8-2019

Effects of Computer Screen Color and Brightness on Student Achievement during Computer-Based Reading Comprehension Assessment

Matthew E. Houselog

Effects of Computer Screen Color and Brightness on Student Achievement during Computer-Based Reading Comprehension Assessment

by

Matthew E. Houselog

A Thesis

Submitted to the Graduate Faculty of

St. Cloud State University

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in Educational Administration and Leadership

August, 2019

Thesis Committee:
Frances Kayona, Chairperson
Thomas Hergert
Jim Johnson

Abstract

This study examined the effect of specific computer screen settings on reading comprehension achievement using a standardized test format in an eighth-grade classroom in Rogers, Minnesota in the spring of 2019. The screen settings tested were a blue-light-minimizing setting, screen brightness, and a combination of both. A blue-light-minimizing setting alters the monitor color settings to minimize the amount of blue used in the projection of the computer screen. The effect of room darkness in combination with these screen settings was also considered. The field experiment had four research questions: Is there a difference in achievement between students who use a (1) blue-light minimizing computer screen setting, (2) darkened computer screen brightness compared to those who do not?; (3) Does the removal of fluorescent room lighting have an effect on achievement?; (4) Is there a difference among the variables in achievement scores when comparing the short-passage scores to the long-passage scores? Thirty randomly samples students from a total population of approximately 170 were used in the study.

The results of the study were analyzed using one-way ANOVA on both the overall reading comprehension scores and also on the reading section scores that differed based on the length of reading passages. A t-test was also done on the reading section scores to see what effect length of passage had on scores overall. The results indicate that none of the screen settings have a conclusive effect on reading comprehension outcomes on tests. However, there were findings that suggest there may be a correlation between the number of words contained in a reading and test performance.

Table of Contents

Page

List of Tables

Table Page

Chapter I: Introduction

Introduction

This study is designed to examine the effect of computer screen settings on student achievement in reading comprehension. Standardized testing has become commonplace in K-12 education throughout the country due to the rise of its use as a political accountability tool in the last twenty-five years (Improving America's Schools Act, 1994; No Child Left Behind Act, 2001; Every Student Succeed Act, 2015). In the last decade there has also been an increase in the use of computer software to implement these standardized tests; however, studies have found that computer-based tests are not consistently performing on the same level of paper-based tests (Myrberg & Wiberg, 2014; Walsh, 2016).

Standardized testing has always been a fixture in education over the past twenty years, but it has taken on a new level of importance in the past two decades. Yearly standardized achievement scores are being used by states to publicly determine school success. Despite the recent reduction in penalties for schools that have not met adequate yearly progress (Every Student Succeeds Act, 2015), there are still ramifications to schools for poor test scores: decreases in community confidence and support, lower student enrollment, pressure on administrators and educators to focus more on tests, and high-stakes pressure on students to meet high expectations (Gewertz, 2018). Questions still remain about the validity of standardized testing being an adequate representation, and sole measure, of accountability for students, teachers, schools, and school districts (Darling-Hammond, 2007; Wiliam, 2010). Despite these questions it is hard to imagine a policy shift happening soon when it comes to computer-based standardized testing due to the ease of administration, relative low cost, and immediate feedback that computer-based testing provides (Herold, 2014).

The question of whether or not computer-based tests are equal to paper-based tests has been a topic of research even before the increased use of computer-based testing over the past decade. Bugbee's (1996) study on computer-based testing compared to paper-based testing revealed deficiencies in the results of individuals who completed the test on a computer. Other studies replicated these findings and pointed to the additional task of operating the digital device as the primary factor in the lowered comprehension scores making it difficult to determine if poor test results from computer-based comprehension testing are due to poor comprehension skills or poor computer-literacy skills (Noyes & Garland, 2003; Noyes, Garland, & Robbins, 2004; Dosch, 2012).

Recent research have found no conclusive differences between digital and print formats (Aydemir, Ozturk, & Horzum., 2013; Myrberg & Wiberg, 2014; Boevé, Meijer, Albers, Beetzma, Bosker, 2015; Niccoli, 2015). Aydemir et al. (2013) conducted a study on fifth grade elementary students and found increases in reading comprehension for students that read from a screen as compared to students who read from paper. Other studies found that while participants in these studies preferred the paper-based test reading method, there was no significant difference in achievement scores (Myrberg & Wiberg, 2014; Boevé et al., 2015). While there are recent studies that have shown differences in achievement between computer-based and paper-based testing (Jeong, 2012; Mangen, Walgermo, & Bronnick, 2012) these findings suggest that increased computer literacy may be limiting and possibly eliminating the operating task's effect on achievement results. The increased exposure to technology could also be a factor as the increases of technology in school and at home have provided more opportunities for students to read material on a digital platform (Noyes & Garland, 2008).

Although there have been mixed results comparing computer-based and paper-based testing, there is definitely a test mode effect occurring. The test mode effect is defined as the situation where “identical paper-based and computer-based tests will not obtain the same results” (Clariana & Wallace, 2002 p. 593). The likelihood of achieving equivalent results between computer-based and paper-based testing is nearly the same as a coin flip (Clariana & Wallace, 2002). What is causing the test mode effect is still greatly debated. Theories on the cause of test mode effect include computer operational skill (Noyes, Garland, & Robbins, 2004; Dosch, 2012), added mental workload of using a computer (Wästlund, Reinikka, Norlander, & Archer, 2004), and technological interferences in the display monitors (Bridgeman, Lennon, & Jackenthal, 2003).

Computer screens have evolved with the advancement of LCD technology, but there are still concerns of what effects computers may have on the human eye. Employers around the world have already begun to pay attention to the effects of screen exposure and certain elements of digital displays that may cause strain to the eyes, increase stress levels, and thus decrease productivity (Meyer & Kollbaum, 2016; Heiting & Wan, 2017). New developments of screen settings may aid in preventing eyestrain and discomfort when using a computer. Technology companies such as Apple and Google have also been developing methods to decrease eyestrain from increased technology use (Bera, 2018; Mulaney, 2018).

One such setting changes the standard computer screen setting to a warmer color, reducing the amount of blue. Blue light refers to the blue wavelengths on the light spectrum that naturally occur with daylight sun and have been recently been simulated in electronic screens through new breakthroughs of LCD screen technologies (*Harvard Health Publishing*, 2018). Repeated exposure to blue light has made it in the news as a potential hazard on sleeping patterns, but it also increases eye fatigue compared to a warmer color setting (Meyer, 2016; Crowder, 2018). Computer companies have responded with pre-formulated screen modes to change the digital screen to a warmer color output that gives the screen a yellow appearance (Apple's "Night Shift" mode and Google Chromebook's "Night Light"). Screen and room brightness should also be considered when using these screen settings to achieve the optimal conditions that minimize eyestrain (Baldwin & Colt 2010; Abrams, 2012; Meyer & Kollbaum, 2016; Heid, 2017).

The level of sustained concentration needed to complete the tasks involving reading comprehension could also add stress to the reader and ultimately inhibit comprehension. As Lipson and Wixson (1986) found, reading ability is difficult to test because several factors are involved including content knowledge, motivation and interest, text organization, the nature and content of the task, and characteristics of the setting in which reading occurs. Reading longer passages on a screen may result in poorer comprehension since computer displays can only comfortably present only about one-third of the information compared to a standard piece of paper (Clariana & Wallace, 2002). It may be that our behavior with computers is also interfering with our ability to comprehend. "When reading screens people seem to reflexively skim the surface of texts in search of specific information, rather than dive in deeply in order to draw inferences, construct complex arguments, or make connections to their own experiences" (Herold, 2014, p. 8).

Despite questions and concerns that have been raised about computer-based standardized assessment, the benefits of administering tests on computers will most likely hinder any future changes. Therefore, it is safe to assume that computer-based testing for reading comprehension is here to stay. However, elements that may have an effect on performance of reading comprehension should be studied in order to optimize achievement levels and lessen the test mode effect. This study seeks to identify if improvements have been made to limit the test mode effect in computer-based, reading comprehension testing to give students and educators strategies to limit the interference of computer-based testing on reading comprehension achievement.

Statement of the Problem

This study examines to what extent computer screen settings impact achievement on reading comprehension tests. Specifically, this study is designed to compare academic achievement among five groups of eighth grade students who are exposed to different computer settings to determine to what extent these computer settings impact reading comprehension results using three separate criterion-referenced tests. This is an experimental study looking at the test scores of 30 random students from a total of approximately 170 participants within a school setting. More specifically, it is an ex post facto research design that was testing an independent variable's effect on the dependent variable. The experiment will use a one-way analysis of variance (ANOVA) to measure differences in outcomes among the student groups and a post hoc test to determine and statistical significance. Data will be gathered during the spring of 2019 from Rogers Middle School in central Minnesota. Conclusions drawn should assist school leaders in determining optimal conditions for administering computer-based tests in an age of high stakes testing accountability.

Purpose of the Study

The purpose of this study is to identify methods to decrease the cognitive load on test takers when completing a reading-comprehension assessment on a computer device. The cognitive load refers to the amount of focus and concentration needed to perform a task (Noyes, Garland, & Robbins, 2004). There are two areas of cognitive load in a computer-based assessment: the cognitive load of reading and answering questions, and the additional cognitive load of running and navigating the computer. If the cognitive load can be decreased, the computer-based testing would be less intrusive regarding focus on the assessing reading comprehension. The results of this study should assist in the preparation and training of students to properly use technological tools and to learn computer-based, test-taking strategies to limit impediments during testing.

Objectives of the Study

The Objectives of the study include:

1. Review the literature on computer-based reading comprehension assessments.
2. Obtain permission from the St. Cloud State University Institutional Review Board to conduct this study.
3. Obtain permission from the Elk River Area School District #728 to conduct this study.
4. Obtain permission from the individuals to participate in the study.
5. Identify the three comprehension assessments to be used in the study.

6. Randomly assign groups to the various lighting situations.
7. Conduct three assessments for a three-week period according to experiment layout.
8. Collect data from assessments and analyze it.

Research Questions

This study primarily focused on the following research questions.

1. Is there a difference in achievement between students who use a blue-light minimizing computer screen setting compared to those who do not?
2. Is there a difference in the achievement when computer screen brightness is darkened compared to when it is not?
3. Is there a difference in the achievement when the fluorescent room lighting is turned off compared to when it is not?
4. Is there a difference among the variables in achievement scores when comparing the short-passage scores to the long-passage scores?

Hypothesis to be Tested

1. The use of a blue-light minimizing computer screen setting during a reading comprehension test will not have a significant effect on achievement scores for the general population of students.
2. Changes in screen brightness settings will not have a significant effect on achievement scores.
3. The lack of room lighting will not have a significant effect on achievement scores.
4. The difference of length of the reading passages will not have a significant

difference in achievement scores.

Assumptions of the Study

Assumptions made for this field study include the following:

1. The assessments created for this test are reliable and valid.
2. The participants of the study have all been taught the same standards, skills, and information that will be included in the assessments.

Delimitations of the Study

This study will be limited by the following factors:

1. The results are limited to Rogers Middle School with no other school participating.
2. The results will focus on reading comprehension achievement only.
3. Data will reflect scores during the 2018-2019 school year.

Human Subject Approval

The data collected for the study will include testing scores and informal test-taker comments. All requirements set forth by the St. Cloud State University Institutional Review Board will be strictly adhered to. All steps will be taken to ensure that the privacy, rights, and welfare of the participants in the study are protected.

Definition of Terms

The following definitions are used for this study:

Blue light: Blue light refers to the blue wavelengths on the light spectrum that naturally occur with daylight sun and have been simulated in electronic screens through new breakthroughs of LCD screen technologies (*Harvard Health Publishing*, 2018).

Blue-light-minimizing computer screen setting: This term is being used by the study's author to describe the computer display setting on Chromebook OS devices (known as "Night Light". This screen setting adjusts "the color emitted from the screen to make reading and viewing easier on the eyes in low light or when you are using your device for extended periods of time" (Brangers, 2018).

Computer-based testing (CBT): "Computer-based testing uses a computer to give exactly the same test as one in a paper-and-pencil format. It has the same test questions and presents them in exactly the same order as the paper-and-pencil version of the test" (Bugbee, 1996, p. 282).

Cognitive workload: "Cognitive (mental) workload has been defined as the interaction between the demands of a task that an individual experiences and his or her ability to cope with these demands. Hence, it arises due to a combination of the task demands and the resources that a particular individual has available" (Noyes, Garland & Robbins, 2004, p. 111).

The Minnesota Comprehensive Assessments (MCA): The MCA (also known as MCA-III) is the yearly state assessment series was created to accommodate changes in academic standards as the reading comprehension was overhauled in 2010 to align with the 2010 Minnesota K-12 Academic Standards in Language Arts (Minnesota Department of Education, 2017).

Test Mode Effect: The test mode effect is defined as the situation where "identical paper-based and computer-based tests will not obtain the same results" (Clariana & Wallace, 2002 p. 593).

Chapter II: Review of Literature

Introduction

This study is focused on the effect computer screen display settings have on reading comprehension assessments. This review of literature is focused on the potential interference computer-based testing may place on assessing reading comprehension. The review of literature is split into two parts. The first part focuses on the test mode effect of computer-based testing compared to paper-based testing. The second part focuses on studies that involved computer screen color and brightness. A full list of the studies and the findings included in this review can be found on Table 1 (computer-based and paper-based testing) and Table 2 (computer screen color and brightness).

The Role of Standardized Testing in K-12 Education

Standardized testing began in the United States during the mid-1800s, replacing oral recitation as the standard practice of assessment. Horace Mann pushed for standardized written essays in 1845 as support for a more objective and efficient form of testing increased (Huddleston & Rockwell, 2015). Sixty years later, the first IQ test was created in 1905 by French psychologists Alfred Binet and Theodore Simon to help identify students that were unable to succeed in school, which was followed by an American version in 1914 named the Stanford-Binet (Wolf, 1973). During this same time Edward L. Thorndike was constructing a series of standardized achievement tests in multiple subject areas such as arithmetic, reading, and language (Wigdor & Gardner, 1982). These tests were the building blocks for future standardized assessment, but the purpose of these tests would shift when used to account for public policy.

The Elementary and Secondary Education Act (ESEA) was signed into law in 1965. The law came after the Civil Rights Act of 1964, and the focus of the law was to provide equal opportunity for all students. The initial purpose was to supply additional funding and services to the most vulnerable students in America by offering federal grants to districts with high populations of low-income students as well as scholarships for low-income college students (Elementary and Secondary Education Act of 1965, 2013). Furthermore, it also established special education policies that would later be mandated in the Education for All Handicapped Children Act (EAHCA) in 1975. The initial ESEA has changed considerably since its inception with the inclusion of standardized test scores to hold school districts and states accountable.

Standardized comprehension assessment was first used as a method of accountability in the reauthorization of Title I of the ESEA in 1968 and continued with the spread of state assessment systems into the 1970s (Sarroub & Pearson, 1998). The purpose of testing was to provide data to ensure that the most vulnerable students that ESEA focused on were performing to accepted standards. The reauthorization of ESEA in 1988 expanded student testing and accountability by requiring districts to review test scores and develop improvement plans if schools were not making adequate progress (Robelen, 2005).

The Improving America's Schools Act (IASA) of 1994 started a new wave of increased student assessment requirements under the ESEA (Elementary and Secondary Education Act of 1965, 2013). Participating states were required to assess mathematics and reading standards for at least one grade level in each of three grade ranges (grades 3 through 5, 6 through 9, and 10 through 13). Nearly ten years later, the No Child Left Behind Act (NCLBA) expanded yearly testing requirements for students from grades 3-8 by the 2005-2006 school year. The biggest impact the NCLB had was the implementation of performance-based sanctions on schools that failed to meet adequate yearly progress (AYP), ushering in a decade of high-stakes standardized testing (Library of Congress, 2008).

The most recent shift in educational policy is the Every Student Succeeds Act (ESSA), signed into law in December of 2015. The framework of required standardized tests established by the NCLB remains unchanged despite the national backlash against them (Gewertz, 2018). This act gave more flexibility to state governments and local school districts by lessening and eliminating some of the accountability measures at the federal level; however, statewide assessments will continue to be used to provide information to educators, families, students, and communities (Library of Congress, 2008).

Reading Comprehension Assessments

Reading comprehension assessments have been a part of American education for over a hundred years, most notably led by Thorndike’s quest to find the best method to test the act of comprehension (Sarroub & Pearson, 1998). Thorndike’s Scale Alpha reading test (1914) included a series of short paragraphs with questions that were limited to a single answer (Huddleston & Rockwell, 2015). This style of test mirrors the modern forms of reading comprehension tests by providing longer reading passages followed by questions for test takers to answer. Tests before it such as the Kansas State Reading Test and the Stanford-Binet relied more on puzzles and shorter passages. These tests also differed from Scale Alpha by putting a time limit on the testing (Huddleston & Rockwell, 2015). As time went on, multiple-choice questions of reading comprehension continued to be the predominant method of assessment (Sarroub & Pearson, 1998).

In the 1980s there was a course change in reading comprehension assessment as the education community pushed for more open and reflective approaches rather than a skills-based approach. States began to alter reading comprehension assessments to address the political shift. The California Learning Assessment System (CLAS) was created allowing students to give open-ended answers to literature questions instead of multiple-choice options (Sarroub & Pearson, 1998). Other states followed suit; however, another shift occurred in the mid-1990s that would curb these efforts due to the high costs of administering the tests, political skepticism of new formats, and equity concerns for vulnerable students (Sarroub & Pearson, 1998; Huddleston & Rockwell, 2015).

Since the 1980s, reading comprehension assessments have consistently regressed back to traditional, multiple choice questions since the IASA (1994) refocused attention towards using reading comprehension assessment to ensure educational equity for all students. There was a window of compromise for open-ended question formats as assessments went to about 80 percent multiple-choice questions and the remaining 20 percent to open-ended questions (Sarroub & Pearson, 1998). There are still state assessments that include open-ended questioning (such as the PARCC and Smarter Balanced tests); however, after five states decided to go away from this form of testing in 2017, only one-third of states are currently using these assessments (Gewertz, 2019).

In the last decade, the greatest shift in reading comprehension assessment has been the increased use of computer-based testing for yearly assessments. One of the greatest benefits of computer-based testing is that it provides immediate results and feedback. The results can often be corrected within minutes of completing the test. It also saves money on materials and personnel needed to score tests, specifically for tests with open-answer formats (Schaffhauser, 2011). Computers have been used for reading comprehension assessments prior to IASA, NCLB, and ESSA, but the volume has increased due to a combination of increased computer access in schools and political factors driving standardized testing as accountability measures (Schaffhauser, 2011).

Test Mode Effect in Computer-Based Testing

Over the years several studies have tested reading comprehension on computers, but some of these studies include technologies or subjects that are not applicable to the formats and practices used in computer-based assessment today (Bugbee, 1996). Some of the studies focused more on reading speed (Dillon, 1992) or reading accuracy that have not shown strong correlation to reading comprehension (Wilkinson & Robinshaw, 1987; Oborne & Holton, 1988). However, there are older studies that look specifically at reading comprehension that most closely resemble the conditions seen in computer-based assessment today. The Mazzeo and Harvey report (1988) was a series of studies on the impact of completing testing tasks on computers that concluded that testers have more difficulty answering questions with graphics or pictures and have more difficulty reading passages on computers compared to paper (Mazzeo & Harvey, 1988).

In 1989, there was another review of computer-based testing done by Bunderson, Inouye, and Olsen. The study focused on computer-adaptive testing (a form of computer-based testing that adapts to the skill level of the test taker during the test based on the responses within the test), but also reviewed previous studies on computer-based testing and found inconsistent results in whether computer-based or paper-based testing resulted in higher scores. Of the 23 studies reviewed by Bunderson, Inouye, and Olsen (1989), three studies concluded computer-based testing produced higher mean scores, nine studies concluded computer-based testing produced lower mean scores, and the remaining eleven were not conclusive.

Studies in the 1990s and early 2000s provided a glimpse into some of the weaknesses inherent with conducting tasks on computers. Computer-based testing was compared to paper-based testing, and the findings revealed deficiencies in the results of individuals who completed a comprehension test on a computer (Bugbee, 1996). The most common conclusions from these studies pointed to the additional task of operating the digital device as the primary factor in the lowered comprehension scores (Noyes et al., 2004). Paper-based tests only require the test taker to have knowledge of basic reading elements that are ingrained in students at an early age, but computers require basic knowledge of computation and additional navigation. Wästlund, Reinikka, Norlander, and Archer (2004) refer to this as the dual processing effect that includes both the knowledge being assessed on the test along with the knowledge needed to control or operate the apparatus being used for testing. The study focused on video display terminals (computer screens) and it showed that participants struggled to gather information presented to them digitally compared to paper form due to the dual task nature of reading the material and operating the screen (Wästlund et al., 2004). Because of this, it is difficult to determine if the lower results from computer-based comprehension testing is due to poor comprehension skills, poor computer-literacy skills, or potentially another factor. Despite improvements in the accuracy of comprehension tasks in more modern computer-based tests, there are still differences in the length of time and workload required for computer-based test takers when compared to paper-based tests (Noyes et al., 2004).

The dual-task nature of computer-based testing does not affect performance the same way for everyone. According to Clariana and Wallace (2002) students identified as “high-attaining” did better on computer-based tests than “low-attaining” students. Students marked “high-attaining” were identified based on survey results on their engagement with the course materials, independence, and competitiveness to do well. “Low-attaining” students did significantly worse taking the computer-based test compared to the paper-based test whereas “high-attaining” students did not see a significant change (Clariana & Wallace, 2002). Noyes et al. (2004) also support this and concluded that individuals with lower comprehension scores were at a disadvantage when using computers during testing. These studies suggest that lower-skilled individuals may have more difficulty on computer-based tests because they are less likely to overcome the added workload required with a computer-based test.

Other studies have attempted to identify additional potential causes for the test mode effect on computer-based assessment. Computer familiarity is one potential cause that has been mentioned, but with mixed support. Clariana and Wallace (2002) conducted a study to identify key factors of test mode effect of computer-based testing. They found that computer familiarity, gender, and competitiveness did not factor into test mode effect; however, content familiarity was related (Clariana & Wallace, 2002).

The effect of task length was highlighted by Daniel and Woody (2013) as they compared student performance between electronic and paper textbooks and found that while test scores did not significantly change, the reading time of individuals using the electronic textbooks were significantly higher than students using paper textbooks. A study by Haas and Hayes (1986) concluded that computer-based tests had lower scores when reordering text structure if the reading required was longer than a page. Ackerman and Goldsmith's research (2011) also suggested that computers were more suited for shorter and less challenging texts. The additional time, focus, and potential interference of scrolling on a computer screen could raise anxiety levels and could have an effect on high-stakes testing (Bridgeman, Lennon, & Jackenthal, 2003).

Bridgeman et al. (2003) studied the effect of screen size, screen resolution, and display rate on computer-based test performance, which produced a number of findings specific to reading comprehension on computer-based tests. The participants were split into groups where the font size was altered for different groups, so that there would be more need to scroll on the computer screen when the text was larger. In the open-ended survey responses, there were complaints (predominantly from the group with the largest text) of the need to scroll through the reading on the screen. The surveys also found that 10% of the student's surveyed complained that the computer screen was difficult to read. The difference with these complaints is that they were evenly split among all screen and resolution conditions grouped in the study (Bridgeman et al., 2003). This suggests that no matter what adjustments are made to computer-based testing some individuals simply prefer paper-based formats.

Test-taking preference has also been suggested as a cause of the test mode effect. Ackerman and Lauterman (2012) studied the role of technology by using three sets of tests in both computer-based and paper-based formats. One testing session was limited to seven minutes (testers knew of the time limit), another had unlimited time, and the final test stopped without testers knowledge after seven minutes. What the study discovered was that the individuals using paper received better scores; however, they did not get better results when under the interrupted time condition (Ackerman & Lauterman, 2012). This shows that technology does not play a significant role in test mode effect, as the results should have been consistent across all testing scenarios. The study concluded that a potential barrier to screen reading might be based more on preference (Ackerman & Lauterman, 2012).

Two years later Lauterman and Ackerman (2014) did another study investigating reading preference. They found that individuals who preferred to read on computers did better and received similar scores to those taking the paper-based tests (Lauterman & Ackerman, 2014). A study on college psychology students showed similar patterns. There was no difference in performance between computer-based and paper-based testing, but the students surveyed preferred the paper-based format, feeling less control when taking the computer-based assessment (Boevé et al., 2015). Their conclusions also suggest that students that have more confidence in their ability to take their preferred testing format produce better results.

While there are more recent studies that have shown differences in achievement results between computer-based and paper-based testing there is evidence that it may lower over time (Jeong, 2012; Mangen et al., 2012). Dosch conducted a study on the impact of practice in computer-based testing regarding the National Certification Examination for nurse anesthetists. He found that students with more experience in computer-based testing earned higher scores compared to those with less experience in computer-based testing (Dosch, 2012). Dosch also points to the potential role of the subject's age on computer-based testing with the assumption that younger test takers typically have had more exposure to technology. There was an 85 percent passing rate of students in their 40s compared to a 94 percent pass rate for students in their 30s (Dosch, 2012). While there is a potential of other factors causing these numbers (such as diminished eyesight or reading skill due to age) it supports the idea that computer literacy factors into computer-based testing success.

The problem of dual-processing theoretically should reduce due to the increase of technology education and availability. Over half of teachers now have 1-1 student-to-device ratios, which is a 10 percent increase from 2015 to 2016 (*EdTech*, 2017). More recent research supports this idea because paper-based testing is not producing better results than computer-based testing on a consistent basis (Aydemir et al., 2013; Myrberg & Wiberg, 2014; Boevé et al., 2015). Aydemir et al. (2013) studied fifth grade elementary students reading comprehension and showed increases in results for students who read from a screen as compared to students who read from paper. Other studies found that while participants in these studies preferred the paper-based test reading method, there was no significant difference in achievement scores (Myrberg & Wiberg, 2014; Boevé et al., 2015).

Computer Monitors and the Test Mode Effect

Computer monitors have also been scrutinized as a potential cause of the test mode effect. A comprehensive review of studies done by Ziefle (1998) concludes that computer-based reading tires the eyes more quickly than paper-based reading due to the display screen qualities of computers. Digital displays have been shown to increase the stress level and tiredness of test takers (Wästlund, Reinikka, Norlander, & Archer, 2004). Wästlund et al. (2004) compared the differences in reading on computer-based and paper-based formats. Subjects reported that using the computer-based formats caused more stress and fatigue than individuals using the paper-based formats (Wästlund, 2004).

Noyes and Garland (2003) also suggest that elements of computer monitors (refresh rate, high levels of contrast and fluctuating luminance) may interfere with reading cognition. A study by Murata, Uetake, Otsuka, and Takasawa (2001) had participants perform tasks that lasted one hour and the results showed weak connection to visual fatigue interfering with task completion. However, there were signs of visual fatigue in both the physiological and psychological measures collected in the study, but an explanation of the lack of influence on task completion points toward the simplicity of the task required of the study (Murata et al., 2001). In other words the users were able to overcome the fatigue to accomplish the task because it was relatively easy. A task requiring complication or advanced skills may be more difficult to overcome.

Blue light emitting from LCD screens has also been shown to cause visual fatigue (Mangen et al., 2013). Blue light is a range on the visible light spectrum between 400-495 nanometers (Nagaraja, 2019). Chang, Aeschbach, Duffy, and Czeisler (2014) concluding that connected blue light exposure to the interruption of sleeping patterns caught media attention and, in response to these studies, technology companies have created blue light reduction displays and brightness settings to combat the growing concerns of the effect of blue light on the human eye (Vimont & Khurana, 2017). The settings shift the screens emitted light to the warm end of the light spectrum (Jabr, 2016).

According to Dr. Gary Heiting, O.D., blue light causes eye fatigue because “short-wavelength, high energy blue light scatters more easily than other visible light”, making it harder to focus on lower-contrasted text (Heiting, 2018, p. 24). Sixty-five percent of Americans surveyed by the Vision Council experienced symptoms of eyestrain when using digital devices

(Meyer & Kollbaum, 2016). Lin, Gerratt, Bassi, and Rajendra (2019) studied the effect of short wavelength-blocking glasses on visual fatigue during computer usage. The experiment had three levels of blue blocking lenses (no block, low block, and high block). They found that the high block glasses, which blocked the most amount of blue light, reduced eye fatigue (Lin et al, 2019). Using some type of blue light blocker or screen setting adjustment seems to lessen the impact of eyestrain when using digital screens.

Background colors on screens have also been tested. A study on elderly screen reading attempted to identify an ideal background color for reading text on screen, but the overall results were inconclusive (Anuardi, Yamazaki, & Eto, 2017). However, the study did point out that white backgrounds resulted in an increase of eye movement, but also an increase of test scores. The study hypothesized that despite the potential for increased eye activity, the subject's familiarity with white backgrounds on computer screens and the lack of difficulty to the questions given may have skewed the results (Anuardi, et al., 2017). Rello and Bigham (2017) determined that the best digital screen background colors for readers and found that warm colors such as peach, orange, and yellow led to faster reading times and less mouse movements (Rello & Bigham, 2017). Warmer color backgrounds such as the ones created by using the

In a study of the effectiveness of iPads Night Shift mode, researchers concluded that screen color alone does not have an impact of blue light levels, and screen brightness most likely also plays an additional role (Nagare, Plitnick, & Figueiro, 2018). The study changed the amount of blue used on the screen, but did not alter the brightness level. Nagare et al. recommended that future blue light testing would factor in screen brightness as well.

The lighting environment can also play a role as different lighting conditions and screen types are shown to cause visual fatigue from the monitor light (Mangen et al., 2012). Room lighting has shown to increase performance on cognitive related tasks in an academic environment (Veitch & McColl, 2001). Fluorescent lighting in classrooms have also been found to cause headaches and obstruct visual performance on reading fluency and mental performance (Winterbottom & Wilkins, 2009; Mott et al., 2012).

Literature Review Summary

This review of literature is focused on the influence of computer-based standardized testing on education and the potential causes of the test mode effect that is apparent in computer-based tests when compared to paper-based tests. Increased access to technology in schools and continued legislation focused on educational accountability has increased the prevalence of computer-based assessment in school settings. Computer-based testing is essential to the data collection needed to support school districts and states as they monitor the progress of students to meet state standards. While computers do offer a better range and ease of accumulating data, there still remain concerns of the effects computer-based assessments have on achievement.

Table 1.

Summary of Studies on Computer-Based versus Paper-Based Testing

Year	*Researcher(s)*	*Findings*
1986	Haas & Hayes	The effects of computer screen versus paper were compared when participants were asked to reorder scrambled text. The study found that participants using the paper copy were able to perform the task faster.
1987	Wilkinson & Robinshaw	This study compared computers and paper with participants attempting to catch proofreading errors in text and concluded that individuals using the computer missed more errors.
1988	Mazzeo & Harvey	Tests that contain reading passages are more difficult when presented on a computer screen compared to paper.
1988	Oborne & Holton	There is no difference in reading comprehension outcomes between computer-based and paper-based formats.
1989	Bunderson, Inouye, & Olsen	Review of computer-based testing that found inconclusive results. Of the 23 studies, three favored computers, nine favored paper-based formats, and the remaining eleven were not conclusive.
1994	Dillon	There is not a single explanation for the reading speed differences that were reported in studies (prior to 1994) and the likelihood of outcomes most likely were based on a combination of fatigue, familiarity with navigation and manipulation of computers, the orientation and size of the text, and computer display characteristics.
1996	Bugbee	Computer-based and paper-and-pencil tests can be equivalent, but just because they look the same doesn’t mean they function the same. It is important that there is adequate preparation of the test and testing items by test creators as well as ensuring that test takers have appropriate computer knowledge.

Table 1 (continued)

1998	Ziefle	This study looked at differing resolutions of computer monitors when testing proofreading speed and accuracy. It found that individuals performed better on paper than the different monitor conditions when testing the speed and accuracy of proofreading.
2002	Clariana & Wallace	The computer-based group outperformed the paper-based group and higher performing students benefited the most from computer-based assessment compared to paper-based assessment.
2003	Bridgeman, Lennon & Jackenthal	Variances in computer screen resolution settings were studied on high school juniors in this study. There was no significant difference on math scores. Verbal scores were better when the display resolution was larger.
2003	Noyes & Garland	There are differences in cognitive processing with memory assimilation when computer-based reading is compared to paper-based reading.
2004	Noyes, Garland, & Robbins	There was no significant difference in comprehension task, but there was a difference in workload with the computer-based test additional effort to complete.
2004	Wästlund, Reinikka, Norlander & Archer	The consumption of information from Video Display Terminals was impaired due to the dual-task nature of reading the information and operating the screen.
2008	Noyes & Garland	This literature review focuses on the equivalency of computer-based and paper-based tests. The authors predict that improvements in computer technology will lead to a more level experience for test takers in the future.
2011	Ackerman & Goldsmith	Test performance did not differ when comparing studying from text on a print hard copy compared to a computer screen.
2012	Ackerman & Lauterman	Test performance did not differ when comparing studying from text on a print hard copy compared to a computer screen. However, print formats were more efficient as assessments done on computer-based lacked self-regulation of learning on screen.
2012	Dosch	This study looked at test mode effect on the computer-based National Certification Examination of Nurse Anesthetists. 40 percent of students reported that their educational programs did not include any computer-based testing. Students that had more extensive experience in computer-based testing achieved higher scores on the test than students with less experience with computer-based testing.

Table 1 (continued)

2012	Jeong	Subjects taking a paper-based test scored better than individuals taking a computer-based test. Test takers specifically struggled in the area of Korean language. This result may be due to the length of the reading required in that section of the test.
2012	Mangen, Walgermo, & Bronnick	This study focused on reading comprehension for digital and paper-based testing for 10th grade students in Norway. There was a significant difference in scores as digital test-takers scored lower in both narrative and expository scores.
2013	Aydemir, Ozturk & Horzum	This study focused on evaluating the effect of reading from screens on reading comprehension tests to reading on paper for 5th grade elementary students. The study found that students that read from the screen did better in informative texts and performed similarly on narrative texts compared to the paper-based method.
2013	Daniel & Woody	Students had significantly higher reading times on electronics compared to paper despite comprehension scores being similar across formats..
2013	Mangen, Walgermo, Bronnick	Students who read text on paper score better than student who read text on computers when tested on reading comprehension, word reading, and vocabulary.
2014	Lauterman & Ackerman	Preference of format plays a role in how individuals perform on computer-based and paper-based assessments.
2014	Myrberg & Wiberg	Participants in a reading study preferred to read on paper compared to computer; however, the study found no support for it being more difficult to read on a digital media. A reader's attitude and preference may play a role in the outcome.
2015	Boevé, Meijer, Albers, Beetsma, & Bosker	This study looked at college students' results between computer-based and paper-based testing. There was no difference in total scores, but the author's indicated that students still preferred the paper-based form and that changes needed to be made in how computer-based tests are prepared for and administered.
2015	Makhoul & Copti-Mshael	Participants did better on shorter, informational questions on tests conducted on computers versus paper. However, performance was worse with reading comprehension questions on computers compared to paper.

Table 1 (Continued)

2015	Niccoli	Adult participants in this study had a greater frequency of high scores for both multiple-choice and short answers when reading off of paper compared to those using tablets. However, there is no significant difference in group test score means between paper and digital results.
2016	Walsh	This literature review concluded that while many of the studies on computer-based testing are out of date due to technological advances, there are still problems that these studies shed light on such as the subjective nature of reading and the unique environmental circumstances of each study.

Table 2.

Summary of Studies on Computer Display and Room Lighting Conditions

Year	*Researcher(s)*	*Findings*
2001	Murata, Uetake, Otsuka, Takasawa	Identified a measurement for visual fatigue with video display terminal tasks. A weak connection between visual fatigue and task completion on computers.
2001	Veitch & McColl	Room lighting has an effect on visual processing and appearance judgements. Full-spectrum fluorescent lighting can improve vision, perception, and performance. However, room lighting solutions are complex than just bulb type.
2003	Bridgeman, Lennon & Jackenthal	Variances in computer screen resolution settings were studied on high school juniors in this study. There was no significant difference on math scores. Verbal scores were better when the display resolution was larger.
2009	Winterbottom & Wilkins	84% of classrooms have fluorescent room lighting that is causing visual discomfort when studying 90 classrooms in the United Kingdom.
2012	Mott, Robinson, Walden, Burnette, & Rutherford	Students increased oral reading fluency performance when in a classroom setting that limited the amount of fluorescent lighting.

Table 2 (continued)

2017	Anuardi, Yamazaki, & Eto	Subjects received more stress when performing a task on a computer with a white background than other colors, with a yellow background providing the least amount of stress. However, reading scores for the white background were the highest among all the colors. The reason for this may be in familiarity with white backgrounds when performing reading tasks.
2017	Rello & Bigham	Background colors such as peach, orange, and yellow significantly improved reading performance.
2018	Nagare, Plitnick, & Figueiro	Anti-blue filter settings (iPad Night Shift mode) had little impact on the reducing eye exposure to melatonin. Screen brightness most likely also plays a role.
2019	Lin	Participants using high-block glasses (highest level of blue-light filtering) had less eye fatigue and reported feeling less pain in and around the eye compared to participants wearing the low-block or no-block glasses.

Chapter III: Methodology

Introduction

The purpose of this experimental study is to examine the impact of certain computer screen settings on student achievement scores and student fatigue during computer-based testing. This study will look at two specific screen settings: Chromebook "Night Light" blue light reduction setting and screen brightness.   While this study will examine and attempt to find an optimal screen setting for enhanced performance, it will also consider the role of personal preference in screen setting options, and darkened room lighting.

Securing Participation in the Study

Permission to administer the experiment was granted by Independent School District #728 in coordination with the St. Cloud State University Institutional Review Board. Parent or guardian permission slips were sent home with students, signed, and returned for participants in the study. Participating students also signed a student assent document indicating their interest in joining the study.

Sampling Technique

The population of the study will be limited to a single school, Roger Middle School, and limited to only eighth grade students. Rogers Middle School is in Rogers, Minnesota, which is located 25 miles northwest of the Minneapolis-St. Paul metropolitan area. It is a part of Independent School District 728, which services the Zimmerman, Elk River, Otsego, and Rogers cities and surrounding townships.

A total of 164 eighth-grade students will be used for the study. They will be divided into nonrandom groups based solely on class period due to the environmental variable of classroom lighting that will affect all participants in the classroom. The five nonrandom groups will be randomly assigned to a group designation (Group A, Group B, Group C, Group D, Group E). Random sampling will be used to select six individuals from each group to compare data.

Experiment Procedure

The experiment will consist of three reading comprehension tests administered over a span of three weeks with tests being administered approximately a week apart. The first test will be used as a baseline test with no variables implemented for any group, and the following two tests will have variables introduced to specific groups. Each of the three tests will consist of two reading sections and a total of twenty questions. The first reading section will be short enough to fit on the screen with minimal scrolling needed to view the entire reading. The reading will then be followed by six questions. The second reading section contains two topic-related readings that require twice the amount of scrolling compared to the first reading section. This reading section will be followed by fourteen questions. The questions will be a mixture of multiple choice, multiple response, and rank or order questions.

Each test will have different readings but will be within the Lexile range appropriate to the eighth-grade reading level. The readings and questions used on each of these tests will be taken from preparatory materials given by Pearson Education, the company implementing Minnesota's statewide assessments (known as the MCA test). The assessments will use the Schoology software program that was purchased by the district and is used to implement curriculum and other assessments for all courses. All participants in the study will use the school-issued 11-inch (model 3189) Chromebook for all three tests. Using these familiar platforms should ensure that results are not affected due to lack of familiarity with the instruments used to give the assessment.

Variables

There will be three primary independent variables used in the experiment: Chromebook's "Night Light" mode (blue-light-minimizing screen setting), screen brightness, and student preference. A secondary independent variable will be adding a darkened room along with the primary independent variables to enhance the effect the screen alterations. The blue-light minimizing screen setting is a standard setting on Chromebooks referred to as "Night Light". This alters the screen color to create a warmer color scheme (white turns into a yellow color). Screen brightness will be altered by adjusting the computer screen brightness to 3/8 power. This setting was chosen based on recommendations suggested during the review of literature (Heid, 2017; Heiting & Wan, 2017; Nagare et al, 2018).

Group A will be the control group and not use any of the variable adjustments to the monitor screen, and the screen brightness will be set at 75% power. Group B will have the blue-light-minimizing screen setting on during Test 2 and 3. Group C will have the screen brightness lowered to 3/8 power for Test 2 and Test 3. Group D will have a combination of the blue-light-minimizing screen setting and screen brightness turned down to 3/8 power for the last two tests. Group E will have a choice to use the blue-light-minimizing screen setting as well as the option to adjust the screen brightness to any level desired. In the final testing stage, Groups B, C, D, and E will have an additional variable of a darkened room as all overhead lighting will be turned off during testing. The dependent variable will be student's scores on the assessments.

Table 3.

Experiment Design

GROUPS	TEST 1	TEST 2	TEST 3
G_A	X	X	X
G_B	X	**1**	**1_5**
G_C	X	**2**	**2_5**
G_D	X	**3**	**3_5**
G_E	X	**4**	**4_5**

Variables:
1 = Blue-light setting used
2 = Reduced screen brightness
3 = Blue-light setting and reduced screen brightness
4 = Student preference for blue-light setting and screen brightness
5 = Darkened classroom combined with other variable

Collection of Data

The results of the tests will be collected using the Schoology assessment software, which will collect the overall achievement score for each test along with itemized results for each question. Student's names will be removed on all data pertaining to this study and each student will be given an identification code that will consist of a letter (the group letter they are apart of) and a randomized, two-digit number. Test data will be stored within the school districts' file space, which is password protected. Any collected paper data will be stored in a locked file cabinet and destroyed upon completion of the study.

Analysis and Treatment of Data

This study will use a one-way analysis of variance (ANOVA) to measure differences in achievement among the student groups on each test. This should identify any significant statistical differences to answer research questions #1, #2, and #3. To answer research question #4, an ANOVA will be used on the two reading section scores to identify any significant statistical differences. A t-test will also be conducted to test for any significant statistical difference from the test results in the reading sections as a whole group.

Research Questions

This study shall primarily focus on the following research questions.

1. Is there a difference in achievement between students who use a blue-light minimizing computer screen setting compared to those who do not?
2. Is there a difference in achievement when computer screen brightness is darkened compared to when it is not?
3. Is there a difference in achievement when the fluorescent room lighting is turned off compared to when it is not?
4. Is there a difference among the tested variables in achievement scores when comparing the short-passage scores to the long-passage scores?

Null Hypothesis

1. Using a blue-light minimizing computer screen setting has no effect on reading comprehension achievement scores.
2. Lowering the brightness setting on the computer has no effect on reading comprehension achievement scores.
3. Turning of fluorescent room lighting has no effect on reading comprehension achievement scores.

4. There is no difference in achievement among the tested variables when comparing short-passage scores to long-passage scores.

Chapter IV: Results

This study was designed to examine what effect blue-light-minimizing computer screen settings, computer brightness, and room lighting have on reading comprehension achievement scores on a standardized test. This field experiment focused on 30 students randomly sampled from approximately 170 students in an 8th grade class setting at Rogers Middle School in Rogers, Minnesota.

The reading comprehension tests were twenty questions in length and consisted of two reading sections (referred to in the results as Reading Section 1 and Reading Section 2). Reading Section 1 had one reading followed by six questions. Reading Section 2 had two readings on a shared topic and required twice the amount of scrolling on the computer screen to complete the reading. It was then followed up by fourteen questions about the texts. Due to the uneven number of questions, the comparison data will be percentages instead of raw scores.

Experiment Results

Means and standard deviations for all three tests are presented in Table 4. The analysis of variance (ANOVA) was conducted for each individual test to determine if there were any statistical differences between variable groups for Test 1 (Table 6), Test 2 (Table 9), and Test 3 (Table 12).

The sample variance between groups and within groups for Test 1 (4.867, 9.660) and Test 2 (6.283, 10.233) were similar with more variance between groups (Tables 5 & 8). However, Test 3 showed more variance between groups (9.533) than within groups (4.427) although this difference did not meet the qualifications for statistical significance, which was at 10.4% (Table 11).

Table 4.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
TEST 1 (Raw)	30	5.00	15.00	10.3667	2.99981
TEST 2 (Raw)	30	6.00	18.00	12.0333	3.11264
TEST 3 (Raw)	30	9.00	17.00	12.8000	2.26518

There was no consistent statistical significance found when comparing groups on all three tests. Howver, there was a 5.6% significance between the control group and Group B - BLM on Test 3 although this did not meet the qualifications of being statistically significant (Table 12). The control group did consistently have a mean score at the bottom of the groups on all three tests whereas Group B - BLM had the highest mean score for Test 1 & Test 2 (Tables 7, 10, & 13).

Table 5.

ANOVA: Test 1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	19.467	4	4.867	.504	.733
Within Groups	241.500	25	9.660		
Total	260.967	29			

Table 6.

Post Hoc Test: Test 1

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Group B - BLM	Group C - Dark	.66667	1.79444	.996	-4.6034	5.9367
	Group D - BLM+Dark	.33333	1.79444	1.000	-4.9367	5.6034
	Group A - Control	2.00000	1.79444	.797	-3.2700	7.2700
	Group E - Preference	1.83333	1.79444	.843	-3.4367	7.1034
Group C - Dark	Group B - BLM	-.66667	1.79444	.996	-5.9367	4.6034
	Group D - BLM+Dark	-.33333	1.79444	1.000	-5.6034	4.9367
	Group A - Control	1.33333	1.79444	.944	-3.9367	6.6034
	Group E - Preference	1.16667	1.79444	.965	-4.1034	6.4367
Group D - BLM+Dark	Group B - BLM	-.33333	1.79444	1.000	-5.6034	4.9367
	Group C - Dark	.33333	1.79444	1.000	-4.9367	5.6034
	Group A - Control	1.66667	1.79444	.883	-3.6034	6.9367
	Group E - Preference	1.50000	1.79444	.917	-3.7700	6.7700
Group A - Control	Group B - BLM	-2.00000	1.79444	.797	-7.2700	3.2700
	Group C - Dark	-1.33333	1.79444	.944	-6.6034	3.9367
	Group D - BLM+Dark	-1.66667	1.79444	.883	-6.9367	3.6034
	Group E - Preference	-.16667	1.79444	1.000	-5.4367	5.1034
Group E - Preference	Group B - BLM	-1.83333	1.79444	.843	-7.1034	3.4367
	Group C - Dark	-1.16667	1.79444	.965	-6.4367	4.1034
	Group D - BLM+Dark	-1.50000	1.79444	.917	-6.7700	3.7700
	Group A - Control	.16667	1.79444	1.000	-5.1034	5.4367

Table 7.

Homogeneous Subsets: Test 1

GROUP	N	Subset for alpha = 0.05 1
Group A - Control	6	9.3333
Group E – Preference of blue-light-minimizing setting and darkened screen brightness	6	9.5000
Group C – Darkened screen brightness	6	10.6667
Group D – Blue-light-minimizing setting and darkened screen brightness	6	11.0000
Group B – Blue-light-minimizing screen setting	6	11.3333
Significance		.797

Note. Means for groups in homogeneous subsets are displayed and use a harmonic mean sample size of 6.000.

Table 8.

ANOVA: Test 2

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	25.133	4	6.283	.614	.657
Within Groups	255.833	25	10.233		
Total	280.967	29			

Table 9.

Post Hoc Test: Test 2

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Group B - BLM	Group C - Dark	.16667	1.84692	1.000	-5.2575	5.5908
	Group D - BLM+Dark	-.33333	1.84692	1.000	-5.7575	5.0908
	Group A - Control	2.16667	1.84692	.766	-3.2575	7.5908
	Group E - Preference	1.16667	1.84692	.968	-4.2575	6.5908
Group C - Dark	Group B - BLM	-.16667	1.84692	1.000	-5.5908	5.2575
	Group D - BLM+Dark	-.50000	1.84692	.999	-5.9242	4.9242
	Group A - Control	2.00000	1.84692	.814	-3.4242	7.4242
	Group E - Preference	1.00000	1.84692	.982	-4.4242	6.4242
Group D - BLM+Dark	Group B - BLM	.33333	1.84692	1.000	-5.0908	5.7575
	Group C - Dark	.50000	1.84692	.999	-4.9242	5.9242
	Group A - Control	2.50000	1.84692	.662	-2.9242	7.9242
	Group E - Preference	1.50000	1.84692	.924	-3.9242	6.9242
Group A - Control	Group B - BLM	-2.16667	1.84692	.766	-7.5908	3.2575
	Group C - Dark	-2.00000	1.84692	.814	-7.4242	3.4242
	Group D - BLM+Dark	-2.50000	1.84692	.662	-7.9242	2.9242
	Group E - Preference	-1.00000	1.84692	.982	-6.4242	4.4242
Group E - Preference	Group B - BLM	-1.16667	1.84692	.968	-6.5908	4.2575
	Group C - Dark	-1.00000	1.84692	.982	-6.4242	4.4242
	Group D - BLM+Dark	-1.50000	1.84692	.924	-6.9242	3.9242
	Group A - Control	1.00000	1.84692	.982	-4.4242	6.4242

Table 10.

Homogeneous Subsets: Test 2

GROUP	N	Subset for alpha = 0.05 1
Group A - Control	6	10.5000
Group E – Preference of blue-light-minimizing setting and darkened screen brightness	6	11.5000
Group C – Darkened screen brightness	6	12.5000
Group B – Blue-light-minimizing screen setting	6	12.6667
Group D – Blue-light-minimizing setting and darkened screen brightness	6	13.0000
Significance		.662

Note. Means for groups in homogeneous subsets are displayed and use a harmonic mean sample size of 6.000.

Table 11.

ANOVA: Test 3

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	38.133	4	9.533	2.154	.104
Within Groups	110.667	25	4.427		
Total	148.800	29			

Table 12.

Post Hoc Tests: Test 3

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Group B - BLM	Group C - Dark	1.33333	1.21472	.806	-2.2342	4.9008
	Group D - BLM+Dark	1.66667	1.21472	.650	-1.9008	5.2342
	Group A - Control	3.50000	1.21472	.056	-.0675	7.0675
	Group E - Preference	2.00000	1.21472	.483	-1.5675	5.5675
Group C - Dark	Group B - BLM	-1.33333	1.21472	.806	-4.9008	2.2342
	Group D - BLM+Dark	.33333	1.21472	.999	-3.2342	3.9008
	Group A - Control	2.16667	1.21472	.405	-1.4008	5.7342
	Group E - Preference	.66667	1.21472	.981	-2.9008	4.2342
Group D - BLM+Dark	Group B - BLM	-1.66667	1.21472	.650	-5.2342	1.9008
	Group C - Dark	-.33333	1.21472	.999	-3.9008	3.2342
	Group A - Control	1.83333	1.21472	.566	-1.7342	5.4008
	Group E - Preference	.33333	1.21472	.999	-3.2342	3.9008
Group A - Control	Group B - BLM	-3.50000	1.21472	.056	-7.0675	.0675
	Group C - Dark	-2.16667	1.21472	.405	-5.7342	1.4008
	Group D - BLM+Dark	-1.83333	1.21472	.566	-5.4008	1.7342
	Group E - Preference	-1.50000	1.21472	.732	-5.0675	2.0675
Group E - Preference	Group B - BLM	-2.00000	1.21472	.483	-5.5675	1.5675
	Group C - Dark	-.66667	1.21472	.981	-4.2342	2.9008
	Group D - BLM+Dark	-.33333	1.21472	.999	-3.9008	3.2342
	Group A - Control	1.50000	1.21472	.732	-2.0675	5.0675

Table 13.

Homogeneous Subsets: Test 3

GROUP	N	Subset for alpha = 0.05 1
Group A - Control	6	11.0000
Group E – Preference of blue-light-minimizing setting and darkened screen brightness	6	12.5000
Group D – Blue-light-minimizing setting and darkened screen brightness	6	12.8333
Group C – Darkened screen brightness	6	13.1667
Group B – Blue-light-minimizing screen setting	6	14.5000
Significance		.056

Note. Means for groups in homogeneous subsets are displayed and use a harmonic mean sample size of 6.000.

Means and standard deviations for Reading Section 1 on all three tests are presented in Table 14 and are presented as percentages. Reading Section 1 of Test 2 had a higher mean score (85.56%) than Test 1 (68.33%) and Test 3 (65%).

A one-way ANOVA was also conducted for Reading Section 1 on all three tests to determine if there were any statistical differences between variable groups (Tables 15-23). Overall, there were no statistically significant comparisons within groups for the shorter reading passages to suggest that any variable treatment alone made an effect on comprehension achievement.

Table 14.

Descriptive Statistics: Reading Section 1

	N	Minimum	Maximum	Mean	Std. Deviation
TEST 1 - SUB 1 – Short (%)	30	.17	1.00	.6833	.22468
TEST 2 - SUB 1 – Short (%)	30	.50	1.00	.8556	.15618
TEST 3 - SUB 1 – Short (%)	30	.33	1.00	.6500	.16580
Valid N (listwise)	30				

Table 15.

ANOVA: Test 1 - Reading Section 1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.200	4	.050	.989	.432
Within Groups	1.264	25	.051		
Total	1.464	29			

Table 16.

Post Hoc Tests: Test 1 - Reading Section 1

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Group B - BLM	Group C - Dark	.08333	.12981	.967	-.2979	.4646
	Group D - BLM+Dark	.13889	.12981	.820	-.2424	.5201
	Group A - Control	.25000	.12981	.330	-.1312	.6312
	Group E - Preference	.13889	.12981	.820	-.2424	.5201
Group C - Dark	Group B - BLM	-.08333	.12981	.967	-.4646	.2979
	Group D - BLM+Dark	.05556	.12981	.993	-.3257	.4368
	Group A - Control	.16667	.12981	.703	-.2146	.5479
	Group E - Preference	.05556	.12981	.993	-.3257	.4368
Group D - BLM+Dark	Group B - BLM	-.13889	.12981	.820	-.5201	.2424
	Group C - Dark	-.05556	.12981	.993	-.4368	.3257
	Group A - Control	.11111	.12981	.910	-.2701	.4924
	Group E - Preference	.00000	.12981	1.000	-.3812	.3812
Group A - Control	Group B - BLM	-.25000	.12981	.330	-.6312	.1312
	Group C - Dark	-.16667	.12981	.703	-.5479	.2146
	Group D - BLM+Dark	-.11111	.12981	.910	-.4924	.2701
	Group E - Preference	-.11111	.12981	.910	-.4924	.2701
Group E - Preference	Group B - BLM	-.13889	.12981	.820	-.5201	.2424
	Group C - Dark	-.05556	.12981	.993	-.4368	.3257
	Group D - BLM+Dark	.00000	.12981	1.000	-.3812	.3812

Group A - Control	.11111	.12981	.910	-.2701	.4924

Table 17.

Homogeneous Subsets: Test 1 - Reading Section 1

GROUP	N	Subset for alpha = 0.05 1
Group A - Control	6	.5556
Group E – Preference of blue-light-minimizing setting and darkened screen brightness	6	.6667
Group D – Blue-light-minimizing setting and darkened screen brightness	6	.6667
Group C – Darkened screen brightness	6	.7222
Group B – Blue-light-minimizing screen setting	6	.8056
Sig.		.330

Note. Means for groups in homogeneous subsets are displayed and use a harmonic mean sample size of 6.000.

Table 18.

ANOVA: Test 2 - Reading Section 1

	Sum of Squares	df	Mean Square	F	Sig.

Between Groups	.133	4	.033	1.452	.247
Within Groups	.574	25	.023		
Total	.707	29			

Table 19.

Post Hoc Tests: Test 2 - Reading Section 1

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Group B - BLM	Group C - Dark	.02778	.08749	.998	-.2292	.2847
	Group D - BLM+Dark	.00000	.08749	1.000	-.2569	.2569
	Group A - Control	.16667	.08749	.341	-.0903	.4236
	Group E - Preference	.11111	.08749	.711	-.1458	.3681
Group C - Dark	Group B - BLM	-.02778	.08749	.998	-.2847	.2292
	Group D - BLM+Dark	-.02778	.08749	.998	-.2847	.2292
	Group A - Control	.13889	.08749	.519	-.1181	.3958
	Group E - Preference	.08333	.08749	.873	-.1736	.3403
Group D - BLM+Dark	Group B - BLM	.00000	.08749	1.000	-.2569	.2569
	Group C - Dark	.02778	.08749	.998	-.2292	.2847
	Group A - Control	.16667	.08749	.341	-.0903	.4236
	Group E - Preference	.11111	.08749	.711	-.1458	.3681
Group A - Control	Group B - BLM	-.16667	.08749	.341	-.4236	.0903
	Group C - Dark	-.13889	.08749	.519	-.3958	.1181
	Group D - BLM+Dark	-.16667	.08749	.341	-.4236	.0903
	Group E - Preference	-.05556	.08749	.968	-.3125	.2014

Group E - Preference	Group B - BLM	-.11111	.08749	.711	-.3681	.1458
	Group C - Dark	-.08333	.08749	.873	-.3403	.1736
	Group D - BLM+Dark	-.11111	.08749	.711	-.3681	.1458
	Group A - Control	.05556	.08749	.968	-.2014	.3125

Table 20.

Homogeneous Subsets: Test 2 - Reading Section 1

		Subset for alpha = 0.05
GROUP	N	1
Group A - Control	6	.7500
Group E – Preference of blue-light-minimizing setting and darkened screen brightness	6	.8056
Group C – Darkened screen brightness	6	.8889
Group D – Blue-light-minimizing setting and darkened screen brightness	6	.9167
Group B – Blue-light-minimizing screen setting	6	.9167
Sig.		.341

Note. Means for groups in homogeneous subsets are displayed and use a harmonic mean sample size of 6.000.

Table 21.

ANOVA: Test 3 - Reading Section 1

	Sum of Squares	df	Mean Square	F	Sig.

Between Groups	.107	4	.027	.973	.440
Within Groups	.690	25	.028		
Total	.797	29			

Table 22.

Post Hoc Tests: Test 3 - Reading Section 1

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Group B - BLM	Group C - Dark	.05556	.09590	.977	-.2261	.3372
	Group D - BLM+Dark	-.08333	.09590	.906	-.3650	.1983
	Group A - Control	-.05556	.09590	.977	-.3372	.2261
	Group E - Preference	-.11111	.09590	.774	-.3928	.1705
Group C - Dark	Group B - BLM	-.05556	.09590	.977	-.3372	.2261
	Group D - BLM+Dark	-.13889	.09590	.604	-.4205	.1428
	Group A - Control	-.11111	.09590	.774	-.3928	.1705
	Group E - Preference	-.16667	.09590	.430	-.4483	.1150
Group D - BLM+Dark	Group B - BLM	.08333	.09590	.906	-.1983	.3650
	Group C - Dark	.13889	.09590	.604	-.1428	.4205
	Group A - Control	.02778	.09590	.998	-.2539	.3094
	Group E - Preference	-.02778	.09590	.998	-.3094	.2539
Group A - Control	Group B - BLM	.05556	.09590	.977	-.2261	.3372
	Group C - Dark	.11111	.09590	.774	-.1705	.3928
	Group D - BLM+Dark	-.02778	.09590	.998	-.3094	.2539

	Group E - Preference	-.05556	.09590	.977	-.3372	.2261
Group E - Preference	Group B - BLM	.11111	.09590	.774	-.1705	.3928
	Group C - Dark	.16667	.09590	.430	-.1150	.4483
	Group D - BLM+Dark	.02778	.09590	.998	-.2539	.3094
	Group A - Control	.05556	.09590	.977	-.2261	.3372

Table 23.

Homogeneous Subsets: Test 3 - Reading Section 1

GROUP	N	Subset for alpha = 0.05
		1
Group C – Darkened screen brightness	6	.5556
Group B – Blue-light-minimizing screen setting	6	.6111
Group A - Control	6	.6667
Group D – Blue-light-minimizing setting and darkened screen brightness	6	.6944
Group E – Preference of blue-light-minimizing setting and darkened screen brightness	6	.7222
Sig.		.430

Note. Means for groups in homogeneous subsets are displayed and use a harmonic mean sample size of 6.000.

Means and standard deviations for Reading Section 2 on all three tests are presented in Table 24 and are presented as percentages. Reading Section 2 of Test 3 had a higher mean score (63.57%) than Test 1 (44.76%) and Test 2 (49.29%).

A one-way ANOVA was also conducted for Reading Section 2 on all three tests to determine if there were any statistical differences between variable groups (Tables 25-33). The post hoc tests done on Reading Section 2 of Test 3 showed a 2.1% statistical difference between the blue-light minimizing screen-setting group when combined with a darkened room and the control group (Table 32). The control group mean score was 27.381% lower than the variable group using the blue-light-minimizing computer screen setting. This shows that the blue-light-minimizng screen setting in combination with a darkened room environment resulted in significantly higher scores than a room without either variable.

Table 24.

Descriptive Statistics: Reading Section 2

	N	Minimum	Maximum	Mean	Std. Deviation
TEST 1 - SUB 2 – Long (%)	30	.14	.71	.4476	.17083
TEST 2 - SUB 2 – Long (%)	30	.21	.86	.4929	.18884
TEST 3 - SUB 2 – Long (%)	30	.36	.93	.6357	.16391
Valid N (listwise)	30				

Table 25.

ANOVA: Test 1 - Reading Section 2

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.038	4	.010	.297	.877
Within Groups	.808	25	.032		
Total	.846	29			

Table 26.

Post Hoc Tests: Test 1 - Reading Section 2

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Group B - BLM	Group C - Dark	.01190	.10378	1.000	-.2929	.3167
	Group D - BLM+Dark	-.03571	.10378	.997	-.3405	.2691
	Group A - Control	.03571	.10378	.997	-.2691	.3405
	Group E - Preference	.07143	.10378	.957	-.2334	.3762
Group C - Dark	Group B - BLM	-.01190	.10378	1.000	-.3167	.2929
	Group D - BLM+Dark	-.04762	.10378	.990	-.3524	.2572
	Group A - Control	.02381	.10378	.999	-.2810	.3286
	Group E - Preference	.05952	.10378	.978	-.2453	.3643
Group D - BLM+Dark	Group B - BLM	.03571	.10378	.997	-.2691	.3405
	Group C - Dark	.04762	.10378	.990	-.2572	.3524

	Group A - Control	.07143	.10378	.957	-.2334	.3762
	Group E - Preference	.10714	.10378	.838	-.1977	.4119
Group A - Control	Group B - BLM	-.03571	.10378	.997	-.3405	.2691
	Group C - Dark	-.02381	.10378	.999	-.3286	.2810
	Group D - BLM+Dark	-.07143	.10378	.957	-.3762	.2334
	Group E - Preference	.03571	.10378	.997	-.2691	.3405
Group E - Preference	Group B - BLM	-.07143	.10378	.957	-.3762	.2334
	Group C - Dark	-.05952	.10378	.978	-.3643	.2453
	Group D - BLM+Dark	-.10714	.10378	.838	-.4119	.1977
	Group A - Control	-.03571	.10378	.997	-.3405	.2691

Table 27.

Homogeneous Subsets: Test 1 - Reading Section 2

GROUP	N	Subset for alpha = 0.05 1
Group E – Preference of blue-light-minimizing setting and darkened screen brightness	6	.3929
Group A - Control	6	.4286
Group C – Darkened screen brightness	6	.4524
Group B – Blue-light-minimizing screen setting	6	.4643

Group D – Blue-light-minimizing setting and darkened screen brightness	6	.5000
Sig.		.838

Note. Means for groups in homogeneous subsets are displayed and use a harmonic mean sample size of 6.000.

Table 28.

ANOVA: Test 2 - Reading Section 2

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.042	4	.010	.263	.899
Within Groups	.992	25	.040		
Total	1.034	29			

Table 29.

Post Hoc Tests: Test 2 - Reading Section 2

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Group B - BLM	Group C - Dark	.00000	.11503	1.000	-.3378	.3378
	Group D - BLM+Dark	-.02381	.11503	1.000	-.3616	.3140
	Group A - Control	.08333	.11503	.949	-.2545	.4212
	Group E - Preference	.03571	.11503	.998	-.3021	.3735
Group C - Dark	Group B - BLM	.00000	.11503	1.000	-.3378	.3378
	Group D - BLM+Dark	-.02381	.11503	1.000	-.3616	.3140

	Group A - Control	.08333	.11503	.949	-.2545	.4212
	Group E - Preference	.03571	.11503	.998	-.3021	.3735
Group D - BLM+Dark	Group B - BLM	.02381	.11503	1.000	-.3140	.3616
	Group C - Dark	.02381	.11503	1.000	-.3140	.3616
	Group A - Control	.10714	.11503	.882	-.2307	.4450
	Group E - Preference	.05952	.11503	.985	-.2783	.3973
Group A - Control	Group B - BLM	-.08333	.11503	.949	-.4212	.2545
	Group C - Dark	-.08333	.11503	.949	-.4212	.2545
	Group D - BLM+Dark	-.10714	.11503	.882	-.4450	.2307
	Group E - Preference	-.04762	.11503	.993	-.3854	.2902
Group E - Preference	Group B - BLM	-.03571	.11503	.998	-.3735	.3021
	Group C - Dark	-.03571	.11503	.998	-.3735	.3021
	Group D - BLM+Dark	-.05952	.11503	.985	-.3973	.2783
	Group A - Control	.04762	.11503	.993	-.2902	.3854

Table 30.

Homogeneous Subsets: Test 2 - Reading Section 2

		Subset for alpha = 0.05
GROUP	N	1
Group A - Control	6	.4286
Group E – Preference of blue-light-minimizing setting and darkened screen brightness	6	.4762
Group B – Blue-light-minimizing screen setting	6	.5119

Group C – Darkened screen brightness	6	.5119
Group D – Blue-light-minimizing setting and darkened screen brightness	6	.5357
Sig.		.882

Note. Means for groups in homogeneous subsets are displayed and use a harmonic mean sample size of 6.000.

Table 31.

ANOVA: Test 3 - Reading Section 2

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.270	4	.067	3.310	.026
Within Groups	.509	25	.020		
Total	.779	29			

Table 32.

Post Hoc Tests: Test 3 - Reading Section 2

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Group B - BLM	Group C - Dark	.07143	.08241	.906	-.1706	.3135
	Group D - BLM+Dark	.15476	.08241	.354	-.0873	.3968
	Group A - Control	.27381*	.08241	.021*	.0318	.5158
	Group E - Preference	.19048	.08241	.175	-.0516	.4325
Group C -	Group B - BLM	-.07143	.08241	.906	-.3135	.1706

Dark	Group D - BLM+Dark	.08333	.08241	.848	-.1587	.3254
	Group A - Control	.20238	.08241	.134	-.0396	.4444
	Group E - Preference	.11905	.08241	.606	-.1230	.3611
Group D - BLM+Dark	Group B - BLM	-.15476	.08241	.354	-.3968	.0873
	Group C - Dark	-.08333	.08241	.848	-.3254	.1587
	Group A - Control	.11905	.08241	.606	-.1230	.3611
	Group E - Preference	.03571	.08241	.992	-.2063	.2777
Group A - Control	Group B - BLM	-.27381*	.08241	.021	-.5158	-.0318
	Group C - Dark	-.20238	.08241	.134	-.4444	.0396
	Group D - BLM+Dark	-.11905	.08241	.606	-.3611	.1230
	Group E - Preference	-.08333	.08241	.848	-.3254	.1587
Group E - Preference	Group B - BLM	-.19048	.08241	.175	-.4325	.0516
	Group C - Dark	-.11905	.08241	.606	-.3611	.1230
	Group D - BLM+Dark	-.03571	.08241	.992	-.2777	.2063
	Group A - Control	.08333	.08241	.848	-.1587	.3254

Note. The mean difference is significant at the 0.05 level.

Table 33.

Homogeneous Subsets: Test 3 - Reading Section 2

		Subset for alpha = 0.05	
GROUP	N	1	2
Group A - Control	6	.5000	
Group E – Preference of blue-light-minimizing setting and darkened screen brightness	6	.5833	.5833

Group D – Blue-light-minimizing setting and darkened screen brightness	6	.6190	.6190
Group C – Darkened screen brightness	6	.7024	.7024
Group B – Blue-light-minimizing screen setting	6		.7738
Sig.		.134	.175

Note. Means for groups in homogeneous subsets are displayed and use a harmonic mean sample size of 6.000.

After reviewing the initial data between variable groups, an additional research question came up as to the possible effects of the outcomes of the study. This question was whether reading passage length alone had an impact on reading comprehension achievement scores. To look at this effect, a t-test was used to compare the difference of scores between all groups from Reading Section 1 and Reading Section 2 (Table 34). The paired samples test showed a statistical difference between scores on Reading Section 1 and Reading Section 2 for both Test 1 and Test 2. The mean score score for Test 1: Reading Section 1 was 68.33% compared to 44.76% for Test 2: Reading Section 2. The mean score score for Test 2: Reading Section 1 was 85.56% compared to 49.29% for Test 2: Reading Section 2. The mean scores for Test 3 did not show any statistical difference between Reading Section 1 (65%) and Reading Section 2 (63.57%)

Table 34. T-Test: Reading Sections 1 & 2

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Test 1	SUB 1 - Short	.6833	30	.22468	.04102
	SUB 2 - Long	.4476	30	.17083	.03119

Test 2	SUB 1 - Short	.8556	30	.15618	.02852
	SUB 2 - Long	.4929	30	.18884	.03448
Test 3	SUB 1 - Short	.6500	30	.16580	.03027
	SUB 2 - Long	.6357	30	.16391	.02992

Paired Samples Correlations

		N	Correlation	Sig.
Test 1	SUB 1 - Short & SUB 2 - Long	30	.227	.228
Test 2	SUB 1 - Short & SUB 2 - Long	30	.367	.046*
Test 3	SUB 1 - Short & SUB 2 - Long	30	-.246	.190

Paired Samples Test

		Paired Differences			
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower
Test 1	SUB 1 - Short - SUB 2 - Long	.23571	.24951	.04555	.14255
Test 2	SUB 1 - Short - SUB 2 - Long	.36270	.19591	.03577	.28955
Test 3	SUB 1 - Short - SUB 2 - Long	.01429	.26027	.04752	-.08290

Table 34 (continued)

Paired Samples Test

		Paired Differences 95% Confidence Interval of the Difference Upper	t	df	Sig. (2-tailed)
Test 1	SUB 1 - Short - SUB 2 - Long	.32888	5.174	29	.000*

Test 2	SUB 1 - Short - SUB 2 - Long	.43585	10.140	29	.000*
Test 3	SUB 1 - Short - SUB 2 - Long	.11147	.301	29	.766

Note. The mean difference is significant at the 0.05 level.

Chapter V: Summary, Conclusions, Discussion, Limitations, and Recommendations

Increased technology access in schools and the recent reauthorizations of the Elementary and Secondary Education Act have increased the prevalence of computer-based assessment in school settings (the IASA in 1994; NCLB in 2001; ESSA in 2015). Computer-based testing is essential to the data collection needed to support school districts and states as they monitor the progress of students to meet state standards. While computers do offer a better range and ease of accumulating data, there have been concerns about the effects of computer-based assessments on achievement.

The purpose of this study was to determine if specific screen settings improve student achievement on reading comprehension assessment. The specific screen settings tested in this study were a blue-light minimizing screen setting (known as “Night Light” on Google Chromebooks), lowered computer screen brightness, a combination of the blue-light minimizing screen setting with computer screen brightness lowered, as well as adjusting the room lighting in combination of the three previously mentioned treatments.

Conclusions

This study was directed by four research questions. They are stated and individually answered in the following paragraphs.

<u>Research Question 1:</u> Is there a difference in achievement between students who use a blue-light minimizing computer screen setting compared to those who do not? The null hypothesis for this research question is, “Using a blue-light minimizing computer screen setting

has no effect on reading comprehension achievement scores." The results of the study failed to reject this null hypothesis.

There was a statistical difference when comparing the group using the blue-light minimizing computer screen setting to the control group for the third test (Table 12). Despite this finding, we cannot reject the null hypothesis because we did not see a consistent result in the second test (Table 10), nor did we see similar results in the group that had the blue-light minimizing screen setting and reduced screen brightness on either the second or third test (Table 12). Therefore, the blue-light minimizing computer screen setting does not result in better achievement scores alone.

Research Question 2: Is there a difference in achievement when computer screen brightness is darkened compared to when it is not? The null hypothesis for this research question is, "Lowering the brightness setting on the computer has no effect on reading comprehension achievement scores." The results of the study failed to reject his null hypothesis because there were no significant findings between groups using a lowered screen brightness to those that did not (Tables 8, 10, & 12).

Research Question 3: Is there a difference in achievement when the fluorescent room lighting is turned off compared to when it is not? The null hypothesis for this research question is, "Turning of fluorescent room lighting has no effect on reading comprehension achievement scores." The results of the study failed to reject his null hypothesis, as there were no consistent, significant findings between groups with a darkened room and the control group that did not.

There was a statistical difference between the blue-light minimizing settings with a darkened room compared to the control group (without a darkened room) (Table 12). However, the other three groups that also had a darkened room environment did not have a similar result (Table 12).

Research Question 4: Is there a difference among the tested variables in achievement scores when comparing the short-passage scores to the long-passage scores? The null hypothesis for this research question is, "There is no difference in achievement among the tested variables when comparing short-passage scores to long-passage scores." The results of the study failed to reject this null hypothesis because there were no significant findings between groups when comparing the scores of shorter passages to longer passages.

There was a difference in mean scores between the short passage questions and long passage questions on the first and second test; however, the third test did not see the same result (Table 34). After inspecting the readings further, it was found that while each of the longer reading passages were nearly the same length on the computer screen requiring the same amount of scrolling, their word count differed due to the formatting of the reading. The long passage readings in the third test had several shorter paragraphs that created more space on the screen and lacked the word density of the other tests. In the first test, the longer-passage section had 1626 words, and the second test had 1330. The third test had only 974 words. This shows that there may be a connection between the number of words required to read and comprehend, and achievement results.

Discussion

Reading comprehension is a unique skill that has several impactful elements such as reading fluency, vocabulary, background knowledge of content, and critical thinking (Kamhi & Catts, 2017). Reading comprehension assessments indirectly assess the actual process of reading comprehension because they all require other tasks such as marking, writing, or speaking (Sarroub & Pearson, 1998). Testing reading comprehension can be a difficult task due to the fact that it is hard to appropriately assess the skill of comprehension without the risk of one of these skills interfering with the results. Current standardized assessments rely heavily on multiple choice or multiple response questions. The structure of reading comprehension assessments that require answers to specific questions can lead to the reading process becoming more of a search for answers instead of reading the passage for understanding (Tenaha, et al., 2018).

Reading and answering questions on a computer screen may require more concentration than reading questions on paper (Jeong, 2012). Ferris Jabr of the *Scientific American* wrote, "compared with paper, screens may also drain more of our mental resources while we are reading and make it a little harder to remember what we read when we are done" (2013, p. 6). As this study has shown, the length of reading seems to have an impact on reading comprehension assessment results. While it makes sense that longer readings would require more concentration and will typically lead to worse comprehension results, more studies should be done to make sure that digital screens do not exacerbate the problem. Further research must also be done to understand why the test mode effect takes place on computer-based tests in an attempt to find a solution to lessen or remove the barrier.

Another consideration to further reading comprehension assessment study is the role of motivation. A study by Kelly and Decker (2009) found that students reading comprehension performance was largely based on intrinsic motivation. Tarchi's (2017) work on expository texts also supported this idea, stating when students were more motivated it can "compensate for lower levels of metacognition, and 'energize' students' approach to the text, with positive effects on their reading comprehension performances" (Tarchi, 2017, p. 174). Noyes, Garland, and Robbins (2004) focused on workload as a test mode effect, but discovered that there were no significant differences in comprehension scores; however, the perception of the difficulty level varied between computer-based and paper-based testers because computer-based testers felt that the task was more difficult. Perception of the task being more difficult could also lead to test-takers being less motivated during the task. Since the growth of computer-based assessment in the early 2000s there has been an increase of use and familiarity with technology, and more studies should be done to see what effect varying computer screen configurations and room lighting may have on motivation.

Computer-based testing must be implemented in a fashion that keeps in mind the assumptions stakeholders (students, teachers, parents, politicians, and community members) make about how computer-based testing is used, examines the educational structure as a whole (not just the test), and identifies potential accessibility problems that may occur when formatting and administering computer-based tests (Thurlow, Lazarus, Albus & Hodgson, 2010). Understanding the limitations of standardized testing (computer-based or paper-based) is important when considering how we should use the tests and the test results as a piece of curriculum, and not the only measure of academic success. The complexity of reading comprehension stresses the need for a diverse approach to assessment.

As Niccoli noted (2015):

> If educators understand the effects of digital reading on the development of deep reading and student's grasp of difficult material, they can formulate instructional decisions. Given the current pace of technological change, educators should seize opportunities to

further advance our understanding of students' learning while using electronic devices. (p. 26)

Limitations

1. The author's definition of what blue-light constitutes is unclearly communicated in the research.
2. Students did not have exposure to the treatments used in the study beforehand, making it the first time that students have used the screen settings, which possibly skewed the results compared to if they were used to the screen settings.
3. The Schoology software used in the study does not have the split-screen format, where a test-taker can view both the reading and the questions at the same time, that most closely resembles the MCA reading test (and most other standardized testing formats like it).

Recommendations for Research

The following items are recommendations for research topics or expansion of this study for further research:

1. Replicate study and expand the sample size in total number and age range.
2. Investigate the effect that reading passage word density has on reading comprehension when reading from a computer monitor.
3. Study the effect of computer screen settings and/or room lighting on test-taker motivation.

Recommendations for Practice

The following items are recommendations for educators administering standardized testing on computers:

1. Keep room lighting configuration in mind when administering a test by limiting the amount of fluorescent light used, and keeping the room brightness below computer screen levels.
2. Inform students of screen setting options and allow them to test their preference before testing begins.
3. Teach students strategies on how to approach longer reading passages on a test (reading questions first, using available on-screen marking tools, etc.)
4. Investigate strategies to combat eye fatigue and motivation such as taking breaks from the screen or taking a minute to stand or step away from the computer.
5. Continue to offer paper-based reading materials a options for students, particularly for reading that is challenging.

References

Abrams, L. (2012). How to Keep Computer Screens From Destroying Your Eyes. *The Atlantic.* Retrieved from https://www.theatlantic.com/health/archive/2012/09/how-to-keep-computer-screens-from-destroying-your-eyes/263005/

Ackerman, R., & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied, 17*(1), 18-32. doi:10.1037/a0022086.

Ackerman, R. & Lauterman, T. (2012). Taking reading comprehension exams on screen or paper? A metacognitive analysis of learning texts under time pressure. *Computers in Human Behavior.* Doi: 10.1016/j.chb.2012.04.023

Anuardi, M., Yamazaki, A. K., & Eto, K. (2017). A pre-NIRS study of background colour effects on the functions of the frontal lobe. *Procedia Computer Science,* 112, 2031-2039. doi: 10.1016/j.procs.2017.08.149

Aydemir, Z., Ozturk, E., Horzum, M. B. (2013). The Effect of Reading from Screen on The 5th Grade Elementary Students' Level of Reading Comprehension on Informative and Narrative Type of Texts. *Educational Sciences: Theory & Practice, 13*(4), 2272-2276. doi: 10.12738/estp.2013.4.1294

Baldwin, R. & Colt, J. (2010). Your PC Is Ruining Your Vision. Here's How to Beat Eye Strain. *Wired.* Retrieved from https://www.wired.com/2013/09/flux-eyestrain/

Bera, A. (2018). How To Use Night Mode in Chrome OS. *Technobezz.* Retrieved from https://www.technobezz.com/use-night-mode-chrome-os/

Boevé, A., Meijer, R., Albers, C., Beetzma, Y., & Bosker, R. (2015). Introducing

Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment. *PLoS ONE 10*(12). doi:10.1371/journal.pone.0143616.

Brangers, G. (2018). Chromebook How-To: Enabling Night Mode. *Chromebook Unboxed*. Retrieved from https://chromeunboxed.com/chromebook-how-to-enabling-night-mode/

Bridgeman, B., Lennon, M. L., & Jackenthal, A. (2003). Effects of Screen Size, Screen Resolution, and Display Rate on Computer-Based Test Performance. *Applied Measurement in Education*, *16*(3), 191-205.

Bugbee, A. C. (1996). The equivalence of paper-and-pencil and computer-based testing. *Journal of Research on Computing in Education*, *28*(3), 282-290.

Bunderson, C.V., Inouye, D.K., & Olsen, J.B. (1989). The four generations of computerized educational measurement. *The American Council on Education/Macmillian series on higher education. Educational measurement*, 367-407.

Chang, A. M., Aeschbach, D., Duffy, J. F., & Czeisler, C. A. (2014). Evening use of light-emitting eReaders negatively affects sleep, circadian timing, and next-morning alertness. *Proc Natl Acad Sci USA*, *112*(4), 1232-1237. doi: 10.1073/pnas.1418490112.

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*(5), 593-602.

Crowder, C. (2018). Night Mode: 8 Benefits for Your Health. *Digital Care.* Retrieved from https://www.digitalcare.org/night-mode-health-benefits/

Daniel, D. B., & Woody, W. D. (2013). E textbooks at what cost? Performance and use of electronic vs. print texts. *Computers in Education, 62*, 18-23. doi: 10.1016/j.compedu.2012.10.016.

Darling-Hammond, L. (2007). Evaluating No Child Left Behind. *The Nation*, 11-18. Retrieved from https://www.thenation.com/article/evaluating-no-child-left-behind/

Dillion, A. (1992). Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics, 35*, 1297-1326.

Dosch, M. (2012). Practice in Computer-Based Testing Improves Scores on the National Certification Examination for Nurse Anesthetists. *AANA Journal, 80*(4), S60-S66.

EdTech Staff. (2017). More than 50 Percent of Teachers Report 1:1 Computing. *EdTech.* Retrieved from https://edtechmagazine.com/k12/article/2017/02/more-50-percent-teachers-report-11-computing

Gewertz, C. (2018). In the Testing Arena, a Wait-and-See Attitude As States Eye ESSA's Offer of New Leeway. *Education Week, 37*(25), 21-22.

Gewertz, C. (2019). What Tests Does Each State Require? *Education Week.* Retrieved from https://www.edweek.org/ew/section/multimedia/what-tests-does-each-state-require.html

Haas, C., & Hayes, J. (1986). What did I just say? Reading problems in writing with the machine. *Research in the Teaching of English, 20*(1), 22-35.

Heid, M. (2017). You Asked: What's the Best Way to Stare at Screens All Day? *Time.* Retrieved from http://time.com/4789208/screens-computer-eye-strain/

Heiting, G. (2018). Blue Light: It's Both Bad And Good For You. *All About Vision.* Retrieved from https://www.allaboutvision.com/cvs/blue-light.htm.

Heiting, G., & Wan, L. (2017). Computer Eye Strain: 10 Steps For Relief. *All About Vision.*

Retrieved from https://www.allaboutvision.com/cvs/irritated.htm

Herold, B. (2014). Screen Reading Poses Learning Challenges. *Education Week*, *33*(30), 1-25.

Huddleston, A. P., & Rockwell, E. C. (2015). Assessment for the Masses: A Historical Critique of High-Stakes Testing in Reading. *Texas Journal of Literacy Education*, *3*(1), 38-49. Retrieved from https://files.eric.ed.gov/fulltext/EJ1110955.pdf

Jabr, F. (2013). The Reading Brain in the Digital Age: The Science of Paper versus Screens. *Scientific American.* Retrieved from https://www.scientificamerican.com/article/reading-paper-screens/

Jabr, F. (2016). Blue Light Blues. *Scientific American*, *315*(5), 24-25. doi: 10.1038/scientificamerican1116-24

Jeong, H. (2012). A comparative study of scores on computer-based tests and paper-based tests. Behavior & Information Technology, 33(4), 410-422. doi:10.1080/0144929X.2012.710647.

Kamhi, A., & Catts, H. W. (2017). Epilogue: Reading Comprehension Is Not a Single Ability - Implications for Assessment and Instruction. Language, Speech, and Hearing Services in Schools, 48, 104-107.

Kelley, M. J., & Decker, E. O. (2009). The current state of motivation to read among middle school students. Reading Psychology, 30, 466-486. doi:10.1080/02702710902733535

Lauterman, T., & Ackerman, R. (2014). Overcoming screen inferiority in learning and calibration. Computers in Human Behavior, 35. doi: 10.1016/j.chb.2014.02.046.

Library of Congress, Congressional Research Service. (2008). Overview of No Child Left Behind. *Congressional Digest 87*(5), 131-137.

Lin, J. B., Gerratt, B. W., Bassi, C. J., & Rajendra, S. A. (2019). Short-Wavelength Light-Blocking Eyeglasses Attenuate Symptoms of Eye Fatigue. *Investigative Ophthalmology & Visual Science, 58*(1), 442-447.

Lipson, M. Y., & Wixson, K. K. (1986). Research on reading disabilities: An interactionist perspective. *Review of Educational Research, 56*, 111-136.

Makhoul, B., & Copti-Mshael, T. Reading Comprehension as a Function of Text Genre and Presentation Environment: Comprehension of Narrative and Informational Texts in a Computer-Assisted Environment vs. Print. *Psychology, 6*(8), 1001-1012. doi: 10.4236/psych.2015.68097

Mangen, A., Walgermo, B., & Bronnick, K. (2012). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Education Research, 58*, 61-68. doi: 10.1016/j.iher.2012.12.002

Mazzeo, J., & Harvey, A. L. (1988). The equivalence of scores from automated and conventional educational and psychological tests. *ETS Research Report Series*, 1-27. doi: 10.1002/j.2330-8516.1988.tb00277.

Meyer, D., & Kollbaum, P. (2016). Doctor, My Eyes...Are Tired! *Review of Optometry*. Retrieved from https://www.reviewofoptometry.com/article/doctor-my-eyes-are-tired

Minnesota Department of Education. (2017). Minnesota Assessments. *Minnesota Technical Manual*. Retrieved from https://www.leg.state.mn.us/webcontent/lrl/guides/AcademicStandards/MCAHistory.pdf

Mott, M. S., Robinson, D. H., Walden, A., Burnette, J., & Rutherford, A. S. (2012). Illuminating the Effects of Dynamic Lighting on Student Learning. *SAGE Open.* doi:

10.1177/2158244012445585

Murata, A., Uetake, A., Otsuka, M., & Takasawa, Y. (2001). Proposal of an Index to Evaluate Visual Fatigue Induced During Visual Display Terminal Tasks. *International Journal of Human-Computer Interaction, 13*(3), 305-321.

Myrberg, C., & Wiberg, N. (2013). Screen Vs. Paper: What Is the Difference for Reading and Learning? *Insights 28*(2), 49-54. doi:10.1629/uksg.236

Nagare, R., Plitnick B., & Figueiro, M. G. (2018). Does the iPad Night Shift mode reduce melatonin suppression?. *Lighting Research Center.* doi: 10.1177/1477153517748189

Nagaraja, M. P. (2019). Tour of the Electromagnetic Spectrum. *NASA Science.* Retrieved from https://science.nasa.gov/ems/09_visiblelight

Niccoli, A. (2015). Paper or Tablet? Reading Recall and Comprehension. *Educause Review.* Retrieved from https://er.educause.edu/articles/2015/9/paper-or-tablet-reading-recall-and-comprehension

Noyes, J. M, & Garland, K. J. (2003). VDT versus paper-based text: Reply to Mayes, Sims, and Koonce. *International Journal of Industrial Ergonomics, 31*, 411-423.

Noyes, J. M., & Garland, K. J. (2008). Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics, 51*(9), 1352-1375. doi: 10.1080/00140130802170387.

Noyes, J. M., Garland, K. J., & Robbins, L. (2004). Paper-based versus computer-based assessment: is workload another test mode effect? *British Journal of Educational Technology 35*(1), 111-113. doi: 10.1111/j.1467-8535.2004.00373.x

Oborne, D. J., & Holton, D. (1988). Reading from Screen versus Paper: There is No Difference. *International Journal of Man-Machine Studies, 28*, 1-9.

doi:10.1016/S0020-7373(88)80049-X.

Rello, L., & Bigham, J. (2017). Good Background Colors for Readers: A Study of People with and without Dyslexia. *Carnegie Mellon University.* doi: 10.1145/3132525.3132546.

Robelen, E. W. (2005). 40 Years After ESEA, Federal Role in Schools Is Broader Than Ever. *Education Week, 24*(31), 1,42. Retrieved from https://www.edweek.org/ew/articles/2005/04/13/31esea.h24.html

Sarroub, L., & Pearson, D. P. (1998). Two Steps Forward, Three Steps Back: The Stormy History of Reading Comprehension Assessment. *Clearing House, 72*(2), 97-106. doi:10.1080/00098659809599604

Schaffhauser, D. (2011). High-Stakes Online Testing. *T H E Journal, 38*(6), 28-39.

Tarchi, C. (2017). Comprehending Expository Texts: The Role of Cognitive and Motivational Factors. *Reading Psychology, 38,* 154-181. doi: 10.1080/02702711.2016.1245229.

Thurlow, M., Lazarus, S. S., Albus, D., & Hodgson, J. (2010). *Computer-based testsing: Practices and considerations* (Synthesis Report 78). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

United States. (1965). Elementary and Secondary Education Act of 1965 : H. R. 2362, 89th Cong., 1st sess., Public law 89-10. Reports, bills, debate and act.

Veitch, J. A., & McColl, S. L. (2001). A critical examination of perceptual and cognitive effects attributed to full-spectrum fluorescent lighting. *Ergonomics, 44*(4), 473. doi: 10.1080/00140130121241.

Vimont, C., & Khurana, R. (2017). Should You Be Worried About Blue Light? *American Academy of Ophthalmology*. Retrieved from https://www.aao.org/eye-health/tips-prevention/should-you-be-worried-about-blue-light

Wästlund, E., Reinikka, H., Norlander, T., & Archer, T. (2004). Effects of VDT and paper presentation on consumption and production of information: *Psychological and physiological factors. Computers in Human Behavior, 21*, 377-394. doi: 10.1016/j.chb.2004.02.007

What is blue light? (2012). The effect blue light has on your sleep and more. *Harvard Health Publishing.* Retrieved from https://www.health.harvard.edu/staying-healthy/blue-light-has-a-dark-side

Wigdor, A. & Garden, W. (1982). Ability testing: Uses, consequences, and controversies. *Washington, DC: National Academy Press*.

Wiliam, D. (2010). Standardized Testing and School Accountability. *Educational Psychologist, 45*(2), 107-122. doi: 10.1080/00461521003703060

Wilkinson, R.T., & Robinshaw, H.M. (1987). Proof-reading: VDU and paper text compared for speed, accuracy, and fatigue. *Behaviour and Information Technology*, *6*(2), 125-133.

Winterbottom, M., & Wilkins, A. (2009). Lighting and discomfort in the classroom. *Journal of Environmental Psychology*, *29*(1), 63-75. doi: 10.1016/j.jenvp.2008.11.007

Wolf, T.H. (1973). Alfred Binet. *Chicago: University of Chicago Press*.

Ziefle, M. (1998). Effects of display resolution on visual performance. *Human Factors, 40*, 554-568. doi: 10.1518/001872098779649355

Appendix A: Institutional Review Board Approval

ST. CLOUD STATE UNIVERSITY
EDUCATION FOR LIFE.

Institutional Review Board (IRB)

720 4th Avenue South AS 210, St. Cloud, MN 56301-4498

Name: Matthew Houselog

Email: mehouselog@stcloudstate.edu

IRB PROTOCOL DETERMINATION:
Expedited Review-1

Project Title: Effects of Computer Screen Color and Brighteness on Student Achievement during Computer-Based Reading Comprehension Assessment

Advisor Kayona Frances

The Institutional Review Board has reviewed your protocol to conduct research involving human subjects. Your project has been: **APPROVED**

Please note the following important information concerning IRB projects:
- The principal investigator assumes the responsibilities for the protection of participants in this project. Any adverse events must be reported to the IRB as soon as possible (ex. research related injuries, harmful outcomes, significant withdrawal of subject population, etc.).

- For expedited or full board review, the principal investigator must submit a Continuing Review/Final Report form in advance of the expiration date indicated on this letter to report conclusion of the research or request an extension.

-Exempt review only requires the submission of a Continuing Review/Final Report form in advance of the expiration date indicated in this letter if an extension of time is needed.

- Approved consent forms display the official IRB stamp which documents approval and expiration dates. If a renewal is requested and approved, new consent forms will be officially stamped and reflect the new approval and expiration dates.

- The principal investigator must seek approval for any changes to the study (ex. research design, consent process, survey/interview instruments, funding source, etc.). The IRB reserves the right to review the research at any time.

If we can be of further assistance, feel free to contact the IRB at 320-308-4932 or email ResearchNow@stcloudstate.edu and please reference the SCSU IRB number when corresponding.

IRB Chair:

Dr. Benjamin Witts
Associate Professor- Applied Behavior Analysis
Department of Community Psychology,Counseling, and Family Therapy

IRB Institutional Official:

Dr. Latha Ramakrishnan
Interim Associate Provost for Research
Dean of Graduate Studies

OFFICE USE ONLY

SCSU IRB# 1887 - 2422	**Type:** Expedited Review-1	**Today's Date:** 3/13/2019
1st Year Approval Date: 3/8/2019	**2nd Year Approval Date:**	**3rd Year Approval Date:**
1st Year Expiration Date: 3/7/2020	**2nd Year Expiration Date:**	**3rd Year Expiration Date:**

Appendix B: Parent or Guardian Consent Form

Educational Study
Parental/Guardian Consent Form

Dear Parents and Guardians,

This form is being sent to ask your permission to allow your child to participate in a study being conducted for my Master's Degree at St. Cloud State University. Two consent forms—one for you, the parents/guardians, and the other for your child— are included with this memo. Both of these forms must be signed and returned prior to the start of the study. If your child is unable to read the student consent form, please take a few moments to read it to him/her and explain it as needed.

Purpose of Study:
The Rogers Middle School 8th-grade teachers will be conducting research regarding the effect of computer screen settings on student achievement during standardized reading comprehension tests. The study seeks to improve testing conditions and student success on computer-based standardized tests.

Procedure:
Three reading comprehension tests will be used to collect data for the study. These tests are part of the planned English curriculum, and students will take the reading comprehension tests whether or not their data is used in the study.

Confidentiality:
We are asking for your permission to use your child's data in the study. Students' names and personal information **WILL NOT** be used in the study's data collection. All data collected from these tests will be linked to a randomized identification number to ensure confidentiality.

Risks:
Screen and lighting conditions will be subtle so there is minimal discomfort risk to students during testing. If a student is feeling discomfort, they will be removed from the study and placed in a location with suitable lighting conditions.

Contact Information:
If you have any questions or would like to discuss the study further, please contact Mr. Matt Houselog matthew.houselog@isd728.org or my advisor, Dr. Frances Kayona fakayona@stcloudstate.edu.

Acceptance to Participate:
Your signature indicates that you and your child have read the information provided here and have decided to participate. You or your child may withdraw from the study at any time without penalty after signing this form. I look forward to having your children participate in this innovative study and I thank you in advance for your cooperation as I continue to complete my graduate study at St. Cloud State University.

Student's Name (Print): ______________________________

Parent/Guardian (Print): ______________________________

Parent/Guardian Signature: X ______________________________

______ Yes, you may use my child's data in the study.

______ No, I would prefer if my child's data is not used in the study.

St. Cloud State University
Institutional Review Board
Approval date: 3-8-19
Expiration date: 3-7-20

Appendix C: Student Assent Form

Educational Study
Student Assent

Students,

You are invited to be part of a study to see the effect of screen monitor and lighting settings on reading comprehension tests. This study will take three weeks to finish and will happen during our regularly scheduled class time.

Purpose of Study:

The Rogers Middle School 8th-grade teachers will be conducting research regarding the effect of computer screen settings on student achievement during standardized reading comprehension tests. The study seeks to improve testing conditions and student success on computer-based standardized tests.

Procedure:

Three reading comprehension tests will be used to collect data for the study. **These tests are part of the planned English curriculum, and students will take the reading comprehension tests whether or not their data is used in the study.**

Confidentiality:

Your name and personal information **WILL NOT** be used in the study's data collection. All data collected from the MCA practice tests will be linked to a randomized identification number to ensure anonymity and confidentiality. Screen and lighting conditions will be subtle so there is minimal discomfort risk to you during testing.

Risks:

Screen and lighting conditions will be subtle so there is minimal discomfort risk to students during testing. If you are feeling discomfort, we can move you to a location with suitable lighting conditions.

Contact Information:

If you have any questions or experience any discomfort during the study, please contact Mr. Houselog.

Acceptance to Participate:

When you sign your name on the line with the "X" it means you understand this information and have agreed to be a part of the study. If you do not like being in the study at any time, you may tell Ms. Cahill and she will see that you are put in a different class.

Student's Name (Print): ______________________________

Signature: x ______________________________________

St. Cloud State University
Institutional Review Board
Approval date: 3-8-19
Expiration date: 3-7-20