The Effects of Technology on Second Grade Students in Improving Language Arts Knowledge

Emily C. Huddleston

Submitted in Partial Fulfillment of the Requirements for the

Degree of Master of Education

December 2009

Graduate Programs in Education

Goucher College

**Table of Contents**

# List of Tables

**Abstract**

The purpose of this causal-comparative study was to determine the effectiveness of technology, specifically the Senteos interactive response system, in the classroom in improving students' language arts knowledge. The participants for the study were the second grade students from neighboring elementary schools during the 2008-2009 school year: Nantucket Elementary School in Crofton, Maryland and Four Seasons Elementary School in Gambrills, Maryland. Students at Nantucket Elementary school used more technology in their classroom on a daily basis than the students at Four Seasons Elementary school. The assessment results from Anne Arundel County's Second Grade Language Arts Benchmark 1 and 3 were analyzed and examined at each school in three different categories to determine significant differences: overall performance, differences by gender, and differences between race/ethnicity. The results from research conclude that there is no statistical difference in scores between Nantucket Elementary School's second graders and Four Season Elementary School's second graders. It is important to note, though, that there was a difference, although not statistically significant, when scores were compared by gender. Further research is recommended to determine the impact of interactive response systems on language arts instruction.

# CHAPTER I

# INTRODUCTION

## Overview

The goal of reading is comprehension and understanding.  In their early elementary school years, students are taught strategies to help them decode unfamiliar words and to make sense of what they are reading.  Decoding refers to the ability to break words apart into individual sounds to read and recognize words.  As learners get older, less time is spent on decoding strategies and more time is spent teaching them comprehension skills and techniques.  As a result, the students who do not develop early decoding strategies fall farther behind because their difficulties with decoding interfere with their efforts to read for meaning.

Extensive research has been conducted to gain insight into ways to address this pervasive, nationwide concern.  For example, studies by The National Institute for Literacy (Armbruster, Lehr, & Osborn, 2003) have examined "how text comprehension can be improved by instruction that helps readers use specific comprehension strategies" (p. 49).   Various interventions are used in schools to assist students with the development of important reading skills and strategies.  Among these interventions are programs such as Voyager and SOAR to Success.  Professional development courses, workshops, and resources have been made available for teachers to enable them to help their students develop decoding skills and read more fluently.

There is evidence suggesting that reading fluency and reading comprehension can be linked to overall understanding about what has been read.  Just as interventions have been designed to help students improve reading fluency, strategies and skills are taught to help students develop proficiency with reading comprehension (Armbruster et al., 2003).

Schools are using more advanced technologies in the classroom to help support reading programs.  According to the National Institute for Literacy (Armbruster et al., 2003), good readers are purposeful and active as they read.  Part of being an active reader is being engaged in what is being read.  In the classroom, using technology is one way to engage the technology-driven youth of today in an active manner.

Assessing reading comprehension allows teachers to determine how well their students are reading and to understand what they are reading.  Assessment allows teachers to adjust their instruction to fit the individual needs of the students within their classes.  One piece of technology that is making its way into more elementary school classrooms is the interactive response system, commonly referred to as "clickers."  These handheld devices prompt students to answer questions, and results are displayed immediately for the teacher and class to view.  Although the "clickers" can be used across all subject areas, they can be an invaluable tool for use in language arts instruction.

This researcher became interested in examining the impact of the use of interactive response systems in the teaching of language arts in her role as a second grade teacher at a public elementary school in Anne Arundel County, Maryland. The researcher used Senteos, an interactive response system appropriate for classroom use, for the first time in her second grade classroom during the 2008-2009 school year.  She was not only able to assess her students' understanding of word sounds and comprehension strategies, but most importantly, she received immediate feedback regarding students' responses through using Senteos.  Since results were available instantly, they could be used for immediate reteaching opportunities and discussion with students.  Anecdotally, the researcher observed that because the students in her class were

fascinated by the Senteos, they appeared to be more engaged and actively participated in discussions while using this interactive response system.

## Statement of the Problem

This study examined the effectiveness of technology, specifically the Senteos interactive response system, in the classroom in improving students' language arts knowledge. The study was designed to determine if interactive response systems improve language arts performance as compared to traditional testing methods.

## Hypotheses

Hypothesis #1: The difference between the first and third language arts benchmark scores of second grade students at Nantucket Elementary School, a school using the interactive response system on a regular basis, will be significantly greater than the difference between the benchmark scores of second grade students at Four Seasons Elementary, a school using traditional testing methods.

This study also investigated the use of the Senteo interactive response system and the impact it has on the language arts benchmark scores upon different student groups. Comparisons of student performance on the first and third benchmarks by gender and by race/ethnicity between the two schools also were made. Therefore, the following research hypotheses also were investigated in this study:

Hypothesis #2: The difference between the first and third language arts benchmark scores of male and female students at Nantucket Elementary School will be significantly greater than the difference between the first and third language arts benchmark scores of male and female students at Four Seasons Elementary School.

3

Hypothesis #3:  The difference between the first and third language arts benchmark scores of students at Nantucket Elementary School, disaggregated by race/ethnicity will be significantly greater than that of the scores of corresponding students, disaggregated by race/ethnicity, at Four Seasons Elementary School.

## Operational Definitions

A student's *language arts knowledge* was defined in this study as a student's overall score on the first and third second grade Reading Language Arts Benchmark Assessment in Anne Arundel County Public Schools.  For the purposes of this study, a comparison was made between the achievement of second grade students in two schools based on the results of first and third benchmark assessments.  A *student's score* is percentage based.  Percentages were defined by the school system reading office as follows. Scores of 59% and below are "basic," scores 60% to 79% are "proficient," and scores 80% and above are "advanced."

A *student's performance level* was percentage based.  The percentage of items a student answered correctly was used by the school system reading office to determine the student's performance level.  Scores of 59% and below reflected a "basic" level of performance, scores 60% to 79% reflected a "proficient" level of performance, and scores of 80% and above demonstrated an "advanced" level of performance.

**CHAPTER II**

**A REVIEW OF THE LITERATURE**

This literature review seeks to explore issues related to instruction and assessment in the area of reading comprehension and the impact of using interactive response systems such as Senteos during language arts instruction. The first section of this review describes what reading comprehension encompasses, with attention given to second grade students. Section two describes types of reading comprehension assessments used in second grade classrooms. The final section discusses the interactive response system Senteos and how this technology is used in classrooms.

**Reading Comprehension**

The goal of reading is comprehension, understanding what is read. (Armbruster et al., 2003). Readers need to know how to read words, but they also need to understand what they are reading. According to Armbruster et al., text comprehension has been identified as one of the five critical areas of reading and reading instruction. Reading comprehension strategies should be taught early to build a foundation for purposeful reading. Armbruster et al. advise that the most effective comprehension instruction is taught directly through direct explanation, modeling, guided practice, and application.

Strategy instruction is a necessary component of a language arts curriculum. Research such as that reported by Armbruster et al. (2003) has shown that comprehension instruction "can help students understand what they have read, remember what they read, and communicate with the others about what they read" (p. 48). Reading extensively improves one's ability to

comprehend; however, researchers such as Scharlach (2008) have concluded that there would be a greater improvement in comprehension if all readers were directly taught to use comprehension strategies. Some important strategies to include in comprehension instruction are drawing conclusions, identifying the main idea, and summarizing.

A child learns early how to draw conclusions. For example, when a helium balloon flies away, the child learns that balloons must be held tightly so they do not fly away (McMackin & Witherell, 2005). While reading, children need to learn how to identify these similar cause and effect relationships without always relying on a prior experience. When drawing conclusions, the reader is aware of the result of an event, but he/she needs to work backwards to determine what caused the event. Some researchers, such as McMackin and Witherell, refer to drawing conclusions as "backward inferencing" (p. 245). Use of this strategy helps readers comprehend because it helps them organize their reading into a logical sequence of events.

Identifying the main idea encompasses more than "the key ideas" of a story such as characters or problems. The main idea also includes themes and summaries (Jitendra, Chard, Hoppes, Renouf, & Gardill, 2001). Understanding the main idea of a reading selection has been considered "the most essential of reading 'skills'" (Johnston & Afflerbach, 1985, p.207). When he/she uses this strategy during and after reading, the reader first has to indentify the most important points and the less important points, or summarize what has been read. Only then will the reader be able to determine the main ideas or themes within the selection. Jitendra et al. suggest that "main idea comprehension is critical to becoming a skilled reader and that explicit teacher-mediated instruction can effectively promote main idea comprehension" (p.54).

Summarizing is an important comprehension strategy to learn because it helps the reader determine what is most important in a selection. According to Garner (1984), "if students cannot distinguish what is important from what is unimportant in texts…they will fail to process the written language" (p.304). Through direct instruction, modeling, and guided practice, students can learn how to distinguish what is important to the overall understanding of a text from what is not important.

Language arts instruction is provided in a variety of ways within the nation's elementary schools. For example, in Anne Arundel County Public Schools in Maryland (AACPS), language arts instruction is taught to students in second grade using the *Open Court Reading* program. The reading and responding portion of the program focuses on building reading comprehension through the use of a variety of text selections and focusing on several comprehension strategies. Although this program offers explicit comprehension strategy instruction, "instruction should be tailored to meet the needs of each individual student" (Dewitz & Dewitz, 2003, p. 434), which can be reinforced during the time allotted for small group instruction. It is important for teachers to design lessons that not only fit the individual needs of their students, but also to teach a variety of comprehension strategies to help them comprehend a variety of text (Gill, 2008).

In second grade in AACPS, teachers assess their students' understanding to determine the effectiveness in strategy application through observation and assessments. Comprehension assessments typically are given at the conclusion of a selection in order to determine how well students use the comprehension strategies taught. Second grade students also are given a formal assessment referred to as Anne Arundel County's Language Arts Benchmark Assessment three times throughout the year. The objectives in these assessments align with the comprehension

strategies taught in the *Open Court* Reading curriculum.  These assessments are discussed in further detail in the next section.

## Assessing Reading Comprehension

An assessment is used to measure a student's, or a group of students', knowledge of a specific subject matter.  In elementary classrooms, many different types of assessments are used. For the purposes of this review, the focus will be on reading comprehension assessments and the reading comprehension assessments used currently in AACPS.

The No Child Left Behind Act of 2001 (NCLB) was designed to increase student achievement in all academic areas for all pre-collegiate students in the United States (McMackin & Witherell, 2005).  McMackin and Witherell state that the enactment of this legislation "heightens the need to match students with appropriate skills so they can achieve steady, fast-paced academic growth" (p.243).  These authors further assert that the best way to measure students' academic growth throughout the year is by using formal assessments.

Second grade students in AACPS take a written, formal assessment, commonly referred to as Anne Arundel County's Language Arts Benchmark Assessment, three times throughout the school year. This assessment measures phonics, vocabulary and comprehension skills, and strategies through selected and brief constructed responses and then provides information about the effectiveness of the strategies taught and acquired throughout the school year as compared to other students in the school system and the state.  This assessment is not nationally recognized, but it can be compared with other assessments through using research results from the *Metropolitan Achievement Tests, Eighth Edition*, commonly referred to as the METROPOLITAN 8 ("Metropolitan Achievement Tests, Eighth Edition," 2008).

Assessments have changed in composition and format over time. In prior years, assessments often were comprised primarily of multi-choice or selected response questions. More recently, assessments include more open-ended or constructed response items (Jenkins, Johnson, & Hileman, 2004). Researchers such as Jenkins et al. debate whether a multi-choice or constructed response format is more effective in assessing a "deeper understanding" (p. 127) of a concept. According to Jenkins et al., when assessing reading comprehension, written responses assess only written expression rather than all reading comprehension abilities. If formatted correctly, multi-choice questions can make the testers organize their thoughts as though they were writing a constructed response.

It is important for teachers to create valid and reliable assessments throughout the year to accurately measure progress made by students. In a study conducted by Coyne and Harn (2006), teachers were required to measure student outcomes because "outcome assessments provide a bottom-line evaluation of student performance and the effectiveness of the overall reading program" (p. 40). According to Coyne and Harn, outcome assessments usually occur toward the end of the school year but should reflect what the teacher already knows through "ongoing screening, progress monitoring, and diagnostic assessments" (p. 40). Teachers can create effective assessments throughout the year by looking critically at each day's lesson objectives.

## Using Interactive Response Systems in the School Setting

Interactive response systems are handheld devices used by students to answer a variety of question types electronically; all the information from each student device is transmitted instantly to the teacher for immediate feedback. Using the same strategies and techniques for creating a standard paper and pencil assessment, teachers can create effective assessments using interactive response systems. Electronic assessments allow the teacher to assess his or her

students at their own ability levels. (Kush, Watkins, & Brookhart, 2005).  According to Brown

(2006), because immediate feedback is transmitted to the teacher, valuable instruction time is not

lost grading papers or surveying the class; instruction can be adjusted immediately.

A Senteo is an interactive response system that is being used in many classrooms today;

it is just one example of an interactive response system.  These handheld devices, or "clickers,"

are used by students to answer questions electronically in classrooms across all subject areas.

Each clicker is wirelessly connected to a central receiver which allows teachers to send and

receive information through each student device. Thus, teachers are able to collect data from

each student electronically.  These assessment tools can be used with both formative and

summative assessments and provide teachers with immediate feedback from their students;

results can easily be used to determine student strengths and weaknesses (Draper, Cargill, &

Cutts, 2002).

In a study conducted by Penuel, Boscardin, Masyn, and Crawford (2007), teachers in

elementary and middle school classrooms used the response systems in the same way; that is,

they were used across all subject areas and to help drive instruction through data collection.  In

recent studies, elementary school teachers have reported that their students are more engaged

during instruction when electronic response systems are used (Herreid, 2006) and the students

are "excited" (Brown, 2006, p. 2) about using the technology in the classroom.

Penuel et al. (2006) report that when teachers use Senteo assessments to measure reading

comprehension the procedures they follow are the same as those used with other subject areas.

Teachers use immediate feedback to adjust instruction.  According to McMackin and Witherell

(2005), "When a teacher notices that two or three students may seem confused about a concept,

she works with this small group a little longer" (p. 242) instead of waiting to grade the assignment and meeting with the group later. Instructional time is not lost but can instead be used more effectively.

## Summary

Reading comprehension is a key component of language arts instruction. Without comprehension, reading is not effective. Since passage of the NCLB legislation, increased attention has been given to assessments that measure the impact of reading strategy instruction on students' ability to understand and retain what has been read. Teachers should administer comprehension assessments to their students regularly to help guide their instruction and to pinpoint exact areas of strength and weakness. Interactive response systems such as Senteos allow teachers to assess their students in an engaging manner; in turn, teachers receive immediate data highlighting each student's strengths and weaknesses. These data can be used to drive instruction and better meet students' instructional needs. Based upon the types of research referenced in this literature review, use of Senteos appears to be an effective teaching strategy and assessment tool. With a more effective teaching strategy and assessment tool in place, formative and summative assessment scores in reading comprehension could show significant increases.

# CHAPTER III

## METHODS

The purpose of this study was to determine if there would be a significant difference in students' first and third language arts benchmark scores at two different schools when one school had access to more technology in the classroom, such as the interactive response systems (Senteos) and the other school that uses more traditional teaching methods. If a significant difference in performance was uncovered, the researcher also sought to determine to what degree the treatment was effective.

## DESIGN

This study used a causal-comparative design in order to gain insight into the differences between using technology as part of daily language arts instruction and assessment versus using traditional teaching methods. A comparison was made between the achievement of second grade students at two elementary schools using the overall student performance data collected from the first and third Language Arts Benchmark Assessment. The results from the two assessments at Nantucket and Four Seasons Elementary Schools then were compared to determine if a significant difference occurred. Both the first and third benchmark assessments were administered within the same week during October 2008 and May 2009.

## PARTICIPANTS

The participants for this study were the second grade students from Nantucket Elementary School and Four Seasons Elementary School in Anne Arundel County, Maryland. At Nantucket Elementary, the second grade class is composed of 58 male students and 70 female students ranging from ages seven to eight years. The participants were primarily Caucasian

(about 60%) and African American (about 16%). Other ethnic groups that were represented among the participants were Asian (about 12%), Hispanic (about 9%), and Indian (less than 1%).

The participants were from the researcher's home school and a neighboring school; all participants attended the schools for the entire school year. The sample group consists of 164 students from various ethnic backgrounds, both genders, different academic needs, and two different elementary schools within the county.

Nantucket Elementary is a public school located in Crofton, Maryland. The school opened in August 2008, the first year of data collection for its student population. The school has the most up-to-date technology in each classroom and in the Anne Arundel County Public School System. Each classroom has a set of Senteos that are used daily during instruction, specifically for language arts instruction. At the time of the data collection, the population consisted of 703 students in kindergarten through grade five. There were 359 males and 344 females ranging from age five to eleven. The majority of the population is Caucasian (about 70%) and African American (about 15%). Other ethnic groups represented among the student population are Asian (10%), Hispanic (4%), and Indian (less than 1%).

Four Seasons Elementary is a public school located in Gambrills, Maryland, the town which borders Crofton. When Nantucket Elementary opened in August 2008, half of the students who formerly attended Four Seasons Elementary were reassigned to Nantucket Elementary. Four Seasons Elementary opened in 1974 and has been gradually receiving technology as funds become available. At the time of this study, the school had one classroom set of Senteos that was shared by the entire school. At the time of the data collection, the population consisted of 524 students in grades one through five. Student enrollment in pre-kindergarten and kindergarten was not available. There were 262 males and 262 females ("2009

Maryland report card," 2009).  The majority of the student population is white.  Out of a total enrollment of 524 students, 69.1% (362) were white, 18.1% (95) were African American, 7% (35) were Asian, 4.8% (25) were Hispanic, and 1.0% (7) represented Native American students.

## INSTRUMENTS

This study used two instruments:  the first and third second grade Anne Arundel County Public Schools Language Arts Benchmark Assessments.  The benchmark assessments are whole-group administered, timed assessments.  They were designed to measure students' overall achievement in reading/language arts.  The objectives measured on the first and third benchmarks are similar to and included within the goals of the Maryland Voluntary State Curriculum for Reading.  The assessments are divided into three sections:  word study, vocabulary, and comprehension.  For purposes of this study, the researcher compared the results from all areas of the benchmark assessments, since all areas are taught during language arts instruction.

## PROCEDURE

Students at Nantucket Elementary used the Senteos interactive response system as part of their daily instruction.  Senteos were used for pre-assessments, ongoing formative assessments, and post-assessments.  The researcher, along with the other second grade teachers at Nantucket, embedded multiple choice, true/false, and multiple answer questions during language arts instruction.  Questions posed by the teachers appeared electronically on a Smart Board in the front of the room.  The students used their handheld devices to answer the questions, either anonymously or using their student ID numbers to log in.  Once their answers were "logged in" and all of the responses were collected, the teacher would stop the question and view the results. A pie chart displayed on the Smart Board revealed the correct answer and the percentage correct.

The teacher can export the students' results to a Microsoft excel spreadsheet with a breakdown of each child's score, what he or she answered, his or her total percentage correct, the percentage correct for each questioned posed, and the overall class average. The results can be used to determine what needs to be taught, re-taught, or what students already know. This time effective tool allows teachers to meet the individual and group needs of their class.

The teachers at Nantucket Elementary followed the same language arts curriculum as the teachers at Four Seasons Elementary. However, the Senteo technology allows the teachers at Nantucket to immediately assess the needs of their students and adjust instruction. Such an immediate, comprehensive response is not possible when using traditional methods to determine the level of students' acquisition of new learning.

At both elementary schools, students were given the Anne Arundel County Language Arts Benchmark Assessment 1 in October 2008 and Benchmark Assessment 3 in May 2009 within the same week in response to requirements established by the school system. The three sections within the assessments (word study, vocabulary, and comprehension) are broken into different timed sections. The examiner reads sample questions at the beginning of each set of questions and, using the teacher assessment booklet, guides students on how to answer each question within the section. After each sample question has been reviewed, the students complete the remaining section independently within the given time frame. For the comprehension section, the students read a selection silently and answer the questions that follow, including three brief constructed response questions.

The students recorded their selected-response items on a Scantron answer sheet that was scored using a Scantron machine. The brief constructed responses were answered on a separate paper and were scored by their teachers using the rubrics and sample answers provided in the

teacher's manual.  The scores were calculated into percentages to determine each student's

proficiency.  Anne Arundel County Public Schools' Office of Reading defines these percentages

as follows:  59% and below is a "basic" score, 60%-79% is a "proficient" score, and 80% and

above is an "advanced" score.

# CHAPTER IV

## RESULTS

The primary purpose of this study was to determine whether there was a statistically significant difference in student performance on the Benchmark 3 versus Benchmark 1 reading assessments among second grade students at Nantucket Elementary School who used the Senteos interactive response system when compared to the Benchmark 3 versus Benchmark 1 performance of a similar group of second graders at Four Seasons Elementary School.  In addition to these hypothesized differences in *overall* student performance at Nantucket versus Four Seasons Elementary Schools, the researcher also hypothesized that there would be a statistically significant difference among male students at Nantucket Elementary School who used the Senteos interactive system versus male students at Four Seasons Elementary who did not have use of this system.

A final purpose of this study was to determine whether there was a statistically significant difference in student performance on the Benchmark 3 versus Benchmark 1 reading assessments when the results were disaggregated by race/ethnicity.  Because of the racial/ethnic diversity of both Nantucket Elementary School and Four Seasons Elementary School, the researcher also wanted to determine whether there were statistically significant differences in students' Benchmark 3 versus Benchmark 1 performance when results were disaggregated by race/ethnicity.

Table I below contains results which describe the *overall* performance of second graders at each of the two schools.  It is important to note that only those students who had both Benchmark 1 and Benchmark 3 scores were included in each of the analyses reported in this

study.  Students, therefore, who were missing either a Benchmark 1 or a Benchmark 3 score were not included in the analyses below.   Thus, for example, the results for 13 students who completed only the Benchmark 1 or the Benchmark 3 assessments were removed from the student data for Nantucket Elementary School.  Similarly, the results for eight students who either completed the Benchmark 1 or the Benchmark 3 assessments were removed from the student data for Four Seasons Elementary School.

**Table I**

**A Comparison of Students' Benchmark One versus Benchmark Three Scores by School**

|  | Benchmark 1 | | | Benchmark 3 | | | Percentage Point Difference |
|---|---|---|---|---|---|---|---|
|  | N | Mean | S.D. | N | Mean | S.D. |  |
| **Nantucket Elementary** | 97 | 71.8% | 19.4 | 97 | 81.7% | 15.2 | 10.6 |
| **Four Seasons Elementary** | 67 | 73.9% | 19.1 | 67 | 84.1% | 14.5 | 10.1 |

A comparison of the mean scores of students at Nantucket Elementary School and Four Seasons Elementary School that are reported in Table I suggests that student performance was similar on both the Benchmark 1 and Benchmark 3 assessments.  There also was a substantial amount of variability in student performance on both benchmarks in each school, as indicated by the standard deviations that are reported in Table I above.

Table I above also reports mean differences in student performance between Benchmark 1 and 3 at each school.  The differences reported for each school (10.6 and 10.1, respectively) are similar.

In order to determine whether these differences in student performance on Benchmark 1 versus Benchmark 3 were statistically significant, a *t* test for independent groups was used. The results, [$t(142) = 0.203$, $p > .05$) indicate that differences in student performance on Benchmark 3 versus Benchmark 1 at Nantucket Elementary School versus Four Seasons Elementary School were non-significant.

A second analysis was conducted in order to determine whether there was a statistically significant difference by gender in student performance when comparing the performance of male students at Nantucket Elementary School who used the Senteos interactive response system versus male students at Four Seasons Elementary School who did not have the Senteos system available. The mean performance of both male and female students at each school is reported in Table II below. These initial results suggest that there was comparable growth by both male and female students in both schools. Differences in student performance for both groups of students from Benchmark 1 to Benchmark 3 also were similar.

**Table II**

**A Comparison of Students' Benchmark One versus Benchmark Three Scores by Gender**

|  | Benchmark 1 | | | Benchmark 3 | | |
|---|---|---|---|---|---|---|
|  | Males | | | Females | | |
|  | N | Mean | S.D. | N | Mean | S.D. |
| **Nantucket Elementary** |  |  |  |  |  |  |
| **Benchmark 1** | 56 | 65.5% | 18.9 | 51 | 70.6% | 16.4 |
| **Benchmark 3** | 56 | 74.8% | 15.1 | 51 | 82.2% | 11.8 |
| **Percentage Point** |  | 9.3 |  |  | 11.6 |  |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Difference** | | | | | | |
| **Four Seasons Elementary** | | | | | | |
| **Benchmark 1** | 19 | 67.6% | 23.8 | 49 | 75.7% | 16.4 |
| **Benchmark 3** | 19 | 79.8% | 16.3 | 49 | 85.6% | 13.4 |
| **Percentage Point Difference** | | 12.2 | | | 9.9 | |

In order to determine whether there was a statistically significant difference in the performance of male students at Nantucket Elementary School compared to male students at Four Seasons Elementary School who did not use Senteos, a $t$ test for the independent groups was used. The results [$t(33) = 0.129$, $p > .05$)] indicate that there were no statistically significant differences in the performance of males on Benchmark 1 versus Benchmark 3. A similar procedure was used to compare the Benchmark 1 versus Benchmark 3 performance of female students between the two schools. These results [$t(93) = 0.987$, $p > .05$) suggest that there were no statistically significant differences in the Benchmark 1 versus Benchmark 3 performance of females in each school.

The final analysis of data that is reported in this study investigates whether or not there were statistically significant differences in student performance on Benchmark 1 versus Benchmark 3 among students at Nantucket Elementary School who used the Senteos interactive response system versus a similar group of students at Four Seasons Elementary School who did not have the system available when results are disaggregated by race/ethnicity. The results are reported in Table III below. It is important to note that the data reported in Table III below do not include Asian students because the number of Asian students in grade 2 at Four Seasons

Elementary School (3) compared to Nantucket Elementary School (14) was insufficient to adequately compare the performance of this student group between schools. Similarly, while there were nine Hispanic students in grade 2 at Nantucket Elementary School, there were no Hispanic students at Four Seasons Elementary School.

The results reported in Table III below suggest that both mean performance of white students on the Benchmark 1 versus Benchmark 3 at both Nantucket Elementary school and Four Seasons Elementary School were similar. In addition, the difference in their performance on the Benchmark 1 versus Benchmark 3 assessments (9.7 versus 10.5) was similar.

In order to determine whether differences between the two groups on Benchmark 1 versus Benchmark 3 were statistically significant, a $t$ test of independent groups was used. The results $[(t(74) = 0.747, p > .05)]$ indicate that there were no statistically significant differences in performance on the Benchmark 1 versus Benchmark 3 assessments between the two groups of students.

The results reported in Table III below suggest that African American students at Nantucket Elementary School improved substantially from Benchmark 1 (50.9%) to Benchmark 3 (67.9%) when compared to the Benchmark 1 performance (70.8%) versus Benchmark 3 performance (80.0%) of African American students at Four Seasons Elementary School.

In order to determine whether these differences in performance were statistically significant, however, a $t$ test of the independent groups was used. The results $[(t(27) = 0.147, p > .05)]$ were not statistically significant.

**Table III**

**A Comparison of Students' Benchmark One versus Benchmark Three Disaggregated by**

**Race/Ethnicity**

| Race/Ethnicity of Students | Nantucket Elementary | | | Four Seasons Elementary | | |
|---|---|---|---|---|---|---|
| **White Students** | | | | | | |
| | **N** | **Mean** | **S.D.** | **N** | **Mean** | **S.D.** |
| **Benchmark 1** | 67 | 71.7% | 15.7 | 44 | 73.5% | 19.5 |
| **Benchmark 3** | 67 | 81.4% | 13.7 | 44 | 84.0% | 15.2 |
| **Percentage Point Difference** | 9.7 | | | 10.5 | | |
| **African American Students** | | | | | | |
| | **N** | **Mean** | **S.D.** | **N** | **Mean** | **S.D.** |
| **Benchmark 1** | 16 | 50.9% | 21.9 | 20 | 70.8% | 19.9 |
| **Benchmark 3** | 16 | 67.9% | 17.2 | 20 | 80.0% | 12.9 |
| **Percentage Point Difference** | 17.0 | | | 9.2 | | |

**CHAPTER V**

**DISCUSSION**

The purpose of this study was to determine if there was a significant difference in the Language Arts Benchmark Assessment 1 versus Benchmark Assessment 3 scores of second grade students at Nantucket Elementary School who used the Senteos interactive response system as an instructional tool and an assessment tool when compared to the difference in Benchmark 1 versus Benchmark 3 scores of students at Four Seasons Elementary School who did not use the Senteos interactive response system.   In addition to looking at overall differences between Benchmark 1 versus Benchmark 3 scores between the two schools, the researcher also wanted to determine whether there were statistically significant differences in performance on Benchmark 1 versus Benchmark 3 between male versus female students and students representing different racial/ethnic groups between the two schools.  The results suggest that there was no statistically significant difference between the two schools' language arts benchmark scores in the fall and in the spring (i.e. Benchmarks 1 and 3, respectively).  It is important to note, however, that there was a substantial difference, although not a statistically significant difference, in the gains made by African American students at the two schools. Among African American students at Nantucket Elementary School, there was a 17 percentage point increase in their performance on Benchmark 1 versus Benchmark 3 compared to a 9.2 percentage point increase from Benchmark 1 to Benchmark 3 among African American students at Four Seasons Elementary School.

**THREATS TO VALIDITY**

There are several threats to internal validity in this study that may account for the results reported.  The first and foremost threat is that there may have been a "teacher effect" involved.

The preparation of students prior to the assessments, the scoring of the written responses, and even the administration of the assessments all may have been different.  At Nantucket Elementary School, the second grade teachers designed their instruction to model the type of questioning that would be used on the benchmark assessments to help prepare their students for these assessments.  Specifically, the teachers designed multiple choice questions on the Senteos for the students to practice the process of responding to benchmark questions.  At Four Seasons, the types of techniques used prior to the assessment, if any, are unknown.  All of the participants received non-standardized instruction from different teachers.  There were five second grade teachers at Nantucket Elementary and four second grade teachers at Four Seasons Elementary at the time of the study.

Another related "teacher effect" that may have been a threat to the internal validity of this study was the scoring of students' written responses to the comprehension portion of the Language Arts Benchmarks.  Teachers within each grade at the two schools graded the written brief constructed responses.  The school system provides a rubric (with possible scores of a 3, 2, or 1) and sample responses.  If a student response does not match the sample response provided by the rubric, then teacher judgment is used to determine what score is most appropriate for the student response.  This lack of inter-rater reliability may have had a significant impact upon the results of this study.

The school system provides specific directions to proctors administering the Language Arts Benchmark Assessments; however, since classroom teachers also serve as test proctors, there also could have been motivational techniques and strategies used by any of the nine teachers administering the assessment that may have had an impact on the students'

performance. Using an assessment that is more high-stakes that provided specific test administration rules for teachers to follow during the actual conduct of the benchmark assessments possibly could eliminate some of the last-minute techniques teachers use with their students.

Given the age of the students used in this study, it also is important to consider the amount of time that took place in between the two benchmark assessments when considering possible threats to the internal validity of this study. It is therefore important to consider the different levels of maturity among the students in this study. The first assessment was given in October and the third was given in May; a significant amount of maturing could have occurred during this time, given that the participants in this study were second graders.

There are several threats to the external validity of the study which limits its generalizability to a larger population of second grade students at the district, state, and national levels. The participants in this study were not chosen at random; in fact, they were selected out of convenience of the researcher. Although it can be assumed that the student populations at both schools are similar because of their geographic location, it would have been a more effective study if the participants had been randomly selected from various schools around the school system, or even the nation, to participate in the study. Random selection of students for both groups and random assignment of each student to one of the two groups would have increased the generalizability of the results of this study to the larger population of second grade students nationally. Given the logistical difficulties of randomly selecting and randomly assigning students from intact classrooms in a school, random selection and assignment of schools to reflect greater demographic diversity (such as percentages of students receiving free-

and-reduced lunch and special education students) would have enhanced the generalizabilty of this study to larger populations.

Another possible threat to external validity is the limited availability of the Senteos interactive response system in other school districts.  This study possibly could be replicated in other schools in Anne Arundel County Public Schools**,** but due to the limited availability of the Senteos system nationally, it might not be replicated easily outside of the school system, or even nationally.

## COMPARISON TO OTHER RESEARCH

Although the results of this study did not show a statistically significant difference in the benchmark scores between the two elementary schools, the percentage gained by African American males at Nantucket Elementary School suggests that instructional strategies may have impacted these scores.  The results of this study therefore appear to support the theory that student response systems help engage students during instruction and provide immediate feedback to teachers for further instruction (Herreid, 2006).  In a study conducted by Brown (2006), teachers in elementary classrooms praise technology systems, like Senteos, for immediate access to data.  When data are received immediately, teachers are able to adjust their instruction to fit the individual needs of the students in their classroom instantly.

A study by Penuel et al. (2007) found that teachers who used student response systems as part of their instruction used it to "diagnose student difficulties, asking students to discuss or rethink answers, and using feedback from responses to adjust instruction" (p. 315).  The researcher notes that this study is similar to the methods used by the teachers at Nantucket

Elementary to prepare their second grade students for the language arts benchmark assessments and suggests that this may account for gains made by some groups.

Furthermore, when the researcher compared the analysis of the current study to the discussions of Draper et al. (2002) regarding their research, there were notable similarities. When students use the interactive response systems, more accurate data are gathered due to confidentiality. Traditional assessments in classrooms without this type of technology allow a student's neighbor to see his or her response to a question, potentially causing embarrassment. With response systems, students can be more private with their answers and avoid humiliation in front of their peers.

## RECOMMENDATIONS FOR FUTURE RESEARCH

Future research should include using a more reliable assessment, selecting participants receiving the same language arts instruction, and using an even greater variety of tools to administer the assessment (i.e. Senteos, computer-based assessment, and pencil-and-paper method). Conducting a different experimental study with an assessment that standardizes in greater detail the procedures for administering these assessments would make the scores more reliable and valid. Using an assessment similar to the Maryland State Assessment (MSA), would be a more valid and reliable tool because the test items are changed continually. This would eliminate teacher bias or excessive preparation prior to the assessment. The MSA also has specific rules and regulations for proctors and test administrators, thus minimizing the chance of bias while the assessment is being administered. When a high-stakes test such as the MSA does contain brief constructed response, the teachers do not take part in the scoring process. These items are scored by professionals hired by the scoring contractor. In scoring the brief

constructed responses for the local benchmark assessments, the use of trained scorers would minimize any variability in scoring the brief constructed response items due to teacher biases.

Choosing as participants students receiving the same language arts instruction prior to taking the language arts assessment would be another direction for future research. In this study, scores collected could be contributed to the amount of language arts instruction and/or intervention received throughout the school year, thus impacting the validity and reliability of the scores.

Finally, to determine if Senteos is a more effective assessment tool than traditional assessment methods, a study should be conducted that would compare Senteos with other common test-taking methods. For example, Senteos might be compared with computer-based online assessments and traditional paper-and-pencil methods. Evaluating the effectiveness of these multiple approaches with a large sample of students representing student populations with diverse demographic characteristics would contribute significantly to determining which method(s) are most effective in administering formative assessments to second graders.

## References

Armbruster, B., Lehr, F., & Osborn, J. (2003, June).  Put reading first:  The research building blocks for teaching children to read, kindergarten through grade 3.  *A publication by the National Institute for Literacy.* iii, 48-57.

Brown, J. (2006, August 1).  Mobile assessment:  Working smarter, not harder.  *T.H. E. Journal,* (33), 32-35. (Eric Document Reproduction Service No. EJ762465) Retrieved November 1, 2008, from RIC database.

Coyne, M., & Harn, B. (2006, January 1).  Promoting beginning reading success through meaningful assessment of early literacy skills. *Psychology in the Schools,* 43(1), 33-43. (ERIC Document Reproduction Service No. EJ761857) Retrieved November 1, 2008, from ERIC database.

Dewitz, P., & Dewitz, P. (2003, February).  They can read the words, but they can't understand:  Refining comprehension assessment. *Reading Teacher.*  56(5) 422. Retrieved November 1, 2008 from Academic Search Premier database.

Draper, S., Cargil, J., & Cutts, Q.  (2002). Electronically enhanced classroom interaction.  *Australian Journal of Educational Technology* (18), 13-23. Retrieved November 1, 2008 from http://www.ascilite.org.au/ajet/aject18/draper.html

Garner, R. (1984, January 1).  Rules for summarizing texts:  Is classroom instruction being provided? *Journal of Educational Research.* 77(5) 304-308. (ERIC Document Reproduction Service No. EJ300725).   Retrieved November 1, 2008 from ERIC database.

Gill, S. (2008).  The comprehension matrix: A tool for designing comprehension instruction. *The Reading Teacher.* 62(2) 106-113. Retrieved November 1, 2008 from Academic Search Premier.

Herreid, C. (2006). "Clicker" cases: Introducing case study teaching into large classrooms. *Journal of College Science Teaching, 36*(2), 43-47. Retrieved October 14, 2008 from Academic Search Premier database.

Jenkins, J., Johnson, E. & Hileman, J. (2004). When is reading also writing: Sources of individual differences on the new reading performance assessments. *Scientific Studies of Reading.* 8 (2), 125-151. (ERIC Document Reproduction Service No. EJ683130) Retrieved November 1, 2008, from ERIC database.

Jitendra, A., Chard, D., Hoppes, M., Renouf, K., & Gardill, M. (2001, January). An evaluation of main idea strategy instruction in four commercial reading programs: Implications for students with learning problems. *Reading & Writing Quarterly.* 17(1) 53-73. Retrieved November 1, 2008 from Academic Search Premier database.

Johnston, P. & Afflerbach, P. (1985, September). The process of constructing main ideas from text. *Cognition and Instruction.* 2(3/4) 207. Retrieved from Academic Search Premier database.

Kush, J., Watkins, M., & Brookhart, S. (2005). The temporal-interactive influence of reading achievement and reading attitude. 11 (1) 29-44. *Educational Research and Evaluation.* Retrieved October 13, 2008 from Academic Search Premier database.

2009 Maryland report card. (2009, July 29). Retrieved from http://www.mdreportcard.org/index.aspx

McMackin, M., & Witherell, N. (2005, November). Different routes to the same destination: Drawing conclusions with tiered graphic organizers. *Reading Teacher.* 59(3) 242-252. Retrieved November 1, 2008 from Academic Search Premier.

Merrell, C. and Tymms, P. (2007). Identifying reading problems with computer –adaptive

assessments. *Journal of Computer Assisted Learning.*(23) 27-35.  Retrieved November 1, 2008 from Academic Search Premier database.

Metropolitan Achievement Tests, Eighth Edition.  (1931-2001). Retrieved November 1, 2008, from Mental Measurements Yearbook database.

Penuel, W., Boscardin, C., Masyn, K. & Crawford, V.  (2006). Teaching with student response systems in elementary and secondary education settings:  A survey study. *Education Tech Research Development* 55, 315-346. Retrieved  October 13, 2008 from Academic Search Premier database.

Scharlach, T. (2008).  START Comprehending: Students and teachers actively reading text. *The Reading Teacher.* 62(1) 20-31. Retrieved November 1, 2008 from Academic Search Premier.