the overall test, no significant differences were found between the two groups for the reading percentile score t (634) = 1.32, $p = .19$. However, at the individual schools, reading percentile scores were higher for the paper group at School A (t [182] = 2.14, $p = .03$), School B (t [634] = 1.32, $p = .02$) but lower at School C (t [120] = 2.77, $p = .007$), and at School D, the two groups had similar scores t (173) = 0.52, p= .61.

The primary hypothesis found no significant overall differences at the 5% level of significance ($p = .19$) between those student scores that were in the paper sample and those student scores in the computer sample, t(634) =1.32, p=.19 (Table 5). When the t-Tests were repeated for the student reading scores for each of the four individual schools, significant differences were found between the groups of student scores at three of the four schools. However, the differences, though statistically significant, but were not educationally relevant. There were weak correlations, for the p values were not less than .05 (George & Mallery, 2011).

Table 5 also includes the point-biserial correlations (r_{pb}) (the Pearson correlations between a dichotomous variables and a continuous variables) as a measure of the strength of the relationship. The largest point-biserial correlation was at school C ($r_{pb} = .25$, $r_{pb}^2 = .063$). The r_{pb}^2 (the coefficient of determination) accounted for 6.3% of the variance in

70

the relationship between students' reading score and group (paper versus computer). School D showed no difference in either testing format (p = .61). All five correlations were considered to be weak, based on the relationship between the groups' format of test and student scores (Cohen, 1988). This combination of findings provided partial support to reject the null hypothesis, since collectively the schools represented $p= .19$, but three of the four schools had p values less than .05 (George, 2011). The differences are relevant for looking at statistical information and to make educational decisions per school as it relates to testing format delivery.

Chapter 4 Summary

In summary, this study used archival data from 636 students to examine and analyze the difference between two groups of student achievement scores; one group using the formative assessment, *Progress Assessments*, and the other group using the formative assessment, *ClassScape*, in order to compare the reading achievement scores of the third grade students using those differing testing delivery formatted formative assessments. The primary hypothesis (reading differences based on paper versus computer assessment) received partial support (Table 5). In chapter 5, findings of this study will be compared to the review of literature, and recommendations will be suggested as the studies' conclusions and implications will be determined.

71

CHAPTER 5. CONCLUSIONS AND DISCUSSION

Introduction

This final chapter summarizes the results of the analysis of summative reading achievement data, discusses the findings to answer the research question, presents a summary of the results in relation to the literature, provides limitations, proposes implications of the results for practice, recommends formats for future studies, and provides the final conclusion.

The purpose of this study was to examine and analyze the difference between two groups of student achievement scores; one group using the formative assessment, *Progress Assessments*, and the other group using the formative assessment, *ClassScape*, in order to compare the summative reading achievement scores of the third grade students using these differing testing delivery formatted formative assessments. This study compared the summative reading achievement scores of third grade students taking formative assessments in either a pencil and paper, or computer testing delivery format. Comparisons were measured using the North Carolina End of Grade Assessments from two consecutive school years (2010-2012).

This study was conducted using the comparative descriptive research design of quantitative methodology, and was undertaken to find data that would provide data to compare and analyze summative reading assessment scores when formative assessment delivery formats were different for two sets of third grade students. The catalyst was a change in the formative reading testing format in a North Carolina public elementary school during the 2011-2012 school year, when the formative reading assessment was

changed from a paper and pencil reading benchmark to a computerized online reading

benchmark. During the previous school year, 2010-2011, reading benchmark assessments

were taken by pencil-paper, and results were used to inform stakeholders of student

progress in reading standards and objectives taught quarterly. The change in the delivery

format of the formative reading assessment initiated teacher conversations that

questioned whether students' scores would improve, falter, or not change on the state

mandated End of Grade summative reading assessment.

Summary of the Results

The summative reading achievement scores of two groups of third grade students

(N=636) were analyzed in this study. The sample included the North Carolina End of

Grade reading achievement scores of third grade students who attended four elementary

schools within the same school district during the 2010-2011 and 2011-2012 school

years. North Carolina End of Grade summative reading assessment scores were

compared for the students that completed the paper and pencil delivery format (2010-

2011) versus scores of students who completed the computerized delivery format (2011-

2012), as reported in Table 5. The individual schools' student reading percentile scores

were higher for the paper group at School A (t [182] = 2.14, p = .03) and at School B

(t [634] = 1.32, p = .19) but lower at School C (t [120] = 2.77, p = .007), and similar

scores for school D t (173) = 0.52, p= .61.

The primary hypothesis found no significant overall differences at the 5% level of

significance (p = .19) between those student scores that were in the paper sample and

those student scores in the computer sample, t(634) =1.32, p=.19 (Table 5). When the t -

Tests were repeated for the student reading scores for each of the four individual schools,

73

significant differences were found between the groups of student scores at three of the four schools. However, the differences, though statistically significant, but were not educationally relevant. There were weak correlations, for the p values were not less than .05 (George & Mallery, 2011).

- Statistical differences were found in student scores at School A and School D, with student scores higher on the End of Grade summative reading assessment during the year the formative assessment was completed using a computerized delivery format.
- Student scores at School C were higher on the End of Grade summative reading assessment the year that the formative assessment was in a paper and pencil delivery format.
- Student scores at School B showed no significant difference in the End of Grade summative reading assessment scores whether students completed computerized or paper and pencil delivery formats in formative reading assessments.

Schools A and B independent t-Test showed that students scored significantly higher on the End of Grade summative reading assessment when they took the paper and pencil delivery formatted formative reading assessment. School C showed students scored significantly higher on the End of Grade summative reading assessment when they took the computerized delivery formatted formative reading assessment. School D showed no significant difference in student the End of Grade reading assessment scores between the paper and pencil or computerized delivery formats of the formative reading assessment. Although three schools showed significant differences, there was no

74

educational relevance, for the Pearson's r correlations of the p values of the t-Test were weak (George, 2011).

Discussion of the Results

There was a difference in the End of Grade reading achievement scores between the third grade students who received *Progress Assessments*, paper and pencil, formative assessment preparation and third grade students who received *ClassScape*, computerized, formative assessment preparation for the End of Grade reading achievement testing, as measured by the North Carolina End of Grade Assessment. The comparison of the independent two tailed t-Test results showed that there were significant differences between student summative reading achievement scores on the End of Grade reading assessment during the paper-pencil (2010-2011) and computerized (2011-2012) formative assessment years. As a result, this study rejected the null hypothesis.

The final results for this sample showed that the individual schools' student reading percentile scores were:

- higher for the paper and pencil student groups at School A (t [182] = 2.14, p = .03) and at School B (t [634] = 1.32, p = .19) but

- lower at School C (t [120] = 2.77, p = .007) and

- School D was about the same.

The variables identified were the changes in the delivery format of the third grade formative reading assessment testing. The population sampled had similar demographic characteristics but showed variances in summative assessment scores, and the overall statistical differences could not be generalized to all populations of third grade students who had a difference in the delivery format of formative assessment testing.

75

There was a statistical difference in reading achievement for students taking the formative reading assessment in two different delivery formats, however, the student summative scores were not consistent with the delivery format that produced higher student scores. So, there is conflicting evidence on the formative assessment delivery formats. This study compared the summative End of Grade reading scores as students took the formative assessments in differing delivery formats; paper-pencil (2010-2011) and computerized (2011-2012).

The standard deviation showed the variance was in a similar range, but it is still unknown what caused certain schools to fall within the various levels utilizing the paper and pencil or computerized formative assessment delivery format. Without additional study, an advantage cannot be concluded for students receiving instruction using either formative assessment delivery formats. Possibly the summative reading assessment scores are based on what the students learned or retained versus the format of the assessment.

Discussion of the Results with Relation to the Literature
Constructivism and Formative Assessments

Constructivism is theoretical framework for the study, as learning is student centered and encompasses building of prior knowledge, and conceptualizing the old knowledge with new experiences to enhance learning (Yurdabakan, 2011). The relationship between this study and the literature reviewed were similar as the benefits of the formative assessments aided in the growth of student scores on the summative assessments, schools A, B, C, and D all utilized a practice formative assessment paper-pencil (2010-2011) and computerized (2011-2012), and all schools scored within a 2%

76

difference in the mean score when compared (46.84 paper-pencil) and (44.14 computerized), as displayed in Table 5. The results of this study were aligned to the research in that studies showed that the format of the assessment had no effect on student outcomes (Korbin & Young, 2003).   The use of formative assessments has benefits that lead to test preparation, as immediate feedback can be given using computerized formats which teachers can provide direct instruction to target deficits (Lane, 2007).

Paper-Pencil and Formative Assessments

The paper-pencil formats also provide assessment practice, and can provide a construct to build a construct of their own knowledge base by the process of accommodation and assimilation (Piaget, 1953). The results of this study supported the literature as it is related to the effects of format of formative assessments effect on reading summative assessment scores. In this study, it was found that the format of the third grade reading formative assessment (paper-pencil or computerized) did not have a major effect on the outcomes of the reading summative assessment. Formative assessments can be utilized as frequent measures to obtain quantitative data on specific content area, and redirect the teaching focus to improve long-term learning outcomes (Dorn, 2010). In the study conducted by Boo and Vispoel, results found that "test takers preferred the computerized format of assessments, but the computerized assessment scores were lower than the paper-pencil assessments given" (2012, p. 445). In this research study, the results of school C were higher on summative reading assessment when the paper-pencil formative assessment was given, as reported in Table 5.

77

Computerized Assessments

The research of Pomplun and Custer supports computerized assessments as the advantages include "reduction of test taking time, immediate feedback, increased test security, increased examinee interest and improved administration-scoring" (2005, p. 153). The advantages to computerized assessments also included "flexible scheduling, elimination of test booklets, focused observation of examinee's test taking pattern, and reduced measurement of errors" (Boo & Vispoel, 2012, p. 443). This research study results found that schools A and D scored higher on the summative reading assessment when computerized formative assessments were utilized, according the study conducted by Poplun and Custer, "the results of the paper-pencil compared to the computerized format produced lower scores for K-3 reading testing" (2005, p. 153). There were variables that may have promoted threats to score equivalence such as, "computer delivery of passages, difference in student computer familiarity, and teacher delivery method" (Poplun & Custer, 2005, p.154).

Quantitative Research

Quantitative data has had a tremendous impact on the educational systems across the world. Numerical data has informed school reforms, policies, hiring procedures, funding, and assessments. Schools systems rely heavily on the data to improve student achievement in reading, math and science. No Child Left Behind policy declares scientific based research as empirical, and evaluated using experimental or quasi-experimental methods. "The research utilizes data analysis to test hypothesis, and justify general conclusions about educational issues" (Horn, 2004, p. 196).

78

In order for stakeholders to have a voice in the educational reforms of present times, quantitative research is the anchor of all educational studies that motivate political pressure for educational reform. It is imperative for educators to implore the application of systematic scientific based research activities in lesson planning, activities, professional development, and continuing education demands. According to Horn (2004), teachers, administrators, parents, policy makers, and students are all "epistemological bases for educational decision making" (p. 197), but the chosen research methodology of 21st century reform employs statistical analyses of casual relationships that affect education interventions. The task of schools making improvements is a joint effort between educators, policy makers, and stakeholders within the close proximity of every child's educational environment.

Limitations

The limitations of this study were:

1.   There were no pre-assessments to compare the student scores in second grade to third grade.

2.   It is not known if students were properly trained or taught strategies to take the computerized formative assessments or paper and pencil formative assessments.

3.   There are environmental factors that may have affected the reading summative assessment scores utilizing the computer format, for students may not have been able to manipulate the computer formatted testing materials as readily as that of a paper pencil format. The computer has the ability to freeze, log the student off line or may not show all available

79

options on one screen. Students also did not have the ability to write on the computerized screen, use strategies such as writing summaries, eliminating wrong answers and underlining or circling important information.

4.	There was a limit to the number of schools that could be sampled, for not all schools in the county opted to take the computerized formative assessment delivery format, so the study data was limited to four schools within one school district, which limited the sample size.

Implication of the Results for Practice

In the area of curriculum and instruction, teachers should be provided with greater knowledge of how students construct knowledge in the area of test taking and how application of test taking skills can be affected by utilizing specific forms of formative assessments. Therefore, school districts should provide professional development training to teachers to familiarize them with differing formative and summative assessment delivery formats, so they can teach test-taking strategies to students in preparation for all academic content assessments. The goal would be for students' achievement scores to represent their level of learning, rather than do poorly due to uncertainty with the delivery format of assessments (formative or summative).

Recommendations for Further Research

Specific recommendations for future research include:

- Conduct research on the differences between the paper and pencil delivery format and computerized delivery format of mathematics formative assessments on mathematics summative assessments for third grade students.

المنارة للاستشارات

www.manaraa.com

- Conduct researched to find out if there would be differences in reading achievement at other grade levels, such as fourth or fifth grade, as the testing format changed, utilizing the third grade summative assessment scores as a pre-test.

- Conduct a mixed method study with a qualitative survey for parents and teachers sharing their experiences with differing delivery formats on formative assessments in preparation for state mandated summative assessments.

- Conduct research using a larger sample for comparison of student reading achievement data.

Conclusion

This study provided data that rejected the null hypothesis, and found there was a significant difference in the End of Grade reading achievement scores between the third grade students who received *Progress Assessments*, paper and pencil, formative assessment preparation and third grade students who received *ClassScape*, computerized, formative assessment preparation for the End of Grade reading achievement testing, as measured by the North Carolina End of Grade Assessment. The comparison of the independent two tailed t-Test results showed that there were significant differences of the results of the End of Grade reading assessment during the paper-pencil formative assessment year (2010-2011) and the computerized formative assessment year (2011-2012). As a result, this study rejected the null hypothesis. However, the study was not able to conclude which formative assessment format is assisting students in summative assessment test scores more effectively.

81

The final results for this sample showed that the individual schools' student reading percentile scores were higher for the paper group at School A (*t* [182] = 2.14, *p* = .03) and at School B (*t* [634] = 1.32, *p* = .19) but lower at School C (*t* [120] = 2.77, *p* = .007), and similar scores at School D t (173) = 0.52, p= .61. The variables were identified as the change in the delivery format for testing. The population sampled had similar demographic information, but showed variances in summative assessment scores, so the overall statistical differences could not be generalized to all populations of third grade students who had a difference in the delivery format of formative testing. For student achievement scores that remained the same on summative assessments after taking paper pencil versus computerized assessments, the delivery format of the formative assessment had little or no effect on the students' summative scores.

Although there was not a large difference in students' summative scores, the educational relevance of the study encourages schools to investigate other factors at their sites. It would be valuable for educators to monitor instruction and technology to discover why certain schools had little or no difference in student summative reading assessment scores by inspecting the use of computers, looking at the length of time students are engaged in reading instruction, and reviewing student preparation for assessments utilizing delivery formats of both paper and pencil and computer assessments.

REFERENCES

Anakwe, B. (2008). Comparison of student performance in paper-based versus computer-based testing. *Journal of Education for Business, 84*(1), 13-17.

Almala, A. H. (2006). Applying the principles of constructivism to a quality E-learning environment. *Distance Learning, 3*(1), 33-40.

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

Becker, H.J. (2001). *How are teachers using computers in instruction*? Paper presented at the annual meeting of the American Educational Research Association (AERA).

Becker, H.J., & Ravitz, J. (1999). The influence of computer and Internet use on teachers' pedagogical practices and perceptions. *Journal of Research on Computing in Education, 31*(4), 356–384.

Bergan, R., & Bergan, R (2009) Benchmark Assessment in Standards-Based Education. Assessment Technology Incorporated. ati.pearson.com/.../Interim-Benchmark-Assessments-Paper

Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-74.

Boo, J., & Vispoel, W. (2012). Computer Versus Paper-And-Pencil Assessment Of Educational Development: A Comparison Of Psychometric Features And Examinee Preferences. *Psychological Reports*, *111*(2), 443-460. doi:10.2466/10.03.11.PR0.111.5.443-460

Brewer, J., & Daane, C. J. (2002). Translating constructivist theory into practice in primary-grade mathematics. *Education, 123*(2), 416-421,426. Retrieved from http://search.proquest.com/docview/196435163?accountid=27965

Burston, J., & Monville-Burston, M. (1995). Practical design and implementation considerations of a computer adaptive foreign language test: The Monash/Melbourne french CAT. *CALICO Journal, 13*(1), 26.

Bush, G. (2006). Learning about learning: From theories to trends. *Teacher Librarian, 34*(2), 14- 18.

83

Catterick, D. (2007), "Do the philosophical foundations of online learning disadvantage non-western students?", In Edmundson, A. (Ed.), Globalized E-learning Cultural Challenges, Information Science Publishing, Hershey, PA, pp. 116-29.

Carpentier, V. (2008). Quantitative sources for the history of education. *History Of Education*, *37*(5), 701-720. doi:10.1080/00467600802056963

Cauley, K. M., & McMillan, J. H. (2010). Formative assessment techniques to support student motivation and achievement. *The Clearing House, 83*(1), 1-6.

ClassScape, (2011). Center for Urban Affairs & Community Services. https://w8ww.classscape.org/ClassScape3/

Clariana, R., & Wallace, P. (2002). Paper-based versus computer-based assessment: Key factors associated with the test mode effect. British Journal of Educational Technology, 33, 593-602

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). New Jersey:NY Lawrence Erlbaum.

Conderman, G., & Strobel, D. (2006). Problem solving with guided repeated oral reading instruction. *Intervention in School and Clinic, 42*(1), 34-39.

Creswell, J. (2009) Research, Design, Qualitative, Quantitative, and Mixed Methods Approaches, 3rd ed. 151-153 Los Angeles: Sage.

CTB/McGraw-Hill; research demonstrates technology-enhanced formative assessment promotes student achievement. (2008). *Education Business Weekly,* 15.

Deville, C., & Chalhoub-Deville, M. (2011). Accountability-assessment under No Child Left Behind: Agenda, practice, and future L*anguage Testing, 28*(3), 307-321. doi:http://dx.doi.org/10.1177/0265532211400876

Dennis, D. (2008). Are assessment data really driving middle school reading instruction? What we can learn from one student's experience. *Journal of Adolescent & Adult Literacy, 51*(7), 578-587.

Diamond, J. (2007). Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction. *Sociology of Education, 80*(4), 285-313.

Dorn, S. (2010). The political dilemmas of formative assessment. *Exceptional Children, 76*(3), 325-337.

Dougherty, Stahl, K.A., & Schweid, J. (2013). Beyond March Madness: Fruitful Practices to Prepare for High-Stake ELA Test. *Reading Teacher*, 67(2), 121-125.

Duffy, M., Giordano, V., etl. (2008). No Child Left Behind: Values and Research Issues in High-Stakes Assessments. *Counseling and Values, 53(1), p 53.*

The Education Alliance at Brown University. (2008). "Outcome measures and data sources." Evaluation Toolkit for Magnet School Programs. Retrieved from: www.evaluationtoolkit.org/resources/66/.../434_sam_reliability.pdf?.

Educational Broadcasting Corporation, (2004). Constructivism as a Paradigm for Teaching and Learning. "How does this theory differ from traditional ideas about teaching and learning?" *Concept to Classroom, 1-2.* Retrieved from http://www.thirteen.org/edonline/concept2class/constructivism/index_sub1.html

Egan, K. (2005). Students development in theory and practice: The doubtful role of research. *Harvard Educational Review, 75*(1), 25-41

Fawcett, L. M., & Garton, A. F. (2005). The effect of peer collaboration on children's problem-solving ability. *British Journal of Educational Psychology, 75*(00070998), 157-169.

Felton, K.A, & Akos, P. (2011). The Ups and Downs of 3rd Grade. *Educational Leadership*, 68(7), 28-31

George, D., & Mallery, P. (2011). *SPSS for windows step by step: A simple guide and reference 18.0 update* (11th ed.). Boston, MA: Allyn& Bacon.

Hagstrom, F. (2006). Formative learning and assessment. *Communication Disorders Quarterly, 28*(1), 24-36, 61-62

Hatch, J. A. (2010). Rethinking the relationship between learning and development: Teaching for learning in early childhood classrooms. *The Educational Forum, 74*(3), 258-268.

Hicks, M. (2013). Interim Assessments – What they are and how to use them to benefit student learning?. https://www.nwea.org/blog/2013/interim-assessments-use-benefit-student-learning/#sthash.d5ZpEjp2.dpuf

Horn, R.Jr. (2004). The new federal definition of educational research: Implications for educators. *The Teacher Educator, 39*(3), 196-211.

Huitt, W., & Hummel, J. (2003). Piaget's theory of cognitive development. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. Retrieved from http://www.edpsycinteractive.org/topics/cognition/piaget.html

Jian-Wei Lin, & Yuan-Cheng Lai. (2013). Harnessing collaborative annotations on online formative assessments. *Journal of Educational Technology & Society, 16*(1), 263-n/a.

Johnson-Glenberg, M.C. (2010). Embedded formative e-assessment: who benefits, who falters. Educational Media International, 47(2), 153-171. doi: 10.1080/09523987.2010.492681

Kamil, M. L. (2004). The current state of quantitative research. *Reading Research Quarterly, 39*(1), 100-107.

Kathleen, M. C., & James, H. M. (2010). Formative assessment techniques to support student motivation and achievement. *The Clearing House, 83*(1), 1-6.

Katherine, C. P., & Cody, J. K. (2009). Cognitive and social constructivism: Developing tools for an effective classroom. *Education, 130*(2), 241-250

Kelle, U. (2006). Combining qualitative and quantitative methods in research practice: Purposes and advantages. *Qualitative Research in Psychology, 3*(4), 293-311. doi:10.1177/1478088706070839

Korbrin, J.L., Young, J.W. (2003). The Cognitive Equivalence of Reading Comprehension Test Items Via Computerized and Paper-and-Pencil Administration. *Applied Measurement In Education, 16(2), 115-140.*

Lane, A. (2007). Comparison of teacher educators instructional methods with the constructivist ideal. *The Teacher Educator, 42*(3), 157-184. Lai, M., & McNaughton, S. (2009). Not by achievement analysis alone: How inquiry needs to be informed by evidence from classroom practices. *New Zealand Journal of Educational Studies, 44*(2), 93-108.

Lodico, M.G., Spaulding, D., & Voegtle, K. (2010). Methods in educational Research: From Theory to Practice. Jossey-Bass 2nd ed., 201-213.

Lomax, R. G. (2004). Whither the future of quantitative literacy research? *Reading Research Quarterly, 39*(1), 107-112.

Lunz, M.E. and Bergstrom, B.A., (1991). "Comparability of Decisions for Computer Adaptive And Written Examinations." Journal of Allied Health. 1991, 15-23.

Lunz, M.E., J.A. Stahl, and Bergstrom, B.A., (1993). *Targeting, Test Length, Test Precision, and Decision Accuracy for Computerized Adaptive Tests.* Paper

86

presented at the Annual Meeting of the American Educational Research Association, Atlanta.

Marzano, R. J., & Kendall, J. S. (2006). *The New Taxonomy of Educational Objectives*. Corwin Press.

Mayer, D. E. (2006). Research funding in the U.S.: Implications for teacher education research. *Teacher Education Quarterly, 33*(1), 5-18.

McKeown, M., Beck, I., & Blake, R. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading Research Quarterly, 44*(3), 218-253.

Miranda, M. V. (2009). Creating the successful community college student: Using behaviorism to foster constructivism. The Community College Enterprise, 15 (1), 21-38.

Moore, K. (2009*).* Effective instructional strategies. *Thousand Oaks, CA: Sage Publications, Inc.*

North Carolina Department Public Instruction (NCDPI). (2003). End-Of-Grade Test Scores NCDPI-Division of Accountability Services/Testing Section. http://www.ncpublicschools.org/accountability/testing/eog/

No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002).

Overbay, A., Patterson, A.S., Vasu, E. S., & Grable, L.L. (2010). Constructivism and technology use: findings from the Impacting Leadership project Educational Media International, 47(2), 103-120. Doi:10.1080/09523987.2010.492675

Ozer, O. (2004). "Constuctivism in Piaget and Vygotsky" The Fountain on Knowledge, Life, and Belief. Issue 48.

Peterson, E., & Siadat, M. V. (2009). Combination of formative and summative assessment instruments in elementary algebra classes: A prescription for success. *Journal of Applied Research in the Community College, 16*(2), 92-102.

Piaget, J. (1953). *The origins of intelligence in children*. New York, NY: Basic Books.

Porcaro, D. (2011). Applying constructivism in instructivist learning cultures. *Multicultural Education & Technology Journal, 5*(1), 39-54. doi:10.1108/17504971111121919 \

Pomplun, M., & Custer, M. (2005). THE SCORE COMPARABILITY OF COMPUTERIZED AND PAPER-AND-PENCIL FORMATS FOR K-3 READING TESTS. *Journal Of Educational Computing Research*, *32*(2), 153-166.

Powell, K. C., & Kalina, C. J. (2009). COGNITIVE AND SOCIAL CONSTRUCTIVISM: DEVELOPING TOOLS FOR AN EFFECTIVE CLASSROOM. *Education, 130*(2), 241-250.

Reis, S., & Field, G. (2007). Exploring the new literacies using two new approaches: The schoolwide enrichment model in reading and renzulli learning. *New England Reading Association Journal, 43*(1), 30-35, 90-91

Roehrig, A., Duggar, S., Moats, L., Glover, M., & Mincey, B. (2008). When teachers work to use progress monitoring data to inform literacy instruction. *Remedial and Special Education, 29*(6), 364-382.

Rehfeldt, R. A., & Hayes, L. J. (1998). The operant-respondent distinction revisited: Toward an understanding of stimulus equivalence. *The Psychological Record, 48*(2), 187-210.

Ronan, Amanda. (2015) Every Teacher's Guide to Assessment. Edudemic, April 29, 2015. Retrieved from: http://www.edudemic.com/summative-and-formative-assessments/.

Sly, L. (1999). Practice tests as formative assessment improve student performance on computer-managed learning assessments. *Assessment and Evaluation in Higher Education, 24*(3), 339-343.

Smith, T. M., Desimone, L. M., & Ueno, K. (2005). "Highly qualified" to do what? the relationship between NCLB teacher quality mandates and the use of reform-oriented instruction in middle school mathematics. *Educational Evaluation and Policy Analysis, 27*(1), 75-109.

Stecker, P., Lembke, E., & Foegen, A. (2008). Using progress-monitoring data to improve instructional decision making. *Preventing School Failure, 52*(2), 48-58.

Stuart, E. (2007). Estimating causal effects using school-level data sets. *Educational Researcher, 36*(4), 187-198.

Stahl, K.A., Schweid, J. (2013). Beyond March Madness: Fruitful Practices to Prepare for High-Stakes ELA Tests. *The Reading Teacher. 67(2) p.121-125.*

Weiss, D.J. (1983). "Introduction." New Horizons in Testing: Latent Trait Test Theory And Computerized Adaptive Testing, edited by D.J. Weiss, 1-8. New York: Academic Press.

Weiss, D.J. (1982). "Improving Measurement Quality and Efficiency with Adaptive Testing." Applied Psychological Measurement 6, 4, 473-492.

White, G. (2003). Standardized mathematics scores as a prerequisite for a first programming course. *Mathematics and Computer Education, 37*(1), 96-104.

Windschitl, M. (2002). Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of Educational Research, 72*(2), 131-175.

Yurdabakan, I. (2011). The view of constructivist theory on assessment: Alternative assessment methods in education. *Egitim Bilimleri Fakultesi Dergisi, 44*(1), 51-77.

APPENDIX A. FREQUENCY HISTOGRAM AND BOX PLOT FOR THE READING PERCENTILE SCORE (N = 636)

www.manaraa.com
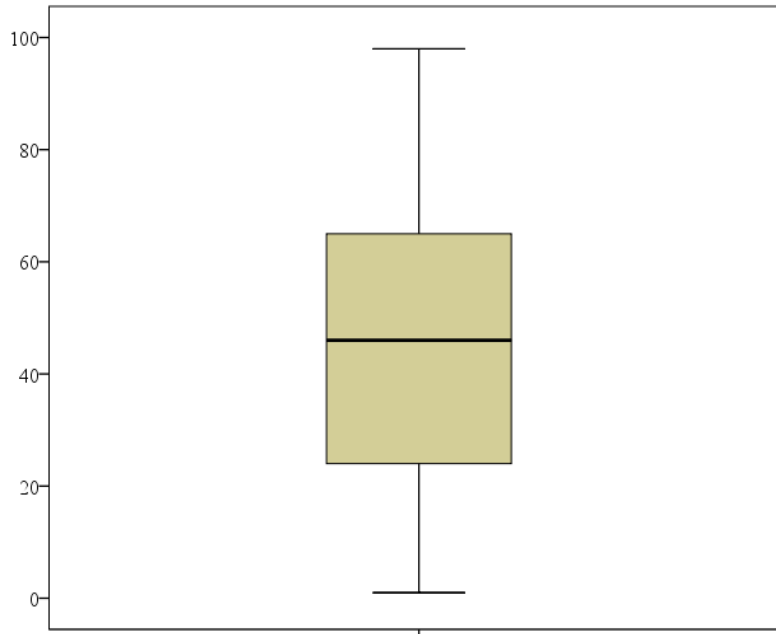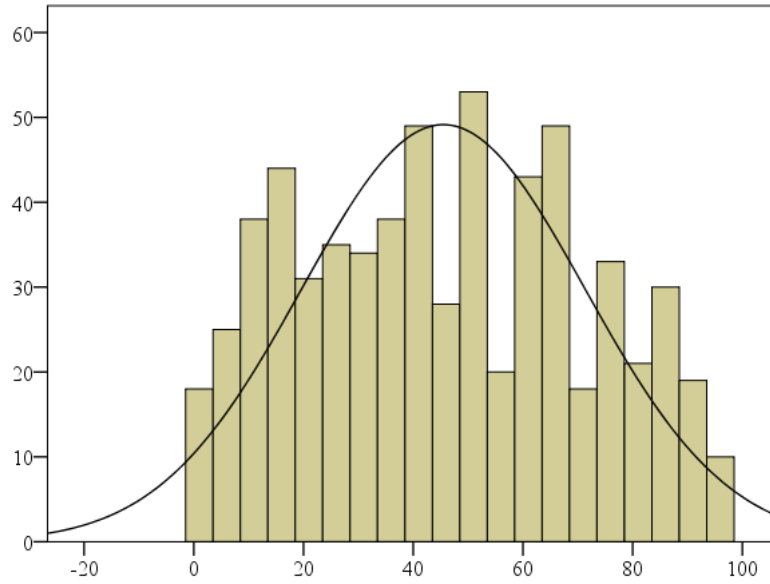
APPENDIX B. LEVENE'S TEST FOR EQUALITY OF VARINACES FOR THE t TEST COMPARISONS

Sample	F	p
Total Sample (N= 636)	2.93	.09
School 338 ($n = 184$)	1.93	.17
School 401 ($n = 155$)	0.03	.87
School 410 ($n = 122$)	0.07	.79
School 414 ($n = 175$)	1.49	.22

APPENDIX C. STATEMENT OF ORIGINAL WORK

Academic Honesty Policy

Capella University's Academic Honesty Policy (**3.01.01**) holds learners accountable for the integrity of work they submit, which includes but is not limited to discussion postings, assignments, comprehensive exams, and the dissertation or capstone project.

Established in the Policy are the expectations for original work, rationale for the policy, definition of terms that pertain to academic honesty and original work, and disciplinary consequences of academic dishonesty. Also stated in the Policy is the expectation that learners will follow APA rules for citing another person's ideas or works.

The following standards for original work and definition of *plagiarism* are discussed in the Policy:

> Learners are expected to be the sole authors of their work and to acknowledge the authorship of others' work through proper citation and reference. Use of another person's ideas, including another learner's, without proper reference or citation constitutes plagiarism and academic dishonesty and is prohibited conduct. (p. 1)

> Plagiarism is one example of academic dishonesty. Plagiarism is presenting someone else's ideas or work as your own. Plagiarism also includes copying verbatim or rephrasing ideas without properly acknowledging the source by author, date, and publication medium. (p. 2)

Capella University's Research Misconduct Policy (**3.03.06**) holds learners accountable for research integrity. What constitutes research misconduct is discussed in the Policy:

> Research misconduct includes but is not limited to falsification, fabrication, plagiarism, misappropriation, or other practices that seriously deviate from those that are commonly accepted within the academic community for proposing, conducting, or reviewing research, or in reporting research results. (p. 1)

Learners failing to abide by these policies are subject to consequences, including but not limited to dismissal or revocation of the degree.

I have read, understood, and abided by Capella University's Academic Honesty Policy (3.01.01) and Research Misconduct Policy (3.03.06), including the Policy Statements, Rationale, and Definitions.

I attest that this dissertation is my own work. Where I have used the ideas or words of others, I have paraphrased, summarized, or used direct quotes following the guidelines set forth in the APA *Publication Manual*.

Learner name
and date

Nerrissia Kemp-Murray   July 29, 2015

Mentor name
and school

Susan Slater, EdD School of Education