THE EFFECTS OF STUDENT CHOICE, INTEREST, AND PERFORMANCE ON STANDARDIZED READING ASSESSMENTS

_____

A Dissertation

Presented to

The Faculty of the School of Teaching and Learning

Sam Houston State University

_____

In Partial Fulfillment

of the Requirements for the Degree of

Doctor of Education

_____

by

Faida A. Stokes

May, 2020

THE EFFECTS OF STUDENT CHOICE, INTEREST, AND PERFORMANCE ON STANDARDIZED READING ASSESSMENTS

by

Faida A. Stokes

_____

APPROVED:

Chase Young, PhD
Dissertation Director

Melinda Miller, PhD
Committee Member

Patricia Durham, PhD
Committee Member

William Blackwell, EdD
Committee Member

Stacey Edmonson, PhD
Dean, College of Education, PhD

DEDICATION

This dissertation is dedicated to my parents, Kenneth and Hellena O. Stokes, and my brother, Imarogbe C. Stokes, who have been proud of my efforts long before this dissertation process and will continue to be **my best cheerleaders**. I **thank you** for your care, concern, and confidence throughout **all** my years of education.

Moreover, this is for my nephew, Kenneth K. Stokes, who was my source of joy and reprieve throughout this dissertation process.

Finally, to my guy,
William A. Howard, who has provided **boundless** patience
and an **abundance** of love.

Cheers!

ABSTRACT

Stokes, Faida A., *The effects of student choice, interest, and performance on standardized reading assessments*. Doctor of Education (Literacy), May, 2020, Sam Houston State University, Huntsville, Texas.

This study examined the effects of choice on student achievement using a criterion-referenced assessment with a quasi-experimental design and a non-experimental design to investigate the relationship between choice, interest, and performance on standardized reading assessment. Furthermore, the study explored the relationship between measures of choice (e.g., interest) and the standardized reading comprehension assessment with fourth and fifth grade students. The research was guided by two questions. Are there significant differences in reading comprehension assessment scores when reading 'teacher-selected' passages versus reading 'student-selected' passages? Is there a relationship between interest and performance on standardized reading comprehension assessments?

Based on the literature review, it was predicted that the provision of choice could increase the reading performance outcome on a standardized reading passage. Specifically, it was thought that choice would be a motivating factor in the students' comprehension and result in a higher score, when compared to the no-choice performance. However, findings from question one established the provision of choice was not a significant variable impacting students' scores on the standardized reading assessment, and findings from question two revealed no correlation between student preference for reading and achievement score. Implications for research, future research recommendations, and practical applications of choice through autonomy-supportive environments are discussed.

KEY WORDS: Student choice, Interest, Reading comprehension, Standardized reading assessments, Quasi-experimental design, Self-determination theory, Motivation, Reading engagement

ACKNOWLEDGEMENTS

I thank God for **all** blessings: direct, indirect, family, colleagues, mentors, and my daily bread. I am grateful for this life and the opportunity to carry out your plans.

I am thankful for my chair, Dr. Chase Young, for your direction(s), certitude, and boldness! I am beholden to my committee members: Dr. Patricia Durham, Dr. William Blackwell, and Dr. Melinda Miller, and professors, Dr. Nancy Votteler, Dr. Hannah Gerber, and Dr. Lory Haas for accepting me into the fold, supporting my thoughts and actions, and being exemplars of educational scholarship.

I acknowledge the teachers, principals, and friends of education who believed in this project and demonstrated that through a multitude of means. I am honored to know you and deeply touched by your favor.

Lastly, I want to uplift my different family circles (church, Conroe, Nashville, Dallas/Fort Worth, immediate and extended) in recognition of both the tangible and intangible ways that you facilitated the navigation of my past and present educational endeavors.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

CHAPTER I

Introduction

Reading is critical to student success in school and in life-long learning. Chambers et al. (2018) noted the improvement of adolescent reading is crucial in the comprehensive efforts to increase high school completion and in the preparation of students for college and career. Reading comprehension is a major predictor of readiness and success within these life steps and as a measure of reading achievement. Overall, children's reading performance and mastery of comprehension are important predictors of their school success (Wigfield & Guthrie, 1997) and, therefore, how reading achievement is measured is essential to this process.

Background of the study

Reading proficiency and student achievement are two ways to measure student preparedness, nationally and internationally (e.g. National Assessment of Educational Progress [NAEP], 2017; Organisation for Economic Co-operation and Development [OCED], 2003; Programme for International Student Assessment [PISA], 2015; Afflerbach, 2002). Student preparedness for college and career readiness is an important agenda item on the national educational circuit, as evidenced by the adoption of the common core curriculum or equivalent at the state level. Standardized assessment of reading comprehension has been common practice, nationally, since the adoption of the No Child Left Behind Act [NCLB] in 2002 (NCLB, 2002) and its subsequent reaffirmation through Every Student Succeeds Act (ESSA) in 2015, which upheld the fundamentals of standardized testing of students in grades 3-8. College and career readiness are designated markers of instructional efforts, interventions, and by-products

of literacy programs implemented from grades kindergarten through 12. Nationally, only 36% of students read at or above the 'proficient' level. (NAEP, 2015). On average across OECD countries, students' mean reading proficiency has not improved since 2000 (OECD/PISA, 2015, p. 148). Thus, 64% of eighth grade students do not have marketable and demonstrated skills in summarization, inferencing, analyzing textual evidence, and/or content analysis.

According to the Nation's Report Card, Texas students have been virtually stagnant in their reading growth over the last 5 years. In 2017, the average score of eighth-grade students in Texas was 260 on the reading state assessment. This was lower than the average score of 265 for public school students in the nation. The average score for students in Texas for 2017 (260) was not significantly different from their average score in 2015 (261) and in 1998 (261). The percentage of students in Texas who performed at or above the NAEP 'proficient' level was 28 percent in 2017. During those same years, average scores for fourth grade students decreased. The average score of fourth-grade students in Texas was 215. This was lower than the average score of 221 for public school students in the nation. In 2017, 35 states scored significantly higher than Texas, two performed significantly lower, and 14 were not statistically significant for fourth grade reading on average, based on the scaled scores of all students. Based on this data, students are not any more prepared for college or career and adolescent readers have not shown an increase in their academic performance, despite current, research-based, instructional practices. Studies have confirmed a gradual decline in reading motivation across grade levels (Marinak & Gambrell, 2008; Parsons, et al., 2018; Wigfield & Guthrie, 1997); however, the cause of this decline is not clear.

One contributing factor could be the lack of choice or limited choice in reading materials and/or instructional practices, including assessment (Gambrell et al., 2017; Gritter et al., 2017; Worthy, et al., 1999). Instructional choices surrounding reading affects the motivational context of the classroom (Gambrell et al., 2017) and attitudes around reading in the near and distant future. A clearer and stronger devotion toward "access to books, reading tasks that are relevant and opportunities for student choice have been linked to reading engagement and achievement" (Gambrell et al., 2017, p. 32) as these elements support intrinsically engaged reading and perpetuate reading habits.

Over a decade ago, Worthy et al. (1999) considered choice in reading materials a viable option for improving reading motivation, spawning interest, and raising competence in middle school students. They described the importance of understanding the preferences and interests of middle school students within the larger constructs of motivation, engagement, classroom environment, instructional materials, and student-centered learning. Their study found student preferences and interests in reading materials were misaligned with the reading materials available for reader consumption. Overall, students' choices were limited, motivation was stymied, and interests were not fostered, all of which contribute to a decrease in reading-based activities.

Gritter et al. (2017) conducted a critical content analysis of the picture books listed on the Children's Choices reading lists from 2000-2014 published by the International Literacy Association. Their analysis probed how deficit views of male protagonists' literate identities manifested in the literature contributed to boys' literary underachievement. The findings uncovered a misalignment between school and non-school settings of male character's relations and behaviors. Unlike in non-school

settings, male characters placed in school settings were often portrayed through negative stereotypes, undervalued, or lacked the opportunity to experience a positive change. These missed literary opportunities could be viewed as a factor of the overall lower literacy assessment scores for males as they fail to motivate and engage young males to participate in the learning experience, when compared with literate activities they choose to do outside of school (Gritter, et al., 2017).

Several have studied direct and indirect choice in literacy (Fisher & Frey, 2018; Khan et al., 2013; Mohr, 2003; Mohr, 2006). Mohr (2003) contends the motivation to read is strongly associated with the opportunity and ability to self-select books. Khan et al. (2013) found choice generated stronger narrative skills, including story grammar knowledge, retell, and production, of preschool students compared to a matched, no-choice group. In addition, post-test comprehension scores for the choice group were greater for questions targeting story elements and sequence. Results support the need for choice, as a facilitator of engagement, attention, motivation, and learning.

Fisher and Frey (2018) designed a 12-week intervention to increase reading volume for six schools of odd-numbered grades by providing student choice of literature, access, classroom discussions, and book talks during school. They theorized an increase in books read would effectively snowball into an increase in oral language skills, spelling, reading comprehension, general knowledge, and interest over the course of a student's career. The residual impact of the study included an increase of library checkouts by close to 10%; higher writing scores on district benchmarks (4%) when compared to other districts; an increase (2%) of individual fluency rates; and a change in school morale around books, including discussions, anecdotes, and pedagogy. The

authors were not able to limit the study's effects to the odd-numbered grades, which compromised their findings. Overall implications of the study suggested a focus on student choice of literature, access, classroom discussions, and book talks produced tangible and intangible, positive outcomes for the schools targeted in this study.

Mohr (2006) explored the direct choices of first graders' (*N*=190) preferences among a set of nine, pre-selected, picture, fiction and non-fiction books, and informed students they were choosing a book for personal ownership. Among the sample, 64% (*N* = 122) of the students were interviewed to solicit a rationale for their selection process, an understanding of gender differences within the book choices, evidence of the preference for particular genres, and evidence of their preference for books that represented gender, first language, or racial identity. Results indicated 84% of the students preferred non-fiction, informational texts. Additional findings showed students made choices based on their interests in the topic and potential means to share texts with friends and family, rather than text difficulty or perceived teacher approval of their text.

Most research on choice is viewed as a variable of a larger construct (Deci & Ryan, 1985; Guthrie, et al., 1996; Guthrie & Wigfield, 2000; Flowerday, et al., 2004; Iyengar & Lepper, 2000; Schraw, et al., 1998). For example, some ways to interpret choice are through autonomy, perceived control, and as a residual factor of intrinsic motivation (Cordova & Lepper, 1996; Flowerday, et al., 2004; Iyengar & Lepper, 2000). Generating student autonomy in the classroom includes multiple opportunities for students to exercise their decision-making power through choice. Additionally, choice created a sense of ownership and increased perceived freedom (Iyengar & Lepper, 2000). Research from Cordova and Lepper (1996) demonstrated choice increases intrinsic

motivation, indicated by prolonged engagement in an activity on their own accord, i.e. personal interest, personal fulfillment, internal enjoyment.

Supplemental research found that students who were more internally motivated were more likely to demonstrate a concerted effort toward task completion (Ryan & Deci, 2000). Over time, an individual's choices become more self-fulfilling as one is inclined to spend more time and effort on tasks that can be enjoyed (Bandura & Schunk, 1981) and/or those that lead to additional opportunities. Given a limited choice, Iyengar and Lepper (2000) found people used their right to choose and performed better when given an opportunity. Given the specificity and individuality of children's needs, it is important to know the types of choices made around literacy to meet the larger goals of overall reading comprehension, competence, and achievement.

Statement of the Problem

Choice is an important factor in personal satisfaction leading to more intensified feelings of determination, motivation, and learning (Fraumeni-McBride, 2017). Research has shown the beneficial properties of providing students with choices in classroom instructional material. Current assessment structures are based on assigned readings, although research has confirmed that choice is a significant factor in improving reading comprehension scores (Fraumeni-McBride, 2017).

Standardized testing practices have not deviated from historical positivist perspectives, while curricular and pedagogical methods have morphed and been influenced by cognitive and social constructivist theories (Slomp, 2008). Scores from standardized tests are strongly correlated to family income, occupation, and education; school location; and school quality (Mulvey, 2009; Wildman, 2007), spurring questions

related to the fidelity of current tests and the need for alternative assessments for economically disadvantaged students (Wortham, 2005). However, the testing culture has not waned in its influence on instructional practices and perceptions of student and teacher success (Afflerbach, 2002, 2017).

Sanacore (2002) suggested state mandated exams were inherently harmful to the literacy lifestyle crusade. He drew specific attention to the lack of affective measurements by these exams; the alignment of the exam to state standards rather than student needs; and the narrowed focus on student testing achievement over other measures of student growth (Sanacore, 2002). Despite the influences and biases generated from tests, standardized testing has been used as a barometer to assess reading comprehension skills.

A search on the provision of choice in high-stakes assessments yielded one study conducted by Campbell and Donahue in *The NAEP Reader* (1997). Developed from the idea that student self-selection of texts had positive effects on engagement and hopeful about the prospect of funneling student reading interests into a more permanent enterprise, the study identified the effects of choice on student performance in an assessment of reading comprehension. While the findings revealed no significant differences between the two groups of assessed subjects (choice vs. non-choice), differences were reported in student perceptions of the assessment and other affective measures. Given the limited information on this topic, there is a need for more research related to the permeable nature of choice, as well as, the role of choice in standardized reading assessments, reading comprehension and motivation.

This study addressed three areas of need: the relationship of choice in relation to reading comprehension; the impact of choice on standardized reading comprehension tests; and, the conditions of choice within the educational structure. As mentioned, there is limited research on the impact of choice on standardized reading comprehension assessments. Presently, a lack of choice in curriculum and instructional methods have led to a narrowed focus on testing strategies, which suppresses student engagement (Sanacore, 2002). New information about effective ways to implement choice in the classroom could stir student engagement and motivation. Results of the current study may potentially inform educators, test administrators and related district personnel on alternative means of sustaining student interest, motivation and engagement through choice under testing conditions.

Purpose of the Study

The purpose of this study was to evaluate how measures of choice ('teacher-selected' and 'student-selected') compared as predictors of reading comprehension and how these vary as a function of achievement for fourth and fifth-grade students. Additionally, the study quantitatively examined the relationship between measures of choice and the standardized reading comprehension assessment with fourth and fifth grade students. Lastly, the study examined the relationship between interest and reading, secondary to the provision of choice.

Significance of the Study

Success in reading is important and predicated on enjoyment in reading (Fraumeni-McBride, 2017); therefore, it is important to focus research on ways to improve these areas. Data from standardized tests have shown either stagnation and or a

decline in the achievement of students across multiple grade levels, nationally (e.g. NAEP). More research to uncover new approaches to improve literacy outcomes is important (Fraumeni-McBride, 2017). Research on student choice and standardized testing is over 25 years old. This study added to the field by focusing on non-traditional means of offering choice in the classroom.

In addition, the study explored choice as a predictor of reading comprehension on standardized assessments for fourth and fifth-grade students. Research has indicated that adolescents' academic motivation, including reading, is liable to digress over time (Hidi & Harackiewicz, 2000) due to shifts in attitudes, access, peer/social groups, and maturation. This research on the provision of choice contributed to the field of information on the decline in academic motivation for adolescents. Research on the impact of choice on reading comprehension assessments has not included these grade levels. The results may be beneficial to educators, test developers, and test administrators interested in ways to increase student productivity, increase motivation, and yield more accurate performance outcomes on standardized reading assessments.

Theoretical Framework

The current study relied on several theories to explain the processes that have influenced student learning and reading behaviors. First, Piaget (1948/1973) wrote extensively about the development of the child across multiple stages of functional and academic learning. One key component of his work was his conceptualization of autonomy. Intellectual autonomy, as described by Piaget, is the ability to independently manage, regulate, and coordinate individual decisions. Autonomy is developed through learned experiences, wherein a child may initially follow the directives of an adult, they

should, eventually, develop their own vision of the world. Over time, the child is less reliant on an outsider to lead them because of exposure to the same if not different outcomes of a task. Eventually, the child can choose and reconstruct an experience, based on their own preferences and the conditions they find desirable and enjoyable. The option of choice is an expression of their autonomy. The choice construction is developed from knowledge built from learned experiences and reflections on previously made decisions.

Another aspect of intellectual autonomy is the activation of knowledge as an outgrowth of human curiosity. Following the accumulation of knowledge stores, persons form schemas or catalogs of information and actively collect more data to these existing schemas to construct more complex concepts and thoughts (Fitzpatrick & Muelemans, 2011; Rumelhart, 1981; Tracey & Morrow, 2012), fueled by their own choices and interests. Schemata are individualized, abstract, knowledge structures created through experience, perceptions, and interpretations of lived and learned concepts (Anderson & Pearson, 1984; Tracey & Morrow, 2012). Schema Theory is a constructive theory established to explain how persons sift and organize information into chunks of summarized understandings (Anderson & Pearson, 1984; Rumelhart, 1981; Tracey & Morrow, 2012). In relation to a reading comprehension event and learning experience, schema expand and contract, as readers interpret, reject, modify, and analyze new information with that of previously established schema, dependent on lived experiences (Anderson & Pearson, 1984; Rosenblatt, 1978, 1994). Rosenblatt (1978, 1994) added the unique contributions of individualized schema results in different reader responses to literacy. A veteran researcher would have a more in-depth article critique than that of a

novice researcher, for example. The convergence of all these ideas occurs at the point of choice and personal connection. As mentioned, a person instinctively accumulates knowledge and, when given the choice, will seek information that is useful or personal to that individual, which is nuanced by their comprehension of this information.

According to Dewey (1990), learning drives student choice. He wrote, "the primary root of all educative activity is the instinctive impulsive attitudes and activities of the child, and not the presentation and application of external material" (p. 117). Dewey reasoned that children were naturally curious and that the school would provide the platform to act out their inquiries. He envisioned the role of teachers as facilitators rather than directors for children to imitate. Further, children create activities, while teachers are guides to assist with concept mastery. Conversely, an overreliance on direct explicit teaching or other external source is a loss of control, which runs contrary to the purpose of education and learning. A child's curiosity should direct their attention on a given subject or course of study. Their attraction to the subject matter leads to the creation of their own ideas. Rosenblatt (1978, 1994) ascribed to Dewey's teachings and conceptualized a framework for this natural approach to reading. She described the reading process as cyclical transactions of reading and reacting to words. Readers seek out information (text) that is known and habitual, as a choice. Control and power become the outgrowth of these experiences for the child and comprehension increases through stronger text connections (Tracey & Morrow, 2012).

Conversely, constraints on choice result in a reduction of knowledge, confidence, and critical reasoning abilities (Kamii, 1991). Kamii clarified Piaget's writings on autonomy suggesting that autonomy is the ability to make an intellectual decision, void of

external stimuli. Using this operational definition of autonomy, Kamii challenged teachers who may be inclined to apply restrictive or compliance-based practices in their classrooms. She argued that children in these restrictive environments are stifled in their moral and intellectual growth.

Similarly, self-determination theory (SDT) presented by Ryan and Deci (2000) describes how environmental conditions promote, balance, or suppress a person's instinct to pursue their own curiosities. This theory established three individual, psychological needs for the basis of self-motivation and self-preservation, which included the need for autonomy. Aligned with Piaget's ideas regarding the natural curiosity of children, Deci and Ryan defined the constructs of intrinsic motivation within SDT to explain the innate principles of competence and autonomy. They posited that in the absence of autonomy, feelings of competence have no effect on intrinsic motivation. In other words, without self-determination a person will not feel supported in his endeavors, even if presented with external rewards. Ultimately, they suggested that through choice, emotional awareness, and opportunities for independent expression, intrinsic motivation is advanced.

Practical applications of SDT in the classroom include a hierarchy of autonomy (organizational, procedural, and cognitive) support, which yield different learning outcomes (Patall et al., 2010; Stefanou et al., 2004). Self-determination and autonomy are cultivated through the acknowledgement and encouragement of student choice (Kohn, 1993; Sanacore, 2002) of texts, leading to deeper comprehension and increased engagement. Later this engagement is sustained through intrinsic motivation, the ideal form of autonomous motivation (Patall, et al., 2010). When teachers create autonomous

environments, Stefanou et al. theorized that students become empowered, motivated, and entrenched in their learning.

The theory of self-efficacy (Bandura, 1977) is connected to self-determination theory and motivation. Self-efficacy is the belief in one's own potential to successfully complete a task or goal. Bandura's self-efficacy theory emphasized the impact that self-efficacy has on learning; specifically, when children monitor their learning progress, their self-efficacy increases. Self-monitoring aids students to gain information independently, while self-efficacy comes with familiarity, interest, and confidence. In addition, a stronger belief is associated with more devotion and persistence toward a task. Overtime, an individual's choices become more self-fulfilling as one is inclined to spend more time and effort on tasks that can be enjoyed (Bandura & Schunk, 1981) and/or those that lead to additional opportunities.

Choice is tantamount to individual growth. These theorists agreed that along a developmental continuum, children have an innate curiosity to explore and discover natural consequences through decision-making. Choices made are an expression of a desired outcome. Students should have the freedom to choose, even under contrived conditions, as these events contribute to their learning and understanding of their environment and actions. Opportunities to increase self-direction and self-determination are through progressive choices, based on various goals, interests, and motivation (Evans & Boucher, 2015). The conversion of choice through acts of autonomy funneled through learning and instructional design has led to more engaged students, who have a more active role in the classroom (Evans & Boucher, 2015). This study examined the impact of choice on standardized reading comprehension assessments. The established ideas

surrounding choice, reading comprehension, reading motivation, and student growth through autonomy generated from these theories and theoretical framework guided the research procedures, analysis, and interpretations of the findings. The following questions led the research:

Research Questions

1. Are there significant differences in reading comprehension assessment scores when reading 'teacher-selected' passages versus reading 'student-selected' passages?
2. Is there a relationship between interest and performance on standardized reading comprehension assessments?

Delimitations of the Study

One delimitation of the study was the comparability of readings. The collection of stories covered a variety of topics and were selected from released STAAR assessments deemed appropriate at each grade level. The length of the four stories ranged from 524 to 730 words at grade 4 and 557 to 834 words at grade 5. The study could not rule out the possibility that 'student-selected' passages were chosen based on peer pressure and/or were motivated by the desire to not appear different among classmates (NAEP, 1997).

This study focused on the impact of choice on individual reading performance for fourth and fifth grade students. Importantly, the students had to be able to read independently. Another delimitation of the study was the exclusion of students identified with a disability that affects reading. According to IDEA 2004, a specific learning disability (SLD) is:

> a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which disorder may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations. Such a term includes such conditions as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. Such a term does not include a learning problem that is primarily the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage (IDEA 2004, Definitions).

The implication is a SLD would impair a student's ability to complete some instructional tasks (Flanagan & Alfonso, 2011) and in this case reading. There are a variety of skill deficits that could explain the problems of these readers, including: decoding, fluency, inferencing, and applying self-regulation (de Milliano et al., 2014). It is also plausible that the student could exhibit struggles with comprehension, which were independent of decoding and vocabulary skills (Seipel et al., 2017). Others are less efficient in their strategy usage or lack the reasoning experience to aid in comprehension (de Milliano, et al., 2017). Seipel et al. (2017) speculated that poor comprehenders struggle with comprehension because of difficulties with making causal, coherent inferences or connections between events in a story or passage. Efficiency with reading comprehension-based tasks, including knowledge of when and how to use text resources and features is another area of difficulty for this group (Vidal-Abarca et al., 2010).

In summary, the profile of a low-achieving reader is complicated and complex. Readers who have found phonological awareness, automaticity, orthographical and phonological representation, and/or a combination of basic reading skills difficult may struggle to read the material contained in the STAAR passages without assistance. Therefore, students receiving special education services, including those diagnosed with a learning disability (e.g., dyslexia, specific learning disability) were not included in the

sample in order to control for cognitive variables. For this study, it was assumed that students with reading-based learning disabilities were identified by the teacher and, therefore, did not participate in this study.

Definition of Terms

1. Reading comprehension: "the process of simultaneously extracting and constructing meaning through interaction and involvement with written language" Research and Development Reading Study Group (RRSG, p.11, 2002).
2. Choice: the act of selecting or deciding when faced with two or more possibilities, without external pressure or reward (Haworth, 1986).
3. Perceived choice: An interpersonal environment that allows individuals decision-making flexibility and opportunities to choose among options (Flowerday, et al., 2004).
4. Autonomy: the ability to think for oneself independent of reward and punishment, and to decide between right and wrong, and between truth and untruth. The ability to govern oneself. (Kamii, 1991; Piaget, 1948/1973).
5. High stakes testing: many refer to standardized assessments as *high-stakes tests* as they "carry serious consequences for students or for educators" (American Educational Research Association [AERA], 2000, p. 24).

Organization of the Study

This study includes five chapters. Chapter I introduces the background, significance, and objectives of the study. Chapter II is the Literature Review, which has three major parts. In the first section, the construct of choice is reviewed. Most of the studies featured in this section have explored choice through interpretations of self-

determination theory (SDT) (Deci & Ryan, 1985). The second section, provides a historical overview including theoretical models of reading comprehension, structural components of reading comprehension, and assessment in reading comprehension. The third section covers the history of assessment, the role of assessment in curriculum and instruction, and the use of choice in high-stakes assessment. The methodology is explained in Chapter III. Critical components of this chapter include a description of the purpose, research design, participants, subjects, procedures, instrumentation, data collection, and data analysis. The results and an analysis are in Chapter IV, followed by a discussion of the findings and their practical significance in Chapter V.

CHAPTER II

Literature Review

The National Reading Panel (NRP, 2000) concluded in their analysis of reading research that effective reading instruction should be comprised of explicit instruction in phonemic awareness, systematic phonics instruction, and strategies for reading fluency and comprehension. Comprehension of written words and materials is one of the five parts of reading; other parts being phonemic awareness, vocabulary, fluency, and phonics. The Research and Development Reading Study Group (RRSG, 2002) expanded on reading comprehension as an application and process. The group described comprehension as "the process of simultaneously extracting and constructing meaning through interaction and involvement with written language" (RRSG, 2002, p. 11). The Texas Education Agency (TEA) characterized this process as the manipulation of skills, including letter-sound awareness, knowledge, and correspondence (TEA, 2017). By this definition, when a student can read, they have demonstrated their ability to decipher the codes embedded within the English written language (Aronesty, 2015).

Mastery of comprehension is relative to the type of material being read, the individual reading at a coherent and functional rate (fluency), the maintenance and retention (memory) of vocabulary, and knowledge to navigate the passage (Young & Rasinski, 2017). Beyond a complex piece of literature, the reader must be a literate, complex thinker to decode the message being conveyed (Young & Mohr, 2018). To foster reading comprehension requires more than a crafted piece of literature and cognitive skills, students must, also, be actively engaged in the reading and learning process (Lin, 2015; Stokes & Young, 2018).

Nejadihassan and Arabmofrad (2016) inferred that an effective means of improving reading comprehension and academic life was through affective variables e.g., choice, motivation, and interest. Throughout the remainder of this review, the relationship between reading comprehension and choice as an affective variable of academic achievement will be magnified. Choice is generally positioned as a behavioral response, based on intrinsic motivations (Deci & Ryan, 1987). A review of literature for this study gleaned a limited pool of research, which directly supported the provision of choice in the classroom, as a determinant for academic achievement and reading comprehension. From the research that was uncovered, choice was defined and applied most frequently through the application of SDT, which created a focal point for the review.

In the interest of creating a foundational understanding of choice, it is discussed in general terms and followed by its applications across genres and content areas. Specifically, the focal point of the research that contributed to this review were limited to studies that intersected choice and education through applications of autonomy, motivation, interest and reading engagement. This section is followed by an explanation of theoretical models of reading comprehension to be considered within the context of achievement and performance. Next, there is a discussion about assessments, beginning with a brief historical overview and ending with a convergence of choice and assessment, in relation to achievement and performance. It should be noted historical content used within this review served to provide foundational references and guiding points, therefore, these sections are limited in scope and depth. Broadly, a general timeline is

presented as a guide and outline, which has led to present trends in education. Lastly, the purpose and significance of the study are recalled which finalized the review of literature.

Choice

Choice is the act of deciding between two or more prospects, which leads to a given outcome (Haworth, 1986). Trivial and non-trivial choices are presented throughout routine and non-routine practices (Stefanou et al., 2004). In a democratic society, people are reminded of their on-going right to choose, express their desires through choice, and the direct and indirect consequences linked with matters of choice (Wehmeyer, et al., 2017). Choices are offered with such frequency that an inherent value has been placed on this decision (Iyengar & Lepper, 2000). Conversely, when choices are deemed to be stifled or non-existent for a given event or task, it is viewed as an infringement of rights and incitement to stir fear or anger (Patall et al., 2008). In both cases, the value and provision of choice elicited feelings of personal responsibility and directed human behavior through differing positive and negative connotations (Patall et al., 2008). More notably, the choice was influenced by confounding factors (motivation, interest, autonomy, and prior knowledge) of the individual (Deci & Ryan, 1985; Ryan & Deci, 2000; Patall et al., 2008).

A brief history of choice. A choice is an action and actions are observable and habitual; they are thoughts and behaviors made visible with each individual exchange. Glasser (1998) formulated Choice Theory as an applicable theory to everyday life, which explained why and how people make decisions. Within the context of choice theory, behavior is described as the culmination of a person's actions, thoughts, feelings, and physiological manifestations that cycle into a choice. According to this theory all

behaviors are chosen based on internal motivations. In fact, Glasser does not consider external stimuli a component within the decision-making process. Even when an action is forced, the person made a choice, therefore, the assumption of responsibility remains that of the decision maker. Through choice theory, Glasser rejected the theory of external control and embraced an internal control psychology. This ideology of choice theory replaces the reactionary mindset of people who feel helpless, vulnerable, and powerless with that of freedom, control, and self-regulation. While Glasser's theory of choice has been used predominately to treat individuals therapeutically, such as reality therapy, he also has demonstrated the impact of choice theory in the educational setting.

In the school setting, Glasser argued students are not fully invested in education because they are being coerced into completing tasks that hold little to no value to the student. Students are not in control of what is taught, which has left them disengaged and unmotivated to produce their best work (Glasser, 1998). Structurally, the school has not provided choice, supported autonomy, or encouraged free thinking within school activities and environmental interactions. Instead, students are coerced into learning and regurgitating information that is not valued by the students. Glasser suggested education should be about information manipulation, rather than acquisition. Like Vygotsky, Glasser envisioned an operational educational environment used to facilitate learning and sustain social transactions (Brown & Dryden, 2004).

Wehmeyer et al. (2017) referred to self-determination as a construct by which people take actions, manifested from their own will or desire. They made a choice determined by their own mind, or free from restraint. These acts of choice are independent and autonomous in nature. Philosophers have debated the idea of choice as

an extension of self-determination, mainly as an internal battle between the self and a higher pressure or authority. For example, Locke argued that choices were based on life experiences, thereby placing actions and thoughts on a developmentally based continuum, rather than an innate action predicated by a higher power. Moreover, the freedom of expression through choice is a personal decision, justified and defensible as a right and a responsibility governed by human nature (Wehmeyer, et al., 2017).

Deci and Ryan (1985) documented the history of choice from behaviorists and other theorists, who claimed that behavior could be predicted from learned experiences and outcomes. The assertion was that behavior was modified through external stimuli and reinforced in the same manner, and learning was based on either positive or negative realities that stemmed from those contacts. Cognitive theorists in the 1950s and 1960s suggested an alternative to behaviorists' theory, concluding that information processing was determined by behavior, which was a function of possible future outcomes. Combined with cognitive competence, a person had to believe in the causal relation between their behavior and the outcomes of their decisions, which fueled the foreseeable results (Deci & Ryan, 1985). Behavior was repositioned as a relative state of being, predicated on how to navigate a state of awareness. It became an expression of a more forward-thinking person, able to gauge potential outcomes, based on their own ability to regulate their environment, a standard, or other personal status. The decision to manipulate their outcomes further demonstrated the person's intention and plan of action. The deliberation and purposefulness from which the decisions are made is inherent in the choices made to yield certain outcomes.

Application of choice. Studies have used definitive pieces of choice to assess its effects under various circumstances, mainly either as a reinforced or controlled variable. Iyengar and Lepper (2000) completed a series of studies on consumers and undergraduate students which questioned the provisions of choice and the conditional circumstances of the presentation of choice. In the first study, they tested the limits of choice by presenting consumers with two options: an extensive-choice (24 different jams) condition and a limited-choice (6 different jams) condition. Despite the array of choices, both groups sampled, on average, the same amount of jams. The researchers speculated that consumers may have either been under personal time constraints, equated the limited choice assortment to a higher quality grade of jam, or became disinterested when presented with so many choices. Findings indicated that choice may lead to saturation and a decrease in motivation.

Intrigued by this discovery the researchers explored the limited-choice condition with undergraduate students (N = 197) in a yoked design, where the limited choice-set was rotated with an extensive-choice set (Iyengar & Lepper, 2000). After showing a film in class, students were given either 6 or 30 potential essay topics to respond to, in writing, for extra credit. Here, again, the results demonstrated that limited choice was more favorable than choice overload. Indirectly, the extra credit writing assignment performance was reflective of intrinsic motivation, as students' motivation stemmed from their choice and drive to gain extra points.

As a follow-up to these results, Iyengar and Lepper (2000) conducted a final experiment to examine chooser satisfaction and understand whether previous conditions were based on a concession of choice or an optimal decision. In this study university

students (N = 134) were randomly assigned to one of three conditions: limited-choice, extensive-choice, and no-choice. This third condition added depth to the inquiry of whether changes in motivation and satisfaction as a result of choice. The findings were insightful. Persons in the extensive choice group reported more frustration and dissatisfaction with the decision-making process than the other groups, because of the overwhelming prospects of deciding and the pressures associated with finding the best choice. The researchers concluded that too many choices might have a diminishing return on motivation and satisfaction. The arbitrary and trivial nature of the choice scenarios may have, further, compounded the results.

Katz and Assor (2007) asserted that many of the studies on choice have misinterpreted SDT, resulting in mixed or inconsistent interpretations of choice. In their clarification, they noted that SDT has three tenets: the need for autonomy, relatedness, and competence. While autonomy is stressed and applied more reflexively as people interpret choice, Katz and Assor argued that all three provisions must be satisfied before the benefits of choice are actualized. The need for personal relevance should not be overlooked, in any case. Choice must be an expression of the self, rather than merely an opportunity to choose, without a personal connection.

According to the self-regulatory perspective, Patall et al. (2008) noted acts of control stem from the same family as self-regulation and volition, and, therefore, are generated from a centralized source. Consequently, a multitude of decisions are generated by a centralized source, which could result in fatigue and overload (Patall et al., 2008). The person's ability to make a choice or self-regulate would and could be diminished or depleted under certain circumstances. Moller et al. (2006) challenged this

theory in a series of experiments and found that making a choice is not always depleting. Instead, the conditions by which the choice was provided would impact their levels of depletion. Therefore, autonomous choice does not result in the diminishment of regulation or vitality and energy (Moller et al., 2006).

These qualifiers of choice include several themes: free will, self-regulation, forethought, and the recognition of human reasoning. There is a personal fulfillment component attached to each decision that is simultaneously consequential and liberating for an individual. Failure to recognize these attributes in a study's design effects whether the impact of choice will be minimized or masked. The next section highlighted studies that interpreted choice through constructs of autonomy, interest, motivation and reading engagement.

Choice as autonomy. Deci and Ryan (1987) considered choice as either decided under pressure (coerced) or decided with free will and with an exertion of control. They described the latter as an expression of autonomy and regulation (Deci & Ryan, 1987). Regulation connotes the absence of constraint. Autonomy is viewed as a ruling made by oneself and rooted in the self to govern the result (Katz & Assor, 2007). It is contingent upon the level of personal value and relevance appointed to an activity e.g., personal, functional, educational, practical or work-related. Reasonably, autonomy increases as alignment of personal desires, interests, and needs are fulfilled by the task (Katz & Assor, 2007). Deci and Ryan explained that an autonomous action or choice is grounded in the assumptions of personal experience, desired effect, cognitive conceptualization, and regarded benefit. Moreover, the decided person is self-determined. Moller et al. (2006) found self-determination to be a regulating factor

between controlled versus autonomous choices. A self-determined person is resolute and assured in their choices, because they made the choice by their own accord and volition.

Skinner et al. (1990) designed an experiment to test whether a perception of control would carry over into performance on cognitive tasks. They reasoned that when a student has the assurance that they play a part in their success through their own decision-making and witness this level of success, academic outcomes become more contingent on the self rather than chance. In the Skinner et al. study, a process model was generated, assessed and analyzed to determine the direct and indirect relationship between children's perceived control, teacher involvement, and academic performance, which was measured by grade and achievement. Findings revealed that teacher behavior influenced children's perceived control which, in turn, affected engagement and academic performance. Critically, the student perceptions of classroom interactions were linked to extended acts of motivation (Stefanou et al., 2004). These results underscore previous work on the contributions of self-efficacy (Bandura, 1977), expectancy-value model (Wigfield & Eccles, 2000), self-regulation theory (Zimmerman, 1989, 2002) and SDT (Deci & Ryan, 1987) toward the academic outcomes of students. The Skinner study, also, demonstrated the impact of autonomy and the necessity of opportunities for students to exercise control in their learning environments.

Stefanou, et al. (2004) concurred with Skinner and her colleagues. In recognition of the teacher's role within student engagement, Stefanou et al. provided more context for teachers to support student autonomy using the pillars of SDT. Included in their research are the tenets of basic needs required for learning: competence, relatedness, and autonomy (Deci & Ryan, 1987) which they claim are upheld, diluted or dissolved within

the classroom setting. In this instance, competence is defined as the level of comprehension of schoolwork; relatedness is equated to belonging; and autonomy is relayed through levels of control over classroom activities, including the attached boundaries, policies, and procedures of the task. This approach assumes that teachers take a facilitative role in the classroom with the goal of supporting students on their journey toward autonomy, self-determination, interest, and an incarnation of school values. The inverse of which would spurn low achievement, uncertainty, low self-efficacy, and a predilection for easy work (Stefanou et al., 2004). While previous researchers had used the tenets of learning to translate choice into practice, Stefanou and her colleagues found that many practitioners were using diluted adaptations of choice. Choice must hold some value with students in order to be impactful toward learning. Both researchers and teachers, alike, have misinterpreted the interplay of choice on both external and internal constructs of the classroom, which they argue has tainted the results of choice provisions (Stefanou et al., 2004).

In their reinterpretation of choice practices, Stefanou and her colleagues envisioned more inclusive practices that are satisfied through authentic student choices. More importantly, they argued these choice extensions should lay a foundation for consistent student motivation and engagement habits. The propositions amount to a continuum of activities which they consider autonomy-supportive and is the gauge for teacher supported autonomy practices. Specifically, when teachers extend provisional choices related to the environment; they promote organizational autonomy support. At this level, students are given the opportunity to collaborate on classroom rules, select due dates for assignments, and choose their group members. These are considered lower

level affordances, but they still play a role increasing autonomy support within the classroom. Student ownership is spurred from choice in materials, presentation of ideas, and policies, at the procedural level of autonomy. Stefanou et al. petitioned teachers to balance their behaviors and shift ideas about control and choice toward the central matter of instructional outcomes. When the goal is learning, procedural and organizational autonomy are supports are not sustainable for student motivation, learning partnerships, critical thinking, self-regulation, and academic engagement (Stefanou et al., 2004).

On the high end of the continuum, cognitive autonomy is achieved when students have the flexibility to discuss, debate and justify strategies and solutions to problems; formulate personal goals; ask questions; and realign tasks to their personal interests. Cognitive autonomy is transformative. At this pinnacle level of autonomy support students are provided opportunities to build intrinsic motivation, engage in collaborative instructional practices, and make choices related to educational outcomes. These inclusive practices by the teacher and student connections lead to deeper learning and educationally relevant changes. Cognitive autonomy is the catalyst for student ownership and investment within the learning environment (Stefanou et al., 2004) and, therefore, carries more influence than organizational and procedural autonomy.

The provision of choice is not a one-size-fits-all construct. McKool (2007) surveyed and interviewed fifth graders (N = 199) from different socio-economic backgrounds and cross-sectioned them into groups of avid and reluctant readers. Avid readers from this study reported choice in reading material and in-class reading time influenced their reading habits, both in- and out-of-school (McKool, 2007). When asked how teachers could promote voluntary reading and increase motivation among students,

these avid readers recommended teachers offer more time for self-selected reading and endorse more specific book titles.

Teachers can access multiple points of entry to the choice continuum, but cognitive autonomy would be the most beneficial to all members of the classroom. Stefanou et al. created a playbook for the SDT framework and envisioned a systemic means of choice to foster student engagement, interest, and motivation. The operationalized definition of cognitive autonomy is applied through the selection of reading materials, grades, and project-based learning. In keeping with the theoretical definition, students in this study were provided a choice in assessment, which is related to an educational outcome. Likewise, this study used the same principles of SDT to examine how an autonomous choice would impact a student's performance on a high-stakes, standardized reading assessment.

Choice as interest. Interest is viewed as a state and a disposition, and, therefore, has cognitive and affective components (Hidi & Harackiewicz, 2000). The construct of interest has been researched through two entities: individual or personal and situational. Individual interest is defined by a person's willingness to engage in a topic, object, activity or text (Renninger, 1990; Schiefele, 2009; Springer et al., 2017; Worthy et al., 2002). It is established as a constant within the context of motivation and is associated with increased knowledge, value and positive feelings toward a topic or domain (Hidi & Harackiewicz, 2000; Renninger, 1998; Schiefele, 1991). Situational interest is more reliant on environmental factors, intermittent, and is less stable than individual interest (Hidi & Harackiewicz, 2000). Given the conditional nature of situational interest, it is more representative of an affective reaction to an event. Despite

these differences, research has determined that all types of interest tend to facilitate reader's comprehension and recall (Hidi, 2001). Interest in reading predicates students' reading comprehension (Gambrell, 2011; PISA/OCEP, 2003) and is critical to learning (Dewey, 1990; Dobrow et al., 2011)

A choice is an expression of preference or interest between two or more items. In the classroom, opportunities to express your interest need to be fostered, which has proven a challenge. Dobrow et al. (2011) hypothesized that exercising choice over grades would positively influence student's triggered situational interest and maintain this interest over the duration of a college course. To test this hypothesis, they conducted an experimental field study with four undergraduate business management classes from two universities). The study was designed as a crossover, where one group from each campus served as a control, while the other was designated as the experimental group. The experimental groups (choice) could apportion the weights of three class assignments (up to 75%) toward their final grade, while the control group was told the same three assignments would account for 25% of their grade. Supplemental data points were collected from an end-of-year survey which included measures for triggered and maintained situational interests, course evaluation, the registrar, and four manipulation checks for the choice intervention.

Students from the Dobrow et al. control condition demonstrated feeling more autonomous than their counterparts, in relation to the grading procedures. Analysis of the manipulation checks revealed that 67% would be in favor of implementing the grade allocation in other courses; 50% of the choice group responded that the grade allocation greatly impacted the amount of time they spent on an assignment, 34% reported

sometimes, and 16% reported that it had no effect. Data from a two-sample t test and multiple regression analyses (controlling for gender and university) supported their hypothesis showing that exercising choice was associated with an interest level 0.5 points ($p < .05$) higher than not exercising choice. Additional support for their first hypothesis was relayed through the end-of-course survey which indicated that 83% of the 36-item sign test were rated more positively in the choice condition than in the control condition, where equal preference would have been indicated by 50% of the items (Dobrow et al., 2011).

Dobrow and her colleagues reported that both conditions created high levels of maintained situational interest. The maintenance of situational interest was supported when students in the choice group indicated (0.7 points; $p < .05$ on a 7-point scale) that they would be interested in future classes that provided this same opportunity, which increased to 0.8 points when a multiple regression analyses was conducted to control for gender and university. Their second hypothesis was upheld with these results. The researchers were able to foster interest through a grade-related choice intervention (Dobrow et al., 2011). The findings substantiated research on the positive relationship between choice and interest. In this case, the students devoted more time and effort to their assignments, were involved in the learning outcomes, and maintained a level of control over the management of grades. The buy-in of choice over the grade allotment fueled and sustained their interest over the course.

Flowerday and Schraw (2001) contend that the effect of choice on learning is mainly mediated by interest and several studies have been conducted pursuant to this principle. For example, Flowerday and Shell (2015) hypothesized that choice does not,

singularly, influence learning outcomes; rather choice, topic interest, and situational interest. In addition, aligned with engagement and self-regulation theories, the researchers suggested that engagement should influence learning outcomes, as well (Flowerday & Shell, 2015). Their theoretical model was assessed using path analysis, based on data from 90 randomly assigned undergraduate students placed into three conditions: Choice-High Interest, No-Choice High Interest, or No Choice-Low Interest (Flowerday & Shell, 2015). Students participated in a topic interest survey and were paired with a corresponding reading task, based on this survey and their respective, assigned groups. Next, they read a passage and completed corresponding comprehension pieces (multiple-choice recall test and reader response writing activity), followed by a situational interest inventory, and an attitude checklist. Throughout the reading and writing, the students' behaviors were observed and time intervals were recorded, as a representation of engagement. Results showed that choice only had a significant direct effect on attitude and no direct effect on situational interest.

Overall, situational and topic interest both directly and indirectly influenced all outcome and engagement measures (Flowerday & Shell, 2015). The researchers found that the results were inconsistent with SDT in relation to choice. In a related series of studies, Flowerday & Schraw (2003) found similar results, where choice had a positive effect on attitude, affective engagement and effort and no positive effect on cognitive engagement. Schraw et al. (1998) explained that this may be due more to how cognition is measured, while Patall et al. (2008) surmised that cognitive engagement is best measured overtime. Still, Flowerday and Shell had hoped that choice would increase situational interest by enhancing feelings of autonomy; however, autonomy was not

measured in this study, which presented a flaw in the design. Notably, while they set out to separate choice and interest, the researchers conceded that they were not able to truly separate choice from interest (Flowerday & Shell, 2015).

Similar studies have attempted to separate choice and interest (D'Ailly (2004); Flowerday, et al., 2004; Wilde et al., 2018). Flowerday et al. (2004) isolated choice from interest in a series of experimental studies using undergraduate students to examine the separate effects of choice and interest on measures of learning, engagement and attitude. Results gained were like previous studies on choice and topic interest regarding cognitive engagement (Schraw et al., 1998), as choice, topic interest, and situational interest appeared to have no effect on the multiple-choice test of facts. The authors credited the subjects (college students) for being advanced in their reading acumen, which mediated any impact that choice, topic interest or situational interest could have had on cognitive engagement. The context of the activity may also have played a role in the outcomes of the study. The students could very well have been motivated by the pretense that the research could be used to fulfill a class research requirement, albeit their participation was voluntary. The authors noted, in isolation, situational interest positively affected affective engagement. Essentially, when situational interest was established, topic interest was no longer a factor. The authors concluded that situational interest in the text influences engagement, rather than choice (Flowerday et al., 2004). The findings upheld their hypotheses about choice and concluded that findings from previous research were skewed by the effects of interest.

Wilde et al. (2018) attempted to separate choice and interest as confounding factors when they analyzed the effect of choice on students' quality of motivation within

a science lesson. Researchers created choices for their participants and implemented a value option to distinguish the effects of choosing and being paired or matched with either a preferred or non-preferred choice. Students in the no choosing and match group expressed the highest intrinsic motivation, followed by students in the choosing and match group, while the no choosing and no match group showed the lowest intrinsic motivation. In relation to the quality of motivation, the researchers concluded that being paired with your preference was more important than the act of choosing. Data determined that interest played a more influential role in satisfaction and motivation. However, the researchers' study design replaced autonomy with an illusion of choice; that is, the participants were not actually paired with their choice, which may have derailed their motivation.

D'Ailly (2004) questioned the applicability of SDT across cultural divisions. In the exploration of student and personal choice, this researcher wondered whether the provision of choice was reinforced as a North American phenomenon, rather than a universally applicable theory. This study proposed to test the impact of choice on children's learning to examine students' interest, effort, and performance on a learning task and to investigate how cultural and gender factors may interact with the choice effect. Children from Canada (n = 130) and Taiwan (n = 153) were randomly assigned to a computerized foreign language learning program under four conditions: self-choice, teacher-choice, computer-choice, and no-choice control group. The tasks for all groups was to learn the names of animals (baseline), colors (provision of choice), and numbers (carry-over effect of choice) in a foreign language. In the self-choice group, the students were able to choose eight colors they wanted to learn of a possible twelve colors, while

the choices for other groups varied, respective to their assigned group. The students were assessed on their acquisition of colors and told that they would not be assessed on the animal or number tasks.

Data collected for all students in the D'Ailly study included student-reported levels of self-efficacy, interest, and time spent on the latter two tasks. The results did not support SDT, as evidenced by the manipulation of the choice conditions, when compared to the reported levels of interest and expenditures on the tasks. The author gave several explanations for these results, including how autonomy was measured by the study and perceived by the participants and the high level of intrinsic interest inherent in a computer-based task. The author had not anticipated the confounding impact of interest on choice and had not accounted for these effects in the design (Flowerday & Schraw, 2001). In relation to the researchers' hypothesis regarding SDT as exclusive to North American culture, the results were more aligned. Interestingly, the data indicated that Chinese students were more comfortable with an authority making choices, were more adept to the frequency of testing in the classroom setting compared to their counterparts, and displayed a different work ethic, relative to the pressures of being tested on a task. Lastly, the data reaffirmed the importance of relevance in the provision of choice as a factor in learning (Katz & Assor, 2007).

Choice and interest have been debated, in relation to the impact of learning. Separately, research has established the role of interest for readers of all types to overcome deficits in individual reading-based skills. Interest in reading, for example, predicated students' reading comprehension (Gambrell, 2011; PISA/OCEP, 2003). Accordingly, interest is critical to learning (Dewey, 1990; Dobrow et al. 2011). Interest

has been important to a reader's competence, confidence, and commitment to a piece of literature, as evidence by this analysis of research. In line with this, the present study sought to uncover the role of choice to express stated interests. With the prospect of interesting material and the affordance of choice, students may be inclined to exceed their normative expectations for achievement with a reading comprehension assessment.

***Choice as motivation*.** Research on motivation has explored two divisions, intrinsic motivation and extrinsic motivation. A person is said to be intrinsically motivated when they are engaged in an activity on their own accord, i.e. personal interest, personal fulfillment, internal enjoyment. Intrinsic motivation is maintained only when an individual is interested in the activity, behavior, or domain (Eccles & Wigfield, 2002). By contrast, extrinsic motivation is fueled by a reward or the fulfillment of another individual's want or need.

Ryan and Deci (2000) posited that intrinsic motivation is advanced through a series of transformations or *internalizations*. In their self-determination theory, they argued that a human's natural progression toward basic competence propels them to seek out stimulation and challenges through activities (Deci & Ryan, 1985; Eccles & Wigfield, 2002). In addition, they proposed a reciprocal relationship among these entities, such that moderate difficulty that satisfies the needs for competence and autonomy should, ultimately, improve intrinsic motivation and engagement (Deci & Ryan, 1985; Ryan & Deci, 2000). Additional research found that students who were more internally motivated were more likely to demonstrate a concerted effort toward task completion (Ryan & Deci, 2000).

In the context of literacy, research has shown that when motivated, student reading behaviors are characterized through the common qualities of interest, curiosity, involvement, and preference for challenge (Deci and Ryan, 1985; Guthrie & Wigfield, 2000). A substantial amount of research has correlated reading comprehension and reading motivation for students in the upper elementary grades (Hidi & Harackiewicz, 2000; Wigfield & Guthrie, 1997). In the school context, reading motivation is gauged by time spent on reading and performance on reading comprehension tests (Wigfield et al., 2016).

Taboada et al. (2009) studied how motivation and cognition predicted reading comprehension performance and growth with fourth graders while controlling for student background knowledge activation and student questioning. The researchers performed a series of multiple regression analyses to determine the variance among the dependent variables: reading comprehension at Time 2 (either multiple-text reading comprehension or Gates-Mac-Ginitie Reading Test) and the independent variables: background knowledge, student questioning, and internal motivation. The data collected included results from the teacher survey of motivation, an open-ended constructed-response, self-generated questions, and scaled scores. The subjects' materials varied in content difficulty, text structure, text difficulty, length per section, number of relevant sections and distracters, and number and type of illustrations. The randomly assigned reading materials ranged in grade level reading (Grades 2-6) and were based on life science.

To address the first question of reading comprehension, the students were administered two reading comprehension measures, multiple-text reading comprehension, and the Gates-MacGinitie reading comprehension test twice during the

fall semester (September and December). Analysis of the results confirmed that each of the variables added significantly to the variance in each of the two measures of reading comprehension, after controlling for the other two variables in the regression equation (Taboada et al., 2009). The second question of reading comprehension growth was established from the data and similar results were discovered. Findings supported the relationships of motivational and cognitive processes to reading comprehension performance and growth. Taboada et al. concluded that readers who were more internally motivated would be more devoted to reading, ask higher-level questions while learning and, thus comprehend better. Further, the researchers examined the specific contributions of internal motivation (for reading), rather than motivation as a general concept of intrinsic or external motivation. They determined that internal, reading motivation is constructed from (a) perceived control, (b) interest, (c) self-efficacy, (d) involvement, and (e) social collaboration (Taboada et al., 2009).

Using these same constructs of reading motivation and engagement, Guthrie et al. (2007) broadened the scope of reading motivation and investigated its potential to predict student growth in comprehension among fourth grade students ($N = 31$). One question assessed was the extent to which reading motivation predicted reading comprehension growth and to what extent reading comprehension predicted growth of reading motivation, using measures that vary according to text, source, and context. Variables used in this multiple regression analyses were pre/post achievement data from a standardized test, interview reports, student self-reports, and general motivation variables (interest, perceived control, efficacy, collaboration, and involvement). The data showed a positive mean change from pretest to posttest, representative of a pattern of growth.

Regarding the secondary question, Guthrie and his colleagues did not find the inverse to be true; reading comprehension did not predict growth in reading motivation, using those same variables. Further analysis demonstrated that three of the five motivational variables explained up to 46% of variance, which was statistically significant. These motivational variables were directly linked to patterns of growth in reading comprehension: cognitive (interest, perceived control and efficacy) and conceptual (collaboration and involvement) attributes. In addition, the examiners interviewed the fourth graders to determine the meanings and attributes of these same motivational constructs of reading motivation. The examination of responses from the interviews revealed highly interested students exhibited deep comprehension, complex cognitive command, high coherence of recalled content, and general enjoyment of the texts among students with average and lower than average ability (Guthrie et al., 2007).

Anmarkrud and Braten (2009) used the expectancy-value framework to examine the contribution of motivation to text comprehension when achievement in the domain, topic knowledge, and different forms of strategic processing were controlled. Ninth graders (N = 104) participated in a motivation inventory, topic knowledge measure, the reading of two social studies texts, strategy inventory, and reading comprehension measure for this study. Results of this study indicated that motivation constructs can improve the prediction of text comprehension and performance when the forementioned variables have been controlled. Among a fourth-grade sample of students, Wigfield et al. (2008) compared student engagement and reading comprehension across three different instructional programs. When students' level of reading engagement was statistically controlled, the differences between the treatment groups were not significant. Their

findings confirmed a correlation between reading engagement and comprehension (Wigfield et al., 2008).

Guided by the expectancy-value model of academic motivation (Wigfield & Eccles, 2000) and the understanding that students have different efficacy beliefs and values for comprehension in different domains (Anmarkrud & Braten, 2009), the researchers explored reading comprehension in the domain of science (Braten et al., 2012). Several points of data were collected over two sessions. First, word recognition skills, reading motivation and topic knowledge were assessed (in that order) during a 45-minute session for all participants. Student knowledge was assessed from a multiple choice, pre and posttest items of conceptual and factual knowledge. Students were given 180 seconds to complete a word chain activity to measure word recognition. Students' interest for reading was collected using a 27-item questionnaire, and nine of those items were used to capture expectancy and 18 for motivation. The study showed that participants scored statistically significantly higher on science reading task value (M = 7.51, SD = 1.48) than on science reading self-efficacy (M = 6.21, SD = 1.67), with $t(63)$ = 7.12, p = .000, d = 0.82. Plus, participants' topic knowledge scores were statistically significantly higher after (M = 15.35, SD = 3.00) than before (M = 10.09, SD = 3.14) reading the texts, $t(64)$ = 13.87, p = .000, d = 1.64. This research confirmed the strong contributions of word recognition skills to comprehension and learning. Based on the essays, reading pattern and reading self-efficacy emerged as predictors of comprehension performance when topic knowledge was controlled. The additional variance was generated from reading motivation, as measured by self-efficacy. This would suggest that the value of a task or activity may be more strongly related to a students' choice than

to performance, which is consistent with previous research using the expectancy-value framework (Braten et al., 2012). Similar findings for intrinsic motivation have been established on positive affect, interest, persistence, and attention (Eccles & Wigfield, 2002), according to the expectancy-value framework.

The crux of this research has highlighted the comprehensive role of motivation in reading comprehension. Also evidenced is the reader's cycle of motivation, reading comprehension, and cognition to accomplish reading-based goals. Reading motivation is viewed as an individual's personal goals, values, and beliefs regarding the topics, processes, and outcomes of reading (Guthrie & Wigfield, 2000). As this data has asserted, reading motivation has a multidimensional relationship with reading comprehension. Research from Wigfield et al. (2008) yielded four variables that influence reading motivation, including: (a) autonomy and choice (b) use of interesting texts in instruction (c) having conceptual goals for reading instruction and (d) supporting collaboration in reading.

Similarly, Wilde et al. (2018) claimed that choice must, also, be crucial to the development of positive motivational qualities during the learning process. Specifically, as relayed through SDT, they argued that a human's natural progression toward basic competence propels them to seek out stimulation and challenges through activities (Deci & Ryan, 1985; Eccles & Wigfield, 2002). As well, this theory proposed a reciprocal relationship among these entities, such that moderate difficulty which satisfies the needs for competence and autonomy should, ultimately, improve intrinsic motivation and engagement (Brooks & Young, 2011; Deci & Ryan, 1985, 1987; Patall et al., 2008; Ryan & Deci, 2000). Essentially, a person is deemed to be intrinsically motivated when they

are engaged in an activity on their own accord, i.e. personal interest, personal fulfillment, internal enjoyment. Conversely, when external rewards are paired with an activity used to maintain or sustain intrinsic motivation, intrinsic motivation decreases, as the personal foci and interest are altered by this type of reward (Deci, 1971; Marinak & Gambrell, 2008). In conjunction, the individual's perception of autonomy is lost in this transaction (Deci, 1971).

Studies have demonstrated a direct and lasting relationship between reading motivation and achievement beginning early in a student's school career (Baker & Wigfield, 1999; Vaknin-Nusbaum et al., 2018; Wigfield & Guthrie, 1997). Historically, students with high intrinsic reading motivation have shown a greater affinity toward reading and the practice of reading skills; therefore, intrinsic reading motivation has been directly related to the success of reading achievement (Guthrie, 2004; Vaknin-Nusbaum, et al., 2018; Verhoeven & Snow, 2001).

By contrast, extrinsic motivation is fueled by a reward or the fulfillment of another individual's want or need and, in the case of reading comprehension, has been negatively correlated (Becker et al., 2010; Schiefele et al., 2012). Subsequently, extrinsic rewards are viewed as manipulatives to behavior and are perceived as a loss of control. The loss of control from outside forces is in direct breach of autonomy, as determined by SDT. Patall et al., (2008) deduced that the introduction of rewards would be detrimental to the provision of choice. Dobrow et al. (2011) noted while grades are used to foster student engagement and interest, grades are extrinsic rewards. Paradoxically, grades are counterintuitive to the provision of choice. However, autonomy may be regained,

depending upon the context of the choice presentation and its relevance (Katz & Assor, 2007).

The effects of choice on motivation are dependent on how well they support autonomy, personal values, and interests (Patall et al., 2008). Zuckerman et al. (1978) paired 80 undergraduate students (N = 40 pairs) to assess how intrinsic motivation would differ in a partnership, where one student was able to choose which puzzles he would complete and how much time to devote to each activity while the other student partner was assigned the same puzzles. Intrinsic motivation was designated by time spent on the puzzle, when compared to the average completion times for the puzzle. In addition, choice was rated by a post check survey which tracked individual perceived control over the task. Results indicated that of the pair, the partner who received a task-choice had a greater perception of control and spent more time on the activity, on average, than the no-choice partner (Zuckerman et al., 1978). In this case, choice provided motivation and control over the environment for the students.

Choice has been used as a tool to enhance learning activities. Cordova and Lepper (1996) stated the need for learning to be meaningful and relevant for students to sustain the activity. Cordova and Lepper examined the effects of choice, contextualization, and personalization on intrinsic motivation with a computer-simulated, math-based learning activity. They reasoned that authenticity and functionality is key to the acquisition of skills and determined that the trend of instruction had led to the decrease in intrinsic motivation for adolescents. Further, they deemed these practices of decontextualization to be a contradiction to the pedagogy of Dewey and Bruner, as well

as, those of national commissions that made recommendations regarding standards and instructional practices.

Their experimental study provided fourth and fifth grade students with two features to increase intrinsic motivation, the personalization of the gaming module and the option of choice over instructionally-based aspects of the game which they designed to complement the learning of arithmetical and problem-solving skills within algebraic expressions (Cordova & Lepper, 1996). The students were randomly assigned to five conditions or versions of the math games including: (a) generic fantasy-no choice, (b) generic fantasy-choice, (c) personalized fantasy-no choice, (d) personalized fantasy-choice, and (e) no-fantasy (control). Choice features related to the level of difficulty at which the student wanted to interact with his opponent, the computer simulator. Examples of personalization included insertions of the student's name, birthdate, or favorite foods within the simulation that were collected from a questionnaire the students completed. Other information gathered from the students included a pre and posttest of math skills; a series of motivational questions measuring their enjoyment, willingness to continue with the activity outside of the study, and perceived competence; and records of task involvement such as complexity of skills, number of problems solved, and use of program hints.

All data points were analyzed with a one-way analysis of variance (ANOVA) with four orthogonal contrasts. Further, the researchers conducted a one-way analysis of covariance (ANCOVA) to examine the effects of personalization and choice on students' learning. Findings indicated that students exposed to the enhancing strategies of personalization and choice exhibited higher levels of intrinsic motivation and became

more entrenched in the activity. Consequently, these students used more complex operations, challenged themselves more often, became more proficient, and gained more information from the lesson (Cordova & Lepper, 1996).

Using the SDT construct, Patall et al. (2008) examined the role of choice in intrinsic motivation and other related outcomes through a meta-analysis in which 41 studies were analyzed. SDT studies selected for this analysis employed a manipulation of choice and measured intrinsic motivation in some capacity. The researchers predicted that choice would have a positive overall effect on intrinsic motivation and related outcomes. In addition, their meta-analysis sought to clarify the effects of choice overload, the relationship of choice on extrinsic rewards, the effect of various types of choice, relative to control measures, and the effect of choice as it is presented in the design of the study. The 41 studies generated 290 effect sizes, of which 165 represented the effect of choice on intrinsic motivation. Statistical testing was used to calculate the average effect sizes, identify the distribution of effect sizes, outliers, moderators, and the overall effect of choice on all nine outcomes: intrinsic motivation, effort, task performance, subsequent learning, perceived competence, preference for challenge, pressure or tension, creativity, and satisfaction. Data revealed that 78 of the 91 overall effect sizes related to intrinsic motivation were in a positive direction, while the remaining 13 were in a negative direction. In support of SDT, they were able to reject their null hypothesis that the effect of choice was equal to zero (d = -0.39 to 1.56, after Winsorization). The meta-analysis demonstrated that choice had a positive overall effect on intrinsic motivation and other related outcomes (Patall et al., 2008). The effect of choice on subsequent learning was positive, but it was not statistically different from

zero. The authors suggested that since most of the studies in this meta-analysis were short in duration, the long-term effects of learning were not measured. Subsequently, they called for more research in this area (Patall et al., 2008).

In a separate study, Patall, et al. (2010) found choice to be an important component to creating a classroom supportive of autonomy and intrinsic motivation, with students (N = 207) in grades 9-12 across 14 classrooms in two urban schools. Over four weeks, the students completed a series of questionnaires, unit tests, an inventory to measure intrinsic motivation, and an experimental manipulation of homework choice, to determine the effects of providing choices on motivation and subsequent academic performance. Results compiled from the analysis of student grade point average, unit tests, and data collected from the various questionnaires demonstrated the provision of choice among homework tasks effectively enhanced motivational and performance outcomes (Patall, et al., 2010). The authors made additional note of the authenticity of the study (real classroom, real teachers, and real students) and the practicality of choice integrated into homework, as a naturally occurring classroom objective and research-based academic achievement measure. Likewise, the current study offered the provision of choice within another classroom activity, assessments, to determine its impact on achievement and motivation.

Choice and reading engagement. Engaged readers are like scientists immersed in their discovery and intentional in their craft. Other descriptors of student engagement encompass actions that mimic motivation including self-reliance, self-initiation, and participation (Orkin et al., 2018). Guthrie (2004) characterized the engaged reader as one who proactively and diligently pursues reading and knowledge much like a profession.

The engagement perspective has been linked to motivation (Gambrell, 2011; Guthrie & Wigfield, 2000; Tracey & Morrow, 2006). Motivation and engagement appear along a continuum in education. Where there is motivation, engagement will follow among learners in the classroom. Moreover, the sustained engagement will lead to an increase in time spent on an activity fostering deeper learning (Gambrell, 2011). Student time spent reading will increase reading achievement (Guthrie, 2004) because the reader has chosen to engage in these reading activities.

To explore this relationship between choice and engagement, Schraw et al., (1998) conducted a series of experiments examining the role of choice relative to types of engagement. Using SDT, computerized testing theory, and reader response theory (Rosenblatt, 1994), the researchers studied the effects of choice on different measures of cognitive and affective engagement during reading. The researchers hypothesized the no cognitive hypothesis, that choice increases self-report measures of affective engagement but has no effect on cognitive engagement (Schraw et al., 1998). This hypothesis parallels the arguments of Deci and Ryan (1987) who claim that when the activity is supportive of autonomy, cognitive flexibility is more actualized. Findings from the Schraw et al. first experiment compared cognitive and affective engagement measures among three groups (unrestricted choice, denied choice, and a control group) and showed no difference between the means and standard deviations for the multiple-choice tests, interest questionnaire, essay responses, and desire for control survey. In the case of the denied choice group, participants were told that they were assigned their reading to balance the study, since not enough people in the unrestricted choice group had chosen the text. Patall et al. (2008) cautioned “when participants are explicitly denied a choice

or in which they are aware of the alternatives that they are not allowed to choose, individuals may experience a particularly pronounced decrement in motivation" (p. 274). In the absence of motivation and autonomy, the students may feel stifled in their work (Deci & Ryan, 1987). As for the secondary piece of affective engagement, an analysis of the 12-item attitude checklist reached significance for both the unrestricted-choice and denied choice groups, who were more likely to indicate that they liked what they were asked to do in the study, when compared to the control group (Schraw et al., 1998).

The second experiment conducted by Schraw et al. (1998) replicated most of the first experiment's procedures. However, they only used two groups (choice and a denied-choice group) as they wanted to provide a stronger case for their no cognitive engagement hypothesis by further isolating the denied-choice group. Results from this experiment showed the choice group reported more interest, more favorable attitudes, and reported more positive comments in their essays and general commentary about the provision of choice than the denied-choice group. Additional results confirmed their hypothesis, as choice appeared to have no impact on any of the cognitive variables. However, the results should be interpreted with caution due to the design of the study (Deci & Ryan, 1985; Patall et al., 2008).

Guthrie and Wigfield (2000) provided an engagement perspective of reading comprehension development, which holds that students' reading outcomes are based on the joint functioning of cognitive comprehension strategies, motivational processes, conceptual knowledge, and social interaction among learners. A similar assertion was made by Guthrie et al. (1996) who described engaged readers as motivated to read for a

variety of personal goals, strategic in using multiple approaches to comprehend, knowledgeable in their construction of new understanding of text, and socially interactive in their approach to literacy.

Guthrie and Wigfield (2000) explained that the cognitive aspect of engagement demonstrates how effective readers deliberately make choices within a context and select strategies to comprehend text content. The reader is assumed to have read a piece of text in a strategic, interested, and captivated manner, leading to a deeper understanding of the information. Collaboration and socialization around the readings further solidify the comprehension of the information. This process is then filtered through motivation, as the readers are intentional in their drive to comprehend the material. Altogether, the act of engagement is assumed to be positive in the case of reading and other forms of academically based work (Guthrie & Klauda, 2014), and is, therefore, strongly associated with reading achievement (Guthrie & Wigfield, 2000). Consequently, as students become engaged readers, they provide themselves with self-generated learning opportunities that are equivalent to several years of education (Guthrie & Wigfield, 2000). This perspective is unique in its suggestion that with all three functioning parts (cognition, motivation, and socialization) the engaged reader could become an agent of his own reading growth (Guthrie & Wigfield, 2000).

Several studies have noted the benefits of choice and the freedom associated with decision-making. Fraumeni-McBride (2017) predicted that reading choice would increase reading enjoyment, thereby increasing comprehension. She assessed the effect of choice on children's reading comprehension and enjoyment using regression analysis. The data were collected from 32 randomly selected third and fourth grade students who

participated in 12 independent readings of grade-level passages, along with the corresponding reading comprehension quizzes. During the readings, the researcher collected and coded various observation data points to evaluate perceptions of enjoyment, interest, levels of focus, and comprehension (Fraumeni-McBride, 2017). In addition, the students participated in self-assessments and an end-of-study survey, formulated to capture their reading style (aloud or silent) preferences and interest in the readings. Analysis revealed that children who were given a choice in reading scored higher in reading comprehension than when they were assigned a reading (Fraumeni-McBride, 2017). The researcher noted that observational data was used to compare perceptions of the children's comprehension with actual scores. While the subjectivity of the researcher's observations may need more clarity, the numeric differences between the comprehension scores of the assigned versus the choice readings provided a more objective account of the outcomes. The findings from this study provided further support for the positive effects of choice on learning.

In summary, these studies have related choice as a significant factor in behavioral and academic outcomes. Primarily, when students are given a choice, the effects have indicated increased interest, autonomy, perceptions of control, cognition, and motivation. Schwartz (2007) contends this narrative of students' increased engagement due to the offering of choice is misleading and ambiguous. However, Katz and Assor (2007) and Patall et al. (2008) have provided new interpretations of SDT, clarifying the theoretical model. They have contended with SDT that the provision of choice is contingent on the value and relevance the choice has to the individual. This is a key distinction between the act of choosing and a true provision of choice. The act of choosing is not without

pressure; therefore, the result of this type of choice may not elicit motivation or engagement. Moreover, within the classroom, students exercise choice based on the boundaries established by the teacher and the environment. When that environment supports autonomy, students are more comfortable making choices and generating thoughtful ideas. In turn, the provision of choice is self-fulfilling. The decisions made should foster engagement and motivation, leading to richer learning.

The cycle of student learning is enriched by choice. When students can have deeper and more meaningful encounters with information, they are able to later demonstrate their comprehension. Assessments are tools used to highlight levels of understanding gleaned from instruction. The next section explores assessment, including a brief historical outlook of assessment and its relation to reading comprehension, followed by a review of some studies that have combined choice and assessment.

Assessment in Education

Assessments are prevalent throughout every phase of life. Newborns are routinely assessed at birth for medical anomalies and developmental milestones. Between birth and age three, children are formally and informally assessed with checklists, standardized instruments, visual and verbal cues, play-based assessments, and other normative assessments to ensure they are developing relative to their peers (Wortham, 2005). School-aged children are subjected to a different range of tests, based on educational exposure of instructional materials, both formally and informally. In the educational environment, assessments are used to make diagnoses, instructional and programming decisions, identify and correct developmental problems, progress monitor for skill growth, and determine the level of skill development that a student has acquired

over the course of a unit, semester, or school year (Wortham, 2005). This section briefly details the history of assessments, types of assessments used in education, the impact of assessments on education, and the relation that choice has had with assessments.

History of assessments. The examination of human developmental phases ushered in the first set of studies of people. Specifically, the acknowledgement of child development within the life cycle appeared to be a turning point in education, as well as for the history of assessment (Wortham, 2005). The first child studies date to the 18th and 19th centuries when scholars such as Johann Pestalozzi, John Locke, and Frederick Froebel published child-centered reflections on rearing children. Emile Rousseau recognized the importance of child study to the education system and advocated for the enhancement of both arenas (Wortham, 2005). Wortham recognized the efforts of G. Stanley Hall, Charles Darwin, and Lawrence Frank for their direct contributions to the scientific study methods of children and their mentorship of other major contributors to educational reform such as John Dewey, Arnold Gesell, and Lewis Terman. As these pioneers permeated different sectors of education, between the 1890s and 1950s, the child study movement expanded into various types of children-first schools dedicated to producing literature on best practices and strategies to cultivate students (Wortham, 2005).

Early in this same period, the behaviorist movement evolved and led to more scientifically based studies dedicated to shaping human behaviors, which also influenced learning outcomes and objectives (Tracey & Morrow, 2012). For example, the constructs of learning through association, the principles of direct instruction, and the deconstruction of reading into isolated skills are all an outgrowth of behaviorism (Tracey

& Morrow, 2012). Another culminating factor of this period was the unintended consequences of World War I (WWI), which highlighted reading deficits of soldiers, who represented a byproduct of the educational practices of the era (Smith, 2002). As a result, a flood of reading innovations ensued, along with greater interests in diagnosing and remediating reading difficulties (Smith, 2002). All these early events altered strategies, pedagogies, and investigations about student and school relations and expectations. Current research in brain development and technological advancements of the 20th and 21st centuries have brought additional variations on assessment and imposed new questions on how students are instructed and assessed (Wortham, 2005).

History of standardized testing. Smith (2002) suggested that the first major turning point in the measurement of educational products began with the introduction of the Thorndike scale in 1909. Thorndike's design of the Scale Alpha standardized reading test was one of three competing approaches to reading comprehension assessment and most closely resembles present-day comprehension tests (Huddleston & Rockwell, 2015). Thorndike's testing format consisted of a short reading passage and a mix of literal-level and limited, single answers, which made for an easy conversion into a multiple-choice test. The introduction of a standardized measure for evaluating an instrument sparked new questions around reading outcomes and productivity in the classroom. Specifically, as interests in tests grew, researchers began to question traditional systems of grades on matters of reliability and objectivity (Huddleston & Rockwell, 2015). Mass testing was viewed as both reliable and efficient for the task of tracking students and assessing student learning.

The first known attempt to determine standard scores within the context of reading was recorded in the Fourteenth Yearbook of the National Society for the Study of Education published in 1915 (Smith, 2002). Wortham (2005) asserted that standardized testing derived from the recruitment needs of colleges and universities that needed to assess prerequisites and college readiness standards of potential applicants. Similar measures began to be used within public school systems, as test developers became employed by the school system. Measurement of students moved from an informal to a more formal process after the first intelligence scale, *Stanford-Binet Intelligence Scale,* was popularized. As a national common set of instructional practices had not been implemented and schools offered a greater continuum of education, there was a need to classify and level students to manage the increase in student enrollment. Several authorities have disagreed with this characterization, asserting that these tests were built on a foundation of discriminatory practices that have separated persons by focusing on eugenics, social engineering, and social efficiency (Au, 2013; Huddleston & Rockwell, 2015) resulting in the achievement gaps that are present in standardized testing scoring analyses.

Standardized tests morphed with changes in political, social and institutional models. Information related to disparities caused by poverty led to early childhood programs and an influx of other resources used to combat the educational gaps due to inequities in access to resources. Allington (2002) found that students of poverty tend to demonstrate adequate growth throughout the year, just as their affluent peers would; however, these same students experienced a setback in the summer, which only widened across an accumulation of years. The Elementary and Secondary Education Act (ESEA)

passed in 1965 provided federal funds to districts with high concentrations of families in low-income areas. In return, additional oversight was mandated by the government to monitor the growth of these students. In 1975, Public Law 94-142, the Education for All Handicapped Children Act was passed, marking a new era in measurement of all student progress. Prior to this law being passed, students with disabilities were systematically isolated from their general education peers creating an exemption for this population from pertinent accountability measures.

With the renewal of ESEA in 2001, the act expanded its monitoring system and set expectations, evidenced by high-stakes, standardized testing. Each state receiving funding under this act, No Child Left Behind (NCLB), adopted the higher standards set by the act, thereby shifting toward a testing culture and fueled by test scores. Allington argued that the repackaged NCLB act has not been reformative in its expansion and is misguided, at best. His research presented misinterpretations of information, unreliable research, and politically charged motives, showing a lack of growth in student achievement and ill-prepared teachers, all in need of assistance. Accordingly, the educational system appeared to need an overhaul, which translated into more testing as that was viewed as the most effective means to monitor progress. Allington (2002) has advocated for less testing and more teacher autonomy, given their level of expertise in the classroom.

A brief history of American reading comprehension. Historical descriptions of education in America can be traced back to opportunities for individuals to exercise their religious rights (Smith, 2002), while providing new avenues for innovative and critical thinking. Between the 16th and 17th centuries, the pedagogy of early education was based

on recitation and rote memorization skills. Information was disseminated orally and directly interpreted by clergymen, administrators, teachers and other figures of authority, as these figures were likely to be educated in the community. Likewise, these figures were revered and their interpretations were rarely questioned; therefore, comprehension of the material was not a functional purpose of reading for adults or children. During this period, teachers dedicated most of their instruction toward individual skill sets, e.g., alphabetic principles, phonics and phonemic awareness, while fostering an appreciation for the art of literature. In subsequent decades literacy instruction has moved beyond a simplistic skill-based pedagogy to a broader, more cognitive-focused knowledgebase, which has altered views, goals, motivation, and learning outcomes aimed at comprehension instruction. Reading comprehension is predicated on two prerequisites: decoding (one's ability to decipher, segment, and blend words) and word knowledge (Snow, 2002). Reading has surpassed the historical emphasis of performance (recitation) and regurgitation of learned actions. At present, knowledge related to literacy and learning has provided more opportunities for individualized instruction, better measures of academic success and personalized, learning goals (Dent & Koenka, 2016; Schunk & Zimmerman, 1998; Zimmerman 1989, 2002). Still, wide disparities in reading comprehension abilities exist among students, as documented through theoretical models of reading comprehension, comprehension assessments and individual skill sets.

Overall, reading comprehension is a measurement of both cognitive and affective factors. Research dedicated to affective factors associated with literacy, including self-regulation, motivation, and self-efficacy has provided a wider scope of the untapped riches of an instructional curriculum. Reading materials of various types foster students'

interest in reading and provide opportunities to increase comprehension based on cognitive and negative variables (Sanacore, 2002).

Theoretical Models of Reading Comprehension

The accepted simple view of reading comprehension is described as the interaction between decoding and listening comprehension (Gough & Tunmer, 1986). Based on this view, comprehension is the reader's ability to decode or understand written language. By their definition, literacy is the product of skilled decoding and listening (Gough and Tunmer, 1986). Although the simple view has captured the basic actions behind reading, it has failed to account for the influential reach of vocabulary, prior knowledge, or context, which influences a shift as reading material becomes more complex and the reader reaches middle and high school (Catts et al., 2006; Hoover & Gough, 1990).

Kintsch (1988, 1994) proposed a hierarchical model of reader's comprehension (Construction-Integration model or CI), which has expanded the simple view of reading. Specifically, this model expanded the influences of language and acquired knowledge (domain knowledge) to provide a better explanation of the reader and text interaction as literature increases in complexity. In all, the model described a bigger purpose for reading and promoted the generalization of reading strategies.

Deshler and Hock (2006) proposed the Adolescent Reading Model based on the same integrated variables of word recognition, language comprehension and executive processes (cognitive and metacognitive strategies) and used it to explain the need for instruction in all these areas. This model is derived from the Strategic Instruction Model (SIM) and is a basic outgrowth of the idea that students need reading interventions at

multiple stages in their reading development. Included in this model is a comprehensive instructional environment that simultaneously focuses on decoding, fluency, vocabulary/language development, and reading comprehension strategies. A key distinction of this instructional program is motivation, which increases learning outcomes and goals of students.

Guthrie and Wigfield (2000) provided an engagement perspective of reading comprehension development, which holds that students' reading outcomes are based on the joint functioning of cognitive comprehension strategies, motivational processes, conceptual knowledge, and social interaction among learners. A similar assertion was made by Guthrie et al. (1996) who described engaged readers as motivated to read for a variety of personal goals, strategic in using multiple approaches to comprehend, knowledgeable in their construction of new understanding of text, and socially interactive in their approach to literacy. Guthrie and Wigfield stated that the cognitive side of engagement demonstrates how effective readers are deliberately making choices within a context and selecting strategies for comprehending text content. The reader is assumed to have read a piece of text in a strategic, interested, and captivated manner, which has leads to a deeper understanding of the information. Collaboration and socialization around the readings further solidify the comprehension of the information. In addition, this process is funneled through motivation, as the readers are intentional in their drive to comprehend the material. The act of engagement is assumed to be positive in the case of reading and other forms of academically based work (Guthrie & Klauda, 2014), and is, therefore, strongly associated with reading achievement (Guthrie & Wigfield, 2000). Consequently, as students become engaged readers, they provide themselves with self-

generated learning opportunities that are equivalent to several years of education (Guthrie & Wigfield, 2000). This perspective is unique in its suggestion that with all three functioning parts (cognition, motivation, and socialization) the engaged reader could become an agent of his own reading growth (Guthrie & Wigfield, 2000).

Both the socio-cultural (Snow, 2010) and cognitive views (Kendeou et al., 2014) have described reading comprehension as an event between the reader, the text, and the task. Kendeou et al. described this interaction of persons and processes as a comprehension continuum, contingent on differentiated low and high levels of skills. The former being decoding, fluency, and vocabulary, while the latter is directed toward coherence and sense-making of the text. These same processes are evident in the interactive model, as well. Like the socio-cultural and cognitive views, Rumelhart's (1977) Interactive Model has replaced the deficit model with a progressive view of reading as an interactive process. Additionally, this model explained the intangible and comprehensive cognitive processes that occur as readers use syntactic, semantic, orthographic, lexical, and visual information to comprehend material (Tracey & Morrow, 2012).

It can be reasonably stated that reading comprehension is a complex and active task. The ability to comprehend reading material is influenced by the mastery of cognitive and affective variables. Essentially, these theories and models assume reading comprehension is the interaction of a reader's thoughts and behaviors with text, which is much more complex than simple vocabulary and decoding abilities. These theories suggest that reading comprehension is grounded on the student's ability to maintain attention and engagement while reading, which influences the level of text understanding.

Measuring Reading Comprehension

A reading comprehension assessment is administered to determine mastery of meaning, application of instruction, and leveled skill sets. As has been established, ability is influenced by content, vocabulary, motivation and decoding skills. As students are promoted through the middle grades into secondary and post-secondary education, instruction, intervention and assessment shift from decoding and fluency toward various modes of comprehension, strategies, and content-specific vocabulary.

Reading comprehension is measured through verbal or written responses to a piece of text (NAEP, OCED/PISA, 2015). Commonly, students are provided a passage and asked to answer questions to demonstrate understanding of the information read (Siepel et al., 2017). The purpose for reading, then, is pre-generated from the questions, making the information only relevant relative to the task (McCrudden & Schraw, 2007). The purpose and specificity of the task is dictated by the outcomes of the activity. In the case of reading comprehension activities and assessments, the questions used to demonstrate mastery range in specificity, which requires the reader to locate either implicit or explicit information. Therefore, how reading comprehension is measured becomes a vital part of the reading process and how success is determined in the school setting.

The RAND Reading Study Group ([RRSG], (2002) qualified successful comprehension as the alignment of the demands of the text, the challenges of the task and the skills and proclivities of the reader. As readers proceed through a text, they are engaged in a combination of these skills or cognitive monitoring procedures (Baker, 1979; Billingsley & Wildman, 1990) which are required for basic comprehension.

Overall, the reader should gain information with each interaction with text, but the information could be more factual or entertaining or both. The level of involvement is dependent on the reader's ability to decode the text, translate the literary features, and monitor his understanding of the words read (RRSG, 2002). Pardo (2004) reasoned that comprehension is a process in which readers construct meaning by interacting with text through the combination of prior knowledge and previous experience, information in the text, and the stance the reader takes in relationship to the text. de Milliano et al. (2014) relayed that the success of this process is dependent on the reader being oriented to information in the text as well as the task requirements.

The cycle of reading and answering questions for comprehension is deemed a task-oriented reading activity. Task-oriented reading activities are increasingly important in American society, as individuals need to access very different sorts of documents to achieve very different goals, as recent comprehensive definitions of reading have recognized (OECD, 2003; RRSG, 2002). While it may be more widely or easily measured through a question and answer format, reading comprehension is shown to be an interactive process with multiple influences. Moreover, these influences are elastic, impressionable, and codependent on the assigned value of a task.

High Stakes Testing and Reading Comprehension

Afflerbach (2016) stated, "our conceptualization of reading evolves; so too must our conceptualization of reading assessment" (p. 415). Several have argued there has been an imbalance in policies and practices related to reading comprehension and high stakes standardized assessments, following the Report of the National Reading Panel (NRP) (2000). Allington (2002) and Afflerbach (2016) questioned the foundation of the

report of the NRP and claimed that it is the source of the inequity in reading assessment and reading instruction. For their part, the NRP bounded their research inclusion criteria with the five cognitive strategy and skill areas–phonics, phonemic awareness, fluency, vocabulary, and comprehension--of reading instruction and achievement. Most of the research on these skills used quantitative research to assess the progress of these cognitive outcomes, therefore, many studies fell outside this scope and were not included in the NRP report's analysis. As a result of these exclusionary factors, Afflerbach claimed that NRP virtually ignored the affective variables that are required for reading achievement and development. Afflerbach insisted that NRP's report inflicted more damage on the testing community, as it shifted the value and purpose of the reading assessment away from a balance of cognitive and affective variables toward a subset of cognitive strategies and skills.

To restore a balance, Afflerbach proposed a major reduction in high stakes testing and a more coordinated push toward formative and summative assessments, which offer more checkpoints for student progress and more accurately measures how readers construct and use meaning while reading. A conversion of this magnitude is needed to repurpose assessments as avenues for student-led independence, self-assessment, self-efficacy, motivation, and engagement. Successful reading assessment is reconfigured under these conditions, where evaluation is no longer done to students but done with and by students (Afflerbach, 2016). Like an autonomy supportive narrative (Stefanou et al., 2004), Afflerbach called for teachers and other stakeholders to revisit the function of high stakes assessments, as they do not provide a whole picture of student reading behaviors, achievement, and development.

Carver (1992) questioned the purported nature of standardized reading comprehension assessments. He argued reading ability was based on the rate of reading and conceptual understanding of the reading. Carver qualified understanding or comprehension as the rate at which thoughts are received and understood. Therefore, a reading test should measure how accurately a word is read and the rate at which the word is read. Combined, both accuracy and rate are reflected in a student's level of reading efficiency, which he declared was the most important factor of a reading comprehension test. Theoretically, standardized tests of reading comprehension would demonstrate differences in readers accuracy, rate, and efficiency. He tested his theory on four standardized tests: the Iowa Test of Basic Skills (ITBS), the Rauding Efficiency Level Test (RELT), the Degrees of Reading Power (DRP) test, and the Nelson-Denny Reading Test (NDRT), which were administered to 354 students in grades 3-8 and 150 college students. A principal components analysis was used to interpret the scores from the assessment. Results showed standardized tests of reading comprehension measured general reading ability, which Carver argued was a depiction of a student's reading rate. Given none of the administered assessments were timed, he reasoned that knowledge of word meaning represented the accuracy level of a student's reading. Standardized reading comprehension assessments are not a reflection of instruction. Carver suggested test-taking strategies include measurements of reading rate, which would improve test performance overall. He added without the pairing of a timed element to instruction, teachers may not be properly preparing students for standardized reading comprehension assessments (Carver, 1992).

High Stakes Testing and Choice

Student choice in assessment has rarely been conducted. Barry and Nielsen (1996) completed a content analysis with data (N = 34, 000 papers) from a pilot study using a state mandated writing assessment. Per the assessment procedures, students could select from among seven given prompts or choose their own topic. It was believed that topic choice would result in more authentic writing, greater autonomy, improved self-efficacy, and more developed content (Barry & Nielsen, 1996). The analysis revealed that students selected topics based on classroom context, content knowledge, and the individualized task (Barry & Nielsen, 1996). Thus, student choice was contingent on prior knowledge and instructional opportunity or exposure to a variety of writing forms (Barry & Nielsen, 1996). An improvement in writing (and reading) can only be achieved through testing alterations (Slomp, 2008). The provision of choice in writing, as in reading, would increase the scope and magnitude of student learning.

In a landmark study, *The NAEP Reader* (1997) designed a study to examine the impact of offering test takers a choice in reading material on an assessment of reading comprehension. The study ran simultaneously with the NAEP's annual assessments and compared the participants to test takers who were assigned a story. They reasoned that student choice of reading material should extend beyond instructional purposes and leisure reading. The testing committee recognized that traditional assigned assessments may not capture student reading ability, given the generic nature of reading comprehension tests (NAEP, 1997). Following the groups respective readings, the participants answered the same eleven comprehension questions, regardless of the story they read for the study. Of the eleven comprehension questions, eight were short

constructed response questions requiring a one or two sentence response and three were extended constructed response questions requiring a more developed, reflective response of one or more paragraphs. In addition, the choice group was surveyed on why they chose their story for the assessment. Based on performance, there was no significant difference between the averages of either group of twelfth graders. Among the eighth graders, who chose a passage, the average score for comprehension was one scale-score point lower than their counterparts. Secondly, students in the choice group perceived the task as easier than the non-choice group, for both grade levels (NAEP, 1997). Similarly, twelfth graders in the choice group had more positive predictions about their performance, when compared to the control group. Lastly, no significant differences were observed between either of the groups in relation to motivation for achievement on the assessment.

Although the findings did not support choice in this study, it was a first step along the continuum of student choice in standardized tests. The test was several steps along the continuum of student choice and standardized assessments for several reasons. First, the assessment provided students with an option to consider before beginning the standardized assessment. This was the first opportunity of choice on a standardized exam that had been documented. Lukhele et al. (1994) had previously offered students choice between multiple-choice, constructed-response, and examinee-selected questions on College Board Advanced Placement exams. In the Lukhele et al. study, the choice was based on answer format, as the researchers had hypothesized that constructed response questions are aligned with Glaser's (1985) four dimensions of cognitive achievement test performance. They found that they were able to glean the same amount of information

from a multiple-choice test as they did a constructed response exam with a significant reduction in time and money. Despite these findings, *The NAEP Reader* (1997) study incorporated the constructed response test format into their study.

Another important aspect to assessment is the test taker's familiarity with the testing format. 'Search-and-destroy' is a common test taking strategy term used to describe how students are taught to read the test item questions first, then read the passage and return to answer the questions (Huddleston & Lowe, 2014). The assumption of this strategy is twofold: The test will be multiple choice and the student will accurately answer test items. This practice is common among all students especially reluctant and struggling readers who have not fared well on standardized tests (Huddleston & Lowe, 2014).

In their exploration of student experiences with Georgia's test-based grade-retention policy, Huddleston and Lowe (2014) found search-and-destroy is highly inefficient method and is more suited for students that know what they are looking for in an assessment passage. Huddleston and Lowe cited one study which found this methodology accounted for more than 38% of the errors fourth through sixth graders made on a standardized assessment. This research also found that the search-and-destroy strategy was reinforced by the design of the test questions. Specifically, some questions referred students to a word or sentence, eliminating the need to read the passage. Finally, Huddleston and Lowe attributed the continued use of search-and-destroy to the lack of student choice in reading, too many practice tests, and reading material that was above their reading level. Students were no longer involved in learning and critical thinking about the reading. Instead, they were consumed with the completion of the task with

little regard for comprehension. Contrary to popular understanding of reading assessments, this study revealed that students are not reading passages for meaning and the researchers consider reading to be unnecessary to complete the assessment task. The results of the constructed response NAEP study showed no significant difference between the average reading scores for either group, which could be explained through strategy use or lack of familiarity with the testing format, which are missing from the NAEP data. If the search-and-destroy method is common practice and reading is not necessary to complete the task, then the students may not have had enough information to provide adequate answers to constructed response questions.

A final point to consider in the discussion of the NAEP study results were its limitations. This study was designed to examine the relationship between choice and performance under specified conditions. The study was unable to definitively separate choice from reading performance, because of the multifaceted impact of choice has on interest, motivation, and self-selection of reading materials. All the students were provided the same generically worded questions for scoring uniformity. However, the test evaluators noticed variable response patterns across specific questions and explained that these differences could have been attributed to the various narrative elements in the stories. In other words, some questions could not be retrofitted to the answers that were provided, as they may not have been appropriate for the story. Finally, the researchers implied that time constraints could have been a factor for the choice group, as they were given the same amount of time, 50 minutes, to complete the task; although they had to devote some of their time to choosing a story. The authors suggested that future studies

not use time restraints and attempt to control for the question/test interaction with an adjustment of the questions.

Summary of the Literature Review

The literature review presented choice as the gateway to student learning. In many of the studies included in this review, choice a variable used to uphold or dispel SDT. Patall et al. (2008) produced evidential support of SDT and made a stronger case for the application of choice in the classroom. Stefanou et al. (2004) proposed alternative ways for teachers to infuse choice in the learning environment and encouraged a continued commitment toward autonomous choice that extends beyond mediocre decisions. Research showed the effects of choice are numerous when fostered through teacher behaviors and opportunities for student growth. For example, when students are given a choice in the selection of reading material, they are more motivated to read (Flowerday & Schraw, 2000). In relation to reading comprehension, the provision of choice could lead to stronger displays of comprehension, given links between reading, interest, motivation and engagement (Campbell and Donahue, 1997).

Choice is a precursor to engagement. Performance and academic achievement in relation to reading comprehension are impacted by the provision of choice, as well. For example, the provision of choice could foster deeper connections with text for students, thereby increasing their performance on daily tasks and assessments. Moreover, students are led by their choices to pursue information deemed valuable and enriching to their own goals and interests, working in tandem with cognitive measures of reading ability. Current theories of reading and knowledge acquisition have provided more depth to

instructional practices and called into question the more antiquated systems of assessment.

The field of assessment has not kept pace with the changes in educational theories and instructional approaches (NAEP, 1997). In 1992, the NAEP reading assessment was redesigned to align with the educational trends of that era with the addition of authentic reading material, offering several types or genres of material and constructed response questions, which was an outgrowth of interactive reading and reader response theories. Importantly, these efforts were done in recognition of the need to increase alignment with instructional practices, showcase the differential abilities of students, and mimic the learning experience. Those changes have not all been maintained within the current format of testing and may explain the persistent gaps in achievement present among students in grades 3-8. This research revisited choice in high-stakes reading assessments to evaluate how measures of choice ('teacher-selected' and 'student-selected') compared as predictors of reading comprehension and how these vary as a function of achievement for fourth and fifth-grade students. The study also examined the relationship between interest and reading as secondary to the provision of choice.

Significance of the Study

Over 25 years ago, the NAEP reading assessments were revised following a call for change in traditional assessment approaches. Traditionally, these assessments were designed to uniformly assess student performance, which could only be reliably compared when every participant took part in the same assessment with commensurate conditions. At the time, educators were concerned by both the design of the standardized assessments and the atmosphere created by the test. Specifically, they were apprehensive

that the sterility of the test would diminish both motivation and interest, which are important to the reading comprehension formula (NAEP, 1997). Ultimately, the changes to the assessments were temporary and previous criticisms of standardized assessments have remained. While decades of research have informed the instructional practices implemented to ensure that students are being educated with scientifically based measures proven to increase achievement, standardized assessments have failed to keep up with these same changes in the field.

In response to this need for change in standardized assessments, this study was designed to revisit the question of the previous study conducted by the NAEP in 1994. The purpose of this study was to evaluate how measures of choice ('teacher-selected' and 'student-selected') compared as predictors of reading comprehension and how these vary as a function of achievement for fourth and fifth-grade students. Additionally, the study sought to quantitatively examine the relationship between measures of choice and the STAAR test with students. The next section explains the methodology employed to answer the following questions:

1. Are there significant differences in reading comprehension assessment scores when reading 'teacher-selected' passages versus reading 'student-selected' passages?
2. Is there a relationship between interest and performance on standardized reading comprehension assessments?

CHAPTER IIII

METHODOLOGY

Context of the Study

This study was designed to evaluate how measures of two choices of standardized reading passages ('teacher-selected' and 'student-selected') compared as predictors of reading comprehension and how these vary as a function of achievement for fourth and fifth-grade students. The study also explored the relationship between measures of choice (e.g., interest) and the standardized reading comprehension assessment with fourth and fifth grade students. This examination was conducted through a comparison of correlations and comprehension scores for fourth and fifth grade readers. Outlined in this section are the participants, procedures, materials, design, data collection, and data analysis completed for the study.

Participants

Selection of participants. The principal investigator was interested in fourth and fifth grade teachers currently teaching in Texas. Participants were identified as full-time, licensed teachers, currently employed in a public, private, or charter school in the state of Texas. Non-probability sampling procedures aided in the recruitment of teachers and increased the range of participants for this study. Specifically, the investigator employed a snowball sampling technique to recruit Texas teachers based on one criterion: persons assigned as a fourth or fifth-grade teacher of English/Language Arts students during the 2019-20 school year within their given school setting. Snowball sampling is optimal when a study pre-specifies a criterion of individuals or clusters (O'Dwyer & Bernauer, 2016). Through a system of endorsements, a small number of participants inform and

recruit other potential participants to join in the study, resulting in a gradual increase in study partners (O'Dwyer & Bernauer, 2016). Elementary teachers across Texas were solicited through electronic mail, telephone calls, text message, and social media. (see Appendix A for recruitment flier).

The study was carried out in five schools located within both the Prairies and Lakes and Gulf Coast regions of Texas. Among the five schools, two are in a suburban area, two in an urban area, and one school was in a rural region. One school was classified as a charter school and four were public schools. On average, families at the schools were described as having low to middle income. Participants represented a spectrum of school constructs: public/charter, rural/urban, high/low income, large/small student body. Teaching experience also ranged from 4-25 years of teaching in both the primary and collegiate instructional settings.

Description of teachers. Teachers were considered the participants of the study. Information collected from the teachers included years of experience, highest level of education, current grade being taught, and the classification of their school (i.e., public, charter, private, or other). Teachers provided descriptions of their current school, including socio-economic status, ethnographic information of the students, and a description of their student population. This data was used to confirm teachers met the minimum participation criteria and for the final analysis. This data, also, provided descriptive information for the variables and participants of the study.

Description of the student subjects. All students were enrolled in either fourth or fifth grade. Students participating in the study completed surveys, providing basic descriptions of the students, including current grade, age, and gender. Three of the five

schools had a predominately Caucasian population, while the other two schools served students from Latinx and African American households.

Based on the variables studied (achievement scores generated from passages that were ‘student-selected’ and those that were ‘teacher-selected’) individual reading ability was important to overall performance (Carver, 1992). Students identified with a reading accommodation (i.e., oral administration) by the school district were excluded from the study. TEA categorized oral/signed administration as a designated support that allows test material to be read aloud or signed to a student. This accommodation has been approved for student usage on the test, under the following conditions: (a) he or she routinely and effectively uses it during classroom instruction and classroom testing, and meets eligibility under one of the following: (a) The student is a current English learner and takes a STAAR test in English; (b) The student is identified with dyslexia or a related disorder; and/or (c) The student has documented evidence of reading difficulties. This exclusion of students was deemed necessary because a lack of reading ability was not a variable in this study (McKool, 2007).

Materials

Materials consisted of four different narrative texts, and corresponding multiple-choice tests for each of the four texts. Each of the narrative passages had previously been used on the State of Texas Assessments of Academic Readiness (STAAR) exam between 2015 and 2018. All passages were based on a third or fourth grade readability level. Subjects were enrolled as fourth or fifth graders during the 2019-2020 school year. Individually, subjects enrolled in the fourth grade were administered the third-grade passage, while the fifth-grade subjects were administered the fourth-grade

passages. The reading ability or text difficulty was not considered a factor, given that the students had been exposed to instructional material that would have exceeded a third or fourth grade reading level, respectively.

STAAR test. Reading achievement was assessed using the State of Texas Assessments of Academic Readiness, or STAAR. The STAAR test is the state testing program required for students in grades 3-8. The test was first implemented in the 2011-2012 school year to assess skill acquisition, as determined by the state-mandated curriculum standards, the Texas Essential Knowledge and Skills (TEKS). Input for test items including revisions, guidelines for reporting scores, quality control and preliminary testing was provided by stakeholders at the district, collegiate, and regional levels through a network of committee processes. The Texas Education Agency (TEA) identified the most critical areas which comprise the core skills emphasized on the STAAR test. Other skills and information are not applied with the same emphasis as that of the readiness standards. These readiness skills are designated in the TEKS as essential skills toward grade content mastery. The following criterion has been issued for readiness standards:

- They are essential for success in the current grade level or course.
- They are important for preparedness for the next grade level or course.
- They support postsecondary readiness.
- They necessitate in-depth instruction.
- They address broad and deep ideas.
- These skills are progressively aligned from elementary through high school and postsecondary readiness.

According to TEA, every STAAR question is directly and vertically aligned to the TEKS currently implemented for each grade/subject or course being assessed. Therefore, student progress from elementary through high school is assessed based on the student's performance on this test. The STAAR test has several versions: STAAR (English), STAAR Spanish, and STAAR Alternate 2. All students are required to participate in the STAAR exam, unless they meet the participation requirements of the STAAR Spanish or STAAR Alternate 2. In some minor cases, students have been exempted from the exam. Lastly, the STAAR test is available in different platforms: paper, braille, large print, and online.

The test is divided into subject areas, including Reading, Math, Science, Social Studies, Writing, Algebra I, English I, English II, Biology, and U.S. History. Districts may choose to administer the STAAR Algebra II and English III end-of-course (EOC) assessments, but these are not mandatory. The STAAR Reading assessment passages are criterion-referenced reading measures administered in grades 3-8. Criterion-referenced assessments are designed to measure how an individual performed on a standard or objective, based on a preceding set of instructions (Wortham, 2005). Unlike norm-referenced tests which compare an individual's performance against that of his peers, the results of a criterion-referenced tests are indicative of an individual's abilities in isolation or without the impact of his peers (Wortham, 2005). Although not predictive, the STAAR Reading test, generally, assesses these readiness skills across several different passages, including at least one narrative, one expository, and one poem per exam. Under testing conditions, typical students are given three hours to complete the STAAR Reading test, independently. For this investigation, fourth and fifth grade students were

not given a time limit to read and answer questions for each narrative passage, each day for two days. Teachers reported that students averaged 20 minutes a day on this activity.

TEA enlisted the services of an independent evaluator, Human Resources Research Organization (HumRRO), to evaluate whether the STAAR test scores are reliable and valid for all versions of the test. The most recent evaluation was conducted for the 2016 grades 3-8 STAAR test forms, which supported the validity and reliability of the test. The following is a brief summary of how reliability and validity were established for this assessment.

Reliability of the STAAR test. An instrument is determined reliable when it can generate consistent results with each administration. Other measurement components such as validity, interpretability, and bias are all contingent on a test's reliability. The reliability of the STAAR test is based on internal consistency and margins of error. Internal consistency is represented by the consistency of student's individual responses to the assessment and is measured in coefficients. The Kuder-Richardson 20 ($KR_{20)}$ was used to calculate the reliability coefficients for the STAAR reading tests, given the multiple-choice format. This calculation compares the true score variance with the observed score variance. Generally, as the distance between these two variances decreases, reliability increases. For example, a reliability coefficient of 0.75 is fair and anything over 0.90 would be ideal. The reliability coefficient reported for the third grade STAAR reading test was 0.88 and 0.89, overall, for the fourth grade STAAR reading test. The test is reliable.

Validity of the STAAR test. Validity is the extent that the test items align with the purpose of the test instrument. In this case, the STAAR reading test is aligned with the

TEKS. Other examples of this alignment include a comparison of test content, response processes, internal structure, relations to other variables, and consequences of testing. Procedural steps of test development are an outgrowth of the TEKS process, where Texas educators agree on the content of the standards and rely on these standards to build questions. The average percentage for third grade reading items for content validity was rated "fully aligned" to the testing content, given an average of 86.2% across all four reviewers. The fourth grade reading items were, also, rated as "fully aligned" to content expectation with an average of 91.5% across all four reviewers.

Further evidence of validity is gathered through response processes, the study of how students respond (i.e. task analysis, think-aloud, etc.) to piloted questions in a lab setting. Observations and student rationales are compared to generate rubrics and test measures of reliability. The internal structure of the assessment is referent to the composition of the test's internal consistency, across groups of students. The STAAR test is also compared to other tests that measure educational variables, such as end-of-year tests, standardized language acquisition tests, and different versions of the STAAR test to ensure validity of the assessment. Lastly, the test developers considered the consequences of the test administration by surveying stakeholders to weigh the outcomes of whether the usage and intention of the test is commensurate or outweighed by the negative effects of the test administration. In all, the test is valid.

Text. The four experimental texts used by each grade level were from the STAAR released reading passages and questions administered in years 2014-2018. The assessments were acquired from the Texas Education Agency (TEA) website. The fictional narrative passages used for the fourth-grade subjects are included in Table 1.

Table 1

Narrative passages used with fourth grade subjects

Title	Description	Category	Word Count
Racing Team	This story is about two boys, who made some unexpected friends at the beach.	Student choice	723
Lizard Problems	This is a story about a girl, who overcame her fear of lizards.	Student choice	730
Good Question, Chowderhead	This is a story about a girl who learned something new about herself and how to talk to other people.	Student choice	524
Jake Drake, Teacher's Pet	N/A	Teacher/No-choice	588

Note. STAAR Released narrative passages administered between years 2014-2018.

All fourth-grade passages used for this study were, originally, used as STAAR reading assessments for third grade students during their respective year. For this study, fourth grade students were required to read the passage adapted from *Jake Drake, Teacher's Pet*, as the 'teacher-selected', no-choice text. *Jake Drake, Teacher's Pet* was adapted from a work by Andrew Clements (2007) entitled *Jake Drake, Teacher's Pet*. This is book number three of a four-part series with the main character Jake Drake. The story was chosen as the control since adolescent readers may be familiar with this text. The remaining three texts were provided for them to choose from and answer questions, as well. Similarly, the fictional narrative passages used for the fifth-grade participants

included the three choice passages and a teacher passage. Table 2 lists the titles and the introductory statements used for the corresponding stories, along with the word count.

Table 2

Narrative passages used with fifth grade subjects

Title	Description	Category	Word Count
A Caterpillar's Tale	This story is about a caterpillar who worked hard to build a home. This story is about a girl, who learned a word and a lesson and turned it into a joke.	Student choice	752
My What is Showing?	This story is about a girl who learned a word and a lesson and turned it into a joke.	Student choice	557
The Spelling Test	This story is about a boy who learned new tips to accomplish a goal.	Student choice	858
Mystery at Dad's House	N/A	Teacher/No-choice	718

Note. STAAR released narrative passages administered between years 2014-2018

All fifth-grade passages used for this study were, originally, used as STAAR reading assessments for fourth grade students during their respective year. For this study, all fifth-grade students were required to read *Mystery at Dad's House,* as the 'teacher-selected', no-choice text. The remaining three texts were provided for them to choose from and answer questions. Each test booklet included a cover sheet with a brief descriptor of the story to elicit uniformity across all stories (Schraw et al., 1998).

Multiple-choice tests. As the NAEP assessment framework specified, students are assessed in reading for three different purposes: reading for literary experience, reading to gain information, and reading to perform a task (p. 3). TEA has noted that the design of the STAAR reading test focuses more on critical analysis rather than literal understanding. Following each passage students were required to answer between six and eight multiple-choice test questions that corresponded with their passage. The test items assessed the subject's comprehension of the passage and consisted of literal information questions, main idea questions, inferential questions, and vocabulary questions. For example, a typical vocabulary test question item (*Racing Team*) is shown below:

1 Read the dictionary entry for the word pass.

pass \ıpas\ verb

1. to ignore
2. to move by
3. to throw or hit a ball to a teammate
4. to complete a class successfully

Which meaning best matches the way the word passed is used in paragraph 7?

A. Meaning 1

B. Meaning 2

C. Meaning 3

D. Meaning 4

An example of a main idea question from the story *Racing Team* is:

Which sentence states the main theme of the story?

A. It is important to be on time when others are depending on you.

B. Winning does not matter as long as you try hard.

C. Working with others can be better than working on your own.

D. The best ideas come to those who are patient.

An example of an inferential question from the story *Racing Team* is:

Which sentence from the story contains descriptions that appeal most to the reader's sense of touch?

F. This was the day he had waited for all year—the New Year's Eve sand-sculpture contest at the beach.

G. Then he packed the sand with both hands while Max dumped on more sand and slapped it into a mound.

H. Rising out of the sand nearby were castles, dragons, whales, and mermaids.

J. A photographer hurried over to take a picture for the newspaper—two beautiful racing cars and four smiling sculptors.

Presentation of Ethical Considerations

The principal investigator requested broad consent for this research. While students were used for this study, the names of individual students were not required for this study nor were they solicited. In addition, the investigator did not interact directly with students for the collection of these data. The nature of this study presented no more than minimal risk or harm to the subjects and did not involve procedures which would normally constitute written consent. The teacher was required to read a statement regarding participation in the research and students were presented with their right to

refrain from participation without recourse from the investigator or teacher. Ethics approval and exemptions for the research were provided through the Institutional Review Board (IRB) committee at the educational institution of the principal investigator.

Procedures

Participants received a packet of materials, which included: an introductory letter (See Appendix B), an instruction card that contained instructions for the teacher to read prior to the administration of the test (see Appendix C), demographic information sheets (see Appendix D), a student information card and survey (see Appendix E), copies of all passages (teacher- and student-selections), and a description of the student choice passages. The instruction card was read to maintain consistency in the evaluation process. Each teacher was trained on the administration procedures by the principal investigator prior to the administration of the study.

Teachers were not restricted by the order in which the tests were administered. Each teacher was instructed to read aloud a set of written instructions which were located on their instruction card prior to administering the assessment to their students on each testing day. The instructions differed with the type of assessment being administered and by grade level. For example, the 'teacher-selected' administration was preceded by generic instructions for students to read the passage and answer questions. However, the 'student-selected' (choice) administration included a brief introductory statement of each of the three stories that the student would have to choose from that were read aloud by the teacher. The investigator required the tests be administered on separate days to minimize fatigue for the students.

Following the administration, teachers were asked to complete the demographic sheet and students were asked to complete their information cards and survey. The investigator collected the student passages, teacher demographic sheets, and student information cards with survey from the teacher at the end of the administration.

Choice activity. All subjects read two passages and completed the corresponding criterion-referenced test questions, related to the passage. One passage was pre-selected and administered to all students. The second passage was selected by the participant from a set of three passages. Subjects completed the activity in their classrooms and under the administration of their respective teachers. Subjects were instructed by their teacher to read the passage and answer the corresponding multiple-choice questions. Each subject was given as much time as needed to finish the reading and answer the questions. Choices related to actions such as process, mode, and pace are considered more meaningful and effective, when compared to other instructionally based choices (Patall et al., 2008). Subjects had access to both the passage and the corresponding questions throughout the entire session. Students were not allowed to return to a testing passage from the previous day. All completed items were returned to the investigator and scored.

Interest activity. Students completed a self-report survey with two items, following the completion of their choice passage. The surveys were integrated into the student demographic card for ease and efficiency of materials. The survey items were designed to measure students' attitudes and experience related to the 'student-selected' or choice reading passage, along with their overall attitude toward reading (Guthrie et al., 2009; Wolters et al., 2017). Using a Likert scale, the survey asked students to rate their

overall attitude toward reading and address to what extent they enjoyed their chosen passage. This survey was returned to the teacher, following the administration of the choice passage.

Research Design

An experimental design is a scientific inquiry developed to explain the causal relationship that occurs where one or more variables have been manipulated, under controlled conditions (Key, 1997; Mallette et al., 2013; Vellutino & Schatschneider, 2011). Assumed in the purpose of the experiment is the need to understand or infer how and to what extent an administered treatment affected a group (Mallette et. al., 2013; Vellutino & Schatschneider, 2011). Conversely, the investigator would also argue the importance of the counter-narrative of the experimental effects in the absence of the treatment. Traditional experimental designs draw conclusions from results, based on the extent of control exerted on the remainder of the experimental environment. For example, when conditions are strict, random, and highly controlled, the probability of the cause and effect relationship is more easily explained. When conditions for control become more malleable and random selection is not possible, the investigator could use a quasi-experimental design, where pre-established groups are compared using varied conditions (Key, 1997; Putman, 2017; Vellutino & Schatschneider, 2011). The advantage of the more controlled condition is in the interpretability and generalizability of the experimental effects with the significant reduction of causal explanations for any occurrence of change in the group. Ultimately, with either design, the proposed question of the study is formulated to understand a causal relationship and sought through a scientific method (Vellutino & Schatschneider, 2011).

History of experimental designs. A historical recollection of the experimental design is traced back to the 16th and 17th centuries, when scholars like Galileo, da Vinci, Gilbert, and Copernicus inquired about naturally occurring phenomena through active queries and experiments (Vellutino & Schatschneider, 2011). These scholars formed the foundational criterion of the scientific methodology, including the standards and expectations of this systematic design (Vellutino & Schatschneider, 2011). Originally established and adopted in the natural sciences, this methodology was ushered into education with the outgrowth of associationism, behaviorism, and structuralism (Tracey & Morrow, 2012). Further influences in basic and applied sciences throughout the 20th century have expanded and diversified literacy research methods, including brain-based and cognitive processing theories (Tracey & Morrow, 2012; Vellutino & Schatschneider, 2011). Recently, Mallette et al. (2013) called for the use of different methodologies to explore the complexities of literacy development. They concluded the field of literacy is enrichened when multiple methodologies, such as quasi-experimental designs, are used to address a single topic.

Quasi-experimental design in the field of literacy. Several in the field of literacy education have used quasi-experimental research to study literacy-related interventions and reading comprehension (Bayless et al., 2018; Fitzpatrick & Meulemans, 2011; Lai et al., 2016 Mitchell et al., 2016; Soltero-Gonzalez et al., 2016; Walpole, et al., 2017). Fitzpatrick and Meulemans (2011) engaged a set of college students (N = 157) enrolled in an Introduction to Developmental Psychology course in an experiment using a quasi-experimental design to determine the impact of an assignment and workshop intended to increase students' information literacy skills. By combining Vygotskian and

Piagetian models of learning, they hypothesized student learning of literacy skills would increase. Results supported their claims, as students who participated in the workshop showed significant improvements in their posttest scores, as measured by knowledge of APA citations, the online library catalog, and a rating scale.

Both Lai et al. (2016) and Soltero-Gonzalez et al. (2016) conducted quasi-experimental studies to collect longitudinal data related to program effectiveness, following an increased focus on professional development. Lai et al. (2016) questioned the effectiveness of school-based literacy initiatives for native New Zealanders, who consistently scored lower in reading comprehension measures than their counterparts, despite having received foundational skills and demonstrated growth potential. They relied on a quasi-experimental design in their longitudinal study, as it provided the flexibility to systematically replicate processes across settings/schools and accounted for the variability of the student population, while comparing program effectiveness. Further, they favored this design over that of a design-based research approach, for example, because they were measuring effectiveness of an intervention as a rate of gain (Lai et al., 2016). Soltero-Gonzalez and her colleagues compared whether the sustained effects of students' Spanish and English reading and writing outcomes in sequential literacy groups and paired literacy groups across 13 schools. Results upheld previous research which showed the paired literacy instructional design promoted reading achievement in English and in biliteracy development through a holistic framework, teaching methods authentic to the Spanish language, and text-based literacy instruction in English (Soltero-Gonzalez et al., 2016). The authors provided recommendations for

schools to transition from a sequential to a paired literacy program, as a follow-up to their findings.

Walpole et al. (2017) and Bayless et al. (2018) both tackled literacy reform initiatives by challenging traditional literacy programs. Walpole and her colleagues compared students receiving traditional guided reading with those who received a comprehensive school reform (CSR) program in grades three to five. Based on the fluency and comprehension measures from DIBELS Next and the computer-adaptive version of the Scholastic Reading Inventory, the treatment group experienced significantly more growth at all three grades. Bayless et al. (2018) combined a book choice and distribution program, small group intervention and one-to-one tutoring in a four-year after-school program to achieve similar results with students across a total of six public housing projects. Again, a quasi-experimental design was determined best to compare improvements in reading proficiency among program participants and a comparable group of students residing in public housing neighborhoods without access to similar after-school initiatives (Bayless et al., 2018).

Mitchell et al. (2016) explored the degree of writing ability among second-grade students, who were paired with fourth graders in a cross-age tutoring model. Among the three groups of second graders were the control group; the teacher/researchers' class; and two intervention groups. The fourth graders were randomly assigned to be in either the trained, untrained or control group of tutors. Further pairing was done by the researchers based on personality traits among the tutor and tutee within the newly created groups. For this study, a pre/post-assessment quasi-experimental design captured their data set while responding to the environmental conditions, i.e. school setting, pre-established

classes, and reducing the bias or conflict of interest with the teacher/researcher class as the control group. The researchers' exploration of the questions revealed a need for additional research in the area of cross-age tutoring, using specific skill achievement. The findings from this study indicated that students demonstrated an increase in the amount of the words in their writing.

A quasi-experimental design was deemed most appropriate to draw comparisons between groups for the aforementioned studies. Several of the authors used quasi-experimental methods to analyze changes in literacy skills over time (Bayless et al., 2018; Lai et al., 2016; Soltero-Gonzalez et al., 2016), while others challenged traditional instructional programs using repurposed theories (Fitzpatrick & Muelemans, 2011; Walpole et al., 2017). Collectively, the studies used practical design elements on intact groups to uncover relationships among literacy-based variables.

Present study. This study used a quasi-experimental design to ascertain the effects of choice on student achievement using a criterion-referenced assessment and a non-experimental design to investigate the relationship between choice, interest, and performance on standardized reading assessments. The quantitative data collected was used to answer the first research question and analyzed using an analysis of covariance (ANCOVA) which compared the variances between two groups (Larson-Hall, 2016). The non-experimental approach examined correlations between variables of a single group between interest, as evidenced by choice, and performance.

Data Collection

The first step in the data collection was to conduct an a priori power analysis. Power is the probability of detecting a statistical result when there are differences

between groups or relationships between variables (Larson-Hall, 2016). Sufficient power is needed to ensure that real differences are discovered and a lack of power leads to instability in results. For this study, power was calculated to make an informed guess about how effect size and sample size would affect the results. Based on these calculations, full power (100%), required the investigator to include a minimum of 250 subjects.

The next step included the identification of Texas teachers who were currently assigned as a fourth or fifth grade teacher. Teachers were recruited through social media posts, e-mail, text message, and via telephone. When applicable, a cover letter and flier explaining details of the study were provided to colleagues and friends of the principal investigator. These materials were forwarded to potential participants to recruit and inform new persons of the details of the study. Participants were encouraged to contact the principal investigator by phone or e-mail. Follow-up contact was made by the same modes of communication and through face-to-face meetings. Testing materials were copied and distributed to teachers through face-to-face meetings once informed consent was collected from participants.

On average, participants were given two weeks to complete the testing with their classes and return the completed items. All data were collected from test scores and surveys gathered over a total of six weeks in the fall of 2019. The number of items answered correctly on each test represented the subject's score. The principal investigator and the respective subject's teachers scored all tests. Teachers were not required to score testing materials; however, some teachers chose to grade some assessments. A completed testing packet comprised: two completed reading

comprehension tests (one categorized as the ‘student-selected’ passage and another categorized as the ‘teacher-selected’ passage), a student demographic inventory sheet and an interest survey. Reading comprehension tests that were missing answer choices were counted as incorrect. Demographic information missing from the profile was followed up on with the teacher. The scores and survey responses were input into a spreadsheet and assigned codes and values, based on variable information. Finally, this coded data was imported for statistical analysis.

Data Analysis

The data collected for this study were used in three ways. To address the first question concerning the impact of choice on reading comprehension scores, an ANCOVA was conducted to compare the impact of choice and no-choice (independent variables) on the reading assessment scores (dependent variable) for fourth and fifth grade students. An ANCOVA is a statistical procedure used to compare the variances of groups while controlling for potentially confounding variables. Specifically, the variance within a group to the variance between the groups to determine if that value is large enough to designate the group from different populations (Larson-Hall, 2016). In this case, the impact of either choice would be assessed by the significance in the difference of scores for the groups of students. This type of statistical analysis produces an F-ratio, which compares the amount of systematic and unsystematic variance in the data (Larson-Hall, 2016).

Further analysis included post-hoc tests or one-way ANOVA with planned comparisons to determine whether any of the mean scores were different. The investigator was interested in comparisons between grade, teacher, and gender. Follow-

up one-way ANOVA testing aided in the isolation of grade and teacher as covariates in a final comparison to determine the impact of choice. These quantitative data were analyzed with an analysis of covariance (ANCOVA) of the choice condition scores, controlling for grade and teacher.

For practical significance, a *t*-test was conducted in order to calculate mean difference effect sizes (Cohen's *d*). A paired samples *t*-test was used to analyze dependent measures. For this analysis, the no-choice and choice groups scores were treated as the pre-and post-measurements and the scores were compared. This information was used to demonstrate the magnitude of the effect on choice. It is calculated as the difference between the pooled mean scores of the groups divided by the amount of variation across groups by using the standard deviation (Larson-Hall, 2016). An effect size is reported as Cohen's *d*. A large effect is considered greater than .80, a medium effect is .50, and greater than .20 is a small effect. Values below .20 are considered negligible effects.

The final step in the analysis addressed the relationship between interest (independent variable) and performance (dependent variable) on the reading passages. With the data from the student survey, the investigator analyzed whether these variables are related based on Pearson's *r*. A correlation test is appropriate when there could be a pattern of relationships among data and the variables are continuous (Larson-Hall, 2016). In the case of the survey, a Likert scale was used to collect information about levels of student interest in reading and a correlation test was conducted to determine if there was a relationship between reading preference and performance.

Summary of Methodology

This section outlined an account of the participants, subjects, procedures, materials, methods, data collection, and analysis utilized for this study. Specifically, this study employed a statistical test (ANCOVA) to determine statistical differences of choice on reading comprehension scores from the STAAR test passages of students in fourth and fifth grades. A discussion of the validity and reliability of the STAAR instrument was presented. Additionally, the study explored the relationship between interest, as evidenced by choice, and performance on standardized reading comprehension tests using a correlations test. Participants were selected among Texas teachers, using a snowball sampling technique. The final sample of teacher participants was described, along with the individual subjects that completed the study. A series of statistical procedures were used to analyze data collected in this study.

CHAPTER IV

Results

Presentation and Analysis of Data

The purpose of this study was to evaluate how measures of choice ('teacher-selected' and 'student-selected') compared as predictors of reading comprehension and how these vary as a function of achievement for fourth and fifth-grade students during the 2019-2020 school year. The study quantitatively examined this relationship between measures of choice and the standardized reading comprehension assessment with fourth and fifth grade students using an ANCOVA, post-hoc tests, a mean difference effect size comparison, and descriptive statistics. In addition, the study examined the relationship between interest and reading, secondary to the provision of choice with a correlational test. Data analysis was presented in the methodology section of chapter III. This chapter provides the results of the analysis for the following research questions:

Research Questions

1. Are there significant differences in reading comprehension assessment scores when reading 'teacher-selected' passages versus 'student-selected' passages?
2. Is there a relationship between interest and performance on standardized reading comprehension assessments?

Results are described in four sections. In the first section, descriptive statistics are provided to understand the sample used in this study. The second section contains an analysis for research question one, including (1) an examination of normality of the data, (2) an examination of pre-experimental differences between groups, and (3) an analysis of covariance (ANCOVA), used to determine differences between the choice conditions

on the subjects. The third section contains pretest to posttest mean difference effect sizes for all outcome measures. Lastly, the fourth section is aimed at answering research question two using a correlation to explain the relationship between interest, as evidenced by choice, and performance on standardized reading comprehension assessments.

Description of the Sample

Participants. The participants of the study were Texas teachers who were employed in the role of English/Language Arts, fourth or fifth grade teacher for the 2019-20 school year. Teachers were recruited or contacted through social media, e-mail, text message, and by telephone. Twenty teachers initially responded and/or expressed an interest in the study. Nine of these teachers were unable to participate due to various time restraints, curricular obligations, and/or district plans to participate in other curriculum-based assessments. A total of 11 teachers consented to participate and distributed the assessments and surveys to their respective students across five schools (see Table 3). Sixty three percent of the teachers were Caucasian, 36% were African American, and all teachers were female. Two of the teachers reported having 25 and 26 years of experience and another two teachers had 18 years of experience. An additional two teachers reported having taught for 12 and 13 years, respectively. The remaining five teachers had less than ten years of experience. Seven of the teacher participants held bachelor's degrees, while the remaining four had advanced degrees. Of the 11 teachers, only the student data from 10 of the teachers were included based on packet completion criterion.

Table 3

Demographic Information of the Participants

Teacher	Years Taught	Highest Degree Earned	Current Grade
1	26	Bachelor	5
2	25	Graduate	4
3	18	Bachelor	4
4	18	Graduate	4
5	13	Bachelor	5
6	12	Bachelor	4
7	9	Bachelor	4
8	9	Bachelor	5
9	7	Graduate	4
10	7	Graduate	4
11	4	Bachelor	5

Subjects. A total of 372 students took part in the study through completion of all or some parts of the assessments or survey. Eighty-eight percent of the students attended a public elementary school and the remaining 12% of the students were enrolled at a charter school. From this sample, 254 fourth (N = 160) and fifth grade (N = 94) student responses were analyzed from each group, based on completion of the full assessment packet. A completed testing packet comprised two reading comprehension tests (one designated as the 'student-selected' passage and another designated as the 'teacher-selected' passage), a student demographic inventory sheet and an interest survey. These students ranged in age from eight to twelve years old. These groups provided the quantitative data analyzed for this study. Table 4 includes information about the school sites that participated in this study.

Analysis

Descriptive statistics for choice and no-choice groups. Each type of choice group totaled 254 students. The no-choice population totaled 254 students with a mean score of 73.87 and a standard deviation of 25.54. Students in the choice group, also, totaled 254, with a mean score of 74.77 and a standard deviation of 23.68. Means, standard deviations and confidence intervals are included in Table 4.

Table 4

Descriptive Statistics for Choice and No-choice Groups

Condition	*N*	*M*	CI	*SD*s	Min.	Max.
No-Choice	254	73.87	[70.71, 77.03]	25.54	.00	100
Choice	254	74.77	[71.84, 77.70]	23.68	.00	100

Assumptions. Paramount to statistical assessments is the satisfaction of assumptions of data used in the analysis which would otherwise alter the conclusions of the research. Tests of normality included a Quantile-Quantile (Q-Q) plot, homogeneity of variance (Levene's test and numerical output of the standard deviations), and box plots. Q-Q plots plot the quantiles of the data under consideration against the quantiles of the normal distribution. If both the sampling and normal distribution are similar, the points should fall in a straight line. Visual inspection of the Q-Q plot demonstrated the data were normally distributed around the mean. Similarly, inspection of the box plots provided a secondary way of checking variances. Based on a review of the lengths of the box plots representing the choice and no-choice group, the distributions contained similar

variances. Several scores fell outside of the mean; however, these outliers were not eliminated as it would violate the objectivity of the study, could mask another outlier, and become a violation of the assumption of normality. Finally, elimination of these outliers, based on student performance would not capture the ecological make-up of the classroom. Q-Q plots and box plots are shown in Figure 1 and Figure 2.

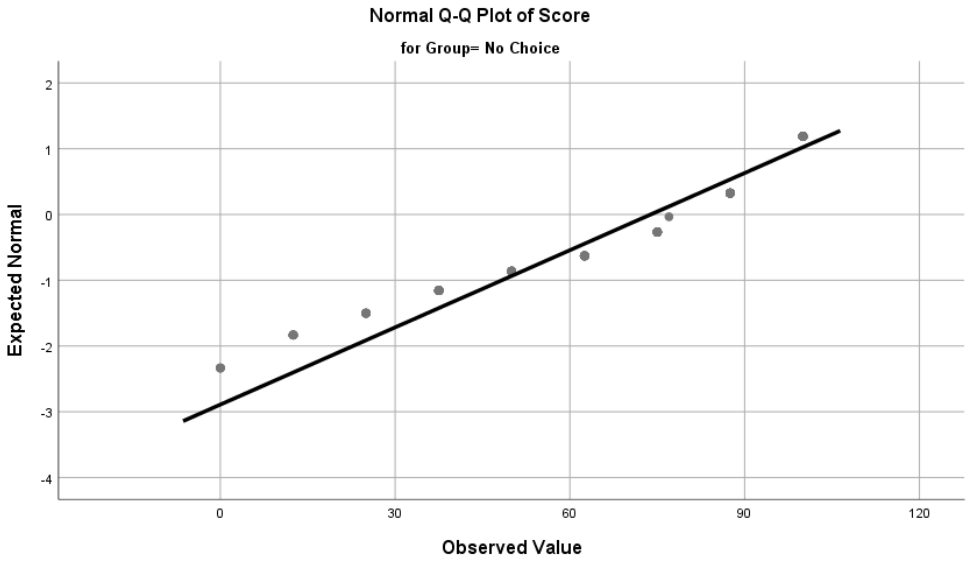


Figure 1. Normal Q-Q Plot of the No-Choice Group.

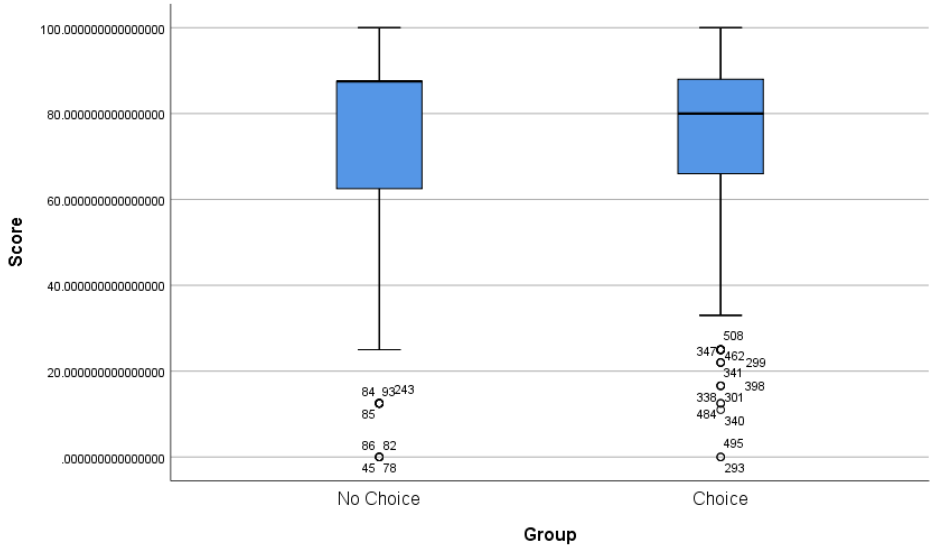


Figure 2. Box plots for No-Choice and Choice Groups.

Homogeneity of variance. A Levene's test was performed to determine homogeneity of variance of the scores under the no-choice condition, as a control or pre-existing condition, within a one-way ANOVA. The Levene's statistic is significant at $p = < .05$. In this case, $p = 0.00$, which means that there was variability in the no-choice condition, and thus results were interpreted using corrected values for homogeneity of variances not assumed.

In order to determine pre-experimental differences based on school, gender, grade, and teacher, a series of one-way ANOVAs were employed. There was a statistically significant difference, based on school, $F(3, 253) = 16.28$, $p = .000$. Post-hoc test analyses using the Tamhane post hoc criterion for significance indicated the average mean scores were significantly lower in school 30 (M= 53.12, SD = 32.72) than in School 13 (M= 65.81, SD = 31.54), School 14 (M = 78.68, SD = 17.97), School 15 (M = 78.12, SD = 22.59) and School 20 (M = 83.27, SD = 13.35. The descriptive statistics are shown in Table 5 and the ANOVA is shown in Table 6.

Table 5

Descriptive Statistics by School

Site	*N*	*M*	*SDs*	*SE*	Min.	Max.
13	58	65.81	31.54	4.12	0.00	100.00
14	202	78.68	17.97	1.26	16.60	100.00
15	120	78.12	22.59	2.06	0.00	100.00
20	62	83.27	13.35	1.69	50.00	100.00

Site	*N*	*M*	*SDs*	*SE*	Min.	Max.
30	66	53.14	32.72	4.02	0.00	100.00
Total	508	74.32	24.61	1.09	0.00	100.00

Table 6

ANOVA by School

	SSs	*df*	*M*	*F*	*p.*
Between Groups	44346.09	4	11086.52	21.21	0.00
Within Groups	262892.27	503	522.64		
Total	307238.36	507			

Table 7

TUKEY HSD Post-Hoc Test by School

School		*M*	*SE*	*p.*	95% CI Lower	95% CI Upper
13	14	-12.86*	4.33	< .05	-25.39	-0.33
	15	-12.31	4.62	0.08	-25.60	0.98
	20	-17.45*	4.47	< .01	-30.36	-4.55
	30	12.67	5.77	0.26	-3.79	29.15
14	13	12.86*	4.33	< .05	0.33	25.39
	15	0.55	2.41	1.00	-6.29	7.39
	20	-4.59	2.11	0.27	-10.61	1.42
	30	25.53*	4.22	0.00	13.37	37.69

15	13	12.31	4.62	0.08	-0.98	25.60
	14	-0.55	2.41	1.00	-7.39	6.29
	20	-5.14	2.67	0.43	-12.71	2.42
	30	24.98*	4.52	< .01	12.03	37.94
20	13	17.45*	4.47	< .01	4.55	30.36
	14	4.59	2.11	0.27	-1.42	10.61
	15	5.14	2.67	0.43	-2.42	12.71
	30	30.13*	4.37	< .01	17.58	42.68
30	13	-12.67	5.77	0.26	-29.15	3.79
	14	-25.53*	4.22	< .01	-37.69	-13.37
	15	-24.98*	4.52	< .01	-37.94	-12.03
	20	-30.13*	4.37	<.01	-42.68	-17.58

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

A one-way ANOVA examined differences in scores within the no-choice condition between subjects with gender as the grouping factor. The sample included 254 (*N* = 120 males, *N* = 132 females, *N* = 3, non-specified) fourth-and fifth-grade students. Table 8 includes means and standard deviations for this group. Results of the one-way ANOVA comparing these groups revealed there was no statistically significant difference on the performance of the no-choice test, based on gender $F(2, 253) = .60$, $p = .54$, shown in Table 9.

Table 8

Descriptive Scores for No-Choice by Gender

					95% CI			
	N	*M*	*SDs*	*SE*	Lower	Upper	Min.	Max.

0	120	73.75	24.16	2.20	69.38	78.11	.0	100.0
1	132	73.68	26.89	2.34	69.05	78.32	.0	100.0
3	2	93.75	8.83	6.25	14.33	173.16	87.5	100.0
Total	254	73.87	25.54	1.60	70.71	77.03	.0	100.0

Table 9

ANOVA for No-Choice by Gender

	SS	*df*	*M*	*F*	*p.*
Between Groups	796.45	2	398.22	.60	.54
Within Groups	164355.39	251	654.80		
Total	165151.84	253			

Table 10

TUKEY HSD Post-Hoc Test for No-Choice by Gender

					95% CI	
Gender	Gender	*M*	*SE*	*p.*	Lower	Upper
0	1	.06	3.22	1.00	-7.54	7.67
	3	-20.00	18.24	.51	-63.01	23.01
1	0	-.06	3.22	1.00	-7.67	7.54
	3	-20.06	18.23	.51	-63.04	22.92
3	0	20.00	18.24	.51	-23.01	63.01
	1	20.06	18.23	.51	-22.92	63.04

A one-way ANOVA examined differences in scores within the no-choice condition between subjects with grade as the grouping factor. The sample included 254 fourth (N = 160) and fifth (N = 94) grade students. Table11 includes means and standard deviations for this group. Results of the one-way ANOVA comparing these groups

revealed there was a statistically significant difference on the performance of the no-choice test, based on grade $F(1, 253) = 14.60$, $p < .01$, shown in Table 12.

Table 11

Descriptive Statistics for No-Choice by Grade

	N	*M*	*SD*	*SE*	95% CI		Min.	Max.
					Lower	Upper		
4	160	78.45	24.23	1.91	74.66	82.23	.0	100.0
5	94	66.09	25.96	2.67	60.77	71.40	.0	100.0
Total	254	73.87	25.54	1.60	70.71	77.03	.0	100.0

Table 12

ANOVA for No-Choice by Grade

	SS	*df*	*M* Square	*F*	*p.*
Between Groups	9045.26	1	9045.26	14.60	< .01
Within Groups	156106.58	252	619.47		
Total	165151.84	253			

A one-way ANOVA examined differences in scores within the no-choice condition between subjects with teacher as the grouping factor. The sample included 254 fourth-and fifth-grade students taught by 11 teachers. Table 13 includes means and standard deviations for this group. Results of the one-way ANOVA comparing the groups revealed there was a statistically significant difference on the performance of the no-choice test, based on the student's teacher $F(9, 253) = 11.36$, $p < .01$. Post-hoc test analyses using Tukey's post hoc criterion for significance indicated the average mean scores was significantly lower for teacher 20 (M= 32.63, SD = 27.83) than teacher 5 (M=

73.56, SD = 19.24), teacher 6 (M = 87.96, SD = 14.07), teacher 7 (M = 73.86, SD = 24.47), teacher 10 (*M* = 87.05, *SD* = 12.01), teacher 11 (*M* = 84.54, *SD* = 20.65) teacher 16 (*M* = 65.08, *SD* = 30.86), teacher 17 (*M* = 75.83, *SD* = 12.90), teacher 18 (*M* = 88.28, *SD* = 13.28), and teacher 20 (*M* = 65.83, *SD* = 34.22). Teachers 18, 10, 6, and 11 have similar means, while Teachers 5, 7, and 17 demonstrated mean scores with the next highest mean. Results of the one-way ANOVA are in Table 14, and the post-hoc results are shown in Table 15.

Table 13

Descriptive Statistics for No-Choice by Teacher

					95% CI			
	N	*M*	*SD*	*SE*	Lower	Upper	Min.	Max.
5	61	73.56	19.24	2.46	68.63	78.49	25.0	100.0
6	27	87.96	14.07	2.70	82.39	93.53	62.5	100.0
7	33	73.86	24.47	4.26	65.18	82.54	12.5	100.0
10	28	87.05	12.01	2.27	82.39	91.71	50.0	100.0
11	12	84.54	20.65	5.96	71.41	97.66	37.5	100.0
16	29	65.08	30.86	5.73	53.34	76.82	.0	100.0
17	15	75.83	12.90	3.33	68.68	82.98	50.0	100.0
18	16	88.28	13.28	3.32	81.20	95.35	50.0	100.0
19	15	65.83	34.22	8.83	46.88	84.78	12.5	100.0
20	18	32.63	27.83	6.56	18.79	46.47	.0	87.5
Total	254	73.87	25.54	1.60	70.71	77.03	.0	100.0

Table 14

ANOVA for No-Choice by Teacher

	SS	*df*	*M*	*F*	*p.*
Between Groups	48788.51	9	5420.94	11.36	< .01
Within Groups	116363.32	244	476.89		
Total	165151.84	253			

Table 15

TUKEY HSD Post-Hoc Test by Teacher

Teacher:	Teacher:	*M*	*SE*	*p.*	95% CI Lower	95% CI Upper
5	6	-14.39	5.04	.12	-30.51	1.72
	7	-.29	4.71	1.00	-15.36	14.77
	10	-13.48	4.98	.17	-29.40	2.43
	11	-10.97	6.89	.85	-32.99	11.04
	16	8.47	4.92	.78	-7.25	24.20
	17	-2.26	6.29	1.00	-22.36	17.83
	18	-14.71	6.13	.33	-34.30	4.87
	19	7.73	6.29	.96	-12.36	27.83
	20	40.92*	5.85	< .01	22.22	59.63
6	5	14.39	5.04	.12	-1.72	30.51
	7	14.09	5.66	.28	-3.99	32.19
	10	.90	5.89	1.00	-17.90	19.71
	11	3.42	7.57	1.00	-20.77	27.61
	16	22.87*	5.84	< .01	4.22	41.52
	17	12.12	7.03	.78	-10.32	34.58
	18	-.31	6.88	1.00	-22.31	21.68
	19	22.12	7.03	.057	-.32	44.58
	20	55.32*	6.64	< .01	34.10	76.54
7	5	.29	4.71	1.00	-14.77	15.36
	6	-14.09	5.66	.28	-32.19	3.99
	10	-13.18	5.61	.36	-31.10	4.72
	11	-10.67	7.36	.91	-34.18	12.83
	16	8.77	5.55	.85	-8.97	26.52
	17	-1.96	6.80	1.00	-23.68	19.74
	18	-14.41	6.65	.48	-35.66	6.82
	19	8.03	6.80	.97	-13.68	29.74
	20	41.22*	6.39	< .01	20.79	61.65
10	5	13.48	4.98	.17	-2.43	29.40
	6	-.90	5.89	1.00	-19.71	17.90
	7	13.18	5.61	.36	-4.72	31.10
	11	2.51	7.53	1.00	-21.54	26.57
	16	21.96*	5.78	< .01	3.49	40.44

	17	11.22	6.98	.84	-11.09	33.53
	18	-1.22	6.84	1.00	-23.08	20.62
	19	21.22	6.98	.07	-1.09	43.53
	20	54.41*	6.59	< .01	33.34	75.48
11	5	10.97	6.89	.85	-11.04	32.99
	6	-3.42	7.57	1.00	-27.61	20.77
	7	10.67	7.36	.91	-12.83	34.18
	10	-2.51	7.53	1.00	-26.57	21.54
	16	19.45	7.49	.22	-4.48	43.39
	17	8.70	8.45	.99	-18.30	35.71
	18	-3.73	8.33	1.00	-30.37	22.89
	19	18.70	8.45	.45	-8.30	45.71
	20	51.90*	8.13	< .01	25.91	77.89
16	5	-8.47	4.92	.78	-24.20	7.25
	6	-22.87*	5.84	< .01	-41.52	-4.22
	7	-8.77	5.55	.85	-26.52	8.97
	10	-21.96*	5.78	< .01	-40.44	-3.49
	11	-19.45	7.49	.22	-43.39	4.48
	17	-10.74	6.94	.87	-32.92	11.43
	18	-23.19*	6.80	< .05	-44.91	-1.47
	19	-.74	6.94	1.00	-22.92	21.43
	20	32.44*	6.55	< .01	11.52	53.37
17	5	2.26	6.29	1.00	-17.83	22.36
	6	-12.12	7.03	.78	-34.58	10.32
	7	1.96	6.80	1.00	-19.74	23.68
	10	-11.22	6.98	.84	-33.53	11.09
	11	-8.70	8.45	.99	-35.71	18.30
	16	10.74	6.94	.87	-11.43	32.92
	18	-12.44	7.84	.85	-37.51	12.61
	19	10.00	7.97	.96	-15.46	35.46
	20	43.19*	7.63	< .01	18.81	67.57
18	5	14.71	6.13	.33	-4.87	34.30
	6	.31	6.88	1.00	-21.68	22.31
	7	14.41	6.65	.48	-6.82	35.66
	10	1.22	6.84	1.00	-20.62	23.08
	11	3.73	8.33	1.00	-22.89	30.37
	16	23.19*	6.80	< .05	1.47	44.91

	17	12.44	7.84	.85	-12.61	37.51
	19	22.44	7.84	.12	-2.61	47.51
	20	55.64*	7.50	< .01	31.68	79.60
19	5	-7.73	6.29	.96	-27.83	12.36
	6	-22.12	7.03	.05	-44.58	.32
	7	-8.03	6.80	.97	-29.74	13.68
	10	-21.22	6.98	.07	-43.53	1.09
	11	-18.70	8.45	.45	-45.71	8.30
	16	.74	6.94	1.00	-21.43	22.92
	17	-10.00	7.97	.962	-35.463	15.46
	18	-22.44	7.84	.12	-47.51	2.61
	20	33.19*	7.63	< .01	8.81	57.57
20	5	-40.92*	5.85	< .01	-59.63	-22.22
	6	-55.32*	6.64	< .01	-76.54	-34.10
	7	-41.22*	6.39	< .01	-61.65	-20.79
	10	-54.41*	6.59	< .01	-75.48	-33.34
	11	-51.90*	8.13	< .01	-77.89	-25.91
	16	-32.44*	6.55	< .01	-53.37	-11.52
	17	-43.19*	7.63	< .01	-67.57	-18.81
	18	-55.64*	7.50	< .01	-79.60	-31.68
	19	-33.19*	7.63	< .01	-57.57	-8.81

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

Summary of Preliminary Differences.

A series of one-way ANOVAs and Post-hoc analyses revealed a significant difference between no-choice condition scores, based on grade $F(1, 253) = 14.60$, $p = .00$ and teacher $F(9, 253) = 11.36$, $p < .001$, as well as school, but school was not included as a covariate as the teacher accounted for these differences. Therefore, teacher and grade were used as covariates in the ANCOVA to answer research question one.

Effects of Choice on Reading Performance

The first question sought to determine statistical differences in reading comprehension scores between a choice and no-choice condition. Results of the first research question determined whether students achieved higher reading comprehension scores on a standardized reading assessment when provided a choice in assessment, compared to a no-choice opportunity of assessment.

Analysis of covariance is used to test the effects of categorical variables on a continuous dependent variable. Therefore, an ANCOVA was conducted to determine the between subject effects for the no choice and choice conditions. The sample included both groups (N = 254 choice and N = 254 no choice). The dependent variable used for the analysis was the standardized reading assessment scores for both groups. The covariates were teacher and grade because these demonstrated statistically significant differences in reading comprehension. The independent variables were the groups for choice and no-choice. Results yielded no main effects for the choice condition, $F(1, 508) = .18$, $p = .67$), after adjusting for grade and teacher, indicating that the provision of choice failed to reach significance on performance outcomes. The results of the ANCOVA are summarized in Table 16 and Table 17.

Table 16

Between-Subjects Effects

	SS	*df*	*MS*	*F*	*p.*
Corrected Model	24082.65[a]	3	8027.55	14.28	< .01
Intercept	86276.31	1	86276.31	153.56	< .01
Teacher	17424.52	1	17424.52	31.01	< .01

Grade	9957.65	1	9957.65	17.72	< .01
Group	102.24	1	102.24	.18	.67
Error	283155.70	504	561.81		
Total	3113505.09	508			
Corrected Total	307238.36	507			

a. R Squared = .08 (Adjusted R Squared = .07)

Table 17

ANCOVA Group Effect

	SS	*df*	*M*	*F*	*p.*
Contrast	102.24	1	102.24	.18	.670
Error	283155.70	504	561.81		

The *F* tests the effect of Group. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

Effect Size Analysis

An effect size is a measure of the weight of the impact of the independent variable on the dependent variable (Larson-Hall, 2016). Effect sizes are important to the question of impact and they demonstrate the size of the impact. Cohen's *d* is an effect size that measures the difference between two independent sample means and expresses how large the difference is in standard deviations. The guidelines for Cohen's *d* are as follows: *d* = 0.2 (small effect), *d* = 0.5 (medium effect) and *d* = 0.8 (large effect).

Q-Q plots were visually inspected for pre and post test data. Differences between the no-choice and choice conditions were considered normally distributed, based on Q-Q plots. Box plots were generated, which showed only four outliers. These outliers were

not eliminated, given the gains demonstrated by three of the four subjects, as reported by differences between the pre and post scores.

Given the practical significance of calculating the effect size, a series of paired samples *t*-tests were conducted to obtain this measure and gain more information about the effects. Like the previous analysis, the no-choice condition was treated as a pre-test, while the choice condition was input as post-test data. The paired samples *t*-test effect sizes were computed from the pooled mean and standard deviations, based on grade, reading preference, gender, school, and teacher. The *t*-test effect size is most often calculated with Cohen's *d* (Larson-Hall, 2016). Results of the effects are shown in table 18.

Table 18

Paired Samples and Effect Sizes by Grade

		Pre/Post Effect Size of Choice/No-Choice
Gender		
	Male	.11
	Female	.00
Teacher		
	1	.18
	2	-.28
	3	-.12
	4	-.26
	5	-.42
	6	.07
	7	.92
	8	-.49
	9	.17
	10	.84
School		
	1	.28
	2	.01
	3	-.17
	4	.12

Grade		
	4	.13
	5	.38
Reading Preference		
	Hate	.13
	Like	.08
	Love	.03

Note: ES determined by Cohen's d, .2 = small, .5 = moderate, .8 = large; Positive and negative effects marked with +/-.

Correlations

The second research question investigated the relationship between reading preferences and interest, as a provision of choice. A correlation measures the linear relationship between two quantitative variables. Data from the student survey was assigned a quantitative value and a simple correlation was conducted. Descriptive statistics were collected on the variables of choice, reading preference (1 = Love, 2 = Like, and 3 = Do not like), school, no-choice, and indicators of student preference for their chosen passage (1 = Yes, 2 = No, and 3 = No response). Mean scores and standard deviations are listed in Table 19.

Table 19

Descriptive Statistics for Correlations

	M	*SD*	*N*
No-Choice	73.87	25.54	254
Choice	74.77	23.68	254
Reading Preference	1.69	.58	254
Chosen passage: Like/Yes	1.10	.34	254
Gender	.54	.54	254

Teacher	10.87	5.58	254
Grade	4.37	.48	254

Relationships. Several correlations were generated from the data, which are shown in Table 21. Notably, there is a significant correlation between choice and no-choice at $p <. 01$, and the relationship was moderate to high. There is a significant correlation ($p < .05$) between no-choice and whether students liked their passage, but there was little to any relationship. Scores for no-choice were negatively correlated with student's grade level, however there was little if any relationship. There was a significant correlation ($p < .01$) between the reading preference and whether the students liked their chosen passage, however, there was little to no relationship. Teacher and scores for no-choice were significantly correlated at $p < .01$, indicating a relationship that is approaching low in strength. In addition, teacher and reading preference were significantly negatively correlated ($p < .01$), but there was little to no relationship between these variables. Scores for choice and teacher were significantly negatively correlated ($p < .01$), as well. However, there was little to no relationship among these variables. There was a significant relationship ($p < .05$) between grade and chosen passage, which is considered minuscule. Students' preference for reading was significantly correlated with the passage chosen, ($p <. 01$), however, according to the Pearson's correlation coefficient, .17 is considered a negligible relationship. The remaining relationships between the variables were not significant.

Table 21

Correlations

Variables	1	2	3	4	5	6	7
1. Scores for No-Choice	-						
2. Reading Preference	-.09	-					
3. Scores for Choice	.64**	-.04	-				
4. Gender	.02	.04	-.05	-			
5. Teacher	-.25**	-.19**	-.17**	.00	-		
6. Chosen Passage Like/Yes	-.14*	.17**	-.18	.05	.06	-	
7. Grade	-.23**	.09	-.05	-.04	-.15*	.13*	-

Note: **. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Summary of Results

This chapter presented an analysis of the data. Data from standardized reading passages, teacher information, and student surveys were used to explore two main research questions. Several statistical tests were used in the analysis, including (1) an analysis of variance to assess pre-experimental differences, (2) an analysis of covariance, with teacher and grade as covariates, (3) an effect size analysis to measure the overall impact of the choice conditions using paired samples *t*-tests, and (4) a simple correlation to assess the relationship between reading preference and interest, as a provision of choice.

The one-way ANOVA operations uncovered statistically significant differences on the pre-existing condition of no-choice at the teacher and grade levels. These variables were compared to the choice scores in an ANCOVA, which yielded no statistically significant differences. Likewise, an examination of the effect size found a

very small effect on choice for fifth grade students. Overall, there was a negligible effect across reading preference, but the effect for students who indicated they did not like reading was slightly greater than those that reported they enjoyed reading to a degree. Lastly, the simple correlation showed a negligible relationship between students' preference for reading.

CHAPTER V

The purpose of this chapter is to summarize the study, discuss and interpret the findings from the data analysis, and provide contextual relevance for this study relative to classroom practices. The first section of this chapter provides an overview of the study, including the literature supporting the ideas for the design of the study. Next, the results are organized by research questions. Finally, the chapter concludes with implications for the classroom and recommendations for future research.

Overview of the Study

Data from standardized tests have shown either stagnation and or a decline in the achievement of students across multiple grade levels, nationally (e.g. NAEP). Presumably this data provides the most widely used marker for student success; therefore, it is important to focus research on ways to improve these areas. Theroux (2010) noted the improvement of adolescent reading is crucial in the overall efforts to increase high school completion and in the preparation of students for college and career. Moreover, Munster and Haines (2019) highlighted the consequences of these depressed data points on schools that included a narrowed curriculum, higher test anxiety, limited professional discretion, and fewer opportunities for instructional decisions. Schools that have perpetually underperformed on standardized assessments have been labeled for their performance, leading to more stringent efforts to reduce instruction to skill-based lessons and school-wide remedial mandates (Munster & Haines, 2019). Accordingly, this narrowed focus on testing strategies suppresses student engagement (Sanacore, 2002).

The theoretical underpinnings of choice connect to the affective variables of reading comprehension, including motivation, interest, and engagement. The provision

of choice in instructional materials, including assessment, is one way to maintain student engagement and motivation to ensure virtual success. Research has shown the beneficial properties of providing students with choices in classroom instructional material (Stefanou et al., 2004). While decades of research have propelled the instructional practices implemented to ensure that students are being educated with scientifically based measures proven to increase achievement, standardized assessments have failed to keep up with these same changes in the field (NAEP, 2017; Sanacore, 2002; Slomp, 2008). In response to this need for change in standardized assessments, this study was devised to revisit the question of choice, like the previous study conducted by the NAEP in 1994. The study added to the field by focusing on non-traditional assessment means interpreted as the provision of choice. As there is limited information on this topic, there is a need for more research related to the permeable nature of choice as well as the role of choice in reading comprehension. The following questions led the research:

1. Are there significant differences in reading comprehension assessment scores when reading 'teacher-selected' passages versus 'student-selected' passages?
2. Is there a relationship between interest and performance on standardized reading comprehension assessments?

The investigator employed a quasi-experimental design to ascertain the effects of choice on student achievement using a criterion-referenced assessment and a non-experimental design to investigate the relationship between choice, interest, and performance on standardized reading assessments. The quantitative data collected were used to answer the first research question and analyzed using an analysis of variance (ANOVA) which compared the variances between two groups (Larson-Hall, 2016). In

addition, the study examined the relationship between interest and reading, secondary to the provision of choice with a correlational test.

The State of Texas Assessments of Academic Readiness (STAAR) released from previous administration years (2013-2018) was used to assess the performance of fourth and fifth grade students. This instrument is a reliable and valid, criterion-referenced test used to assess an individual's abilities in isolation or without the impact of his peers (Wortham, 2005). All subjects read two passages and completed the corresponding criterion-referenced test questions, related to the passage. One passage was pre-selected and administered to all students. The second passage was selected by the subject from a set of three passages. Students also completed a one-question survey related to their preference for their chosen passage. Subjects completed all testing activity in their classrooms and under the administration of their respective teachers.

All results were based on data collected from fourth and fifth grade subjects' (N = 254) standardized reading comprehension scores and survey responses. Data collection extended a six-week span across five total school sites, located in various communities throughout the southeastern regions of Texas. Texas fourth and fifth grade teachers employed in that capacity for the 2019-2020 school year were considered participants and recruited through various snowball sampling techniques. A total of 11 teachers participated in the study, representing urban, suburban, and rural school communities. However, the student data from one teacher was not included in the analysis, due to insufficient data packets.

Performance data from both the 'teacher-selected' (no-choice) and 'student-selected' (choice) standardized reading assessments were used in the statistical analyses.

Descriptive statistics and homogeneity of variance tests were completed as preliminary statistical measures. Further analysis included post-hoc tests or one-way ANOVAs with planned comparisons, to determine whether any of the mean scores were different. The investigator was interested in comparisons between grade, teacher, and gender. Follow-up one-way ANOVA testing aided in the isolation of grade and teacher as covariates in a final comparison to determine the impact of choice. This quantitative data was analyzed with an analysis of covariance (ANCOVA) of the choice condition scores, controlling for grade and teacher. Next, a paired samples *t*-test was conducted in order to calculate mean difference effect sizes (Cohen's *d*). For this analysis, the no-choice and choice group scores were treated as the pre-and post-measurements and the scores were compared. This information was used to demonstrate the magnitude of the effect on choice. The final step in the analysis addressed the relationship between interest (independent variable) and performance (dependent variable) on the reading passages. A correlation test was conducted to determine if there was a relationship between reading preference and performance. The investigator analyzed whether these variables are related based on Pearson's *r*, using data from the student survey.

Discussion of Findings

Research question 1. This study was designed to evaluate how measures of choice ('teacher-selected' and 'student-selected') compared as predictors of reading comprehension and how these vary as a function of achievement for fourth and fifth grade students. In addition, the study quantitatively examined the relationship between measures of choice and the standardized reading comprehension assessment with fourth and fifth grade students. While the testing instrument did not differ much in task and

format between the choice and no-choice administration, it had been predicted that the provision of choice could increase the reading performance outcome on a standardized reading passage based on the literature review. The investigator posited that when students can express themselves through choice, they assign ownership and value to the task, personally invest in the activity, and demonstrate their interest through their decisions. Moreover, interested and motivated readers are more likely to persist and engage with text leading to deeper levels of comprehension (Deci & Ryan, 1985; Fulmer et al., 2014; Krapp, 2002; Reeve & Tseng, 2011; NAEP, 1994). Therefore, choice would be a motivating factor in the students' comprehension and result in a higher score, when compared to the no-choice performance. When teacher and grade level were controlled, there were no posttest mean differences, when comparing the choice and no-choice performances. Likewise, the effect sizes that emerged from these analyses were small, as reported in Table 19.

The overall findings of this study were consistent with the findings of the National Assessment of Educational Progress (NAEP) in 1994, which also was designed to compare the performance of students who selected a story with those assigned a story. The choice group among the eighth graders scored lower than the no-choice group in that study. Similarly, Wolters et al. (2017) found similar results in their investigation of whether a motivational treatment would impact students' performance on a standardized reading performance. In a related study, Marinak and Gambrell (2007) concluded choice of a book or choice of a token neither enhanced nor undermined subsequent reading motivation.

Differences in assessment scores. This study offered some evidence of the impact that choice had across school sites based on teacher and grade level, as evidenced by achievement scores and correlational data. Specifically, it was noted that students at suburban school sites performed significantly higher (50-60%) on both choice and non-choice assessments, when compared to students at urban school sites. One explanation for these differences could be the reading abilities of the students. Students were provided a standardized reading assessment intended for a student in a lower grade, to alleviate the anticipated decoding and comprehension restraints of grade level reading. The data, also, showed the sample included many students who did not obtain minimal proficiency scores on a reading task that was below grade level by design. Despite the purposeful design, students attending schools located in urban areas were not able to obtain a score above 25%, in some cases. Students with a history of reading difficulties and poor performance on tasks that require comprehension skills often do not possess the competence-related motivational beliefs needed to overcome these deficits given such a brief intervention (Wolters et al., 2017).

Similarly, Mulvey (2009) noted the curricular, organizational, and teaching differential practices of educational facilities across divisions of social class, culture, and race. For example, one Head Start program, positioned to mediate the non-cognitive factors of education such as low-socioeconomics, language barriers, and access, offered fewer choices and less autonomy than the preschool program offered within the same area. Teachers from the preschool were more liberal in their teaching style and content, while the Head Start teachers used a more authoritarian approach. The differences between these two approaches was evidenced by lower student outcomes in the Head

Start program. Differences in teaching approaches can affect student outcomes and, in relation to this study and the provision of choice, similar connections could be made for the differences between school sites.

Conversely, Flowerday et al. (2004) explained a lack of variance between choice and no-choice scores among proficient readers would not be uncommon. They reasoned that the reading level of the individual could offset the effects of affective variables such as interest and choice. Proficient readers are capable of advanced planning, goal setting, and analyses resulting in higher levels of achievement (Zimmerman, 2013). Zimmerman and Schunk (2008) described these learners as self-evaluative, as having an established and productive environment for learning, and expending better effort when compared to their counterparts. Becker et al. (2010) confirmed similar patterns among poor readers in their longitudinal study of third graders. Findings demonstrated levels of reading achievement remained constant when reading ability was controlled. Students who were poor readers in third grade, remained poor readers in the sixth grade; however, this relationship was disrupted by intrinsic motivation. In fact, the data confirmed that fourth grade intrinsic reading motivation was positively related to reading achievement in sixth grade (Becker et al., 2010). Consequently, the impact of choice would be minimized by reading ability.

The readability of the instrument (STAAR Reading test) could account for an additional lack of variance and for the larger gaps in achievement scores. Several years ago, Szabo and Sinclair (2012) investigated whether the STAAR released reading passages, intended for the use of third through eighth grades, were written at the appropriate grade levels. Their sample included poems, fictional plays, non-fiction

narratives, and folktales. Using several readability formulas: Fry, Raygor, Gunning Fog, Flesch-Kincaid, and SMOG, they found that, on average, the passages for third and fourth graders were written two and three grade levels too high, respectively. Immediate implications of such a finding demonstrate the difficulty for students to be successful on such standardized tests. However, a pattern of unsuccessful attempts for students could be detrimental to their reading motivation, interest, and engagement regardless of choice. Still, the authors suggested teachers differentiate, provide choice through exposure of reading materials and topics, use project-based learning to promote deeper understanding of topics, resist the urge to "teach to the test", ask richer questions, and teach direct inferencing strategies to prepare them for varying questions offered on the STAAR test.

Revisiting the theoretical framework. Piaget (1973) envisioned an autonomous student who adapts and fluidly manipulates information over time. This experienced student constructs knowledge from learned experiences and materials read, demonstrating comprehension through performance. Following each reading experience, students gain more control over learned and prior knowledge (Dewey, 1990; Rosenblatt, 1994) which is reinforced with choice, given the associated interest and motivation that accompanies choice (Tracey & Morrow, 2012).

When interpreting the results, the researcher is reminded of the principle components of the SDT framework and the principles of the choice provision. Self-determination theory (Deci & Ryan, 2000) highlights self-control and determination, which provides the opportunity to make choices. Self-determination is only accomplished when persons are internally motivated to initiate an activity based on an expected outcome. The regulation and expression of the self, anticipation of the activity

being personally beneficial, and dominion over decisions are factors that make a task agreeable to undertake. Students reported high levels of interest in reading and in the standardized reading passage they chose, in this study. Additionally, the provision of choice was meant to spark interest and motivation toward the activity, however the results are not able to explain why this combination did not lead to better performances on the choice activity. Subsequently, these findings should not be understood to be as contrary to research on motivation and comprehension (NAEP, 1997).

Choice, for learning, must be relevant and meaningful (Cordova & Lepper, 1996; Patall, et al., 2008). Although assessments are normal processes within the educational cycle, perhaps these assessments did not provide enough meaningful relevance, as defined in SDT. One explanation for this misalignment could be the relevance of the choice to the individual. Given this assessment was not a graded item and/or part of the student's instruction, the testing choices may not have been valuable to the extent that students' performance would differ between the choice and no-choice assessments. Katz and Assor (2007) warned of the need for all three tenets of SDT, autonomy, personal relevance, and competence, to be satisfied before the benefits of choice are actualized. From this perspective, these findings underscore the necessity of this trifecta within the design.

The assumptions of choice are synonymous with liberty and freedom. Choices bring about privileges, immunities, rights, and abilities, which are not otherwise available in a non-choice condition. In the context of this study, the provision of choice was expected to enhance the performance of the student on the standardized reading comprehension passage. However, both Iyengar and Lepper (2000) and Brooks and

Young (2013) found some choices could be de-motivating for students, under certain conditions. In the latter study, Brooks and Young (2011) recommended choice-making opportunities remain consistent and equal across multiple classroom activities to bring about positive and beneficial outcomes. For the present study, the provision of choice on passages could be an anomaly for students, altering the value, expectations, and impact of choice.

Research question 2. The second question asked is there a relationship between interest and performance on standardized reading comprehension assessments. The results showed scores for choice and no-choice were moderately to highly correlated, indicating that whether students had a choice, students were likely to perform similarly on both passages. While there were correlations between other variables, the Pearson's correlation coefficients demonstrated low to negligible relationships. Students' gender, teacher, fondness for reading, or interest in the passage had little to no influence on their overall performance on the assessment. In this study, the most meaningful correlation was how well the students performed on the standardized test, relative to the condition, which is more indicative of the validity of the test. All variables considered, these standardized tests are constructed to measure reading ability and a student's ability to use these skills within this assessment format.

Previous studies have suggested that the opportunity to designate choice and preferences (interest) triggers inner motivational resources during the learning activity (Deci & Ryan, 1985; Krapp, 2002; Reeve & Tseng, 2011), and is expected to lead to greater engagement and deeper learning (Fulmer et al., 2014). Likewise, interest facilitates reader's comprehension, attitude and recall (Hidi, 2001; Hidi & Harackiewicz,

2000; NAEP, 1994). Both situational and personal interest were elicited in this study. Categorically, the provision of choice afforded both expressions of interest.

Secondary to the choice provision, the student's rated perception of reading represented their level of interest in the reading task. The relationship between engagement and reading is fueled by interest and desire (Becker et al., 2010). As students' perceptions of reading increases, they read more frequently, deeply, and with more refined skills (Becker et al., 2010). Furthermore, highly interested students exhibit complex cognitive command, high coherence of recalled content, and general enjoyment of the texts among students with average and lower than average ability (Guthrie et al., 2007). Additional evidence has shown that the positive impact of interest on text learning is largely independent of text difficulty, type of learning test (e.g., multiple-choice vs recall), reading ability, and age or grade level (Schiefele, 1996). Schiefele discovered that interest appears to mediate the effect of motivation on reading comprehension. The results from this study were not consistent with the previous studies on reading interest and attitude.

Correspondingly, interest was not correlated to performance on reading assessment. These results more closely aligned with Tarchi's (2017) explorations of the integration of cognitive and motivational factors on reading comprehension, using an expository history text in secondary school students (N = 146). In this study, he defined cognitive factors as inference-making skills and metacognition, while motivational factors included self-efficacy, intrinsic and extrinsic motivation, social aspects of motivation, and topic interest. He hypothesized that (a) topic interest contributed to reading comprehension, independently from reading motivation; (b) topic interest would

mediate the effect of reading motivation on reading comprehension; (c) students' cognitive and motivational skills would independently contribute to reading comprehension, and (d) motivational variables would play an energizing role in the relationship between cognitive variables and reading comprehension (Tarchi, 2017).

The results contradicted the hypothesis and that of previous findings (Schiefele, 1996, 1999), in that interest did not bring any independent contribution to reading comprehension performances and it did not mediate the effect of motivation on reading comprehension (Tarchi, 2017). The author gave several explanations for these results, including how the variable was measured and how it was associated with prior knowledge, which was controlled in this study. In relation to the third hypotheses, the results were more aligned. Lastly, the data reaffirmed the importance of intrinsic motivation as a mediating factor in reading comprehension (Deci & Ryan, 1985). In relation to the present study, motivation may have skewed the results, as well. It is possible that the reading instrument (STAAR test) was not motivating for students, resulting in underwhelming scores.

Reading attitude is important because it affects reading ability and directs reading behavior. A student's perception of reading drives their reading-based tasks. Poor perception of reading will likely lead to less time spent reading (Schiefele et al., 2012). Interestingly, the students who rated reading as a preferred task and liked their chosen passage, did not obtain significantly higher scores on the choice assessment. More research is needed to examine the motivating qualities of standardized assessments.

Classroom applications and implications

Studies on choice have noted multiple ways student choices are inhibited and diminished to trivial pursuits (Katz & Assor, 2007; Rogat et al., 2014; Stefanou et al., 2004), mainly due to practical purposes and misinterpretations of self-determination theory (Deci & Ryan, 1987). Stefanou et al. (2004) recommended a hierarchy of autonomy supportive practices based on the tenets of self-determination theory: competence, relatedness, and autonomy. Specifically, they offered an autonomy-supportive continuum for teachers through three entry points: organizational, procedural, and cognitive. For their part, teachers build the culture of autonomy, scaffolding their roles and transferring responsibility for learning to students. At each level students experience and exercise increasing ownership of their learning. Descriptions of organizational and procedural provisions of choice include student choice of materials, input on procedures and classroom norms, and alternative ways to display learned concepts. A more advanced step toward an autonomy supported classroom includes cognitive autonomy practices (Stefanou et al., 2004). Cognitive autonomy practices are authentic opportunities aligned to the instinctual, individual pursuit of knowledge. Children in this supportive environment steer their learning based on personal interests, intrinsic motivation, and connections they want to form, while teachers facilitate this pursuit with modeling, encouragement, and collaboration (Rogat et al., 2014).

Rogat et al. (2014) provided a rich description of this application by four seventh grade science teachers, using inquiry-based instructional units. The observations yielded evidence of multiple structures of autonomy-supported practices over several modules. Some examples of the range of practices included maintaining open curriculum tasks,

recognizing student contributions of ideas through meaningful feedback, opportunities for self-directions, choice and decision making using open curriculum tasks, and the provision of choice of activity after completing assigned work.

These examples are provided to understand basic and complex ways to promote autonomy and choice. In addition, these examples should provide a barometer for teachers to broaden their instructional practices to more effectively impact educational outcomes through academic choice, engagement, and autonomy. The implication of this promotion of affective variables through instructional changes is that students will make improvements across educational settings. More precisely, with reading-based tasks (assessments, leisure reading, content-based reading) student achievement will improve based on a more balanced approach to education.

Testing Implications

Within this study multiple measures were taken to control for reader experiences, textual elements, personal interest, and reading ability. Additionally, across both choice and non-choice standardized reading passages, individual students were shown to utilize the same if not similar strategies to comprehend both reading passages. Most students indicated they loved reading and liked the passage they chose to read for the study. Results showed students' preference for reading was significantly correlated with the passage chosen, ($p < .01$), although this relationship was considered negligible. Another indication of student preference noted was exhibited by the quantity of students completing specific reading passages, among the set of 'student-selected' standardized reading assessments for their grade level. Still, findings from this study demonstrated little to no differences in performance data, relative to the provision of choice.

Similar experimental controls were elicited in *The NAEP Reader* (1997) study, such as textual elements, comparability of stories, and comprehension questions. Student participants in their study perceived the chosen assessment to be easier, when compared to those that were assigned a story. In addition, students given a choice were more likely to rate their assessment as easier than other reading tests; however, this positive indication did not translate into higher scores for these students (NAEP, 1997). The researchers noted the summation of the difficulty in "comparing the performance of students who read different stories is complicated by the fact of choice itself" (NAEP, 1997, p. 60), which is less stable construct than reading ability. Essentially, student choose a text that is not interpretable (NAEP, 1997), relative to their skill level.

In consideration of both cognitive and affective variables, these standardized reading assessments are constructed to measure reading ability and a student's ability to use these skills within this assessment format. Specifically, the most meaningful correlation of this study was how students performed on the standardized reading passages, which is more indicative of the validity of the test. While the study was designed to uncover the impact of choice, using the standardized reading assessment as the instrument of choice seemed to underscore the strength and effectiveness of these instruments, as a measure of reading ability.

Limitations of the Study

One limitation stemmed from the design of the study. Factored into the research design was the reduction of simple interest matching. For example, for the choice texts, the investigator selected texts in which most of the subjects would not already have strong specific preferences. Additionally, texts students had previous access were

designated as ‘teacher-selected’ text. Lastly, ‘student-selected’ packets were designed to appear uniform to eliminate choice based on length or structure. By design, large-scale assessments do not account for the multitude of interests and preferences of individual students (NAEP,1997), therefore, interest could not be ruled out as a factor of performance.

A secondary limitation of the study is the sample. This study used a non-probability sampling procedure called snowball sampling. In general, non-probability sampling could result in a non-representative sample, which could compromise generalizability (O’Dwyer & Bernauer, 2014). The students sampled were limited to the state of Texas and grades four and five. Likewise, the limits on grade level and texts were offset by the 11 different classrooms and 254 students of diverse backgrounds.

In relation to the testing instrument, standardized tests are designed to measure individual characteristics such as ability, aptitude, interest, achievement, attitude, value, and personal characteristics (Wortham, 2005). In theory, scores can be used to plan instruction, study differences between individuals, and groups. However, these tests are limited in their use of outcome measures, following a brief intervention, as they are not as sensitive to small changes in raw scores (Orkin et al., 2018).

Lastly, this study used a multiple regression analysis to determine the relationship between interest and performance. Results of correlational studies demonstrate that a relationship exists; however, it is not predictive of causation (Larson-Hall, 2016). Therefore, interpretations of the findings are limited.

Recommendations for Future Research

Despite research advocating choice in the classroom, there were very few studies examining choice and assessment. Teachers are critical to the provision of choice, as they facilitate and steer the culture, motivation, and academic tenants of the classroom (Brooks & Young, 2011). It would be beneficial to understand the impact of choice on assessment for classrooms that actively encourage choice or foster autonomy-supportive classrooms. In the present study, information related to teacher perspectives on choice was not considered, but future research should explore this variable.

The authors of the NAEP perceived students' reading abilities as a factor in comparing students on an assessment that involves choice. They claimed that students select stories based on their reading ability. More research is needed to clarify the role of reading ability in relation to the choice provision. Even more research could be conducted with a different grade level.

Mohr (2006) found that non-fiction texts, especially those that featured animals, were preferable over fiction texts amongst a sample (N = 190) first grade boys and girls who were provided an opportunity to self-select or choose a book to own. The choice of text is indicative of the level of interest and intrinsic motivation. Indeed, an active and pursuant interest in reading materials is a demonstration toward relatedness and a deliberate attempt by the reader to seek out enjoyable literature. Reciprocally, when children choose their reading materials, they are likely to be more motivated and engaged readers (Mohr, 2006). Replication of this study using non-fiction passages might provide information on student reading preferences and the impact of choice on genre, topic and types of interest.

Lastly, Siepel et al. (2017) reanalyzed think-aloud data from 138 students in grades 3-5 classified as both good and poor comprehenders. With the dataset, they explored how and when good and poor comprehenders differed in the processes they used while reading. Overall, they found good and poor comprehenders engaged in diverse comprehension processes; however, when viewed through a moment-by-moment continuum, strategic differences between the groups were more readily identified. Based on preliminary data generated from this study, future research might uncover information along divisions of proficiency for the subjects that participated in this study.

Conclusions

This study applied the provision of choice to standardized assessments, thereby extending the research on larger constructs such as student engagement, motivation, interest, assessment, and achievement. It is important to continually refine ways to foster these affective variables. Although the results did not yield a statistical significance for the choice condition, it provided a platform for alternate ideas for teachers to consider along the choice continuum. In the most practical sense, the research contributed to the understanding of one attempt at promoting and applying cognitive autonomy with fourth and fifth grade students. Furthermore, this study used a different lens to investigate the integrated concepts of student motivation, interest, autonomy, and achievement, based on the theoretical principles of self-determination theory.

This study added to the field by focusing on non-traditional means of offering choice in the classroom. Scores for choice and no-choice were moderately to highly correlated, indicating that whether students had a choice, students were likely to perform similarly on both passages. Overall, while there were correlations between other

variables, the Pearson's correlation coefficients demonstrated low to negligible relationships. Conclusively, students' gender, teacher, fondness for reading, or interest in the passage had little to no influence on their overall performance on the assessment.

References

Afflerbach, P. (2002). The road to folly and redemption: Perspectives on the legitimacy of high-stakes testing. *Reading Research Quarterly, 37*(3), 348-360.

Afflerbach, P. (2016). Reading assessment: Looking ahead. *The Reading Teacher, 69*(4), 413-419. https://doi.org/10.1002/trtr.1430.

Allington, R. L. (2002). *Big brother and the National reading curriculum: How ideology trumped evidence*. Portsmouth, MH: Heinemann.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading. In P. D. Pearson (Ed.), *Handbook of reading research* (Vol. 1, pp. 185-224). Longman.

Anmarkrud, O., & Bråten, I. (2009). Motivation for reading comprehension. *Learning and Individual Differences*, *19*(2), 252–256. doi:10.1016/j.lindif.2008.09.002.

Aronesty, J. (2015). *Deciphering the English code: The untold history of the words we speak.* Los Angeles, CA: Two Candle Press.

Au, W. (2013). Hiding behind high-stakes testing: Meritocracy, objectivity and inequality in U.S. education. *The International Education Journal: Comparative Perspective, 12*(2), 7-19.

Baker, L. (1979). Comprehension monitoring: Identifying and coping with text confusions. *Journal of Reading Behavior*, *11*, 366–374.

Baker, L., & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly, 34*(4), 452–477. doi:10.1598/RRQ.34.4.4.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral

change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality & Social Psychology, 41*(3), 586-598. https://doi.org/ 10.1037/0022-3514.41.3.586.

Barry, A. L., & Nielsen, D. C. (1996). Student choice of writing topic in a state writing assessment. *Reading Improvement, 33*, 19-30.

Bayless, S.D., Jenson, J. M., Richmond, M. K., Pampel, F. C., Cook, M., & Calhoun, M. (2018). Effects of an afterschool early literacy intervention on the reading skills of children in public housing communities. Child Youth Care Forum, 47, 537-561. Retrieved from https://doi.org/10.1007/s10566-018-9442-5.

Becker, M., McElvany, N., & Kortenbruck, M. (2010). Intrinsic and extrinsic reading motivation as predictors of reading literacy: A longitudinal study. *Journal of Educational Psychology, 102*(4), 773-785.

Billingsley, B. S., & Wildman, T. M. (1990). Facilitating reading comprehension in learning disabled students: Metacognitive goals and instructional strategies. *RASE: Remedial & Special Education, 11*(2), 18-31. http://dx.doi.org/10.1177/074193259001100205.

Braten, I., Ferguson, L. E., Anmarkrud, O., & Stromso, H. I. (2012). Prediction of learning and comprehension when adolescents read multiple texts: the roles of word-level processing, strategic approach, and reading motivation. *Read Writ, 26*, 321-348. DOI 10.1007/s11145-012-9371-x.

Brooks, C. F., & Young, S. L. (2011). Are choice-making opportunities needed in the

classroom? Using self-determination theory to consider student motivation and learner empowerment. *International Journal of Teaching and Learning in Higher Education*, *23*(1), 48-59.

Brown, C. A., & Dryden, W. J. (2004). Quality storytents: Using choice cheory to support reading through a community literacy project. *International Journal of Reality Therapy, 24*(1), 3–12.

Campbell, J. R., & Donahue, P. L. (1997). *Students selecting stories: The effects of choice in reading assessment*: *Results from The NAEP Reader special study of the 1994 National Assessment of Educational Progress*. U.S. Dept. of Education, Office of Educational Research and Improvement, Educational Resources Information Center.

Carver, R. P. (1992). What do standardized tests of reading comprehension measure in terms of efficiency, accuracy, and rate? *Reading Research Quarterly*, *27*(4), 346-359.

Catts, H., Adlof, S., & Weismer, S. (2006). Language deficits in poor comprehenders: A case for the Simple View of Reading. *Journal of Speech– Language– Hearing Research*, *49*, 278– 293.

Chambers, S. C., Rintamaa, M., Anderman, E. M., & Anderman, L. H. (2018). Rural adolescents' reading motivation, achievement and behavior across transition to high school. *The Journal of Educational Research*, *111*(4), 417-428. http://doi.org/10.1080/00220671.2017.1284737.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of*

Educational Psychology, *88*(4), 715-730.

Csikzentmihalyi, M., (1991), Flow – the psychology of optimal experience, HarperPerennial ed. New York : HarperPerennial.

D'Ailly, H. (2004). The role of choice in children's learning: A distinctive cultural and gender difference in efficacy, interest, and effort. *Canadian Journal of Behavioural Science*, *36*(1), 17-29.

de Milliano, I., van Gelderen, A., & Sleegers, P. (2014). Types and sequences of self-regulated reading of low-achieving adolescents in relation to reading task achievement. *Journal of Research in Reading*, *39*(2), 229-252.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, *18*(1), 105-115.

Deci, E. L., & Ryan, R. M. (1985). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology*, *53*(6), 1024-1037.

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, *28*, 425-474. http://doi.org/ 10.1007/s10648-015-9320-8.

Deshler, D. D., & Hock, M. F. (2006). Shaping literacy achievement. Guilford Press.

Dewey, J. (1990). *The school and society and the child and the curriculum*. University of Chicago Press.

Dobrow, S. R., Smith, W. K., & Posner, M. A. (2011). Managing the grading paradox:

Leveraging the power of choice in the classroom. *Academy of Management Learning & Education*, *10*(2), 261. https://doi-org.ezproxy.shsu.edu/10.5465/AMLE.2011.62798933.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review Psychology*, *53*, 109-32.

ESSA (2015). Every Student Succeeds Act of 2015, Pub. L. No. 114-95 § 114 Stat. 1177 (2015-2016).

Evans, M., & Boucher, A. R. (2015). Optimizing the power of choice: Supporting student to foster motivation and engagement in learning. *Mind, Brain, and Education*, *9*(2), p. 87-91.

Fisher, D., & Frey, N. (2018). Raise reading volume through access, choice, discussion, and book talks. *The Reading Teacher*, *72*(1), 89–97.

Fitzpatrick, M. J., & Meulemans, Y. N. (2011) Assessing an information literacy assignment and workshop using a quasi-experimental design, *College Teaching*, *59*(4), 142-149, DOI: 10.1080/87567555.2011.591452.

Flanagan, D. P., & Alfonso, V. C. (2011). Essentials of specific learning disability identification. John Wiley & Sons, Inc.

Flowerday, T., & Schraw, G. (2001). The role of choice and interest in reader engagement. [Paper presentation]. American Educational Research Association, Seattle, WA, United States.

Flowerday, T., & Schraw, G. (2003). Effect of choice on cognitive and affective engagement. *The Journal of Educational Research*, *96*(4), 207-215.

Flowerday, T., & Shell, D. F. (2015). Disentangling the effects of interest and choice on

learning, engagement, and attitude. *Learning and Individual Differences, 40*, 134-140. http://dx.doi.org/10.1016/j.lindif.2015.05.003.

Flowerday, T., Schraw, G., & Stevens, J. (2004). The role of choice and interest in reader engagement. *The Journal of Experimental Education*, *72*(2), 93-114.

Fraumeni-McBride, J. P. (2017). The effects of choice on reading engagement and comprehension for second- and third- grade students: An action research report. *Journal of Montessori Research*, *3*(2), 19-37.

Fuchs, L. S., & Fuchs, D. (2007). A model for implementing responsiveness to intervention. *Council for Exceptional Children, 39*(5), pp. 14-20.

Fulmer, S. M., D'Mello, S. K., Strain, A., & Graesser, A. C. (2014). Interest-based text preference moderates the effect of text difficulty on engagement and learning. *Contemporary Educational Psychology, 41*, 98-110.

Fulmer, S. M., & Frijters, J. C. (2011). Motivation during an excessively challenging reading task: The buffering role of relative topic interest. *The Journal of Experimental Education, 79*, 185-208.

Gambrell, L. B. (2011). Seven rules of engagement. *The Reading Teacher*, *65*(3), 172-178. http://doi.org/10.1002/TRTR.01021.

Gambrell, L. B., Marinak, B. A., & Malloy, J. A. (2017). Yes, Virginia, motivation matters. *Making Literary Connections, 32*, p. 32-38.

Glaser, R. (1985). The integration of instruction and testing. Paper presented at the ETS Invitational Conference on the Redesign of Testing for the 21st century, New York.

Glasser, W. (1998). Choice theory: A new psychology of personal freedom. Harper Collins Publishers.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, *7*(1), 6–10. https://journals-sagepubcom.ezproxy.shsu.edu/doi/pdf/ 10.1177/074193258600700104.

Gritter, K., Van Duinen, D. V., Montgomery, K., Blowers, D., & Bishop, D. (2017). Boy troubles? Male literacy depictions in children's choices picture books. *The Reading Teacher*, *70*(5), 571–581. https://doi-org.ezproxy.shsu.edu/10.1002/trtr.1559.

Guthrie, J. T., & Klauda, S. L. (2014). Effects of classroom practices on reading comprehension, engagement, and motivations for adolescents. *Reading Research Quarterly*, *49*(4), 387–416.

Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (3rd. ed., pp. 403–422). Longman.

Guthrie, J. T. (2004). Classroom contexts for engaged reading. Motivating reading comprehension. *Concept-Oriented Reading Instruction*, 1-24.

Guthrie, J. T., Hoa, A. L. W., Wigfield, A., Tonks, S. M., Humenick, N. M., & Littles, E. (2007). Reading motivation and reading comprehension growth in the later elementary years. *Contemporary Educational Psychology*, *32*(3), 282–313.

Guthrie, J. T., Wigfield, A., Cambria, J., Coddington, C. S., Klauda, S. L., Morrison, D.

A. (2009). Motivations for reading information books among adolescent students. http://cori.umd.edu/real/ Motivation_Report4.27.09.pdf.

Haworth, L. (1986). *Autonomy: an essay in philosophical psychology and ethics*. New Haven: Yale University Press.

Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review*, *13*(3), 191-210.

Hidi, S., and Harackiewicz, J. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *American Educational Research Association*, *70*(2): 151-179.

Hoover, W.A., & Gough, P.B. (1990). The simple view of reading. *Reading and Writing*, *2*, 127–160. https://doi.org/10.1007/BF00401799.

Huddleston, A., & Lowe. T. N. (2014). "I skim and find the answers": Addressing search-and-destroy in reading. *The Reading Teacher*, *68*(1), 71.

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.

Individuals with Disabilities Education Act. (2004). Public Law 108-446.

Iyengar, S.S., & Lepper, M.R. (2000). When a choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, *79*(6), 995-1006. http://doi.org/10.1037//0022-351479.6.995.

Kamii, C. (1991). Toward autonomy: The importance of critical thinking and choice. *School Psychology Review*, *20*(3), 382-389.

Katz, I., & Assor, A. (2007). When choice motivates and when it does not. *Educational*

Psychology Review, 19(429). https://doi.org/10.1007/s10648-006-9027.

Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice, 29*(1), 10–16.

Key, J.P (1997). Experimental design and research. http://www.okstate.edu/ag/agedcm4h/academic/aged5980a/5980/newpage2.htm.

Khan, K., Nelson, K., & Whyte, E. (2013). Children choose their own stories: The impact of choice on children's learning of new narrative skills. *Journal of Child Language, 41*, 949-962. http://doi.org/10.1017/S0305000913000160.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*(4), 294–303. https://doi-org.ezproxy.shsu.edu/10.1037/0003-066X.49.4.294.

Kintsch, W. (1998). Comprehension: A paradigm for cognition. Cambridge University Press.

Kohn, A. (1993). Choices for children: Why and how to let students decide. *The Phi Delta Kappan, 75*(1), 8-16, 18-20.

Krapp, A. (2002). Structural dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction, 12*, 383-409.

Lai, C., & Hwang, G. (2016). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers & Education, 100*, pp. 126-140. Retrieved from

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using*

SPSS. Routledge.

Lin, Y. (2015). Using readers theater as a facilitator in elementary school English training. *Journal of Education and Learning*, 4(2).

Lukhele, R., Thissen, D., & Wainer, H. (1994). On the relative value of multiple-choice, constructed response, and examinee-selected items on two achievement tests. *Journal of Educational Measurement, 31*(3), 234-250.

Mallette, M. H., Duke, N. K., Strachan, S. L., Waldron, C. H., Watanabe, L. M. (2011). Synergy in literacy research methodology. In N. K. Duke & M. H. Mallette (Eds.), *Literacy research methodologies* (2nd ed., pp. 91-128). Guilford Press.

Marinak, B. A., & Gambrell, L. B. (2008) Intrinsic motivation and rewards: What sustains young children's engagement with text? *Literacy Research and Instruction, 47*(1), 9-26, http://doi.org/10.1080/19388070701749546.

McCrudden, M. T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review*, *19*, 113–139. http://doi.org/10.1007/s10648-006-9010-7.

McKool, S. S. (2007). Factors that influence the decision to read: An investigation of fifth grade students' out-of-school reading habits. *Reading Improvement, 44*(3), 111-131.

Milliano, I. D., Gelderen, A. V., & Sleegers, P. (2014). Types and sequences of self-regulated reading of low-achieving adolescents in relation to reading task achievement. *Journal of Research in Reading, 39*(2), 229-252.

Mitchell, R.J., Morrison, T.G., Feinauer, E., Wilcox, B., & Black, S. (2016). Effects of fourth and second graders' cross-age tutoring on students' spelling. *Reading*

Psychology, *37*: 147-166. http://doi.org/10.1080/02702711.2015.1025164.

Mohr, K. A. J. (2003). Children's choices: A comparison of book preferences between Hispanic and non-Hispanic first-graders. *Reading Psychology*, *24*(163), 163-176. http://doi.org/ 10.1080/02702710390197462.

Mohr, K. A. J. (2006). Children's choices for recreational reading: A three-part investigation of selection preferences, rationales, and processes. *Journal of Literacy Research*, *38*(1), 81-104.

Moller, A. C., Deci, E. L., & Ryan, R. M. (2006). Choice and ego-depletion: The moderating role of autonomy. *Personality and Social Psychology Bulletin*, *32*(8), 1024-1036.

Mulvey, M. E. (2009). Characteristics of under-prepared students: Who are "*the under-prepared*"? *Reading & Teaching in Developmental Education, 25*(2), 29-58.

Munster, C., & Haines, C. (2019). "Students get what flows downward": District leaders' rationalizations of the standardized testing of children. *The Educational Forum*, *83*(2), 160-180. http://doi.org/10.1080/100131725.2019.1567891.

National Assessment of Educational Progress (1997). The nation's report card. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

National Center for Education Statistics. (2017). *National assessment of educational progress: An overview of NAEP*. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education.

National Reading Panel (2000). Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and its Implications

for Reading Instruction. *Reports of the Subgroups*, NIH Publication No. 00-4754, National Institute of Child Health and Human Development.

Nejadihassan, S., & Arabmofrad, A. (2016). A review of relationship between self-regulation and reading comprehension. *Theory and Practice in Language Studies*, *6*(4), 835-842. http://dx.doi.org/10.17507/tpls.0604.22.

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

O'Dwyer, L.M., & Bernauer, J.A. (2014). *Quantitative research for the qualitative researcher*. Sage.

Organization for Economic Cooperation and Development (OECD). (1999). Measuring student knowledge and skills: A new framework for assessment. Author.

Organisation for Economic Co-operation and Development (OECD) (2000) Measuring student knowledge and skills: the PISA 2000 assessment of reading, mathematical and scientific literacy. OECD, Paris.

Organisation for Economic Co-operation and Development (OECD) (2003) Education at a glance: OECD indicators 2003. OECD, Paris.

Organisation for Economic Co-operation and Development (OECD). 2003. *The PISA 2003 Assessment Framework. Mathematics, Reading, Science and Problem Solving Knowledge and Skills*. OECD.

Orkin, M., Pott, M., Wolf, M., May, S., & Brand, E. (2018). Beyond gold stars: Improving the skills and engagement of struggling readers through intrinsic motivation, *Reading & Writing Quarterly*, *34*(3), 203-217. http://doi.org/10.1080/10573569.2017.1387834.

Pardo, L. S. (2004). What every teacher needs to know about comprehension. *The*

Reading Teacher, *58*(3), 272. https://doi-org.ezproxy.shsu.edu/10.1598/RT.58.3.5.

Parsons, A. W., Parsons, S. A., Malloy, J. A., Marinak, B. A., Reutzel, D. R., Applegate, M. D., Fawson, P. C., & Gambrell, L. B. (2018). Upper elementary students' motivation to read fiction and nonfiction. *Elementary School Journal*, *118*(3). 505-523.

Patall, E. A., Cooper, H., & Wynn, S.R. (2010). The effectiveness and relative importance of choice in the classroom. *Journal of Educational Psychology*, *102* (4), 896 –915.

Patall, E. A., Cooper, H., & Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: A meta-analysis of research findings. *Psychological Bulletin*, *134*(2), 270-300.

Piaget, J. (1948/1973). To understand is to invent New York: Viking (Original work published 1948).

Putnam, R. S. (2017). Technology versus teachers in the early literacy classroom: An Investigation of the effectiveness of the istation integrated learning system. *Education Tech Research Development*, *65*, 1153-1174. http://doi.org/10.1007/s11423-016-9499-5.

RAND Reading Study Group (RRSG) (2002). Toward an r & d program for reading comprehension. Santa Monica, CA. Retrieved from: https://www.rand.org/content/dam/rand/pubs/monograph_reports/2005/MR1465.pdf.

Reeve, J. & Tseng, C.-M. (2001). Agency as a fourth aspect of students' engagement

during learning activities. *Contemporary Educational Psychology*, *36*, 257-267.

Renninger, K. A., (1990). Children's play interests, representation, and activity. In R. Fivush & J. Hudson (Eds.), *Knowing and remembering in young children* (pp. 127-165). Cambridge University Press.

Rogat, T. K., Witham, S. A., & Chinn, C. (2014). Teachers' autonomy-relevant practices within an inquiry-based science curricular context: Extending the range of academically significant autonomy-supportive practices. *Teachers College Record*, *116*, 070305, 1-46.

Rosenblatt, L. M. (1978). *The reader, the text, the poem: The transactional theory of literacy work*. Southern Illinois University Press.

Rosenblatt, L. M. (1994). The transactional theory of reading and writing. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 1057-1092). International Reading Association.

Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance* (Vol. 6, pp. 573-603). Erlbaum.

Rumelhart, D. E. (1981). *Understanding understanding* (Technical report No. CHIP 100 - ONR-TR-8101). https://files.eric.ed.gov/fulltext/ED198497.pdf.

Ryan, R.M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.

Sanacore, J. (2002). Struggling literacy learners benefit from lifetime literacy efforts. *Reading Psychology*, *23*, 67-87.

Schiefele, U. (1991). Interest, learning, and motivation, *Educational Psychologist*, *26*(3-4), 299-323, DOI: 10.1080/00461520.1991.9653136.

Schiefele, U. (1996). Topic interest, text representation, and quality of experience. *Contemporary Educational Psychology*, *21*, 3-18. http://doi.org/10.1006/ceps.1996.0002.

Schiefele, U. (1999). Interest and learning from text. *Scientific Studies of Reading*, *3*, 37-41. http://doi.org/10.1207/s1532799xssr0303.

Schiefele, U. (2009) Situational and individual interest. In K. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 197-222). Erlbaum.

Schiefele, U., Schaffner, E., Moller, J., & Wigfield, A. (2012). Dimensions of reading motivation and their relation to reading behavior and competence, *Reading Research Quarterly*, *47*(4), 427-463.

Schraw, G., Flowerday, T., & Reisetter, M. (1998). The role of choice in reader engagement. *Journal of Educational Psychology*, *90*, 705–714.

Schunk, D. H., & Zimmerman, B. J. (1994). Self-regulation in education: Retrospect and prospect. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance*. Erlbaum.

Schwartz, B. (2007). *The paradox of choice: Why more is less*. New York, NY: HarperCollins Publishers Inc.

Seipel, B., Carlson, S.E., & Clinton, V. E. (2017). When do comprehender groups differ? A moment-by-moment analysis of think-aloud protocols of good and poor comprehenders. *Reading Psychology*, *38*, pp.39–70. http://doi.org/10.1080/02702711.2016.1216489.

Skinner, E. A., Wellborn, J. G., & Connell, J. P. (1990). What it takes to do well in school and whether I've got it: A process model of perceived control and children's engagement and achievement in school. *Journal of Educational Psychology*, *82*(1), 22-32.

Slomp, D. H. (2008). Harming not helping: The impact of a Canadian standardized writing assessment on curriculum and pedagogy. *Assessing Writing, 13,* 180-200. http://doi.org/10.1016/j.asw.2008.10.004.

Smith, N. B. (2002). *American reading instruction*. International Reading Association.

Snow, C. E. (2010). Reading comprehension: Reading for learning. *International Encyclopedia of Education* 5, 413-418.

Soltero-Gonzalez, L., Sparrow, W., Butvilofsky, S., Escamilla, K., & Hopewell, S. (2016). Effects of a paired literacy program on emerging bilingual children's biliteracy outcomes in third grade. *Journal of Literacy Research*, *48*(1), 80-104. http://doi.org/10.1177/1086296X16653842.

Springer, S. E., Harris, S., & Dole, J. A. (2017). From surviving to thriving: Four research-based principles to build students' reading interest. *The Reading Teacher*, *71*(1), 43-50.

Stefanou, C. R., Perencevich, K. C., DiCintio, M., & Turner, J. C. (2004). Supporting autonomy in the classroom: Ways teachers encourage student decision making and ownership. *Educational Psychologist, 39*(2), 97-110.

Stokes, F., & Young, C. (2018). Research on readers theater. *Association of Literacy Educators and Researchers Yearbook, 40,* 133-144.

Szabo, S., & Sinclair, B. (2012). STAAR reading passages: The readability is too high.

Schooling, 3, 1-14.

Taboada, A., Tonks, S.M., Wigfield, A. (2009). Effects of motivational and cognitive variables on reading comprehension. *Reading and Writing, 22*(85) https://doi.org/10.1007/s11145-008-9133.

Tarchi, C. (2017). Comprehending expository texts: The role of cognitive and motivational factors. *Reading Psychology, 38*, 154-181. http://doi.org/ 10.1080/02702711.2016.1245229.

Texas Education Agency (2017). Comprehension instruction. Austin, TX.

Theroux, K. (Fall, 2010). Advancing Adolescent Literacy: The Cornerstone of School Reform. Carnegie Review, New York: Carnegie Corporation.

Thorkildsen, T. (2002). Literacy as a lifestyle: Negotiating the curriculum to facilitate motivation. *Reading & Writing Quarterly, 18*, 321-341. http://doi.org/ 10.1080/07487630290061863.

Tracey, D. H., & Morrow, L. M. (2012) *Lenses on reading: An introduction to theories and models* (2nd ed.). The Guilford Press.

Vidal-Abarca, E., Mana, E., & Gil, L. (2010). Individual differences for self-regulating task-oriented reading activities. *Journal of Educational Psychology, 102*(4), 817-826.

Vellutino, F. R., & Schatschneider, C. (2011). Experimental and quasi-experimental Design. In N. K. Duke and M. H. Mallette (2nd Ed.), *Literacy research methodologies* (pp. 155-187). Guilford Press.

Walpole, S., McKenna, M. C., Amendum, S., Pasquarella, A., & Strong, J. Z. (2017). The promise of a literacy reform effort in the upper elementary grades. *The*

Elementary School Journal, *118*(2), 257-280.

Wehmeyer, M. L., Shogren, K. A., & Little, T. D. & Lopez, S.J. (Eds.) (2017). Development of self-determination through the life-course. New York: Springer.

Wigfield, A., & Eccles. J. C. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, *25*, 68–81.

Wigfield, A., & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. Journal of Educational Psychology, 89, 420-432. doi:10.1037/0022-0663.89.3.420.

Wigfield, A., Guthrie, J.T., Perencevich, K., Taboada, A., Klauda, S.L., McRae, S., (2008). The role of reading engagement in mediating the effects of instruction on reading outcomes. *Psychology in the Schools*, *45*(5), 432–445. doi:10.1002/pits.20307.

Wigfield, A. Gladstone, J. R., & Turci, L. (2016). Beyond cognition: Reading motivation and reading comprehension. *Child Development Perspectives*, *10*(3), 190-195.

Wilde, M., Basten, M., GroBmann, N., Haunhorst, D., Desch, I., Struber, M., & Randler, C. (2018). The (non)benefit of choosing: If you get what you want it is not important that you chose it. *Motivation and Emotion*, *42*(3), 348-359. https://doi.org/10.1007/s11031-018-9675-5.

Wildman, L. (2007). The coefficient of determination: Family wealth and student achievement scores. *International Journal of Educational Leadership Preparation*, *2*(1), 1-3.

Wolters, C.A., Barnes, M. A., Kulesz, P. A., York, M., & Francis, D. J. (2017).

Examining a motivational treatment and its impact on adolescents' reading comprehension and fluency. *The Journal of Educational Research, 110*(1), 98-109.

Wortham, S. C. (2005). *Assessment in early childhood education*. Upper Saddle River, New Jersey: Pearson Education.

Worthy, J., Moorman, M., & Turner, M. (1999). What Johnny likes to read is hard to find. *Reading Research Quarterly, 34*(1), 12-27.

Worthy, J., Patterson, E., Salas, R., Prater, S., & Turner, M. (2002). "More than just reading": The human factor in reaching resistant readers. *Reading Research and Instruction, 41*(2), 177-202. Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*(3), 329-339.

Vaknin-Nusbaum, V., Nevo, E., Brande, S., & Gambrell, L. (2018). Developmental aspects of reading motivation and reading achievement among second grade low achievers and typical readers. *Journal of Research in Reading, 41*(3), pp 438-454.

Verhoeven, L., & Snow, C. E. (Eds.) (2001). *Literacy and motivation: Reading engagement in individuals and groups*. Mahwah, NJ: Lawrence Erlbaum.

Young, C., & Mohr, K. (2018). Exploring factors that influence quality literature circles. *Literacy Research and Instruction, 57*(1), 44-58.

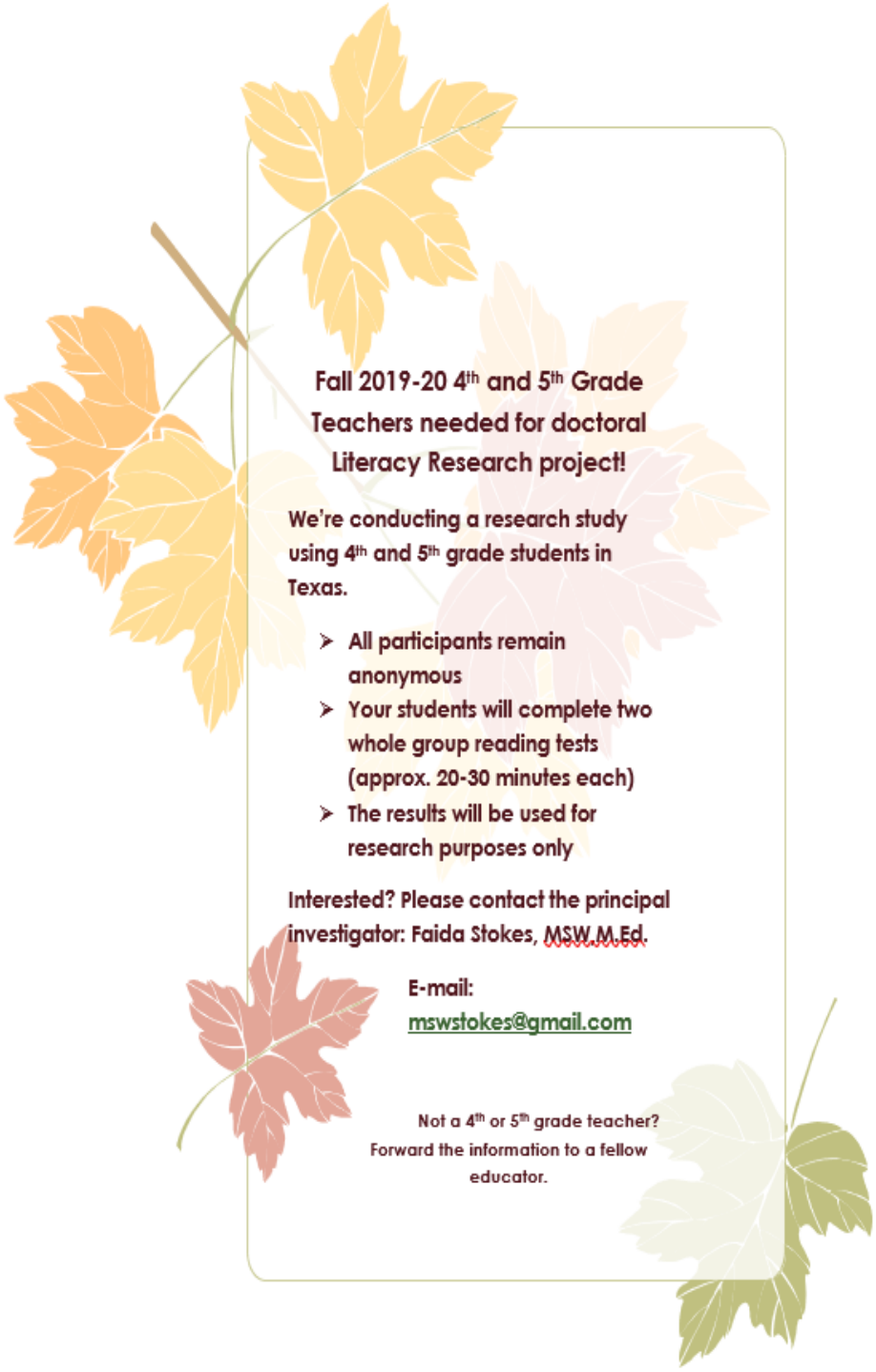
Young, C., & Rasinski, T. (2018). Readers theatre: Effects on word recognition automaticity and reading prosody. *Journal of Research in Reading, 41*(3), 475-485.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive

career path. *Educational Psychologist*, *48*(3), 135-147. http://doi.org/10.1080/0046152.2013.794676.

Zuckerman, M., Porac, J., Lathin, D., Smith, R., & Deci, E. L. (1978). On the importance of self-determination for intrinsically-motivated behavior. *Personality and Social Psychology Bulletin*, *4*(3), 443-446.

APPENDIX A


Fall 2019-20 4th and 5th Grade Teachers needed for doctoral Literacy Research project!
We're conducting a research study using 4th and 5th grade students in Texas.
All participants remain anonymous
Your students will complete two whole group reading tests (approx. 20-30 minutes each)
The results will be used for research purposes only
Interested? Please contact the principal investigator: Faida Stokes, MSW,M.Ed.
E-mail:
mswstokes@gmail.com
Not a 4th or 5th grade teacher? Forward the information to a fellow educator.

APPENDIX B

Dear Teacher,

Thank you for taking this opportunity to partner with me in this study.

My name is Faida Stokes. I am a doctoral student at Sam Houston State University. I am conducting research on the impact of choice on high-stakes reading assessments on fourth- and fifth-grade students. The results of this study will extend our knowledge on the effects of choice in the classroom using various instructional materials. All subjects of the reading assessments are anonymous and confidential. Specifically, I am not asking you to provide the names of any of your students that participate in the assessments. However, I will be asking for demographic information about you, your school, and your students.

Included in this testing packet, you will find

- An instruction card
- A teacher/school demographic card
- A student demographic card
- STAAR reading assessments for your students
- A 1-question interest survey

If any of these pieces are missing, please contact me via e-mail at fas011@shsu.edu. Once all the materials are completed, please return promptly. All data from this study should be completed and returned by December 15, 2019.

Again, thank you for partnering with me in this study.

Sincerely,

Faida Stokes, M.Ed.

APPENDIX C

Instruction Card:

Timeline: Two (2) separate instructional days; approximately 45 min/day

Materials:

4th grade students will use the 3rd grade STAAR reading passages

5th grade students will use the 4th grade STAAR reading passages

Special Notes:

1. Please, do not assist any student with reading the passage or questions. You may provide minor clarifications of words or phrases.
2. Students that have qualified for special education services, a 504 plan, or have any other disability that would inhibit them from independently reading this exam and answering the questions should not take part in this study. Generally, students that require reading assistance of more than 10% of a document should not take part in this exam. Please keep this in mind, when deciding a time to administer the test. In the case that this student is present, I would suggest that you pair him/her with another struggling reader to work on the task. <u>However, do not include his/her packet in the final packet without marking it with a "NO SCORE" label.</u> (Rationale: this study is not meant to harm or frustrate students. - wording)
3. This study is not timed.
4. Please keep all materials safe and confidential.
5. Please check all the student cards for completion and **<u>remove any student names from any documents</u>**. Return all materials, ungraded, to the principal investigator, following the completion of the study.

General Instructions:

This research should be conducted over two separate days. As the instructor, you will decide the sequence of the administration of the reading passages. On one of the days, the students will have a choice of three passages. These should be separated and offered to the students. They are asked to select one passage from the three passages.

On another day, you will "assign" the passage. If you are a 4th grade teacher, you will assign the passage, *Jake Drake, Teacher's Pet*. If you are a 5th grade teacher, your assigned passage is, *Because of Winn-Dixie*.

(**Please read the following instructions before handing out the reading passages**.)

Student choice of passage say: *You are about to take a reading test. For this test, you will choose which passage you would like to read. I will read aloud a short summary about the story and you will decide which story you would like to read.* Each story has a summary attached, which you will now read. When the student indicates their choice in passage, hand the passage to this student. Once all the students have chosen, say: *Please read the passage and answer all the questions by circling your answer to the question. You can write directly on the pages and take any notes that you need to complete the assignment. When you have completed all the questions, please complete the survey in the back of the packet. Be sure to do your best work. When you have completed this, please return the entire packet*.

Teacher assigned passage say: *You are about to take a reading test. Please read the passage and answer all the questions by circling your answer to the question. You may write directly on the pages and take any notes that you need to complete the assignment. Be sure to do your best work. When you have completed this, please return the entire packet*.

APPENDIX D

Please complete and return with testing packets.

Teacher Information

Name: _____

How many years have you taught? _____

Highest Degree Earned: (Circle One)

Bachelor Graduate Doctoral Other (specify):

Which grade do you currently teach? (Circle One)

4^{th} 5th

Type of School: (Circle One)

Public Charter Private Other (specify):

How would you describe your school? (Circle all that apply)

Location: Rural Urban Suburban

Population: >500 500-800 800+ students

How would you describe most of the families at your school? (Circle one)

Income: Low Middle High

Ethnicity: Latinx Caucasian African American Other:

APPENDIX E

Please have each student complete one card. Attach this card to their assessments and return with study materials.

Student Information Card

Directions: Circle an answer for each question that best describes you.

I am ____ years old.

8 9 10 11 12

I consider myself a:

Female Male

I am in the ___ grade.

4th 5th

I _____ to read.

love like do not like

Did you like the passage that you chose?

Yes No

APPENDIX F

Description of stories

Fourth Grade Stories

Racing Team: This story is about two boys, who made some unexpected friends at the beach.

Lizard Problems: This is a story about a girl, who overcame her fear of lizards.

Good Question, Chowderhead: This is a story about a girl, who learned something new about herself and how to talk to other people.

Fifth Grade Stories

The Spelling Test: This story is about a boy, who learned new tips to accomplish a goal.

My What is Showing?: This story is about a girl, who learned a word and a lesson and turned it into a joke.

A Caterpillar's Tale: This story is about a caterpillar, who worked hard to build a home.

VITA

Faida Stokes, M.Ed.

Education

Sam Houston State University
Doctoral Student 2016

West Texas A &M University
M.S. Educational Diagnostician 2015

Tennessee State University
M.S. Curriculum and Instruction 2011
Emphasis: Special Education

University of Houston
Master's in Social Work 2006
Emphasis: Children and Families

University of Houston
B.S. Human Development and Family Studies 2003
Minor: Nonprofit Management

Academic Positions

Educational Diagnostician 2016
Snyder Elementary, Conroe ISD

Educational Diagnostician 2015-16
Sycamore Elementary, Crowley ISD

Department Chair – Special Education 2014-15
Inclusion Teacher – K-2 2013-14
Atwood McDonald Elementary (Title I), Fort Worth ISD

Inclusion Teacher – K-2 2011-13
Ross Elementary (Title I), Nashville Metro Public Schools

Interim Inclusion Teacher – K-2 2011
R. C. Museum Magnet Elementary (Title I), Nashville Metro Public

Honors

Outstanding Student in Doctoral Literacy Program
Sam Houston State University 2018

PUBLICATIONS

Refereed

Young, C., Stokes, F., & Rasinski, T. (2017). Readers Theater plus comprehension and word study. *The Reading Teacher, 71*(3), 351-355. doi: 10.1002/trtr.1629.

Invited/Non-Refereed Publications

Rasinski, T., Stokes, F., & Young, C. (2017). The role of the teacher in readers theater. *Texas Journal of Literacy Education*, *5*(2), 168-174.

Young, C., Rasinski, T., & Stokes, F. (2017, June 15). Finding scripts for readers theater [Web Log Post]. *Literacy Daily.* Retrieved from: https://www.literacyworldwide.org/blog/literacy-daily/2017/06/15/finding-scripts-for-readers-theatre

Professional Presentations

Gerber, H., Stokes, F., Maynard, C., Muse, B., Panozzo, M., Daudey, A., Gates, J., & Rice, R. (2018, November). *"On being a literacy education doctoral student in the digital age: Phenomenological explorations of contemporary doctoral education"*. Roundtable presentation at the National Council of Teachers of English, Houston, TX.

Young, C. Stokes, F., & Maynard, C. (2017, November). *Adding depth to readers theater*. Paper presented at the annual meeting of the Association of Literacy Educators and Researchers, St. Petersburg, FL.

Professional Membership and Affiliation

National Council of Teachers of English

Association of Literacy Educators and Researchers
International Literacy Association
Kappa Delta Pi – International Honor Society in Education
Alpha Chi – National College Honor Society