AN INVESTIGATION OF THE RELATIONSHIPS AMONG HIGH SCHOOL STUDENTS' READING COMPREHENSION STRATEGY INSTRUCTION, STRATEGY USE, ATTITUDES, AND ACHIEVEMENT

by

Jennifer A. Buxton
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
Doctor of Philosophy
Education

Committee:

____________________________________ Chair

____

____

____________________________________ Program Director

____________________________________ Dean, College of Education and Human Development

Date: ______________________________ Fall Semester 2017
George Mason University
Fairfax, VA

An Investigation of the Relationships Among High School Students' Reading Comprehension Strategy Instruction, Strategy Use, Attitudes, and Achievement

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Jennifer A. Buxton
Master of Science
Duquesne University, 2001
Bachelor of Arts
Duquesne University, 2000

Director: Anastasia Kitsantas, Professor
College of Education and Human Development

Fall Semester 2017
George Mason University
Fairfax, VA

Dedication

This is dedicated to my family, whose love and support enabled me to achieve my dream.

Acknowledgements

I want to thank everyone who helped and guided me through this labor of love. Thank you for believing in me and sticking by me over all these years. My deepest gratitude goes to Dr. Anastasia Kitsantas, my dissertation committee chair. From the first course I had you as a professor up and through the completion of my dissertation, your teaching, kindness, assistance, encouragement, and guidance has given me the strength I needed to keep working to reach my goal. I would also like to thank Dr. Sheri Berkeley and Dr. Jehanzeb Cheema, my dissertation committee members, for their guidance time spent helping me throughout my dissertation process.

Table of Contents

List of Tables

List of Figures

List of Equations

List of Abbreviations

National Center for Education Statistics .... NCES
Organization for Economic Cooperation ....OECD
Program for International Student Assessment ....PISA
Reading Comprehension Achievement .... RCA
Reading Comprehension Strategy Use ....RCSU
Reading Comprehension Strategy Instruction ....RCSI
Reading Attitudes..RA
United States .... US

Abstract

AN INVESTIGATION OF THE RELATIONSHIPS AMONG HIGH SCHOOL STUDENTS' READING COMPREHENSION STRATEGY INSTRUCTION, STRATEGY USE, ATTITUDES, AND ACHIEVEMENT

Jennifer A. Buxton, Ph.D.

George Mason University, 2017

Dissertation Director: Dr. Anastasia Kitsantas

This study explored the relationships among high school students' reading comprehension achievement, three latent constructs (reading comprehension strategy use, reading comprehension strategy instruction, reading attitudes), and five control variables (gender, minority status, socio-economic status [SES], class time, class size). The purpose of this study was to examine the effects of the latent and control variables on reading comprehension achievement, direct effects among the latent constructs, effects of the control variables on the latent constructs, and group differences in reading comprehension achievement and the latent constructs across the control variables. Data were obtained from the Program for International Student Assessment (PISA) 2009, including 5,233, 15-year-old students from 165 schools. Structural equation modeling results indicated 33% of student differences in reading comprehension achievement, 17% of differences in reading strategy use, 0.3% of differences in reading strategy instruction, and 15% of differences in reading attitudes were accounted for by their predictors. The results also suggested that all three latent factors predicted reading comprehension achievement, reading comprehension strategy instruction and reading attitudes predicted reading comprehension strategy use, and reading comprehension strategy instruction predicted reading attitudes. Further, the findings revealed socio-economic status predicted all three latent constructs and reading comprehension achievement, gender predicted reading comprehension strategy use and reading attitudes, and minority status predicted reading comprehension achievement and reading comprehension strategy use. Additionally, the results implied class time predicted reading comprehension achievement, reading comprehension strategy use, and reading comprehension strategy instruction; and class size predicted reading comprehension achievement and reading comprehension strategy use. Finally, the results indicated reading comprehension achievement differences across all five groups; reading comprehension strategy use differences across gender, minority status, socio-economic status, and class time; reading attitude differences across gender, minority status, and socio-economic status; and class time group differences in reading comprehension strategy instruction. Implications, limitations, and recommendations for future research were discussed.

Chapter One

Over the past 20 years, preventing reading failure has become a national priority (Al Otaiba & Fuchs, 2006). Currently, improving students' reading proficiency is one of the most important educational initiatives in the United States because the ability to comprehend text is critical to obtaining an education and is strongly associated with academic achievement (Calhoon, 2005). According to Mastropieri and Scruggs (1997), reading comprehension is one of the most vital academic skills students need to learn. It is essential for academic and lifelong success in any literate society, and it is becoming ever more critical considering the increasing reading demands of daily life (Shang, 2010).

Problem and Background

Despite the ongoing national focus on preventing reading failure, there are many students in the US who continue to have difficulty learning to read (Reis et al., 2007). Although there is an extensive knowledge base focused on students' reading acquisition including how to remediate early reading problems, improving students' reading proficiency remains an ongoing challenge for educators. It is particularly challenging as students get older, and the text and content demands become increasingly complex. Nevertheless, it is crucial for teachers to prevail in the face of these challenges, because students' failure to acquire proficient reading comprehension skills can negatively impact their overall school performance (Schenck, Walker, Nagel, & Webb, 2005).

Reading comprehension achievement. Although reading comprehension is the foundational skill on which academic success depends (Fuchs et al., 2002), a large number of students enter high school with significant reading deficits. Sadly, many high school students in the US are unable to comprehend grade level texts. In fact, on the 2015 National Assessment for Educational Progress (NAEP) the average score of the majority of 12th grade students fell below the benchmark for proficiency (National Center for Education Statistics [NCES], 2017). Among the 62% of 12th graders who scored below proficient, 37% scored on the basic reading proficiency level, and 25% scored the below basic level of proficiency (NCES, 2017). Practically speaking, these test scores indicate that 62% of 12th grade students in the US possess little to no mastery of the minimum reading skills necessary to perform everyday grade-level school work (Reis et al., 2007). Additionally, according to Hurst, Franklin, and Scales (2010), some struggling high school students even have difficulty reading above a second grade level.

Student level predictors of reading comprehension achievement. There are several individual characteristics that can impact students' reading comprehension achievement including their gender, minority status, and their socio-economic status (SES). In a review of the literature related to gender based reading disparities, Logan and Johnston (2010) reported that significant gender based reading achievement differences favoring girls have been found with elementary and secondary aged students. Logan and Johnston (2010) also pointed out that girls and boys naturally use different types of strategies to comprehend what they read. While there is a sufficient amount of research evidence to support this claim in terms of elementary aged students, there are relatively few studies that have explored gender based differences in high school students' reading comprehension strategy use (Denton et al., 2015). Similarly, race/ethnicity and SES based differences in high school students' reading comprehension strategy use are also understudied topics (Flowers, 2007). While there is a lack of research examining the influence of gender, minority status, and SES on high school students' reading comprehension strategy use, there is existing research focused on the relationships between these factors and high school students' overall reading achievement.

A common assumption regarding academic gender differences is that boys outperform girls in math and science, whereas girls perform better in reading, social studies, and languages. However, an inspection of the research indicates that these assumptions do not always hold true; studies of gender differences in reading achievement do not consistently produce results favoring girls (Lietz, 2006). Some studies showed that boys outperformed girls, other studies reported that girls outperformed boys, and still other studies reveal no gender differences in reading achievement (Chui & McBride-Chang, 2006). Gender differences in reading achievement favoring girls are often found in large scale studies of national and international reading assessments (Logan & Johnston, 2010). However, it must be noted that the results of studies with very large sample sizes need to be interpreted cautiously because with large samples relatively small, and not necessarily meaningful, performance differences can still yield statistically significant results.

In the research literature, reading achievement is often determined by a single outcome measure and reported in terms of gender without consideration of other influential variables such as a students' minority status or SES. In the 1960s, the Equality of Educational Opportunity Report, also known as the Coleman Report, revealed significant racial/ethnic and SES gaps in student achievement (Coleman et al., 1966). Although some research evidence regarding the effects of race/ethnicity and SES on student achievement existed prior to the Coleman Report, the significant results of the study drew new attention to these areas of student need. The heightened focus on these disparities in student achievement also sparked a host of new concerns that initially lead to a major growth in the amount of research centered on the minority and SES achievement gaps (Lee, 2002). The increased interest was not limited to the effects on students' overall achievement; the body of literature addressing minority and SES reading achievement gaps grew as well.

Since that time, differences in reading achievement have been frequently investigated and consistently found based on students' minority status and SES. For example, the average longitudinal NAEP reading scores from 1992 through 2009, indicated that white students, in every grade tested (fourth, eighth, and 12th), scored higher than their African American and Hispanic peers each year the test was given, and higher than their Asian peers in all but two assessment years (National Center for Education Statistics [NCES], 2011c). More specifically, Reardon, Valentino, and Shores (2012) stated that upon entering high school, the reading skills of African American and Hispanic students are approximately three years behind those of white and Asian students. Similarly, Sulkunen (2013) found that, in reading, students with a higher SES typically outperform students with a lower SES. Indeed, on average, the reading achievement of eighth graders with a low SES is comparable to third graders with a high SES (Buckingham, Beaman, & Wheldall, 2014). In other words, there is about a five year reading gap in favor of students with a high SES. Further, Reardon et al. (2012) reported that the SES reading achievement gap is growing.

Reading comprehension strategies. Numerous high school students struggle with comprehension even though they can accurately decode and understand the meaning of individual words. While these skills are certainly vital to reading comprehension, they are insufficient. The ultimate goal of reading is to comprehend, or understand, the meaning of a text to obtain knowledge from the ideas communicated by the author. While fluency is essential for comprehension, the meanings of individual words are not simply combined to obtain an understanding of the meaning of a complete sentence, paragraph, or an entire text (Klein, 2008; Nation, 2001). Rather, comprehension occurs when a student extends their thinking beyond the individual words in the text to obtain a deeper understanding of the meaningful ideas, and the relationships among those ideas presented by the author (McNamara, Ozuru, Best, & O'Reilly, 2007). This process requires the knowledge of a diverse set of reading comprehension strategies and skills.

Good readers read with fluency and employ the sophisticated skills and strategies necessary to comprehend and retain what they read in texts (Edmonds et al., 2009). For some students, the process of moving from comprehending individual words to meaningful ideas happens effortlessly (Rogoff, 2003). However, for other students, early reading skills do not naturally translate into more complex comprehension skills. As these students advance past elementary and middle school, their comprehension breaks down as the level of texts become more challenging (McNamara et al., 2007). Many of these students, including those with acceptable fluency, struggle to comprehend texts because they lack the knowledge of and/or inadequately use reading strategies to support their comprehension (Cantrell & Carter, 2009).

While reading has been a prevalent topic in academic research over the years, much of the focus has been on early reading skills (Hagaman & Reid, 2008). Research interest in foundational reading skills is understandable because they are integral to successful reading development (National Reading Panel [NRP], 2000). However, due to the concentration on basic reading skills, the body of research centered on improving the reading proficiency of adolescent students is relatively limited (Cantrell & Carter, 2009). Nevertheless, there is evidence within the existing literature that students' use of reading strategies is a significant and positive predictor of high school students' reading comprehension achievement (e.g., Cromley & Azevedo, 2007; Denton et al., 2015; Hong-Nam, Leavell, & Maher, 2014; Shera, 2014).

Reading comprehension strategy instruction. A federal report from the National Reading Panel (NRP) suggested instruction that explicitly taught reading comprehension strategies positively impacts high school students' reading achievement (Biancarosa & Snow, 2006). Further, Ness (2009) reported that the reading achievement of high school students improve when teachers explain and model reading strategies, and provide guided and independent practice with feedback until students begin to use the strategies independently. This suggests that high school students, especially those with reading difficulties, should receive explicit reading comprehension strategy instruction.

However, the lack of a clear understanding about content area reading strategies and high school students' reading strategy use often causes teachers to be unsure about which strategies to teach and how to most effectively teach them. Unfortunately, most of the existing research focused on effective reading comprehension instruction has been done at the elementary level (Cantrell & Carter, 2009). Although the research base regarding the unique needs of struggling high school readers is growing, not enough is yet known about how to best address their reading difficulties (Somers et al., 2010).

School level predictors of reading comprehension achievement. There are also several school level characteristics that can impact students' reading comprehension achievement including class time and class size. Notwithstanding the research evidence confirming the positive relationship between reading comprehension strategy instruction and reading achievement, the likelihood that students receive any type of reading comprehension instruction after fifth grade is rare (Irvin, Meltzer, & Dukes, 2007), this is especially true in high schools. Multiple studies, detailed in the literature review, reveal that little to no time is devoted to reading comprehension strategy instruction in U.S. high schools (e.g., Swanson et al., 2016; Wexler, Mitchell, Clancy, & Silverman, 2016).

Increasing class time has become a major policy initiative in countries that belong to the Organization for Economic Cooperation and Development (OECD). The focus on increasing class time resulted from sizable differences in classroom instruction time repeatedly used to account for considerable achievement disparities across the countries (Huebener, Kuger, & Marcus, 2016). Several studies discussed in the literature review provide evidence of a positive relationship between reading class time and reading achievement (e.g., Cattaneo, Oggenfuss, & Wolter, 2016; Huang, 2015; Kasapoglu, 2014; Lavy, 2015). Although there are some existing studies that have examined this relationship, minimal focus has been given to this topic (Huebener et al., 2016), and additional research in this area is needed.

Class size is another school factor associated with instruction that positively influences academic achievement (Graue, Hatch, Rao, & Oen, 2007). The relationship between class size and academic achievement has been a longstanding topic in educational policy debates. Advocates for reducing class sizes contend that small class size should be a core policy in education. Whereas, opponents of class size reduction (CSR) who are skeptical about its' true advantage to student learning, argue CSR is not cost effective (Harfitt & Tsui, 2015). Unfortunately, inconsistencies within the limited body of class size literature have done little to squelch this debate.

A preponderance of the evidence reported in class size studies conducted over the last 40 years comports with teachers commonly held beliefs that small class sizes positively effect student achievement. Yet, some researchers have found contradictory results in which class size had no impact on achievement (e.g., Hoxby, 2000; Leuven, Oosterbeek, & Rønning, 2008; Milesi & Gamoran, 2006). And other researchers assert findings of significant relationships between large class sizes and student achievement (Johnson, 2000; Shin & Chung, 2009). Most prior class size research has measured its' effect on elementary student achievement, and significantly less has explored its' influence on high school student achievement. Moreover, class size research specifically focused on reading achievement has exclusively studied elementary age students.

Consequently, no research, in the class size or reading literature, was located regarding the role class size plays in the reading achievement of high school students.

Reading attitudes. In addition to the student and school level variables associated with reading comprehension achievement, there is evidence in the literature suggesting that high school students' reading achievement is influenced by their attitudes toward reading. Even though the association between students' reading attitudes and reading achievement has been studied for many years, there is still no general agreement regarding the overall importance of this relationship (Petscher, 2010). Various issues make it difficult to conclusively assess the nature of the relationship between reading attitudes and reading achievement. The main difficulty is that the reported direction and magnitude of this relationship are inconsistent across studies. Specifically, the correlations in earlier research often range between .20 and .40, whereas more recent studies report correlations ranging from .60–.70 (Petscher, 2010).

Although still somewhat unclear, evidence of a relationship between students' reading attitudes and their reading achievement make it important to understand how their attitudes toward reading interact with other individual student characteristics. It is also important to understand how to strengthen high school students' reading attitudes, given the evidence that elementary, middle, and high school students' attitudes toward reading decline over time (Bokhorst-Heng & Pereira, 2008; Gökhan, 2012; McKenna, Kear, & Ellsworth, 1995). Unfortunately, like most other areas of interest in this study, a majority of reading attitude research focuses on elementary age students (McKenna, Conradi, Lawrence, Jang, & Meyer, 2012). Moreover, little research has focused on the relationship between high school students reading attitudes and reading comprehension strategy use (Lim, Bong, & Woo, 2015). Further, while there is existing research that examined the relationships between high school students' gender and reading attitudes (e.g., Bussert-Webb & Zhang, 2016; Gökhan, 2012; Jhang, 2014; Lim et al., 2015) and between their SES and reading attitudes (e.g., Gökhan, 2012), no studies regarding the relationship between their minority status and reading attitudes was located.

Conclusion. The pervasive inability to effectively remediate U.S. high school students reading deficits, as evidenced by their alarmingly low levels of proficiency, has led to an increased number of college students who have difficulty reading. According to Falk-Ross (2002), many new college students arrive ineptly prepared for the stringent reading requirements of most college courses. And the number of freshman who enter college with less than adequate reading comprehension ability is steadily increasing (Falk-Ross, 2002). According to the National Center for Education Statistics (NCES), in the 2011-2012 school year, more than 1.5 million college freshmen had to enroll in a remedial reading course (National Center for Education Statistics [NCES], 2014). The number of students currently entering college in need of reading remediation has grown significantly, from the approximately 950,000 students who needed it in the 2003-2004 school year (NCES, 2014). These statistics confirm assertions made in recent years by numerous colleges/universities and business, that a vast number of high school graduates are unable to effectively comprehend complex texts (Hasselbring & Goin, 2004).

The importance of reading proficiency extends well beyond the classroom. It is a prerequisite for students to become successful, productive adults (Hagaman & Reid, 2008). Reading proficiency has a considerable bearing on success in life because reading difficulties undermine the mastery of language as well as the understanding of general world knowledge (Calhoon, 2005). Reading difficulties are exceedingly detrimental in today's society because, to one degree or another, everyone's life depends on knowledge obtained from text (Savolainen, Ahonen, Aro, Tolvanen, & Holopainen, 2008). This knowledge is extremely valuable in that it fosters a sense of personal, social, and intellectual worth while simultaneously providing the necessary awareness for adequate functioning in our increasingly text oriented society (Harris & Hodges, 1995).

Purpose Statement

Considering the ongoing reading failure of countless U.S. high school students, the overarching purpose of this study was to examine several relationships associated with reading comprehension with the goal of substantiating new paths to increasing proficiency. The specific purpose of this correlational investigation was twofold. The first objective, due to the inconsistent and/or incomplete evidence in the reading research literature defining these relationships, was to examine the relationships among U.S. high school students' reading comprehension strategy use, reading comprehension strategy instruction, reading attitudes, and their reading comprehension achievement (as measured by the PISA 2009 U.S. Reading Literacy achievement scores). Five student and school level control variables (i.e., gender, minority status, SES, class time, and/or class size) were also included in the investigation based on existing research evidence suggesting that they have a significant relationship with one of the latent constructs and/or with reading comprehension achievement. The second objective of this study was to examine group differences in U.S. high school students' reading comprehension achievement, reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes across the five student and school level control variables with the intent of developing a more robust understanding of these unique relationships.

Research Questions

The following research questions were used to guide this investigation:

1. Do the observable indicators selected to measure the three latent constructs (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes) appropriately define them?
2. Are there direct and indirect effects among the five student and school level variables, the three latent variables, and high school students' reading comprehension achievement?
    a. What are the direct and indirect effects of the three latent constructs and the five student (i.e., gender, minority status, and SES) and school (i.e., class time and class size) level variables on high school students' reading comprehension achievement?
    b. What are the direct effects among the three latent constructs (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes)?
    c. What are the direct effects of the five student and school level variables on the three latent constructs?

3. Are there significant group differences in high school students' reading comprehension achievement, reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes across the five student (i.e., gender, minority status, and SES) and school (i.e., class time and class size) level variables?

Hypotheses

Several hypothesized relationships among the control variables, reading constructs, and reading comprehension achievement are proposed for testing in this study. Figure 1 illustrates the hypothesized relationships among the variables, based on a review of relevant empirical and theoretical research. The first hypothesis is that each latent variable will positively correlate with the two other latent variables. Specifically, it is expected that high school reading attitudes will positively correlate with reading comprehension strategy instruction. It also expected that strategy use will be positively correlated with reading attitudes and reading comprehension strategy instruction.

Next, it is hypothesized that each of the five control variables (gender, minority status, SES, class time, and class size) and each of the three latent variables (reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes) will have a direct effect on the reading comprehension achievement. No indirect effects of the eight control and latent variables on reading achievement are hypothesized. However, for exploratory purposes, the indirect effects of these variables in reading achievement will be included in the structural regression model.

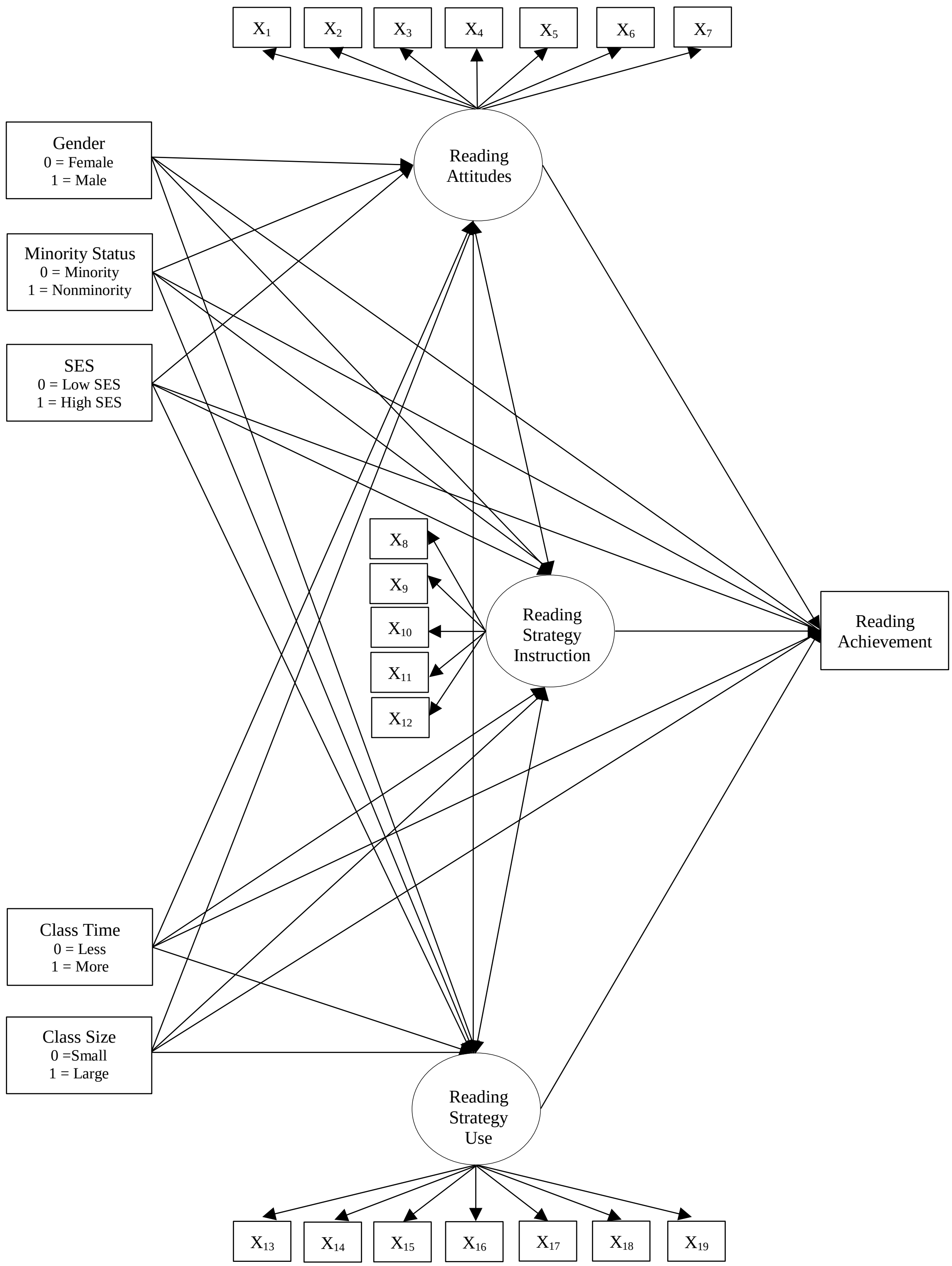


Figure 1. Hypothetical model of the proposed relationships among study variables.

It is also hypothesized that reading comprehension strategy instruction, gender, minority status, and SES will have a positive direct effect on reading strategy use and their attitudes toward reading. Additionally, it is hypothesized that reading attitudes, class time, and class size will each positively impact reading strategy use. Further, direct effects of class size and class time on strategy instruction are hypothesized.

Finally, significant group differences in reading strategy use, reading strategy instruction, reading attitudes, and reading achievement across the five student and school level variables are hypothesized. Considering the results of existing research related to these hypotheses of group differences, it is anticipated that the results of the MIMIC analysis for each of the four variables will favor girls, nonminority students, students with a high SES, more reading instruction class time, and smaller class sizes.

Rationale and Significance

The development of students' reading proficiency has, and will, always be a critical element of education because a students' failure to comprehend text can lead to long-term academic, psychological, social, behavioral, and economic hardships. Obtaining a better understanding of the important relationships among high school students' reading comprehension achievement, strategy use, reading strategy instruction, and reading attitudes was the fundamental motivation for this investigation.

There were several reasons why the population of high school, versus secondary, students were selected as the focus of this study. First, the construct of secondary students has been inconsistently defined across various educational references (both research and non-research). At times, the term was used to refer to students in grades five or six through 12 while other times it referred to students in grades nine through 12. Second, the population of high school students, versus all secondary students, was selected for this study due to the dearth of empirical research specifically targeting the reading needs of high school students (Cantrell & Carter, 2009; Kamil, 2008b). Third, although the sample of students in the U.S. portion of the PISA 2009 dataset represented grades 8-12, only four (0.1%) students were in the eighth grade. Finally, four concerning downward trends from the 1998-2015 NAEP, unique to the reading comprehension achievement of 12th grade students, highlight the importance of specifically focusing on improving the reading proficiency of high school students. First, from 1998 to 2015 the number of 12th grade students who scored at proficient decreased from 36% to 31%. Also, the total number of 12th grade students who scored below proficient increased from 59% in 1998 to 63% in 2015. Further, of the students who scored below proficient, the number of students who scored at the basic level decreased from 39% to 35% while the number of students scoring below basic increased from 20% to 28% (NCES, 2017).

Various studies have reported positive relationships among high school students' reading comprehension achievement, strategy use, reading strategy instruction, and reading attitudes, and between these constructs and various student and school level characteristics such as gender minority status, SES, class time, and class size. However, until recently, reading research has primarily centered on elementary age students, and most research with struggling readers beyond elementary school has only focused on students as old as eighth grade (Kamil et al., 2008b). Thus, less research attention has been given to improving the reading comprehension achievement of high school age students (Cantrell & Carter, 2009). Further, most extant multivariate research on high school students' reading comprehension achievement typically utilize a restricted design that examines a limited number of predictor variables (Silva, Verhoeven, & van Leeuwe, 2011). As such, the simultaneous effects of this set of predictor variables on the reading comprehension achievement of high school students have not yet been investigated.

The inconclusive, inconsistent, and often lacking research focused on these areas coupled with the goal of U.S. public schools to close the persisting reading achievement gap, supports the need for the present study. The results of this research will have important implications for parents, teachers, administrators, educational policy makers, and institutions of higher learning regarding various critical components of high school students' reading comprehension achievement. For example, a better understanding of how individual characteristics such as gender, minority status, and SES affect students' reading achievement, strategy use, and reading attitudes provided by the results of this study will serve to inform high school reading instruction, which is important considering high quality, research based, reading instruction is one key to preventing further high school reading failure. Additionally, the results of this study related to the effects of class time and class size on reading comprehension strategy instruction and high school students' reading achievement will inform educational policy makers interested in making changes necessary to improve the reading achievement of high school students.

Glossary of Terms

The key terms in this study are defined as the following:

Reading comprehension achievement. In this study, reading comprehension achievement (i.e., reading literacy achievement) referred to students' ability to access, retrieve, integrate, and evaluate textual information (Fleischman, Hopstock, Pelczar, & Shelly, 2010). During the study, the terms reading achievement and comprehension achievement were used as synonyms of this term for brevity, to avoid redundancy, and due to limited space in tables and figures.

Reading comprehension strategy use. Reading comprehension strategy use was broadly defined in this study as the intentional application of a reading strategy to repair or improve comprehension. The terms reading strategy use, and strategy use were used as synonyms of this term for the same reasons described above.

Reading comprehension strategy instruction. Reading comprehension strategy instruction was defined in this study as the explicit instruction of comprehension strategies. Explicit reading comprehension strategy instruction requires direct instruction including modeling of strategy use, guided and independent practice, progress monitoring, and feedback (Ness, 2009). Amid the study, this term was replaced by the terms reading strategy instruction, and strategy instruction for the above reasons.

Reading attitudes. Reading attitudes were defined in this study as the spectrum of positive to negative feelings about reading that cause a reader to pursue or avoid reading opportunities and reading related activities (Alexander & Filler, 1976).

U.S. high school students. In this study, this term referred to United States public school students enrolled in ninth through 12th grade.

Gender. The gender of a student, in this study, was defined by whether they self-identified as a male or a female on the PISA 2009.

Minority status. The terms race and ethnicity are often incorrectly conflated. As Hudley, Graham, and Taylor (2007) explained, this is common because although race and ethnicity are theoretically different, they are not mutually exclusive. In theory, race is defined as a group of people who have common genetic, biological, and physical features. Whereas, ethnicity refers to a shared history, nationality/geography, language, and culture. The term minority status was chosen for this study to avoid this common confusion of ideas. Minority status was defined based on the race/ethnicity group with which students self-identified as on the PISA 2009.

Socio-economic status. SES was defined in this study as their family's rank within the social hierarchy, determined by their access to or possession of a combination of wealth, power, and social status (Caro, McDonald, & Willms, 2009).

Class time. For this investigation, class time was defined as the product of the minutes of class per day and the number of class sessions per week.

Class size. In this study, class size referred to the number of student assigned to one teacher's class.

Chapter Two

In his article focused on high school reading instruction, Smith (1976) recounted an interview that was part of a study on social studies teaching conducted by Negley (1975). The interview consisted of a group of 14 high school teachers who worked in an average, middle class neighborhood. One of the social studies teachers revealed that the problem with high school students' reading had gotten so bad that she stopped giving reading assignments. The teacher explained that she felt like it was useless to assign class readings because most of her students couldn't read. She added that she also didn't use the course textbook during classroom instruction. She explained that instead, the textbook was merely used as a reference for students who wanted to learn more about a topic than was covered during class. She also explained that she only used verbal instructional techniques in her classroom teaching (e.g., lectures, discussions, cooperative learning projects, movies). Sadly, eight of the 14 teachers involved in the interview agreed with this teacher's statements (Smith, 1976).

The notion that these high school teachers believed most of their students were unable to learn from reading is unsettling. Yet, it's even more alarming that, according to longitudinal national reading assessment data, the low level of high school students' reading proficiency discussed in this interview has remained relatively unchanged in the more than 40 years since this article was published. Based on the longitudinal NAEP, the

2012 reading proficiency scores of 17-year-old students were not significantly different than their scores in 1971 (National Center for Education Statistics [NCES], 2013).

By nature, reading comprehension difficulties are complex. Reading comprehension is a complicated process involving knowledge, experience, thinking, and teaching (Prado & Plourde, 2011). Proficient comprehension is a critical skill students need to master because virtually all academic learning requires successful comprehension of information read in various texts (Hulme & Snowling, 2011). Reading comprehension demands for high school students are increasingly challenging due to the heightened difficulty of content area texts, and because students are expected to read a wider variety of genres and subject areas that require advanced skills. In order for students to read the required high school texts with proficiency, they must possess and effectively utilize cognitive and metacognitive reading comprehension strategies. High school students' reading comprehension achievement is also influenced by their positive attitudes toward reading and the flexibility to read for a wide variety of purposes across the spectrum of media. The review of literature presented in this chapter is a convergence of research from the bodies of literature addressing the following topics: reading comprehension, reading comprehension strategies, reading comprehension strategy instruction, instructional class time, instructional class size, and reading attitudes.

Reading Comprehension Achievement

Comprehension is the desired outcome of reading (Oakley, 2011). To comprehend, or understand the information described in a text, readers must simultaneously extract, synthesize, and integrate new ideas in the text with prior knowledge in order to construct meaning (Snow, 2002). The construction of meaning is an intentional, interactive, process between a reader and a text. During this interaction, comprehension is influenced by readers' experiences, abilities, motivation, and goals for reading (Anastasiou & Griva, 2009; Artelt, Schiefele, & Schneider, 2001). This definition of reading comprehension is based on Kintsch's (1988) construction-integration (CI) model which, according to Andreassen and Bråten (2010), is currently recognized as the most complete framework of reading comprehension.

The CI model views reading comprehension as the product of construction and knowledge driven integration processes. It is a systematic reading process that integrates basic as well as higher-order reading skills (Kintsch, 1998). The CI model extended Gough and Tunmer's (1986) often cited Simple View of Reading (SVR) theoretical framework. The SVR framework narrowly defined reading as the product of decoding and listening comprehension. Kintsch expanded the scope of the comprehension component of the SVR framework to include the text processing and prior knowledge aspects of comprehension. Kintsch further extended the SVR framework by incorporating the role of reading strategies in comprehension. Kintsch's rationale for including reading comprehension strategies was that students must be taught how to use cognitive and metacognitive strategies to support their understanding when their lack of prior knowledge presents roadblocks to comprehension (Deshler & Hock, 2007).

The CI model draws a distinction between two essential layers of mental representation of a text, the textbase and the situational models (Andreassen & Bråten, 2010). The textbase model is an account of what the text says (i.e., the written words in the text organized into sentences and paragraphs); whereas, the situational model is an account of what the text means (Palinscar & Schutz, 2011). While a textbase level of understanding occurs when the reader comprehends the text literally, in order to develop the situation model, readers must integrate the textbase level information in the working memory with prior knowledge retrieved from the long-term memory.

Constructing a situation model is central to reading comprehension. It allows readers to develop new knowledge from the text by combining the new information they read with their existing knowledge base (Duke, Pearson, Strachen, & Billman, 2011). Struggling readers who lack the prior knowledge and/or word-level skills necessary to develop these models and comprehend what they read need specific strategies to make up for their deficiencies (Deshler & Hock, 2007). According to Kintsch (2005), when comprehension breaks down students need to engage in strategic reading. Thus, they must be aware of the cognitive and metacognitive strategies that can be used to support their comprehension (McNamara et al., 2007). As such, reading comprehension strategy instruction, which is supported by the CI model, is necessary to teach students how to effectively comprehend text (Kintsch, 2005).

Theoretical foundation – Social cognitive theory. This conception of reading, based on the CI model, is grounded in Bandura's (1986) social cognitive theory of learning. Reading comprehension is a cognitive and social process in which students interact with each other, their teachers, the text, and the socio-cultural context for reading to achieve specific goals (Griffiths, Sohlberg, & Biancarosa, 2011). Most learning occurs in a social environment and according to social cognitive theory, cognitive and metacognitive reading comprehension skills are partially acquired through social processes (Schunk, 2011). However, readers rely on more than cognitive and metacognitive reading skills to construct meaning. The entirety of a readers' experiences and background knowledge, as well as individual psychosocial and biological factors also influence a readers' understanding of a text (Artelt et al., 2001; Griffiths et al., 2011).

Development of reading comprehension. Although formal instruction in reading does not begin until kindergarten, children can develop early literacy skills prior to the start of formal schooling by observing others reading behaviors. As they observe others reading, children become acquainted with letters, words, and books and they begin to learn the importance of reading (Jacobs, 2008). Children typically begin to receive formal instruction in reading around the age of five, when they enter kindergarten. In the primary grades (K-3), they receive direct instruction to learn a myriad of basic reading skills. According to Reardon et al. (2012), by third grade most students can read using basic word reading skills (e.g., sounding out words, sight reading, comprehending words in context, and basic inferencing). However, the goal of teaching students to read is to advance them beyond basic skills so they can learn how to focus their limited cognitive processing capacity on comprehension (Randi, Newman, & Grigorenko, 2010).

As students enter the upper elementary grades they likely experience a drastic change in reading expectations put upon them. The focus of reading instruction quickly transitions from learning to read to reading to learn. Rather than a focus on basic reading skills they were accustomed to in the primary grades, students are expected to decode and comprehend difficult text with virtual independence (Hasselbring & Goin, 2004).

Reading comprehension in the upper elementary grades (and beyond) involves knowledge of content area vocabulary, background knowledge, understanding of complex text structures, and critical thinking (Wanzek et al., 2013). These tasks require students to not only understand the meaning of text but to also evaluate, retain, and apply new knowledge obtained from text (Alfassi, 2004). Throughout the middle and high school years, students are required to independently analyze and synthesize content area texts that contain "multiple layers of meaning from multiple points of view that often contrast and conflict" (Jacobs, 2008, p. 15). Therefore, students need extensive background knowledge as well as knowledge of reading comprehension strategies to effectively comprehend text.

The important role of reading comprehension in academic success increases as students progress through school. It lays the foundation for the acquisition of disciplinary knowledge (Valencia, Pearson, & Wixson, 2011), and according to Hamilton (2009), students who read well perform better in other academic content areas. Unfortunately, learning to read for understanding can be a challenging task for all children. Even fluent readers are challenged by the complex cognitive demands of reading comprehension as text difficulty increase (Randi et al., 2010). As students move through the grades, the gap between strong and struggling readers increases (Hasselbring & Goin, 2004). In fact, Ness (2009) reported that a large number of middle and high school students struggle with the complex literacy tasks required in the content area classes. Many struggling readers can decode what they are reading, but they get little out of their reading beyond that (Prado & Plourde, 2011). Comprehension deficits make it difficult, if not impossible, for readers to benefit from instruction in most content areas because information derived from text is the primary source of knowledge (Hagaman & Reid, 2008). In other words, students with reading difficulties are unable to access all the information taught, which often leads to poor performance in the academic content areas (Savolainen et al., 2008).

Importance of reading comprehension. The importance of preventing reading failure cannot be overstated (Mathes, Grek, Howard, Babyak, & Allen, 1999). Compared to skilled readers, students with poor reading skills are unfortunately less likely to complete high school, and those who do are less likely to pursue a post-secondary education (Fuchs et al., 2001; Wanzek, Wexler, Vaughn, & Ciullo, 2010). Indeed, there is a significant correlation between low reading achievement and high dropout rates (International Reading Association [IRA], 2002). This is likely because reading difficulties make the transition to increasingly challenging academic work almost insurmountable (Reis et al., 2007). As a result, students with reading difficulties typically choose less academically demanding educational programs. In turn, this increases the likelihood that students will end up in a lower educational trajectory, which is often associated with dropout rates (Savolainen et al., 2008).

The impact that proficient reading comprehension has on students' post-secondary readiness and performance is also significant. According to Savolainen et al. (2008), students with reading comprehension difficulties are less likely to graduate college, and those who do are more likely to graduate with lower qualifications. Reading comprehension plays a key role in preparing students for higher education because the ability to read complex text independently and proficiently is essential for their success (National Governors Association Center for Best Practices & Council of Chief State School Officers [NGACBP & CCSSO], 2010; Valencia et al., 2011). Reading proficiency is an even greater necessity at the collegiate level because college students, in comparison to high school students, are expected to read more complex texts with greater levels of independence. In college, students are expected to be able to learn independently from the required (out of class) readings assigned to supplement the in class instruction. The material in the out of class readings is often not explicitly taught in class, yet students are required to demonstrate their understanding of the material through tests, papers, and their participation in class discussions and presentations (NGACBP & CCSSO, 2010).

In addition to the poor academic achievement outcomes related to reading failure (Hitchcock, Prater, & Dowrick, 2004), the negative psychosocial and behavioral effects often associated with reading failure can have a detrimental impact on students' overall educational experience. There is evidence that students with poor reading skills often have negative attitudes towards school as well as lower academic self-concept and lower motivation to learn (Savolainen et al., 2008). Many of these students are known to struggle with low self-confidence and low self-esteem, which compound their predisposition towards learned helplessness (Fuchs et al., 2001; Schenck et al., 2005). Further, students with reading difficulties often demonstrate increased behavioral issues and pose greater disciplinary problems than their peers (Hitchcock et al., 2004). In fact, according to The Melissa Institute for Violence Prevention and Treatment (2010), "Up to

80% of violent juvenile offenders are functionally illiterate. And, in general, the lower the reading comprehension rate, the more violent the behavior" (p. 1).

Reading comprehension is an important prerequisite for lifelong success. The skills developed in school, which rely on the ability to independently and proficiently read complex text, have important bearing on success in adulthood. These skills overlap with those required for the reading tasks presented in everyday life (Kaestle, Damon-Moore, Stedman, Tinsley, & Trollinger, 1991; NGACBP & CCSSO, 2010). Unfortunately, the results of several longitudinal studies of reading difficulties indicated that early reading problems often persist into adulthood (Savolainen et al., 2008). The negative impact reading deficits have on academic achievement mirrors the negative impact it can have on lifelong success; limited reading proficiency can negatively affect the economic, social, and personal aspects of one's adult life.

Reading comprehension proficiency plays a key role in preparing students for success in the work place (Spörer & Brunstein, 2009; Valencia et al., 2011). As with higher education, the ability to read complex text independently and proficiently is essential for high achievement in the workplace. Like all aspects of life, reading demands in the workforce, specifically in training programs, have largely increased over the last fifty years (NGACBP & CCSSO, 2010). Thus, reading is a fundamental skill on which most successful, secure employment depends, and people with low levels of reading ability do not do as well in the job market (Calhoon, 2005). Poor reading skills correspond directly with lower levels of job success and are often associated with limited employment opportunities and/or fewer chances for advancement (Hitchcock et al., 2004). Moreover, poor reading skills are also heavily correlated with lower levels of financial success including lower wages, unemployment, and a greater likelihood of living in poverty (Fuchs, Fuchs, & Kazdan, 1999; IRA, 2002). Effective communication, another important job-related skill, also corresponds directly to reading comprehension proficiency (Hamilton, 2009). According to Randi et al. (2010), understanding language is essential for effective communicative interactions. The importance of effective communication is not only limited to job success, it is fundamental to successful everyday functioning.

The fact that reading proficiency increases the likelihood of academic and economic success is frequently reported. Yet, the equally crucial role it plays in successful personal, social, and civic functioning is not given as much attention. Reading is a foundational skill on which personal autonomy depends, and corresponds directly to one’s ability to achieve personal fulfillment (Calhoon, 2005). As Hamilton (2009) pointed out, reading skills correlate with almost every measure of personal and social behavior. Poor reading skills correspond directly to one’s ability to be an informed citizen (Hamilton, 2009). They are also associated with a lack of civic awareness and involvement, meaning that deficient readers are less likely to become active in civic and cultural life (Fuchs et al., 1999; Hamilton, 2009). Of potentially greater importance is the fact that juvenile and adult reading deficits are also associated with poor health maintenance, substance abuse, violence, and incarceration, which are all currently major issues of national concern (Hamilton, 2009; National Institutes of Health [NIH], 1999).

Student level predictors of reading comprehension achievement. Although reading comprehension difficulties are relatively common, they often go unrecognized (Hulme & Snowling, 2011). Most middle and high school teachers presume that students have adequately mastered the fundamental reading skills and are able to read text with basic comprehension (Ness, 2009). In reality, a great number of students fail to learn basic reading skills in the primary grades, and enter the upper elementary grades with significant difficulties reading (Manset-Williamson & Nelson, 2005; Therrien, 2004). According to Al Otaiba and Fuchs (2006), as many as 30% of students at risk for reading difficulties may not have benefitted from primary reading instruction. Reading comprehension is a process that develops over time and requires mastery of a number of different skills. Thus, in order to learn how to effectively support struggling readers it is important to identify all the variables that can impact the development of reading comprehension. Extensive evidence exists supporting the importance of basic reading processes such as decoding and fluency in developing reading comprehension. The correlations among these skills vary between 0.3 and 0.6 indicating the need to explore additional factors involved in the variance of reading comprehension (Anastasiou & Griva, 2009). Gender, minority status, and SES are three factors known to play an important role in students' comprehension.

Reading comprehension achievement and gender. Gender plays an important role in reading comprehension. Gender patterns are not innate, they develop over time as a consequence of individual life experiences (Hyde & Durik, 2005). Reading researchers have posited various explanations for gender differences in reading. They are frequently attributed to differing cognitive processing, socio-cultural experiences, and/or reading behaviors typically associated with males and females (Singh, 2008). Brain development researchers have also contributed to the understanding of gender differences in reading achievement. According to Prado and Plourde (2011), girls are advantaged in reading acquisition because, when compared to the brain functioning of boys, they typically utilize more of their brains. Girls are also at an advantage in terms of reading achievement because they tend to have greater neural connectivity in their cerebral cortex which allows them to process information more quickly than boys. Further, girls' hippocampus is typically larger than boys, thereby increasing their capacity to remember information (Gurian, 2010). Additionally, evidence from Position Emission Tomography (PET) scans and Magnetic Resonance Imaging (MRI) studies have indicated the female brain, at rest, is more active than the activated male brain (Gurian, 2010). Since the male brain is less active, it is more easily overwhelmed by stimulation, which disadvantages boys as they attempt to begin comprehending text (Prado & Plourde, 2011).

There is ample support in the literature that comports with the notion that girls outperform boys in reading comprehension. For example, Lietz (2006) conducted a meta-analysis to examine gender differences in the reading achievement of secondary school students. The sample included 139 large scale national and international studies completed between 1970 and 2002. The estimated grand mean effect size of the gender differences from the hierarchical linear modeling was positive and small, $g = 0.19$. In other words, females performed 0.19 standard deviations higher than males.

Similarly, the OECD stated that for each Program for International Assessment (PISA) administration, girls demonstrated higher reading achievement scores than boys (Organization for Economic Cooperation and Development [OECD], 2011). The OECD is the organization that initiated and continues to manage the PISA, a triennial international assessment of 15 year-old students' academic achievement (Fleischman et al., 2010; Organization for Economic Cooperation and Development [OECD], 2010d). The OECD (2010d) also reported that since the test began in 2000 the reading achievement gender gap has increased over 20%.

Chui and McBride-Chang (2006) examined gender differences in reading using PISA 2000 scores from 43 countries. The results of their hierarchical linear modeling analysis revealed that girls outperformed boys in every country. Further, gender differences favoring girls were statistically significant in all countries except Romania and Peru. They also found that, on average, girls outscored boys by 22.73 points, and gender explained about 2% of the differences in reading achievement. Likewise, the results of several studies conducted in individual OECD countries revealed statistically significant differences between girls and boys reading scores favoring girls (e.g., Huang, 2015; Kasapoglu, 2014; Shera, 2014; Singh, 2008).

While a majority of the research examining gender differences in reading supports the existence of a gender gap in reading performance in favor of girls, some studies of adolescent reading dispute this finding. In 2007, White investigated gender differences in reading on the Ontario Secondary School Literacy Test. The sample included data on nine sub-scores of reading achievement from 113,505 10th grade students. Neither the effect size of the total reading score ($d = 0.15$), nor any of the sub-scores indicated substantive gender differences. White (2007) reported significant differences, in favor of girls, on two of the nine sub-scores, however, the magnitudes of the effect sizes were negligible ($.02 < d > .11$). Additionally, Duncan, McGeown, Griffiths, Stothard, and Dobai (2016) examined gender differences in reading using data from the York Assessment for Reading Comprehension (YARC). The results of the two-way between participants ANOVA indicated no overall gender difference in reading achievement, however, males outperformed females in non-fiction reading achievement and females outperformed males in fiction achievement.

Reading comprehension achievement and minority status. The important findings of the 1966 Coleman Report raised early awareness in the education community regarding the significantly negative influences of race/ethnicity on students' academic achievement (Lee, 2002). Regrettably, the achievement gap between minority and nonminority students continues to be a critical impairment to the academic achievement of many students in the US. Indeed, Snow and Biancarosa (2003) reported that students who belong to a minority group are more likely to have difficulty learning to read, which results in their consistently lower reading achievement than nonminority students throughout middle school and high school. Buckingham et al. (2014) echoed this sentiment stating, upon entering high school, the average reading achievement of African American and Hispanic students is three years behind that of their white and Asian peers.

In 2012, Reardon et al. conducted a qualitative review of studies using data from large scale national and international literacy assessments (ECLS-K, NAEP, PISA, &

PIRLS) to examine various factors associated with U.S. students' reading proficiency. The authors discussed three specific areas related to the minority status reading gap: the African American-white gap, the Hispanic-white gap, and the Asian/Pacific Islander-white gap. They explained the fluctuating African American-white achievement gap, stating that the large gap of 1970s decreased in the 1980s, increased in the early 1990s, and since the late 1990s it has continually decreased. Likewise, the pattern of the Hispanic-white reading gap basically followed the same temporal pattern as the African American-white gap. Reardon et al. (2012) also pointed out important nuances within this pattern based on the origin of Hispanic students (e.g., scores of students from Central America and Mexico are frequently lower than those of students from Cuba, Puerto Rico, or South America). Finally, the review indicated the Asian/Pacific Islander-white reading gap is small and has barely changed over the past 30 years.

Brown-Jeffy (2006) also investigated the reading achievement disparity between minority and nonminority students in the US. This study utilized data from the High School Effectiveness Study (HSES) assessment, a subset of the original National Educational Longitudinal Study of 1988 (NELS: 88), administered at 219 high schools to 4,065 10th and 12th grade students. Hierarchical linear modeling (HLM) was used to examine the influence of the schools' racial makeup regarding the African American-white reading achievement gap. The results indicated that African American-white reading achievement gaps were larger in schools with less minority students. Specifically, the African American-white reading gap was higher in schools with less than 10% African American, Hispanic, and/or Native American students compared to schools with 25-54% African American, Hispanic, and/or Native American students.

Reading comprehension achievement and socio-economic status. The effects of a student's minority status on their reading achievement are closely related to the effects of their socio-economic status (SES). Hudley et al. (2007) believed that most of the disparities associated with the racial/ethnic achievement gap are a function of social and economic inequities. This notion is supported by empirical evidence that the relationship between student achievement and SES is moderated by a student's minority status (Kitsantas, Ware, & Cheema, 2010). In their meta-analytic investigation, described below, Sirin (2005) found that minority status significantly predicted the relationship between SES and student achievement, Q (1,35) = 164.86, $p < .001$. As such, a clear understanding of the impact SES has on achievement can be confounded by the large concentration of racial/ethnic minorities often associated with low SES because, as Maerten-Rivera Myers, Lee, and Penfield (2010) explained, the important effects of minority status and SES are difficult to disentwine. Thus, the effect of a student's SES on reading achievement is another important factor associated with the reading achievement gap in need of further attention.

SES is among the strongest predictors of academic achievement (Brown-Jeffy, 2006). As a result, it is one of the most commonly used context variables in educational research (Sirin, 2005). SES is defined as the rank of an individual or family within a social hierarchy determined by their access to or possession of a combination of wealth, power, and social status (Caro et al., 2009). Although is there no one accepted method of measuring students' SES, it is typically measured using their family income, parents' education level, and parents' occupation (Jeynes, 2002).

In 2005, Sirin conducted a meta-analysis to explore the relationship between SES and academic achievement. This study analyzed a sample of data including 101,157 students (6,871 schools) obtained from 58 articles published between 1990 and 2000. Fixed and random effects modeling were used for the main effect size analysis. Overall, the results revealed a medium to strong average correlation, $r = .29$, between student achievement and SES and an effect size, ES = .25, for the 22 analyzed correlations of high school students. Students from low SES backgrounds demonstrate lower achievement as a result of various factors such as lack of economic resources, lower parental involvement, and limited access to high quality educational opportunities (Huang, 2015). Sirin (2005) also analyzed the moderating effects of six aspects of SES (family income, parental occupation, parental education, student's free or reduced lunch status, home resources, and neighborhood characteristics) and found that student's home resources had the largest moderating effect on achievement (ES = .51).

In addition to the research evidence demonstrating the significant relationship between students' SES level and overall academic achievement, evidence exists in the reading research literature that supports the relationship between students' reading achievement and their SES status. Buckingham et al. (2014) stated that there is substantial research evidence indicating that SES is positively related to reading from before students enter school up and through high school. They explained that is likely because the development of early reading skills is less common among students from disadvantaged backgrounds. Further, students with a low SES are less likely to successfully transition to higher reading levels than their peers with a high SES (Sonnenschein, Stapleton, & Benson, 2010). According to Buckingham et al. (2014), upon entering high school, the average reading achievement of students with a low SES is five years behind students with a high SES.

A multitude of researchers have studied the relationship between high school students' SES and reading achievement using a variety of quantitative methods to analyze data from different versions and years of the PISA. For example, Chui and McBride-Chang (2006) used hierarchical linear modeling (HLM) to analyze data from the PISA 2000**.** Their sample consisted of 199,097 15-year old students from 43 of the OECD countries. The results of their HLM analyses indicated that the relationship between students' SES and reading achievement was statistically significant. They also found that students' SES levels accounted for 24% of the total variance in reading achievement and that for every 10% increase in SES students' reading scores, on average, increased 4.7 points. Shera (2014) also investigated this relationship using HLM to analyze PISA data. The data used in this study was obtained from the Albanian PISA 2009, which included 4,596 students from 181 schools. The results of the analysis of an individual country using HLM were consistent with the results of the previous multi-country study. Shera (2014) found that students with a low SES performed significantly lower than students with a high SES (β = 11.01/ES = .20). In other words, students with a high SES demonstrated higher reading achievement than students with a low SES.

Singh (2008) also examined PISA data from one individual country, Canada, although t-tests were used to analyze the PISA 2000 test data, rather than HLM. Again, the results of this study indicated a statistically significant relationship between SES and reading achievement $t(259{,}017) = 13.54$, $p < .01$, $d = 0.17$ (Singh, 2008). Likewise, Özdemir and Gelbal (2014) analyzed PISA 2009 data from one individual country. They employed canonical commonality analysis to examine the data of 4,496 students from Turkey. The results indicated that SES was highly correlated ($r = .67$) with students' reading achievement and contributed to 84% of the variance in students' reading scores. These findings align with the results of the previous three studies.

Finally, Huang (2015) used a combination of statistical methods to analyze the relationship between SES and reading achievement. This study examined the PISA 2012 U.S. dataset of 4,978 15-year old students from 162 schools using both ANOVA and HLM fixed-effect analyses. The results of ANOVA employed first indicated that students with high SES performed significantly higher than students with low SES. Specifically, students with high SES scored an average of 86 points more than students with low SES. The ANOVA results also revealed that SES accounted for 14% of the variance in students' reading achievement. Due to the nested structure of the data and knowledge of the student and school control variables related to SES and achievement, HLM was then utilized to reexamine the SES reading gap initially estimated using ANOVA. The results of the HLM confirmed the ANOVA estimates yielding similar findings concerning the statistically significant relationship between SES and reading achievement. Specifically, the HLM results indicated that a one-unit increase in SES was associated with an increase of 20.14 points ($p < .01$) in reading achievement while controlling for student and school level variables.

Understanding the impact of SES on U.S. students' reading achievement is critical for a variety of interrelated reasons. Specifically, their persistently low levels of reading proficiency, the increasingly strong association between family income and achievement (Reardon et al., 2012) the large percentage (18%) of children in the US under 18 living in poverty (United States Census Bureau, 2017), and the growing socio-economic achievement gap that has exceeded the achievement gap between minority and nonminority students (Reardon et al., 2012).

Conclusion. Reading comprehension demands increase as students progress through school and are required to read more advanced texts. One key to learning how to read challenging texts is knowledge of how and when to use different reading strategies to aid comprehension (Paris, Lipson, & Wixson, 1983). Proficient reading comprehension requires the successful selection, application, and monitoring of understanding (Gersten, Fuchs, Williams, & Baker, 2001). Unfortunately, struggling high school students often lack a sufficient repertoire of reading strategies necessary to help them comprehend what they read.

Reading Comprehension Strategies

Students' use of reading comprehension strategies is another variable that can significantly impacts their reading comprehension achievement. Reading comprehension strategies are defined as the cognitive and behavioral processes selectively employed and intentionally applied by a reader to maintain or repair text comprehension (Graesser, 2007). Over and above basic reading skills and fluency, proficient comprehension requires the knowledge of and ability to effectively use a wide range of comprehension strategies (Andreassen & Bråten, 2010). Unfortunately, many high school students have not acquired this strategic knowledge and are thus unable to effectively utilize reading comprehension strategies to adequately comprehend what they read (Cantrell & Carter, 2009). The lack of comprehension strategy awareness is particularly concerning because, according to Cromley and Azevedo (2011), high school students' reading comprehension strategy use positively predicted their reading comprehension achievement.

Proficient readers use a combination of strategies before, during, and after reading to maximize their comprehension of a text. For example, before reading, skilled readers identify their purpose for reading and determine the specific goals needed to accomplish their desired learning from the text. They also preview the text to identify the text structure and specific sections that may be most relevant to their learning goals (Anastasiou & Griva, 2009). After previewing the text, strategic readers predict what they think will occur in the text (Woolley, 2011).

During reading, strategic readers evaluate whether they are meeting their established goals, monitor their understanding, recognize breakdowns in their comprehension, and make adjustments to their reading when necessary (Denton et al., 2015). They also distinguish between important information and supporting details to determine what to read carefully, what to skim, and what they need to reread (Gersten et al., 2001). While reading, competent readers use context clues more effectively to identify the meaning of unfamiliar words and concepts in the text. Additionally, they are better able to recognize deficiencies in their understanding of the text and employ the appropriate strategies to support their comprehension. After reading, strategic readers identify the main idea of the text and construct summaries of what they read (Duke & Pearson, 2008). Overall, proficient readers are more cognizant of the strategies they use, and they use them more efficiently than poor readers (Hong-Nam et al., 2014).

Indeed, the results of various experimental, correlational, and observational studies indicate that reading comprehension strategy use has a positive relationship with students' reading achievement. For example, Shera's (2014) investigation revealed that students' use of reading comprehension strategies was associated with increased reading achievement. Data from the PISA 2009 Albanian dataset, which included 4,596 students, was used in this previously mentioned study. The results of hierarchical linear modeling indicated that students' use of reading comprehension strategies significantly and positively predicted their reading achievement ($\beta = 12.89/ES = .24$).

Similarly, in their study to test the validity of three researcher developed reading comprehension strategy use measures, Cromley and Azevedo (2007) examined the relationship between reading comprehension strategy use and reading achievement. The researchers also examined the effects of vocabulary, prior knowledge, inferencing, and word reading on students' reading achievement. The participants of the study included 177 (44% girls, and 56% boys) ninth grade students (*M* age = 14.20, *SD* = 0.55). Data for the investigation was obtained from the following instruments: student demographics form, 23 of 45 odd numbered items of the Gates-MacGinitie Vocabulary subtest (GMRT-V; K-R 20 = .90-.92; MacGinitie, MacGinitie, Maria, & Dryer, 2001), the 48 item Gates-

MacGinitie Reading Comprehension subtest (GMRT-RC; K-R 20 = .91-.93; MacGinitie et al., 2001), researcher developed 13 item background knowledge test based on the GMRT-RC content, researcher developed 10 item multiple-choice inference test, researcher developed 10 item multiple-choice measure of strategy use, a 1-minute timed individual passage reading task, and the Letter-Word Identification and Word Attack subtests of the Woodcock Diagnostic Reading Battery (Woodcock, 1997).

Cromley and Azevedo (2007) used canonical commonality analysis was to identify the unique and shared proportions of variance of the five variables of interest (i.e., vocabulary, prior knowledge, inferencing, word reading, and reading achievement). The statistically significant results indicated that the set of five predictors account for 67% of the variance in students' reading comprehension. As far as the shared variance explained by the five predictors, vocabulary accounted for 60% shared variance, background knowledge accounted for 51%, inferencing accounted for 44.0%, word reading accounted for 35%, and comprehension strategy use accounted for 34.0% of the shared variance. Inferencing is typically considered a reading comprehension strategy; thus, it is possible to conclude that reading comprehension strategy use accounted for 78% of the shared variance in students' reading comprehension.

Likewise, Hong-Nam et al. (2014) investigated the relationship between students' reading comprehension strategy use and reading achievement. The study participants consisted of 2,789 students (47% male and 53% female) ranging in age from 14 to 20 years old (*M* age = 16 years old). The three following measures were used to collect data in this study: researcher developed Individual Background Questionnaire (IBQ), the

Metacognitive Assessment of Reading Strategies Inventory (MARSI; Alpha = 0.89 to 0.93; Mokhtari & Reichard, 2002), and the Texas Assessment of Knowledge and Skills (TAKS), reliability not reported. The Pearson *r* correlation coefficient was calculated to determine the correlation between reading achievement (based on TAKS reading scores) and students' overall use of reading comprehension strategies (based on MARSI scores). The results revealed a statistically significant correlation between students' reading achievement and reading comprehension strategy use ($r = 0.91$, $p < 0.05$), indicating a strong positive linear relationship. A one-way ANOVA was conducted to determine the variation in strategy use related to reading achievement. The results indicated a statistically significant difference in strategy use ($F = 11.79$, $p = 0.00$) among the high, medium, and low achievement level groups. Students with high TAKS scores reported using reading strategies more than students in the medium and low achievement groups.

In addition to the positive effects of overall strategy use, studies have investigated effects of individual and sets of reading strategies. For example, in their investigation of differences in adolescents' use of reading strategies in terms of proficiency, grade level and gender, Denton et al. (2015) specifically explored four groups of strategies: integration, regulation, note taking, and help seeking. This study included 1134 students (51% boys and 49% girls) in grades 7–12 (*M* age = 15.07 years old, *SD* = 1.77). Data was collected using the GMRT-RC (K-R 20 = .91-.93; MacGinitie et al., 2001) and the researcher developed self-report Contextualized Reading Strategy Survey (CReSS) which demonstrated adequate reliability and validity (Denton et al., 2015). The two, three, and four way interactions among reader status, grade level, gender, and strategy use were analyzed using mixed model ANOVA. The results indicated statistically significant two-way interactions between reader status (adequate and struggling) and reading strategy use, $F(3, 1088) = 13.68$, $p < .001$. Tukey's post hoc pairwise analysis was also utilized in this study. The results of the follow up analyses indicated that skilled readers reported using integration and regulation strategies more frequently than struggling readers, $t(1088) = 3.39$, $p < .001$, $d = 0.24$, $t(1088) = 3.36$, $p < .001$, $d = 0.24$), respectively. The post hoc results also revealed no significant differences in the use of note taking strategies, $t(1088) = -1.89$, $p < .05$, or help seeking strategies, $t(1088) = -1.85$, $p < .05$, between the two groups of more and less skilled readers.

Evidence in the literature also indicates gender differences in reading strategy use. Awareness of gender differences in reading comprehension strategy use is important information teachers can use to better help students improve their comprehension. Several studies comparing the reading strategy use of male and female students showed that girls use reading comprehension strategies more often than boys (e.g., Bouchamma, Poulin, & Ruel, 2014; Cantrell & Carter, 2009; Griva, Alevriadou, & Semoglou, 2012; Sheorey & Mokhtari, 2001). Although limited, a few existing studies specifically focused on gender differences in high school students' reading comprehension strategy use, obtained the same results (e.g., Denton et al., 2015; Lim et al., 2015; Shera, 2014).

While awareness of differences in high school students' use of reading strategies in terms of minority status would also be beneficial, no prior research was found. Research focused on the relationship between high school students' reading strategy use and their SES was also extremely limited. One previously discussed study, conducted by

Lim et al. (2015), was the only existing study located that investigated this relationship. Their findings revealed high school students' SES had a direct effect on the use of three different types of reading comprehension strategies, including memorization, elaboration, and control strategies ($\gamma = .12$, $\gamma = .18$, and $\gamma = .21$, respectively).

Cognitive and metacognitive reading comprehension strategies. Cognitive and metacognitive strategies are two of the six types of reading comprehension strategies commonly discussed in the literature. Each of the six types of reading strategies (i.e., memory, compensation, cognitive, metacognitive, affective, and social) can offer valuable support to students' comprehension (Chamot & O'Malley, 1987; Oxford & Burry-Stock, 1995). However, Khezrlou (2012) explained that among the array of strategies students use to understand what they read, cognitive and metacognitive strategies are the most vital to successful comprehension.

A clear distinction between the meaning of the terms cognitive and metacognitive strategies is often difficult to differentiate due to the overlapping nature of the concepts cognition and metacognition. An easy way to discern the concepts is to think about metacognition as knowledge of cognition, or thinking about thinking. More concretely, cognitive reading comprehension strategies are general cognitive processes intentionally performed by students that involve direct interaction with the text with the goal of strengthening their comprehension (Gajria, Jitendra, Sood, & Sacks, 2007). On the other hand, metacognitive reading comprehension strategies are higher order cognitive processes that require awareness and conscious control of text comprehension. This includes anticipating, monitoring, identifying, and repairing breakdowns in comprehension, as well as evaluating the overall success of achieving specified goals for reading (Anastasiou & Griva, 2009). Simply stated, while cognitive strategies are the skills necessary to comprehend a text, metacognitive strategies are the skills required to understand how effectively the text was comprehended.

Information regarding reading comprehension strategies is abundantly available via scholarly peer reviewed journals, websites, teacher education textbooks, reading textbooks, content area textbooks, and so on. However, the overall body of literature is not very well organized. Overall, there is a consensus in the myriad of empirical and non-empirical resources that reading comprehension strategies play an important role in developing and sustaining students' ability to successfully comprehend texts. However, there is still not one universally accepted set of recommended reading comprehension strategies that should be taught to students. Furthermore, the suggested amount and specific strategies important for students to learn are inconsistent across resources. For example, in their quintessential report on reading comprehension instruction, the NRP (2000) stated that students need to learn six strategies including predicting, imagining, summarizing, question generation, and clarifying and repairing that promote successful reading comprehension. Whereas, in an investigation conducted to establish guidelines for a U.S. Department of Education program to improve students' reading proficiency the RAND Reading Study Group recommended teaching 11 strategies (Snow, 2002). Some, but not all, of their recommended strategies overlapped with those suggested by the NRP (2000). They proposed teaching the following 11 strategies: setting a purpose for reading, activating prior knowledge, identifying text structure, inferencing, concept mapping, story mapping, identifying main ideas, summarizing, question generation, comprehension monitoring, and clarifying and repairing (Snow, 2002).

Recommended sets of reading comprehension strategies have even differed in research published by the same principal investigators. In their 2008 study, prominent reading researchers Duke and Pearson recommended the seven strategies of predicting, thinking aloud, identifying text structure, visualizing, summarizing, comprehension monitoring, and question generation. Yet in their 2011 book chapter, Duke et al. suggested 12 strategies including setting a purpose for reading, previewing, predicting, activating prior knowledge, thinking aloud, retelling, visualizing, inferencing, summarizing, comprehension monitoring, question generation, clarifying and repairing. Between the two studies six of the strategies were the same, one of the earlier strategies was removed, and six new strategies were added.

Despite the lack of a clear consensus in the literature to confirm which cognitive and metacognitive strategies most effectively support students' reading comprehension, the following nine strategies are commonly recommended: previewing, identifying text structure, activating prior knowledge, predicting, identifying main ideas, summarizing, comprehension monitoring, question generation, clarifying and repairing. The first six strategies described below are cognitive reading comprehension strategies, and the last three are metacognitive strategies.

Previewing. Previewing the text is one of the most well-known cognitive reading strategies. This strategy, sometimes referred to as skimming and scanning, is a quick way for students to develop a basic idea of what the text they are going to read may be about (Hougen, 2015). Previewing a text involves briefly skimming its key components such as the title, table of contents, introduction, subheadings, key sentences (e.g., the first sentence of each paragraph), bold or italicized words, figures/tables/graphs, pictures, and the conclusion. Previewing aids students' comprehension in many different ways. First, familiarizing themselves with the key components before reading helps students focus better on the important aspects of the text while reading. It also gives them an opportunity to think about and plan which reading strategies they will likely need to use. Lastly, it motivates students' interest in the topic of what they are about to read. As reported in McNamara et al. (2007), the results of empirical research indicated that fourth-12th grade students who have used the previewing strategy demonstrated significant improvements in their reading comprehension achievement.

Identifying text structure. Identifying text structures is an important cognitive reading comprehension strategy because text structures help students understand the way information is organized in a text. Text structures are associated with the two types of genres, narrative and expository. There are different text structures within the two genres. Narrative texts have a single structure, whereas there are six main expository text structures. The text structure of narrative texts is referred to as the story grammar. Identifying elements of the story grammar (i.e., setting, main characters, important conflicts, and the resolution) helps students comprehend the main ideas in fictional texts (Meyer & Wijekumar, 2007). The six main expository text structures are compare/contrast, cause and effect, description, listing, sequence, and problem solution.

Understanding the structure of a text is critical for comprehension. Identifying text structure is important because each type of text structure has distinct characteristics that impact comprehension. The differences among text structure have implications for students' reading comprehension strategy use. For example, identifying that a text is organized using a cause and effect text structure should alert a student to the importance of employing the predicting comprehension strategy while reading to better understand upcoming events/details in the text. Text structures reflect logical connections among ideas in the text (Williams, 2007). Therefore, identifying text structures can help students connect main ideas in the text to organize the content and facilitate their comprehension (Meyer & Wijekumar, 2007). Identifying text structure is also beneficial to students when they read challenging texts. Understanding the structure of a difficult text allows students to logically organize the information which makes comprehending the complex material much easier (Meyer & Wijekumar, 2007).

Activating prior knowledge. The activation of prior knowledge (i.e., background knowledge) is another key cognitive strategy that aids reading comprehension. Activating prior knowledge requires students to connect to their existing knowledge, experiences, and opinions about the world to the information in the text (Keene, 2006). Making these connections is important to students' text comprehension because it increases their ability to remember the information they read. Students' comprehension can be negatively affected if they do not possess, or are unable to activate important prior knowledge (Prado & Plourde, 2011). Cromley and Azevedo (2007) explained that low levels of prior knowledge exacerbate high school students', specifically ninth grade students, comprehension difficulties. To enhance their comprehension and facilitate the activation of prior knowledge, high school students, especially those with reading difficulties, may need to be briefly taught important background knowledge before reading an assigned text. This supplemental instruction should be brief. Typically, only a few minutes of extra instruction to build students' background knowledge is necessary to enhance their comprehension (Hougen, 2015).

Predicting. Making an educated guess as to what the text they are going to read is about (i.e., predicting), is one of the first steps in students' successful reading comprehension. Students can use information obtained from previewing the text, such as the title, table of contents, subheadings, pictures, etc., as clues to help them predict what a text is about (Pressley & Harris, 2008). Predicting before reading helps students' comprehension because it helps students activate their prior knowledge. Predicting upcoming events in the text while reading, aids comprehension because it increases students' engagement with the text by motivating them to check whether their predictions were accurate. Comprehension is also supported using this cognitive strategy when students monitor their understanding to check the accuracy of their predictions and adjust their expectations of the text, when necessary (Woolley, 2011).

Identifying main idea. A main idea, is the central idea in a text that is supported by all the important details presented in the narrative. Wang (2009), stated that students' accurate identification of the main idea of a text is indicative of effective reading comprehension. The cognitive strategy of identifying the main idea aids comprehension because it makes students aware of the key ideas and details they need to pay attention to as they interpret the meaning of the text (Pressley & Harris, 2008). It is also helpful to comprehension because it helps students evaluate their understanding, differentiate between important information and irrelevant details in the text, and remember and recall important information (Wang, 2009). Additionally, identifying the main idea and supporting details helps students summarize the text more succinctly.

Summarizing. The cognitive summarizing strategy is closely related to the cognitive strategy of identifying the main idea. Students' ability to write a summary is also indicative of how well they comprehended the text (Caccamise, Franzke, Eckhoff, Kintsch, & Kintsch, 2007). First, summarizing requires students to identify the main ideas and important supporting details in the text. Then, students must compose a coherent statement integrating this important information (Pečjak, Podlesek, & Pirc, 2011). While summarizing is one of the most difficult reading comprehension strategy for students to learn, it is also one of the most important. The summarizing strategy supports comprehension because it forces students to ignore unimportant details in the text while succinctly communicating the most important ideas in their own words (McNamara et al., 2007). Struggling readers often have difficulty summarizing a text because they have difficulty differentiating between the more and less important information in the text. Consequently, they often end up retelling rather than summarizing the text because they tend to focus on the most interesting ideas, rather than the most important (Clark et al., 2004).

Comprehension monitoring. According to Perfetti, Landi, and Oakhill (2005), there is evidence of a relationship between lower levels of reading comprehension achievement and poor comprehension monitoring for students in all grade levels. As students read, they need to pay attention to and think about their understanding of the text in terms of the purpose and goals of the reading activity. Comprehension monitoring is an important metacognitive reading strategy because it enables students to immediately identify comprehension difficulties and adjust their behavior accordingly to repair their understanding of the text. All metacognitive reading comprehension strategies are related to comprehension monitoring in some way, because inherently, they all require students' to be cognizant of their understanding while reading (Pintrich, 1999). Likewise, all metacognitive strategies involve students' use of self-regulatory processes. In addition to being aware of their understanding while reading, students are required to regulate their behavior when they notice a road block to their comprehension (Shang, 2010).

Question generation. The question generation metacognitive strategy is one of the specific techniques students can use to monitor their comprehension (Pintrich, 1999). Kamil (2008a) stated that question generation is an extremely valuable strategy that has more influence on students' reading comprehension than many other strategies. The purpose of this strategy is to increase students' understanding and retention of the information they read (Vaughn et al., 2011). This strategy requires students to develop questions about the text material that can be used to independently monitor their comprehension. These self-assessment questions can be composed during or after reading. Students can use the questions they generate during reading to test their comprehension throughout the reading process. The review questions students create after reading can be used to verify that they have understood the material and to check if they have retained the important information from the text (McNamara et al., 2007). Students can use the following strategy during or after reading if they notice a comprehension breakdown using the question generation strategy.

Clarifying and repairing. A students' ability to identify when their comprehension fails and make the appropriate adjustments, is key to increasing their reading proficiency (Paris et al., 1983). Perfetti et al. (2005) explained, that students who want to make sense of what they read need to be able to clarify and repair their understanding throughout the reading process. This metacognitive strategy requires a student to take some type of action to repair their understanding when they detect a breakdown in their comprehension. Students can utilize an assortment of reading comprehension strategies to clarify/repair their understanding when something in the text doesn't make sense (NRP, 2000). For example, students do not know the meaning of a word, they can use strategies like rereading, identifying context clues, identifying word parts (e.g., prefixes or suffixes), and/or using a dictionary to clarify their understanding of the word meaning and repair their comprehension of the text (Vaughn et al., 2011).

Conclusion. Reading comprehension strategies are essential for students to successfully learn to read, and to effectively understand what they read. Students with proficient reading comprehension know how to consistently and appropriately use reading comprehension strategies to comprehend text, whereas, students who struggle with reading comprehension typically do not (Prado & Plourde, 2011). Students' use of reading comprehension strategies can be influenced by a variety of factors including, but not limited to, their gender, minority status, and SES. One approach to countering these influences and to increasing high school students' reading comprehension proficiency is to explicitly teach them how to effectively utilize the reading comprehension strategies necessary to best support their understanding of what they read.

Reading Comprehension Strategy Instruction

According to national longitudinal data, a majority of students who have difficulty reading at the end of third grade continue to read poorly in high school (Snow, 2002). Ideally, reading interventions would not be necessary after elementary school, however, approximately 70% of secondary students require some form of reading remediation (Edmonds et al., 2009). Additionally, students who can read on grade level in the early elementary grades do not necessarily maintain this proficiency as they progress through school (Biancarosa & Snow, 2006). One reason, according to Sturtevant (2003), is that many students do not smoothly transition from the simplistic stories they learn to read in the elementary grades to the more complex content area texts they encounter in secondary school. Thus, reading comprehension instruction beyond the elementary grades is critical because successful acquisition of early literacy skills does not always equip students with the necessary skills needed to maintain reading proficiency in the upper grades (Swanson et al., 2016).

Reading deficits that persist into middle and high school are extremely difficult to correct. Teaching older students with persistent reading difficulties to read is uniquely challenging for a multitude of reasons. First, many believe that these students are past the age when reading skills are most efficiently acquired (Manset-Williamson & Nelson, 2005). Based on brain development research there is a limited window of time for the development of certain cognitive functions such as reading (Denton, 1999). This research aligns with academic reading research which suggests that it becomes exceedingly more difficult to learn to read after approximately age 10 (Denton, 1999).

Second, the longer reading difficulties persist the harder they are to remediate because the older and further behind a student is, the more learning they must make up (Denton, 1999). Older students reading deficits are sometimes so extreme that they not only struggle with comprehension, they also lack fundamental skills such as automatic word identification and fluency (Manset-Williamson & Nelson, 2005). Unfortunately, some of these students were not taught, or were insufficiently taught, the basic skills required to read fluently and with comprehension. Whereas, some older struggling readers may have received effective early reading instruction, yet they still have difficulty reading fluently and/or comprehending what they read. These students present additional instructional challenges because they likely have difficulty with more than one aspect of reading (Roberts, Torgesen, Boardman, & Scammacca, 2008). Understandably, persistent reading failure can frustrate students. Frustration with reading can create negative attitudes toward reading and cause students to begin to avoid engaging in reading activities (Hurst et al., 2010). Thus, helping students overcome years of frustration presents added challenges for teachers (Manset-Williamson & Nelson, 2005).

Although improving the reading comprehension of older struggling students is challenging, it is not impossible. Several intervention research reviews have reported positive outcomes for high school students with reading difficulties that were taught to use reading comprehension strategies (e.g., Scammacca et al., 2007; Swanson, 1999; Edmonds et al., 2009; Wanzek et al., 2013).

Effective reading comprehension strategy instruction. Reading comprehension is not an automatic skill students simply pick up through the act of reading, it must be taught (Moats, 1999). Reading comprehension strategy instruction is one of the core components of reading comprehension instruction; students must be taught the reading strategies they need to construct meaning from what they read (Prado & Plourde, 2011). In fact, strategy instruction is one of the five practices recommended by the NRP to improve reading proficiency (NRP, 2000). Reading comprehension strategy instruction is defined as the explicit teaching of reading comprehension strategies including what comprehension strategies are, why they are important, and how, when, and why to apply specific strategies (Ness, 2009).

The philosophy behind the explicit teaching of reading comprehension strategies is that comprehension can be enhanced by learning specific strategies to use when comprehension difficulties are encountered (NRP, 2000). The goal of reading comprehension strategy instruction is to teach students how to think while they read by encouraging them to actively participate in the comprehension process (Gersten et al., 2001). This is accomplished through the explicit instruction of reading comprehension strategies (Kamil et al., 2008b). Explicit reading strategy instruction involves direct explanation of a strategy, modeling its' use, scaffolding student learning by providing support during guided practice, increasing student responsibility through independent practice, giving students opportunities to discuss the text and use of the strategy, and monitoring student progress by checking for understanding and providing feedback (Houtveen & van de Grift, 2006; Prado & Plourde, 2011).

Explicit reading comprehension strategy instruction has shown to have a positive effect on the reading comprehension achievement of struggling students (Woolley, 2011). In 2011, Prado and Plourde conducted a quasi-experimental study to explore whether the explicit teaching of reading comprehension strategies increased students' reading comprehension. The sample for their study included 57 fourth grade students. Pretest and posttest data was obtained using a 41-question reading test, the Northwest Evaluation Association (NWEA), that assessed five reading subskills: word recognition and vocabulary, literal reading comprehension, inferential reading comprehension, evaluative reading comprehension, and literary response and analysis. Students received explicit instruction of the following reading comprehension strategies: visualization, activating prior knowledge, inferencing, question generation, drawing conclusions, determining the main ideas and themes, synthesizing information, and clarifying and repairing. A matched t test was used to analyze the pretest and posttest data. The results indicated a statistically significant increase in the mean scores of students' posttest (M = 207.14) compared to their pretest scores (M = 201.82), $t(56) = 4.59$, $p < .001$. The findings suggest that explicit reading comprehension instruction positively effects fourth grade students' reading comprehension achievement (Prado & Plourde, 2011).

According to Scammacca et al. (2007), reading comprehension instruction is also associated with improvements in the reading comprehension achievement of high school students. Their assertion was supported by findings of Sari's 2015 study that investigated the relationship between reading comprehension instruction and students' reading comprehension achievement. Data for this study was obtained from the PISA 2009 Turkey dataset, which included 4,996,15-year old participants. Structural equation modeling was used to examine the relationships among high school students' reading achievement, reading strategy instruction, reading attitudes, and study habits. In terms of the relationship between reading strategy instruction and students' reading comprehension achievement, Sari used two latent variables in the SEM model to measure different aspects of reading instruction, 'teacher strategies' and 'motivating students'. Both variables were measured by students' responses to items obtained from questions on the PISA 2009 Student Questionnaire.

Sari (2015) utilized students' frequency ratings of the following nine items to measure the 'teacher strategies' variable: the teacher poses questions that motivate students to participate actively in reading class, the teacher explains before reading assignments what is expected of students, the teacher tells students in advance how their assignments will be assessed, the teacher gives students the chance to ask questions about the assignments, the teacher asks whether every student understands how to complete reading assignments, the teacher checks that students are concentrating while working on reading assignments, the teacher assesses students' reading assignments when completed, the teacher discusses students' work individually with them after they have finished reading assignments, and the teacher tells students how well they did on reading assignments immediately after they have finished reading assignments. Likewise, the 'motivating students' variable was measured by students' frequency ratings of the following five items: the teacher asks questions that challenge students to get a better understanding of a text, the teacher gives students enough time to think about their answers, the teacher encourages students to express their opinions about a text, the teacher helps students relate the stories they read to their lives, and the teacher recommends books or authors to students to read. The SEM results revealed a significant correlation between high school students' reading achievement and both aspects of reading instruction, 'teacher strategies' ($\gamma = 0.33$) and 'motivating students' ($\gamma = 0.26$). Thus, indicating a positive relationship between reading comprehension instruction and high school students' reading comprehension achievement.

Multiple strategy instruction. Early reading comprehension strategy instruction in the 1970s and 1980s focused on teaching individual strategies in isolation. However, over time, as teachers and researchers realized that good readers do not rely on individual strategies, rather, they use a repertoire of strategies throughout the reading process the focus transitioned to multiple strategy instruction (Pressley & Afflerbach, 1995). The predominant belief became that while it is important to explicitly teach students each individual strategy, they must also be taught to think of strategies cohesively as a tool box of skills that can be used flexibly and integrated as needed (Palinscar & Schutz, 2011).

An eclectic approach to reading comprehension strategy instruction to teach students a variety of cognitive and metacognitive strategies positively impacts students' reading achievement more so than teaching strategies in isolation (Logan & Johnston, 2010). This may be because the effectiveness of the strategies varies based on the individual reader, the specific reading activity, the content of the text, and the reading environment (McNamara et al., 2007). Consequently, students not only need to learn how to correctly use reading strategies, they must also learn when to use which strategies. Students also need to understand the importance of using comprehension strategies. Even when students are taught when to perform the strategies, they are rarely explicitly taught why they are important (Paris et al., 1983). Students need to be convinced the strategies they are learning are truly useful and necessary to support their comprehension.

Furthermore, while having a robust repertoire of reading comprehension strategies is beneficial, students must also learn how to regulate the strategies they use to aid their comprehension during the reading process (McNamara et al., 2007). **S**tudents need to learn the conditions under which certain strategies are applied. Dewitz, Jones, and Leahy (2010) stated, teachers need to provide clear, direct explanations about what the strategies are (i.e., declarative knowledge), how the strategies work (i.e., procedural knowledge), and why/how they should be utilized (i.e., conditional knowledge).

Multiple strategy instruction is useful for teaching students how to become self-regulated learners in that it helps them learn how to independently determine what strategies to use and when (Cantrell, Almasi, Carter, Rintamaa, & Madden, 2010). Multiple strategy instruction training programs include sets of cognitive and metacognitive reading comprehension strategies taught with a focus on teaching students how to flexibly use different strategies to accomplish their established goals for reading (Cantrell et al., 2010). Multiple strategy instruction programs such as Reciprocal Teaching, Collaborative Strategic Learning, and Peer Assisted Learning Strategies have proven to effectively improve students' reading comprehension across various grades.

Reciprocal teaching, as described by the creators Palinscar and Brown (1984), involves four comprehension strategies, including predicting, summarizing, question generation, and clarifying and repairing. The main features of reciprocal teaching are instruction of the four strategies, extensive guided practice applying the strategies, scaffolding, and cooperative learning (Snow & Biancarosa, 2003; Takala, 2006). Similarly, the Collaborative Strategic Reading (CSR) program is comprised of the following four main strategies: preview, click and clunk, get the gist, and wrap up and review. The four main strategies correspond to eight separate reading comprehension strategies: previewing, predicting, activating prior knowledge, identifying the main idea, summarizing, monitoring comprehension, question generation, and students are taught to use these strategies at different stages of the reading process to increase text engagement and reading comprehension (Vaughn et al., 2011). Like reciprocal teaching, CSR teaches students how to monitor and repair their comprehension. Also, similar to reciprocal teaching, cooperative learning is a vital element of CSR.

Predicting, identifying the main idea, and summarizing are the three reading comprehension strategies included in the Peer Assisted Learning Strategies (PALS) program. Like reciprocal teaching and CSR, PALS initially requires extensive teacher involvement as they explain the individual strategies, model appropriate strategy use, and implement guided practice. PALS also utilizes independent practice and cooperative learning in the form of tutoring sessions where pairs of students follow a specific protocol to practice using the comprehension strategies. Teachers continue to provide direct support to students during the tutoring sessions by helping them to fix any mistakes, guiding their appropriate strategy use, and rewarding them for accurate strategy use (Liang & Dole, 2006). The guidelines for PALS programs implemented with high school students encourage teachers to select expository text covering topics that are applicable in the students' lives such as social relationships, sports, entertainment, and youth employment opportunities (Fuchs et al., 2001).

Countless studies demonstrating the effectiveness of these three multiple strategy instruction techniques have been conducted with general and special education elementary and middle school students (e.g., Boardman et al., 2016; Klingner & Vaughn, 1996; Mathes et al., 1999; Mathes, Howard, Allen, & Fuchs, 1998; McMaster, Kung, Han, & Cao, 2008; Morgan, Young, & Fuchs, 2005; Spörer & Brunstein, 2009; Takala, 2006; Vaughn, Klingner, & Bryant, 2001; Vaughn et al., 2011). However, far fewer studies have examined their effectiveness with high school students.

One example of a study focused on examining the efficacy of multiple strategy instruction with high school students is Alfassi's two-part investigation of reciprocal teaching (2004). The first of the two studies in this investigation compared the effects of reciprocal teaching to the effects of traditional methods of reading strategy instruction on students' reading comprehension achievement. The participants in the study included 49 ninth grade students from two English language arts classes. Each class was randomly assigned to one of the two conditions. The treatment group included 29 students and the control group included 20 students. Data collection measures consisted of eight 10-item researcher developed reading comprehension tests (Alpha = .71-.85) as well as the GMRT-RC (K-R 20 = .91-.93; MacGinitie et al., 2001). The intervention was conducted for 20 consecutive school days. Both treatment and control group class sessions lasted approximately 90 minutes. The 20-minute reciprocal teaching intervention was incorporated into the treatment group's 90-minute class.

Alfassi (2004) explained that the reciprocal teaching the intervention included the following three steps: direct instruction, guided practice, and group sharing. During the direct instruction of the four reciprocal teaching strategies (predicting, summarizing, question generation, and clarifying and repairing) the teacher modeled and practiced each of the strategies for two consecutive days. In the guided practice phase, students gradually increased their independent use of the strategies. Finally, in the group sharing stage of the intervention, the class was divided into small groups where students took turns leading discussions on portions of an assigned text while practicing the four strategies. Students in the control condition were not explicitly taught the reading comprehension strategies and did not practice applying them. Univariate ANCOVA tests were employed to determine any differences between treatment and control groups. The findings indicated that students in the treatment group performed higher on reading comprehension measures than the students in the control group, $F(2, 44) = 4.08$, $p < .05$.

Alfassi (2004) then conducted a second study to test the generalizability of the findings from the first study. The generalizability of the results from the first study was limited because it was conducted with a small sample of students in one subject area. The participants in the second study included 277 10th grade students across science, art, social studies, and math classes. The data measures included four 10-item researcher developed reading comprehension tests (Alpha = .73-.84) that were administered before and after the intervention. The intervention in this study was the same as the intervention in the first study. In addition to the 20 days of intervention, students received an additional three consecutive days of the intervention every other month of the school year. A 2 x 2 repeated measures ANOVA was used to examine performance differences before and after the intervention. The results indicated a significant improvement in student performance from the intervention, $F(1, 276) = 6.73$, $p < .05$ (Alfassi, 2004).

The only study located that examined the effects of the Collaborative Strategic Reading (CSR) program with high school student was a small mixed methods study conducted by high school teacher researchers (Shook, Hazelkorn, & Lozano, 2011). The participants of the study consisted of 26 students (14 males and 12 females) in a high school biology class. The intervention began with 30-minute training sessions implemented for three consecutive days to introduce students to the four CSR strategies and the intervention procedures.

After training was completed, the CSR intervention was implemented for 30 minutes, two times a week, for eight weeks. Before reading, students began by using the ‘preview’ strategy which entailed previewing the text, brainstorming, and discussing (activating) their prior knowledge of the topic. The ‘preview’ strategy also required students to predict what they think will happen in the text and what they will learn from reading. While reading, students monitored their comprehension using the ‘click and clunk’ strategy to identify portions of the text that they do (‘click’) and do not (‘clunk’) understand. Next, the ‘get the gist’ strategy was employed to help students with identifying the main ideas and summarizing. Finally, after reading the text, students used the ‘wrap up and review’ strategy to review what they learned and generate questions that will help them review what they learned from the text.

Data for this investigation was obtained from 20-item teacher developed weekly quizzes. To determine differences in student learning, three quizzes taken prior to the intervention, considered baseline data, were compared to the quizzes from the last five weeks of the intervention. The average baseline quiz score for the class was 75 points and the average score of the final five weeks of the intervention was 94. Although Shook et al. (2011) reported using the Wilk’s Lambda test to obtain the result of statistically significant differences in student scores, this test is not the appropriate for this investigation. Despite the lack of accurate statistical results, the increase of almost 19 points from baseline to intervention data points supports the authors assertion that the CSR program positively impacted students’ performance.

Finally, a study investigating the effectiveness of the PALS reading comprehension strategy instruction program with high school students in remedial and special education classes, conducted by Fuchs et al. (1999), was found in the literature. The participants of this study included 119 (49 female/70 male) ninth grade students from 18 intact classes, with reading levels ranging from second to sixth grade. Students’ comprehension was measured before and after treatment using the Comprehensive Reading Assessment Battery (CRAB; Fuchs, Fuchs, & Hamlett, 1989). The 18 classes were divided into nine treatment and control groups. In the treatment condition, the PALS intervention was implemented five times every two weeks over a 16-week period.

Each PALS session consisted of three partner activities: partner reading, paragraph shrinking, and prediction relay. For the first activity, partner reading, each student read aloud for 5 minutes, for a total of 10 minutes of sustained oral reading. If the reader incorrectly pronounced a word, the partner stopped them and prompted them to correct the error. After both students finished reading, the lower performing student had two minutes to retell the sequence of what occurred in the text. In the second activity, paragraph shrinking, the students continued taking turns reading aloud with their partner. At the end of each paragraph the reader stopped to identify the main idea. After five minutes, the students switched roles. The last activity, prediction relay, extended paragraph shrinking to larger portions of text and required students to practice predicting and summarizing. This activity entailed the following four steps: first, the reader predicted what they thought they would learn on the next half page of text; then, they read aloud the half page while their partner identified and corrected any reading mistakes; next, the reader confirmed/disconfirmed their prediction; and finally, they briefly summarized the main ideas of the half page. If their prediction was unreasonable, the partner prompted the reader to try again. Finally, after 5 minutes, the students switched roles and repeated the steps of the activity (Fuchs et al., 2001).

Students in the control group received the typical reading instruction, which did not include any PALS activities. A one-way ANOVA was used to analyze the growth from students' pre- and post-treatment reading comprehension performance. Fuchs et al. (2001) reported that the students who received the PALS training demonstrated statistically significant growth in reading comprehension compared to the students who did not receive the training, $F(1,118) = 5.84$, $p < .05$, ES = .34. Overall, the findings of these four studies demonstrate that multiple strategy instruction programs can effectively impact high school students' reading achievement.

Content area reading comprehension strategy instruction. Reading comprehension strategy instruction positively impacts students' overall reading achievement as well as their performance in the various content areas (e.g., science, social studies, math). Indeed, according to Anastasiou and Griva (2009), teaching students to utilize strategies during reading has been the focus of many studies and it has proven to affect both reading performance and strategy use of poor readers in a positive way. Reading comprehension strategy instruction has also been shown to significantly increase students' retention and comprehension of texts in various high school content area domains such as science and social studies (Duke et al., 2011; Ness, 2009). However, an increased focus on research regarding high school content area reading interventions is greatly needed (Vaughn et al., 2015b).

Berkeley, Marshak, Mastropieri, and Scruggs (2011) conducted one such study with middle school students, in an inclusive classroom, to test the effectiveness of explicitly teaching a reading strategy to improve their comprehension of social studies texts. They conducted a pre-post randomized experimental study to examine the effects of a self-questioning strategy on students' ability to comprehend social studies texts. The sample of students participating in this study included 57 (29 female/28 male) seventh grade students. In order to assign students to the two conditions, students were first stratified by their class, disability status, and English Language Learner status. Then, students were randomly assigned to the treatment and control groups. Finally, teachers were randomly assigned to teach the conditions in two separate classrooms. The self-questioning reading comprehension strategy intervention consisted of three 20 minute lessons taught over three days. In the treatment group, explicit instruction, as detailed above, was used to teach students how to create and answer appropriate questions for individual sections of text using the headings and subheadings in the textbook. Whereas, students in the control group were simply told to read a certain number of sections from the textbook while trying to retain as much information as possible.

Three separate measures were used to collect data for this investigation including a 20-item multiple-choice researcher developed social studies content knowledge test (Alpha = .62), a 13-item open ended researcher developed social studies content knowledge test (Alpha = .82), and a strategy awareness survey. Berkeley et al. (2011) reported that students in the treatment group outperformed students in the control group on multiple-choice social studies content knowledge test (M = 10.30, SD = 3.54) and (M = 7.70, SD = 2.11), respectively. Likewise, the results indicated that students in the treatment group outperformed students in the control group on the open ended social studies content knowledge test (M = 7.03, SD = 3.16) and (M = 2.98, SD = 1.87), respectively. Large effects were associated with the performance difference between the treatment and control groups on the multiple-choice and the open-ended social studies content knowledge tests (ES = 1.61, ES = 0.92, respectively).

Further, Berkeley et al. (2011) employed an independent *t*-test to analyze the posttest scores on the multiple-choice test to examine whether differences in specific content knowledge between the groups existed following instruction. Based on the results of this analysis, the specific content knowledge of the treatment group following instruction was significantly greater than the control group, $t(55) = 3.40$, $p = .001$. An additional independent *t*-test was used to analyze the posttest scores of the open-ended test to determine if there were groups differences in main idea knowledge following instruction. The results indicated that following instruction, the main idea knowledge of the treatment group was greater than the control group, $t(55) = 5.96$, $p < .001$. Finally, the results of the strategy awareness survey revealed that an average of 63% of all participants reported using one or more reading comprehension strategies on the strategy awareness survey, with more students in the treatment group (83%) reporting strategy use than students in the control group (40%). Overall, the results of this study support the notion that reading comprehension strategy instruction positively effects students' reading comprehension as well as their reading comprehension strategy use.

Lim et al., obtained similar results in their 2015 study of reading comprehension strategies. The purpose of this study was to examine the relationships among high school students' reading strategy instruction, reading strategy use, reading attitudes, gender, and SES. The dataset from the Korean PISA 2009 including data of 4,988 students from 157 schools was used in this investigation. Structural equation modeling (SEM) was employed to test the relationship between reading strategy instruction and reading comprehension strategy use. The SEM results indicated that reading strategy instruction directly and positively predicted students' use of memorization ($\gamma = .07$), elaboration ($\gamma = .12$), and control ($\gamma = .12$) reading comprehension strategies.

In order to improve students' content area knowledge and overall reading comprehension ability, high school content area teachers need to be held responsible for teaching them the strategies necessary to effectively comprehend their content area texts (Scammacca et al., 2016). In successful high school classrooms, teachers teach content area knowledge in tandem with requisite literacy skills. Integrating reading comprehension strategy instruction into high school content area instruction is necessary because meeting the increased knowledge, strategy, and discourse requirements of high school content areas requires students' awareness and use of more sophisticated comprehension strategies (Conley, 2008). Specifically, it is critical that high school teachers teach students' content area specific reading comprehension strategies because certain reading strategies are considered unique to each subject area, due to the different reasoning processes and presentation of material in each discipline (Wexler et al., 2016).

A study conducted by Shanahan and Shanahan (2008) offers specific examples of these domain specific reading comprehension strategies. In 2008, Shanahan and Shanahan began a multi-year study to reevaluate the basic curriculum of adolescent literacy instruction, specifically in terms of disciplinary reading comprehension strategy instruction. Their study was of many conducted as part of the Carnegie Corporation's effort to identify effective adolescent literacy instructional practices and subsequently develop higher education course curricula to help future teachers learn how to integrate literacy instruction into content area instruction.

In the first year of their investigation Shanahan and Shanahan (2008) consulted with a host of experts in math, chemistry, and history to identify the sophisticated specialized reading skills pertinent to their specific areas of expertise. The math experts identified close reading and rereading as two of the most important reading comprehension strategies for their domain. They explained that comprehending math text requires an understanding the precise meaning of each word that cannot be accomplished without close reading. Reading comprehension strategies that aided in the translation of information from one form to another were of greatest importance to the chemistry experts. They emphasized the importance of strategies such as visualizing, writing down formulas, and obtaining information from pictures, diagrams, charts or graphs in developing a complete understanding of the concepts. Finally, the historians explained the awareness that their texts are mainly interpretations of historical events makes it necessary to determine what story specific authors want to tell. As a result, they stressed the importance of reading comprehension strategies that focus on the author and source of the text in order to be able to evaluate any potential biases in the text.

Students of all ages and performance levels need the ongoing support of competent teachers to continue developing their reading comprehension abilities. The highest achieving high school students may need assistance with advanced vocabulary, while struggling readers may need extensive support to improve their reading comprehension proficiency and their comprehension of critical content area information (Sturtevant, 2003). The importance of continuing reading comprehension instruction beyond the elementary years to better support the needs of older students is clear (Shanahan & Shanahan, 2008). Specifically, strategy instruction needs to be integrated with content area instruction to help students improve their reading proficiency and to access the knowledge necessary for content area learning (Wexler et al., 2016).

Lack of reading comprehension strategy instruction. Despite agreement in the literature supporting the positive impact students' effective use of reading comprehension strategies has on their reading achievement, as well as the research indicating that reading comprehension strategy instruction supports the development of students' reading comprehension, surprisingly few teachers engage in reading comprehension strategy instruction; especially in high school content area classes. Empirical research indicates that in most classrooms, students receive inadequate instruction regarding reading comprehension strategies and skills (Shang, 2010). There is at least 40 years of evidence highlighting teachers' resistance to reading comprehension instruction in both reading and content area classes. Historically, the results of observational research conducted in elementary through high schools indicate that most content area teachers provide little to no strategy instruction (Durkin, 1978; Ness, 2009; Pressley, Wharton-MacDonald, Mistretta-Hampton, & Echevarria, 1998; Quirk, Trismen, Nalin, & Weinberg, 1975; Swanson et al., 2016; Wexler et al., 2016). The following three studies provide a detailed explanation of the reading comprehension instruction practices of high school teachers.

Ness (2009) conducted a mixed methods study to measure the frequency of reading comprehension strategy instruction in secondary social studies and science classrooms. Data was collected through classroom observations and interviews of eight teachers (four middle school and four high school). Of the 2,400 total minutes of classroom observations a total of 82 minutes (3%) of reading comprehension strategy instruction was recorded. The results revealed that reading comprehension strategy instruction was only implemented for 60 minutes (10%) of the total time observed in the middle school social studies classes, and no explicit reading strategy instruction was observed in the high school social studies classrooms. While Ness stated that the instruction observed in the middle and high school social studies classrooms far exceeded that observed in the science classrooms, no empirical results for the science classrooms were provided or possible to deduce from the information presented.

In a similar study, Swanson et al. (2016) investigated the reading comprehension instruction in middle and high school English language arts (ELA) and social studies classes. The researchers observed 20 middle and high school classes (nine ELA and 11 social studies). They conducted 137 classroom observations (58 ELA and 79 social studies) totaling 7,208 minutes (3,283 minutes and 3,925 minutes, respectively). Based on the purposeful sampling criteria established by the researchers, the teachers selected for this study all had a minimum of three years of experience and were considered content area experts by their administrators. The results indicated that reading comprehension strategy instruction was observed in 26% of ELA classes and in 20% of social studies classes. Further, the results revealed that four of the nine ELA teachers were never observed teaching or reviewing reading comprehension strategies and two of the 11 social studies teachers not observed teaching or referring to comprehension strategies during their instruction (Swanson et al., 2016).

In 2016, Wexler et al. investigated the frequency of reading comprehension strategy instruction in high school biology classes. The participants in this study included

10 teachers who taught a total of 198 ninth through 12th grade students, 34% of which were at risk of reading failure. Data was collected using a researcher developed classroom observation protocol adapted from the Writing and Reading Observation Tool (Bryant et al., 2013) designed to observe secondary literacy settings. A total of 3,167 minutes during 40 class sessions (lasting an average of 79 minutes) across the 10 teachers were observed over a three month period. Once all 40 observations were completed, each of the 10 teachers participated in a 45-minute semi-structured interview.

After coding the observations, the research team concluded that the biology teachers integrated essentially no reading comprehension strategy instruction into their content area instruction. In the total of less than 1% of the 1 minute time intervals observed collectively across all 10 teachers, the researchers reported minimal instruction of the following strategies: previewing (0.09%), identifying main idea and summarizing (0.60%), and question generation (0.76%). Qualitative analysis of interviews revealed that teachers were aware that students require reading comprehension strategy instruction to become independent learners and they all agreed with the importance of integrating strategy instruction into their lessons. Nevertheless, eight of the 10 teachers defended their lack of reading comprehension strategy instruction. Four teachers cited limited instructional time as a barrier to appropriately teaching reading comprehension strategies. The consensus among those four teachers was that they only had enough class time to teach basic short-term strategies to provide students a quick fix to comprehend the immediate text. According to Wexler et al. (2016), the other four teachers didn’t explicitly teach reading comprehension strategies during class because they believed that students independently practice reading strategies at home. Although the sample size of the teachers observed in these three studies is relatively small, the agreement of these studies, along with the consistent results of similar studies conducted in elementary and middle classrooms dating back to the 1970s, indicates that there is clearly a pervasive and unsettling lack of reading comprehension strategy instruction in U.S. classrooms.

School level predictors of reading comprehension achievement. There are several student and school level factors associated with classroom instruction that have been found to effect student learning. In addition to the student characteristics previously discussed (gender, minority status, and SES) the school characteristics of class time and class size were included in this investigation. Research evidence exists supporting the positive effects of additional class time and reduced class size on overall academic achievement. However, research in these areas specifically focused on the effects class time and class size on the reading achievement of high school students is limited.

Reading comprehension achievement and class time. Since the early 1980s some educators and policy makers have been pushing for an increase in instructional class time (Kolbe, Partridge, & O'Reilly, 2012) based on the assertion that more instructional time will lead to increased student achievement. In recent years, increasing classroom instruction time has become a common topic of school reform debate. In particular, many argue in favor of increasing classroom instructional time to allow teachers more time to focus on improving students' reading and math performance.

A study conducted by Rivkin and Schiman (2015) lends credence to the argument for increasing class time. Their investigation utilized reading, math, and science achievement data from the PISA 2009. The sample of participants included in the dataset consisted of 253,286 15-year old students across 72 countries. Results of the fixed effects regression analysis used in this study, revealed a positive and statistically significant relationship between class time and student achievement. Specifically, one additional hour of class time per week increased students' achievement by 11% of a standard deviation across all three subject areas. Thus, indicating that increased instructional time can improve high school student achievement. In addition to the benefit of improved student achievement, increased instructional time offers teachers innumerable benefits. For example, additional class time would give teachers the chance to cover more course material, invest more time teaching challenging concepts, effectively individualize and differentiate instruction, ensure that they sufficiently answer student questions, consistently provide individualized feedback to students regarding their performance, implement more project based learning activities, etc.

Despite the growing calls for an increase in instructional class time, the average numbers of days in the traditional public school year, 179, has not changed in at least 15 years. Over that same time period, the average number of hours in the traditional public school day has barely changed. On average, it has increased 4 minutes to the current average 6 hour and 42-minute school day (Kolbe et al., 2012). Although the instructional time in traditional public schools is relatively stagnant, instructional time in charter schools has increased. Since the 1999-2000 school year, charter schools have added one school day, increasing the average number of days in school year from 179 to 180.

Likewise, charter schools have increased the average number of hours per day by approximately 7 minutes (Kolbe et al., 2012).

There is empirical evidence indicating that the increased instructional class time in charter schools positively effects student achievement. Hoxby, Murarka, and Kang (2009) conducted a multi-year project to evaluate New York City Charter schools. They examined longitudinal data of students in grades three through 12 from the 2000-01 school year through the 2007-08 school year. As part of their evaluation, they investigated the effects of increased reading instruction time on student achievement using simple linear regression analysis. According to the results of the analysis, increased reading instruction time is associated with a statistically significant improvement in student achievement (Hoxby et al., 2009).

In addition to positively impacting students' overall achievement, converging evidence from different investigations of PISA reading test data have revealed a positive relationship between increased reading class time and students' reading achievement. For example, Lavy (2015) examined the effects of instructional time on students' reading, math, and science achievement using PISA 2006 data. Although the PISA 2006 international dataset included data from 58 countries, the researcher only used data from the 22 OECD developed countries when estimating the effects of instruction time on reading, math, and science achievement. The results of the fixed effects regression analyses indicated a positive and significant effect of instructional time across all three subject areas. Consistent with, but smaller than the results obtained by Rivkin and Schiman (2015), Lavy (2015) reported, one more hour of instruction per week increased achievement by 6% of a standard deviation across all three subjects. Additionally, the results of the ordinary least squares regression analysis suggested a positive relationship between an increase in reading instruction time and students' reading achievement. Specifically, an increase of one hour of reading instruction per week improved students' reading achievement by 5% of a standard deviation (Lavy, 2015).

Similarly, Huebener et al. (2016) sought to examine the relationship between high school reading instruction class time and students' reading comprehension achievement. Their approach to investigating this relationship differed from the prior studies in two main ways. First, the researchers utilized longitudinal data in this study rather than cross sectional data. Second, instead of only using international PISA data to explore this relationship, the researchers used a combination of international PISA data, and data from a PISA dataset of one individual OECD country, Germany. The researchers examined a total of five sets of PISA data ranging from 2000-2012. Data from 2000, 2003, and 2006 were obtained from German PISA datasets and data from 2009 and 2012 were obtained from international PISA datasets. The sample consisted of 33,217 15-year old students. Ordinary least squares regression analysis was used to analyze the effects of increased reading class time on students' reading achievement. The results revealed a significant improvement in students' reading achievement by an average of 6% of an international standard deviation when reading class time was increased by two hours per week (Huebener et al., 2016). This finding, indicating that increased reading instruction time improves students' reading achievement, is consistent with the findings of Lavy (2015).

Three additional studies investigated the relationship between reading class time and reading achievement (e.g., Cattaneo et al., 2016; Huang, 2015; Kasapoglu, 2014). They each used cross-sectional data from PISA datasets of one individual OECD country (Turkey PISA 2009, United States PISA 2012, and Switzerland PISA 2009, respectively). The findings of all three studies, including a total of 21,407 students, confirm the positive association between class time and reading achievement. Thus, providing additional support for the claim that an increase in reading instruction time can effectively improve students' reading comprehension achievement.

Reading comprehension achievement and class size. In the last two decades, class size reduction has become a popular, albeit controversial, school reform initiative. An array of researchers, teachers' unions, policymakers, and politicians have debated the benefits and costs of reducing class size (Milesi & Gamoran, 2006). At the core of education reform and policy making are major decisions regarding the most effective distribution of school resources to optimize student achievement (Konstantopoulos & Traynor, 2014). This includes decisions pertaining to the assignment of teachers and students to classrooms, which requires a determination of the ideal number of students per classroom (i.e., class size). In addition to its critical role in school reform and policy initiatives, there is evidence that class size effects student achievement.

An investigation of the effects of class size on academic achievement was one component of the earliest empirical study of various educational processes and their effects on student achievement conducted by Rice in 1902. The findings related to class size suggested there was not a strong relationship with student achievement. However, as

Glass and Smith (1979) pointed out, the researcher unfortunately reported practically no quantitative results, making it impossible now to determine if the relationship was genuinely small in that it was based on the subjective judgment of the researcher.

In 1979, Glass and Smith conducted a meta-analytic review of 77 early class size research including studies ranging from 1900 to 1979. The sample of studies included 900,000 students from 5 to 19 years or older (M = 12, SD = 4). The researchers employed ordinary least squares regression analysis to examine the effect of small class sizes on students' reading achievement. The results revealed a strong relationship between class size and achievement in favor of small class sizes by approximately 0.10 standard deviations in reading achievement. Further, Glass and Smith (1979) reported that the relationship between class size and reading achievement was consistently stronger for high school students than elementary students.

In a more recent study, Krassel and Heinesen (2014), examined the influence of class size on the achievement of high school students. Their longitudinal study included a sizeable sample of 29, 184 10th grade Danish students. The researchers obtained four years (2003-2006) of class size and student achievement data from Statistics Denmark and the Denmark Ministry of Education, respectively. Achievement data including students' average ninth grade GPA and their year-end exit exam scores from ninth grade was compared to the class size data from 10th grade. Achievement measure from ninth grade had to be used because in Denmark, that is the final year of compulsory education as well as the last year students a required to all take the same end of year-exit exam. In 10th grade, students have the option of either retaking the ninth grade exit exam or taking the advanced 10th grade exit exam. Regression analysis was used to test the effect of class size on high school students' academic achievement. The results of Krassel and Heinesen's analysis revealed a statistically significant positive effect of small class size on students' academic achievement. When the class size is reduced by 10 students, the estimated effect on achievement is approximately 0.08 standard deviations.

Leuven et al. (2008) conducted a similar longitudinal investigation to evaluate the effects class size on high school students' achievement. Participants in their study included 111,463 ninth grade Norwegian students. Three years (2001-2003) of class size and academic achievement data of these was obtained from the Statistics Norway and Norway's Ministry of Education, respectively. Regression analysis was used to estimate how students' academic achievement is affected by class size. The researchers reported that class size had virtually no effect on student achievement. Based on the results, when class size is reduced by 10 students, achievement is improved by an estimated 1-2% standard deviation (Leuven et al., 2008). The small nonsignificant effect of class size on students' academic achievement found in this study is inconsistent with a lot of existing class size evidence suggesting the need for additional research in this area.

Research investigating the effects of class size on high school students' reading achievement is another area ripe for research. Unfortunately, research investigating the influence of class size on high school students' achievement in reading was not found in the either the class size or reading literature. However, several large scale multi-year studies conducted with elementary students reported a statistically significant positive relationship between smaller class sizes and students' reading achievement. For example, Finn and Achilles (1999) analyzed the effectiveness of the Tennessee Student/Teacher Achievement (ProjectSTAR) which is one of the most prevalent CSR programs in the country. ProjectSTAR consisted of approximately 12,000 students from the 329 classrooms that participated in the four-year intervention. Molnar et al. (1999) investigated a similar but slightly smaller program in Wisconsin entitled, Student Achievement Guarantee in Education (SAGE). Their study evaluated data of 6,308 students from 284 classrooms that participated in the first two of the five-year program. Finally, one of the two main purposes of Chatterji's (2006) study, was to identify student and school level variables that correlated with or moderated kindergarten and first grade students' reading achievement. Data for this investigation was obtained from a subset of the Early Childhood Longitudinal Study (ECLS) that included 2,296 students from 184 U.S. schools. The correlation between class size and reading achievement was one of the relationships examined in this study. Consistent with the findings of the other two large scale multi-year studies, Chatterji (2006) found that class size had a statistically significant and negative correlation with elementary aged students' reading achievement.

Despite the limited and slightly inconsistent class size evidence, a majority of the research literature indicated a significantly negative relationship between the effects of class size and high school students' overall achievement. In addition to the positive effect of smaller class sizes on student achievement, small class sizes offer a myriad of other benefits to students and teachers, similar to those of increased class time, that promote learning. Awareness of these benefits is also important to researchers' understanding of the teaching and learning process because, as explained by Harfitt and

Tsui (2015), in essence, they are contextual and affective variables that provide insight into how class size mediates both teaching and student learning.

In 1994, Blatchford and Mortimore were among the first researchers to investigate the effects of small class size on specific classroom processes. They identified increases in teacher quality, curriculum coverage, and effective classroom management in small classes. They also found that in small classes teachers are better able to more adequately meet the specific needs of each student by individualizing their instruction. Further, additional class time and classroom space, better classroom morale, and improved relationships among the students were also observed in small classes (Blatchford & Mortimore, 1994). Another observational study to investigate the relationship between class size and classroom processes, conducted by Blatchford in 2003, focused specifically on examining student-teacher interactions, student attentiveness and off-task behavior, and student peer interactions. The results of Blatchford's 2003 study revealed that in small classes teachers interacted with students more frequently and students demonstrated increased attention, decreased off-task behaviors, and interacted with peers more often.

In 2011, Blatchford, Bassett, and Brown extended Blatchford's earlier class size research to include high school students. In this study, the effects of small class sizes on pupil classroom engagement and teacher pupil interaction were systematically observed and analyzed to determine whether any observed effects varied across elementary and high school students' achievement or grade levels. The participants in this study included 686 students from a total of 88 1st, 3rd, 7th, and 10th grade classes. The results of this study indicated that both elementary and high school students were more engaged in smaller, rather than larger, classes. Further, Blatchford et al. (2011) explained that the positive effect of small class size on engagement was most notable with low achieving high school students, indicating that struggling high school students would benefit from small class instruction. Small class sizes also led to an increase in individual student attention of both elementary and high school teachers.

Similarly, Harfitt and Tsui (2015) conducted a multiple case study to examine the effects of class size on the teaching and learning processes in high school classes. Participants in the study included four high school teachers (two eighth grade, one ninth grade, and one 10th grade). Each teacher instructed both large and small classes, in the same grade level, with students of comparable ability. Classroom observations and semi-structed student and teacher interviews were used to collect data for this investigation. The researchers conducted 60 observations of 35-40 minute lessons as well as 229 student and 29 teacher semi-structured interviews.

The results obtained from the large amount of qualitative data produced an extensive number of valuable findings. Only the main results of Harfitt and Tsui's 2015 study are reported here. Their overall findings indicated that students in small classes were more motivated and engaged in their learning. Students and teachers in the small classes demonstrated a stronger sense of belonging to the classroom community, developed more meaningful relationships with each other, offered more support to their peers, acknowledged each other's strengths, and had a more positive perspective of the classroom environment. All of which contributed to the increased motivation and engagement demonstrate by the students in the small classes.

Another important benefit of smaller class sizes is the positive impact they have on minority students and students with a low SES. According to Fan (2012), a multitude of researchers who have studied the effects of class size and identified a significantly negative impact of large classes on academic achievement, consistently reported large classes had the most negative impact on minority students and/or students with a low SES. Consequently, class size reduction has been recommended by researchers to lessen the minority and SES achievement gaps (Graue et al., 2007). While relatively little is known about the effects of class size on high school students' reading achievement, the evidence presented here regarding the impact it has on high school students' overall achievement and elementary students' reading achievement, and the extensive benefits of small class sizes to students and teachers, supports the need for further research.

Conclusion. High school students need to be explicitly taught how to improve their reading comprehension by effectively utilizing various cognitive, metacognitive, and content area specific reading comprehension strategies. Specifically, students need to learn how to apply reading comprehension strategies through the teachers' description of a particular strategy, explanation of the thinking process that underlies the use of the strategy, modeling of when and how to use the strategy, and guided practice implementing the use of the strategy. Finally, teachers need to gradually turn over the responsibility of strategy use to the students during independent practice to ensure that students' have developed the necessary knowledge of the reading comprehension strategy and its use. Despite empirical evidence supporting the positive effects of reading comprehension strategies on high school students' reading comprehension achievement, these strategies are rarely taught at the high school level.

Reading Attitudes

Students' positive attitudes toward reading has been consistently found to be associated with higher reading achievement. Students' attitudes toward reading influences their success in reading because reading attitudes are likely to affect students' engagement, persistence, frequency, and enjoyment of reading (Logan & Johnston, 2009; Woolley, 2011). Unfortunately, many teachers are unaware of the influential role students' reading attitudes can have in developing and/or increasing their reading proficiency, and as a result little class time is devoted to activities designed to develop or reinforce students' positive attitudes toward reading (Petscher, 2010).

A host of definitions of reading attitudes have been proposed, with Alexander and Filler (1976) providing the earliest definition of attitudes specific to reading. They defined reading attitudes as the spectrum of positive to negative feelings about reading that motivate a reader to pursue or avoid reading opportunities and reading related activities (Alexander & Filler, 1976). It is important to differentiate here between the concepts of attitude and motivation because the terms are often incorrectly used interchangeably in the literature. There are two common ways that these terms are used incorrectly. First, attitude is conflated with the concept of motivation when motivation is treated as a uni vs multidimensional construct (Petscher, 2010). Second, attitude is used as a synonym for the term motivation (Conradi, Jang, & McKenna, 2014). Although both concepts are affective factors, they are mutually exclusive (Flippo, 2014). While motivation is influential in why we choose to do or not do something, attitudes moderate one's level of motivation (Petscher, 2010). For example, we are more likely to be motivated to engage in a task if it produces positive outcomes that reinforce positive attitudes; likewise, we are less motivated to engage in an unrewarding task that produces negative outcomes and elicits negative attitudes (Bandura, 1999; Ryan & Deci, 2000).

This understanding of attitudes toward reading is based on the McKenna model of reading attitudes (McKenna et al., 1995). This model extended the existing Mathewson and Ruddell-Speaker models of reading attitudes by including a focus on the long-term development of reading attitudes (Mathewson, 1994; Ruddell & Speaker, 1985). In their model, McKenna et al. (1995) posited that, over time, attitudes towards reading develop as a result of the following three interrelated factors: normative beliefs (i.e., subjective norms based on the expectations of others such as parents, teachers, peers, including the motivation to adhere to such expectations), beliefs about the outcomes of reading in terms of the desirability of the outcomes, and beliefs about the outcomes of reading activities when judged against competing activities or the physical/time investment they require.

In addition to the influence these complex dynamic factors have on students' reading attitudes, McKenna et al. (1995) explained that these factors also influence each other. Similar to the Construction Integration model of reading comprehension adopted for this investigation, the McKenna model is grounded in the social cognitive theory. Human functioning, as explained by social cognitive theory, has a multifaceted causal structure. It occurs within a framework of reciprocal influences among environmental, personal, and behavioral factors which Bandura (1986) referred to as reciprocal determinism. In this triadic model of reciprocal causality, each of the factors influence and are influenced by each other. The McKenna model of reading attitude is an example of a model of triadic reciprocal determinism. This multidimensional model of reading attitudes highlights the relationships among social environmental factors, reading attitudes (personal factor), and reading behavior, as well as the relative influences of each factor. For example, a students' social environment can influence their reading attitudes. If a student's environment encourages, models, and reinforces reading, they are more likely to have a positive attitude toward reading (Smith, 1990). Likewise, reading attitudes can influence reading behavior in that students who have a positive attitude toward reading are more likely to choose to engage in reading activities even when competing and desirable activity options are available.

Reading attitudes and reading comprehension achievement. Several models of reading development focus on the cognitive factors that influence reading. However, since reading requires an intentional action of the reader to initiate and sustain effort to complete, cognitive capacity alone doesn't ensure reading success. Understanding this led researchers to explore the influence of affective characteristics on students' reading achievement. Growing awareness of the effect of students' affective characteristics led to the creation of reading models that sought to operationalize the development and role of affective factors (including reading attitudes) involved in the reading process.

Emerging evidence that both cognitive and affective factors are integral elements of reading promulgated a wide-ranging body of research concerning the affective dimension of reading development and ability, which included reading attitude research

(McGeown et al., 2015). While there is now a substantial body of literature demonstrating a consistent relationship between students' reading attitudes and reading achievement, like most of the previously discussed topics, there is a dearth of research in this area investigating high school students. A variety of studies and meta-analyses that have investigated this relationship in elementary and middle school populations have identified a positive relationship between higher reading attitudes and students' achievement in reading (e.g., Bastug, 2014; Kush, Watkins, & Brookhart, 2005; Logan & Johnston, 2009; Martínez, Aricak, & Jewell, 2008; McGeown et al., 2015; McKenna et al., 1995; McKenna et al., 2012; Mullis, Martin, Gonzalez, & Kennedy, 2003; Mullis, Martin, Kennedy, & Foy, 2007; Petscher 2010; Sallabaş, 2008).

To date, the largest investigation of students' reading attitudes in the US was conducted in 1995 by McKenna et al. The purpose of this study was to examine the developmental trends in elementary students' reading attitudes. A national sample of 18,185 first through sixth grade students participated in this study. The Elementary Reading Attitude Survey (Reliability ranged from .74-.89; ERAS; McKenna & Kear, 1990), a 20-item pictorial rating scale, was used to measure students' reading attitudes. Reading attitude scores were analyzed in terms of gender, grade, ethnicity, reading ability, and the use of basal readers. Due to the large number of schools included in this study there was no universal measure of student achievement that could be used for all students. Instead, researchers asked teachers to assign students an above average, average, or below average rating based on the performance of students over the previous five months of school. The results of analysis of variance (ANOVA) testing with respect to relationship between reading attitudes and academic reading ability indicated a significant main effect for academic scores $F(5,18,155) = 29.8$, $p < .001$ (McKenna et al., 1995). Additional findings of this study will be discussed below.

The positive relationship between reading attitudes and students' reading achievement was also supported by the findings of Petscher's (2010) meta-analysis. This study was conducted to provide a clearer understand of this relationship because the magnitude of effect sizes in existing studies was wide ranging. The sample of studies in this meta-analysis included 32 articles with a total sample size of 224,615 students and 118 effect sizes. Hierarchical linear modeling was used to estimate the mean effect size, variability across studies, and moderator effects. The overall results indicated that that the reading attitudes of elementary and middle school students positively correlated with reading achievement. The researcher found a moderate relationship between students' reading attitudes and reading achievement ($Z_r = .32$), confirming findings from previous research in which the strength of the relationship ranged from .20 to .40. Petscher (2010) also reported that the relationship between students' reading attitudes and reading achievement was stronger for elementary students ($Z_r = .44$) than middle school students ($Z_r = .24$). This finding supports McKenna et al.'s (1995) assertion that the strength of the relationship between reading attitudes and reading achievement declines over time.

The relationship between reading attitudes and reading achievement of high school students was examined as part of Kasapoglu's 2014 study to investigate the factors associated with Turkish students who scored above average ($M = 493$) on the PISA 2009 reading assessment. The PISA 2009 Turkish dataset included data of 4,996

15-year old students. The researchers utilized logistic regression analysis to model the data of above average and below average students. In addition to reading attitudes, other variables in the model used to predict the probability of students' scoring above average included class time, gender, school entry age, the mother's and father's level of education, the number of books present in the home, recreational reading time, and extracurricular time learning the test language. The results indicated that students' reading attitudes significantly predicted their probability of attaining higher reading achievement ($\beta = .57$, *OR* = 1.76, Wald statistic = 42.54, $p < .01$). The odds ratio greater than 1 indicated that students' attitudes toward reading were positively correlated with the probability of attaining higher reading achievement. Which means, students with positive reading attitudes are 1.76 time more likely to demonstrate higher reading achievement than students with negative attitudes toward reading (Kasapoglu, 2014).

Sari (2015) also investigated this relationship using the same Turkey PISA 2009 dataset. As part of this study, discussed above, structural equation modeling was used to examine the relationships between high school students' reading achievement and their attitudes toward reading. The results of the SEM model indicated a significant, positive correlation between high school students' reading achievement and their reading attitudes ($\gamma = 0.16$). Thus, confirming the results obtained by Kasapoglu (2014), indicating that reading attitudes of high school students positively influence their reading achievement.

Reading attitudes and gender. Substantial gender differences in reading attitudes, in favor of girls, have been consistently reported in the literature (Logan & Johnston, 2010). Gender effects are important to consider in the relationship between reading attitudes and reading comprehension achievement because the reading attitude gender gap is wider than the gender gap in reading achievement (Lim et al., 2015). Further, there is evidence that reading attitudes gender gaps are present as early as the primary grades and exist up and through the end of high school (McKenna et al., 2012).

Gender differences in elementary students' attitudes toward reading were reported in McKenna et al.'s 1995 study. The results of their study, consisting of 18,185 first through sixth grade students (50% male, 50% female) revealed statistically significant main effects for recreational reading scores, $F(5,18,155) = 29.8$, $p < .001$, and academic reading scores, $F(5,18,155) = 29.8$, $p < .001$. This indicated that girls possessed more positive attitudes toward reading at all grade levels than boys. The results also suggested the reading attitude gender gap increased over time. These findings were consistent with several other studies of elementary and middle school students' reading attitudes (e.g., Kush & Watkins, 1996; Logan & Johnston, 2009; Martìnez et al., 2008; McGeown et al., 2015; Mullis et al., 2003; Şahbaz, 2012; Swalander & Taube, 2007).

Research of high school students' reading attitudes indicates the trend of elementary and middle school age girls demonstrating more positive reading attitudes than boys grade is consistent among high school age students. For example, Gökhan (2012) conducted a study to examine the relationships among high school students' reading attitudes, gender, SES, grade level, and school type. Random sampling was used to select the 426 (237 female, 189 male) Turkish ninth and 12th grade high school students that participated in this study. The 30-item Attitude Scale Towards Reading (Alpha = .88; Gömleksiz, 2004) was administered to collect data about reading attitudes.

The results of independent sample t-tests revealed significant difference in the reading attitudes of girls and boys, $t(424) = -3.348$, $p < .05$, in favor of girls.

Similar results were obtained in a larger study of high school students' reading attitudes conducted by Bussert-Webb and Zhang (2016), including 2,553 (63% male, 37% female) ninth through 12th grade students. Researchers surveyed students' reading attitudes using the Rhody scale, which consisted of 25 Likert-type items ($r = 0.84$; Tullock-Rhody & Alexander, 1980). Regression analysis revealed significant gender differences in students' reading attitudes. Overall, female high school students reported more positive attitudes toward reading ($M = 3.12$) than male students ($M = 2.95$).

Investigations of large-scale international PISA data also consistently revealed that girls demonstrate more positive attitudes toward reading than boys (e.g., Lim et al., 2015; Organization for Economic Cooperation and Development [OECD], 2001; Organization for Economic Cooperation and Development [OECD], 2004). In 2014, Jhang conducted a study to examine high school students' reading attitudes using the international PISA 2009 dataset. The dataset included data from approximately 470,000 15-year-old students from all 65 participating OECD countries. Three level hierarchical linear modeling (HLM) was used in this investigation to examine the relationship between students' reading attitudes and instruction considering country, school, and individual level mediating variables (including gender). The results of the HLM indicated that gender was significantly correlated with reading attitudes (.30), suggesting that girls had more positive attitudes toward reading than boys (Jhang, 2014). Gender differences in reading attitudes were also examined by Lim et al. (2015) in their study of

4,988 Korean high school students' reading attitudes. They utilized structural equation modeling to analyze data from the PISA 2009 Korean dataset. Their results indicated that girls reported more positive attitudes toward reading than boys (F = -.19; $\gamma = -.16$).

Reading attitudes and socio-economic status. There is also evidence suggesting that sociocultural factors play an important role in the development of students' attitudes toward reading. As mentioned above, in addition to examining the gender differences in reading attitudes of 426 high school students, Gökhan (2012) also investigated the effects of students' SES status on their reading attitudes. In this study, students' SES level was measured by the monthly income level of their family. Analysis of variance (ANOVA) was used to test the relationship between students' reading attitudes and their SES level. The ANOVA results indicated that reading attitudes were significantly and positively related to their SES level, $F(2,423) = 542.777$, $p < .05$. In other words, high school students with greater family income levels demonstrate more positive attitudes toward reading. This evidence aligns with the influence of students' social environment on their reading attitudes hypothesized by McKenna et al. (1995).

Reading attitudes and reading comprehension strategy use. McKenna et al. (1995), believed that reading is an ongoing process requiring the initial decision to begin reading as well as a recurring decision to continue reading throughout the reading process. Furthermore, they asserted that the decision to continue reading is collectively influenced by the goal of the reading activity, the reader's metacognitive feedback about their current reading progress, and their attitudes toward reading. Their description illustrates one perspective of the important relationship between students' reading attitudes and their use of reading comprehension strategies.

Although few studies have investigated the nature of the relationship between students' use of reading comprehension strategies and their attitudes toward reading, many researchers believe that students' positive reading attitudes are associated with an increased use of reading comprehension strategies (Kirmizi, 2011). One study that investigated this relationship was conducted by Türkyilmaz (2015). The purpose of this study was to investigate the relationship among the reading attitudes, personality, self-regulation, and metacognitive awareness of high school students use of reading comprehension strategies. Participants in this study consisted of 419 (218 girls, 201 boys) Turkish high school students. Two measures, including the researcher developed Attitude Scale of Reading Attitude (Alpha = .71; Türkyilmaz & Aydemir, 2014) and the MARSI translated into Turkish (Alpha = .89; Mokhtari & Reichard, 2002), were used to obtain the data necessary to analyze the relationship between students' attitudes toward reading and their metacognitive awareness of reading comprehension strategy use. The results of the SEM analysis used in this study indicated a significant and positive relationship between students' reading attitudes and their reading comprehension strategy use. According to Türkyilmaz (2015), increased reading strategy use was associated with increased positive attitudes toward reading. However, specific statistical results cannot be reported here because these values were not reported in the narrative portion of the results section. The values were only presented in the figure depicting SEM results and unfortunately, the variables in the model were not translated from Turkish to English.

In their study of 4,988 Korean high school students' PISA 2009 reading scores, Lim et al. (2015) examined the relationship between students' use of reading strategies and their reading attitudes. The results of their analysis via structural equation modeling indicated that students' attitudes toward reading positively predicted students' use of reading comprehension strategies. Students' reading attitudes were significantly and positively correlated with the use of memorization ($r_s = .25$), elaboration ($r_s = .36$), and control ($r_s = .38$) reading comprehension strategies. The magnitude of paths from positive reading attitudes to elaboration ($\beta = .51$) and control strategies ($\beta = .36$) were substantial. Researchers found an unexpectedly significant path from negative reading attitudes to elaboration strategies. Results indicated that the direction of this path was positive ($\beta = .20$), rather than negative. Lim et al. (2015) attributed this inconsistent result to an issue of multicollinearity considering the negative bivariate relationship between negative reading attitudes and elaboration strategies ($f = -.29$).

Decline of reading attitudes. A multitude of studies investigating students' reading attitudes, conducted over the last 20 years, have found a consistent decline in students' attitudes toward reading as they get older. For example, in their study of elementary students' attitudes toward academic and recreational reading McKenna et al.'s 1995 concluded that students' reading attitudes grew increasingly negative as they transitioned through first through sixth grades. This large-scale study of students' reading attitudes consisted of 18,185 elementary school students, evenly distributed across the grade levels. Researchers employed two separate one-way ANOVA tests to examine the developmental trends in students' attitudes toward recreational and academic reading from first through sixth grade. The results of the *F*-tests for recreational and academic reading attitudes were statistically significant, $F(5,18,155) = 104.1$, $p < .001$ and $F(5,18,155) = 266.0$, $p < .001$, respectively. The effect size for the decline in students' attitudes toward recreational reading (ES = .20) and academic reading (ES = .27) were moderate according to Cohen's (1992) guidelines for interpreting effect sizes. McKenna et al. (1995), also discovered that the decline in students' reading attitudes throughout the elementary school years led to a decrease in reading frequency and an increase in students' avoidance of challenging reading activities.

Parallel results echoing the claim that students' positive reading attitudes decline over time were obtained in Bokhorst-Heng and Pereira's 2008 examination of middle school students' reading attitudes. This study was conducted with 173 (53% female, 47% male) 13-year old students in Singapore. A researcher developed 31-item Likert type Attitudes Toward Reading Survey was administered to students at the beginning and end of the school year to measure changes in their reading attitudes. The Cronbach alpha for the first and second administrations of the survey were .88 and .83, respectively. Paired *t*-tests comparing the students beginning and end of year scores revealed a decline in reading attitudes ($M = .36$), $t(172) = -10.880$, $p < .001$. The effect size (ES = -1.0) of the decrease in reading attitudes was large (Bokhorst-Heng & Pereria, 2008).

Finally, Gökhan (2012) obtained evidence that the reading attitudes of high school students also decline over time. Differences across grade levels were examined as part of the reading attitudes study including 426 ninth and 12th grade students, 226 (53%) and 200 (47%), respectively. The results of the independent samples *t*-test indicated a statistically significant difference $t(424) = 2.275$, $p < .05$ between the ninth and 12th grade students, in favor of the ninth grade students (Gökhan, 2012). In other words, students' positive reading attitudes declined between ninth grade and 12th grade.

There are a number of possible reasons for the decline of students' reading attitudes over time. First, for struggling readers, the cumulative effect of long term reading difficulties has a negative impact on high school students' attitudes toward reading (Fuchs et al., 1999; McKenna et al., 1995). Second, according to the last tenet of McKenna's model of reading attitudes, students' beliefs about the outcomes of reading are partially formed in unison with their beliefs about the competing outcomes of simultaneously available activities (McKenna et al., 1995). As students get older, the increasing availability of potentially more desirable recreation activities can negatively influence their attitudes toward reading. According to Bokhorst-Heng and Pereira (2008), when more desirable activities counter the option of reading, students' positive attitudes towards reading decline, even for proficient readers.

Reading attitudes and reading comprehension strategy instruction. The decline in students' reading attitudes over time, coupled with the high rate of struggling high school readers, makes it exceedingly important for high school teachers to invest instructional time into developing and reinforcing students' positive attitudes toward reading. Evidence in the literature indicating the positive relationship between reading instruction and students' attitudes toward reading supports this notion.

As part of a study of students' reading attitudes, Jhang (2014) examined the relationship between reading instruction and high school students' attitudes toward reading. This study, mentioned earlier, utilized the PISA 2009 international dataset including data from approximately 470,000 15-year-old students across all 65 OECD countries. Three level hierarchical linear modeling (HLM) was employed to estimate the relationship between reading instruction and students' reading attitudes at the country, school, and student levels. The results revealed that instruction was significantly and positively correlated with students' attitudes toward reading (.10), which indicated that reading instruction positively influences students reading attitudes.

Lim et al. (2015), also used data from the PISA 2009 to explore the relationship between reading instruction and high school students' reading attitudes in their examination into different aspects of Korean high school students' reading achievement. As mentioned above, the PISA 2009 Korean dataset included data of 4,988 students from 157 schools. The researcher used structural equation modeling (SEM) to investigate the relationship between reading strategy instruction and reading attitudes. Results of the SEM analysis revealed a statistically significant and a positive path from instructional strategies in reading to positive attitudes toward reading ($\gamma = .07$), suggesting that reading instruction positively influences high school students reading attitudes.

Conclusion. It is important that high school teachers understand the relationship between reading comprehension achievement and reading attitudes, including the factors that influence this relationship. This understanding is important because it informs teachers of the need to explicitly identify their students' attitudes toward reading. Obtaining this concrete knowledge of students' reading attitudes is necessary for teachers to plan instruction that may potentially improve students' negative reading attitudes, and subsequently increase their reading comprehension achievement.

Conclusion

High school students' lack of reading comprehension strategy knowledge, and their inability to use these strategies effectively could inhibit their overall academic achievement by limiting the knowledge they can independently obtain via text. The assertion that high school students who appropriately utilize comprehension strategies demonstrate higher reading achievement than students who do not, is consistently supported by the evidence discussed in the literature review (e.g., Cromley & Azevedo, 2007; Denton et al., 2015; Hong-Nam et al., 2014; Shera, 2014). Despite this evidence, many high school students often do not possess the knowledge of how to, or do not appropriately use the strategies to effectively make sense of what they read. This suggests that high school content area teachers need to teach struggling readers how to effectively use the strategies used by proficient readers. However, most of the studies presented in the literature review indicated that very little, if any, high school instructional class time is spent teaching reading comprehension strategies (e.g., Ness, 2009; Swanson et al., 2016; Wexler et al., 2016).

In addition to improving students' reading achievement by teaching them reading comprehension strategies, the results of research discussed in this chapter indicated that improving students' attitudes toward reading can also positively impact achievement. A majority of the research conducted to examine students' reading comprehension difficulties has focused on the cognitive rather than the affective aspects (i.e., attitude, motivation, engagement, etc.) of the reading process. However, some studies have provided evidence to support the unique variance in reading comprehension accounted for by students' attitudes toward reading (e.g., Kasapoglu, 2014; Sari, 2015). Additional investigations of high school students' reading attitudes have the potential to contribute to a more complete understanding of the affective components of the reading process.

Many factors that contribute to high school students' reading difficulties including poor reading fluency, limited background knowledge, limited strategy knowledge, inadequate strategy use, ineffective reading instruction, and/or negative reading attitudes were described in this literature review. The influences of student (i.e., gender, minority status, and SES) and classroom (i.e., class time and class size) variables on reading difficulties were also explained. However, exactly how these variables affect high school students' comprehension, specifically the simultaneous effect of these factors on reading comprehension, and the extent to which differences in the three latent variables are accounted for by student and classroom factors, are still not fully understood.

Considering the existing research discussed in this chapter, the purpose of the present study was to first examine the relationships among U.S. high school students' reading comprehension strategy use, reading strategy instruction, reading attitudes, and their reading comprehension achievement including any effects of gender, minority status, SES, class size, and class time. The subsequent purpose was to investigate group differences in students' reading comprehension achievement, strategy use, strategy instruction, and reading attitudes across the five student and school level variables.

Reading comprehension is a complex cognitive process and the ability to understand text involves a convergence of different correlated variables. Consequently, it is virtually impossible to identify an individual cause of high school students' reading comprehension failure. Indeed, an assortment of recent studies have examined the simultaneous effect of multiple variables on students' reading comprehension. Several of the multivariate studies conducted to investigate high school students' reading comprehension proficiency, discussed in the literature review, employed a number of different types of analytical methods (please see Table 1). Many of these studies also utilized PISA datasets from various countries and various years.

However, a study utilizing the PISA 2009 U.S. dataset to examine the relationships among the specific combination of variables identified above on U.S. high school students' reading comprehension achievement via structural equation modeling was not identified in the existing literature. This study seeks to add to the limited research base on high school students reading comprehension including the influences of gender, race/ethnicity, SES, reading attitudes, reading strategy use, and reading strategy instruction have on their reading comprehension achievement. Specifically, concurrently accounting for a variety of factors known to contribute to reading comprehension achievement can add to the existing empirical research by potentially explaining a larger portion of the variance in achievement among students.

Table 1

Multivariate Studies of High School Students' Reading Achievement

Author/Year	PISA Data	Analysis Method(s)	Variables
Brown-Jeffy, 2006	N	HLM	RCA, MS, SES, School Environment
Chui & McBride-Chang, 2006	Y	HLM	RCA, Gender, SES, RE
Cromley & Azevedo, 2011	N	CCA	RCA, RCSU, Prior Knowledge, Inference, Vocabulary, Decoding
Denton et al., 2015	N	ANOVA	RCA, RCSU, Gender, Grade
Duncan et al.,2016	N	ANOVA	RCA, Gender, SES, Word Identification, Fluency, Print Exposure
Hong-Nam et al.,2014	N	ANOVA	RCA, RCSU, Grade
Huang, 2015	Y	ANOVA/ HLM	RCA, Gender, SES, Parental Involvement, Immigration, Language, T Certification, Major T, ST Shortage, S Size, S Climate, S Location, S Selectivity, S Leadership, S Discipline, Persistence, CT, Out of CT
Kasapoglu, 2014	Y	LR	RCA, RA, Gender, CT, S Age, Home Resources, Recreational Reading, Parental Education, Time Learning Test Language
Lietz, 2006	N	HLM	RCA, Gender, Age, Instruction Language
Ozdemir & Gelbal, 2014	Y	CCA	RCA, SES, Self Confidence
Sari, 2015	Y	SEM	RCA, RCSI, RA, Study Habits
Shera, 2014	Y	HLM	RCA, RCSU, Gender, SES, Reading Engagement, Class Environment, Family Structure, S Type, S Sector, S Location
Singh, 2008	Y	*t*-tests	RCA, Gender, SES, RE, Reading Diversity
Sirin, 2005	N	FEM/REM	RCA, MS, SES, Grade, School Location

Note. HLM = Hierarchical Linear Modeling; CCA = Canonical Commonality Analysis; ANOVA = Analysis of Variance; LR = Logistic Regression; FEM = Fixed Effects Modeling; REM = Random Effects Modeling; RCA = Reading Comprehension Achievement; RSCU = Reading Comprehension Strategy Use; RCSI = Reading Comprehension Strategy Instruction; RA = Reading Attitudes; MS = Minority Status; SES = Socio-Economic Status; CT = Class Time; S = School; T = Teacher; RE = Reading Enjoyment.

Chapter Three

This study employed a cross-sectional correlational research design to explore the relationships among students' reading comprehension achievement, reading strategy use, reading comprehension strategy instruction, and reading attitudes. The purpose of this chapter is to describe the design of the study including the data source, data collection and processing procedures, study sample, and data measures. The statistical techniques used to analyze the data including the preliminary data analyses, confirmatory factor analysis, structural regression analysis, and MIMIC analysis procedures are also described in this chapter. According to the George Mason University Office of Research Integrity and Assurance (ORIA), this study was classified as exempt because it did not meet the requirements for human subject research. The Institutional Review Board letter of exemption is provided in Appendix A.

Data Source

The U.S. subset of the PISA 2009 dataset was used in this study (National Center for Education Statistics [NCES], 2011a). PISA, a standardized assessment coordinated by the OECD, was first issued by the organization in 2000 (Fleischman et al., 2010). This internationally comparative assessment, designed to measure the reading, mathematics, and science literacy achievement of 15-year-old students, is conducted every three years.

Although reading, mathematics, and science are always tested in the PISA assessment, only one of the three subject areas is designated as the main area of focus for the assessment cycle. Every three years two-thirds of the total testing time is devoted to one main subject area, while the other two are considered the minor subject areas for that cycle (Fleischman et al., 2010). The main subject area focus of the assessment rotates with every administration cycle. In the PISA 2009, the main subject area was reading literacy (OECD, 2010d). The PISA 2009 dataset was utilized because it was the most recent administration cycle that focused on reading achievement.

The PISA 2009 consisted of several components including, but not limited to, the cognitive test of reading literacy, student background questionnaires, a parent questionnaire, and a school questionnaire. Data from the cognitive test of reading literacy and student questionnaires were obtained for this study from the U.S. portion of the PISA 2009. Data from the two student questionnaires were utilized because they surveyed various aspects of students' background relevant to this investigation. According to the PISA 2009 Assessment Framework, the purpose of surveying these background factors was to associate them with student's reading achievement outcomes and offer assessment stakeholders insight into both the development of students' skills and attitudes at school and at home, and into how the factors interact with each other (Organization for Economic Cooperation and Development [OECD], 2010a). The combination of student achievement and background data available in the PISA dataset made it an excellent match for answering the questions of this investigation.

The target population for this study was the entire population of students between the ages of 15 years and 3 (complete) months and 16 years and 2 (complete) months at the start of the testing window. All students were in grades 7-12 and attended private or public schools in the US (Organization for Economic Cooperation and Development [OECD], 2010c). A representative sample of 5,233 students from 165 schools was drawn using a two-stage systematic stratified sampling procedure (Hopstock & Pelczar, 2011). First, schools were sampled using the Probability Proportional to Size sampling design (Organization for Economic Cooperation and Development [OECD], 2012). This design was used to account for the varying size of U.S. schools to ensure that small schools were as equally likely to be selected to participate. In the second stage, PISA eligible students (students born in 1993 and in seventh grade or higher) were randomly selected from the identified schools. If the school had less than the established U.S. target cluster size of 42 PISA eligible students, then all PISA eligible students were selected (OECD, 2012).

Data Collection and Processing Procedures

Once the data source was selected, the PISA 2009 U.S. dataset was extracted and processed. Data processing included analyzing the dataset for missing data, standardizing and rescaling the reading achievement variable, normalizing sampling weights, and recoding the control variables and negatively worded latent indicators.

Data extraction. Extracting the data was the first of the five data collection and processing procedures conducted. Data extraction involved retrieving the PISA 2009 cognitive test of reading literacy dataset and examining it to identify the information relevant to this study. The raw PISA 2009 U.S. dataset, and the corresponding SPSS control files, the student codebook, and the PISA 2009 U.S. Student Questionnaire were downloaded from the NCES website (NCES, 2011a). The PISA 2009 Reading for School Questionnaire was then located on and downloaded from the OECD PISA 2009 database website (Organization for Economic Cooperation and Development [OECD], 2015). After downloading the data, the SPSS control files were used to open the dataset in the IBM SPSS Statistical Software, Version 23 (hereafter referred to as SPSS).

Next, the original dataset (consisting of 450 variables), the student questionnaires, and the codebook were substantively reviewed to identify potential items for selection. The review was guided by the tentative model developed based on the review of theoretical and empirical evidence. The tentative model, like the hypothetical model illustrated in Figure 1, included the hypothesized latent constructs, relationships among the latent constructs, and their relationships with other observed variables in the model.

A total of 26 variables were selected for analysis including a reading comprehension achievement variable and the corresponding sampling weight. Seven items that could potentially define the reading comprehension strategy use factor, five items to measure the reading comprehension strategy instruction factor, and seven items to define the reading attitudes latent factor latent factor were also selected. Finally, three student level (i.e., gender, race/ethnicity, and SES) and two school level (i.e., class time and class size) control variables were selected for analysis. A new data file was created in SPSS including only the 26 variables selected for analysis. Editing the new data file was the final step in the data extraction process. This included revising the variable names and labels, as well as reordering the variables to better facilitate data analysis.

Missing data. After processing the dataset, it needed to be analyzed for missing data. SPSS was used to obtain the descriptive statistics for all variables in the model. Of the total sample (*N* = 5,233), 1014 cases had missing values on one or more items. The total attrition of the missing data was determined by calculating the proportion of missing data (1,014/5,233 = .19 or 19%). Thus, listwise deletion, a commonly used method of handling missing data, would result in the removal of 19% of the cases in the sample. According to Cheema (2013), unless missing data are missing completely at random, this approach to handling missing data could cause the sample to be no longer representative of the original population. This is problematic because any results based on a non-representative sample cannot be generalized to the target population.

After reviewing the missing data and considering the large sample size of the dataset as well as the nature of the variables of interest, is was assumed that the missing data was missing at random. In other words, it was assumed that the data missing on a certain variable, *Y*, are probably missing due to a different observed variable, not due to *Y* itself, after controlling for *Y* (Allison, 2000; Organization for Economic Cooperation and Development [OECD], 2010b). To avoid the possibility of working with an unrepresentative sample, the Expectation Maximization method (EM Imputation) was selected to impute the missing data. The EM Imputation method was selected because it outperforms other missing data handling methods such as mean imputation, regression imputation, and listwise deletion by providing the largest accuracy gains in parameter estimation (Cheema, 2012). Missing data was imputed using SPSS, which created a new data set that was used in all further data analyses. After imputing missing data, the final sample still had 5,233 cases which were representative of the 4,103,738 15-year-old high school students in the target population (OECD, 2010c).

Standardizing and rescaling data. The next step in processing the PISA 2009 data was to standardize and rescale the reading comprehension achievement plausible values to improve the interpretation of the results. The plausible values were standardized by calculating a z-score for each value using SPSS. The z-scores were then rescaled to have a mean of 500 and a standard deviation of 100.

Normalizing sampling weights. Next, the sampling weights needed to be normalized. Complex sampling techniques used to collect the samples of large scale data sets tend to produce population subsets (e.g., race/ethnicity) with larger numbers of students in some categories of the subsets (e.g., Caucasian, African American, etc.) than others (Hahs-Vaughn, 2005). The subsets that have a lesser number of students are then oversampled to equalize the groups, thus creating a distorted picture of the population. Therefore, sampling weights are applied to the raw data to ensure that the sample is representative of the population. Sampling weights for the PISA 2009 dataset were provided to accommodate the issue of oversampling (Organization for Economic Cooperation and Development [OECD], 2009).

However, the weights provided for the PISA 2009 dataset were for the population rather than the sample. Using population sampling weights is problematic because any estimates that are sensitive to sample size, such as standard error estimates and test statistics, will be incorrect (Hahs-Vaughn, 2005). Consequently, the population sampling weights needed to be normalized. SPSS was employed to calculate the normalized sampling weights using *Equation 1*, $W_n = W_o/\mathbf{\Sigma} W_o \times N$; where W_o was the population weight provided by PISA, $\mathbf{\Sigma} W_o$ was the sum of the population weights, and N was the total sample. The normalized sampling weights were subsequently used in all structural regression and MIMIC analyses.

Recoding. The final step in processing the data was to recode the control variables and the negatively worded latent variable indicators. Dichotomous dummy variables were created for the three student level control variables (gender, race/ethnicity renamed minority status, and SES) and two school level control variables (class time and class size). It was necessary to create these dichotomous dummy variables because in MIMIC analyses the construct of interest is regressed on a dummy variable which assumes values of 0 and 1 (Dimitrov, 2009). The dummy variables created were gender (0 = female, 1 = male), minority status (0 = minority, 1 = nonminority), SES (0 = low SES, 1 = high SES), class time (0 = less class time, 1 = more class time), and class size (0 = small class size, 1 = large class size). The reference categories for these five variables were female, minority, low SES, less class time, and small class size.

Next, the responses of the negatively worded latent variable indicators were reverse coded to create a uniform direction of responses (negative to positive) across all the student questionnaire items. Indicators 1, 3, 5, and 7 of the reading attitudes construct and all seven indicators of the reading comprehension strategy use construct were reverse coded. As this was the last step in the data processing, the SPSS data file was saved as a Fixed ASCII file so it could be read in future Mplus analyses.

Study Sample

After missing data imputation, the final sample, *N* = 5,233, selected from 165 public and private schools in the US, was representative of the 4,103,738 students that comprised the target population. As the student selection criteria specified the inclusion of all students born between July 1, 1993 and June 30, 1994, the age and grade levels of the students varied (Fleischman et al., 2010). The age of the student sample ranged from 15.25 years to 16.33 years (M = 15.79, SD = .30) and the grade level of the sampled students ranged from grade eight to grade twelve (M = 10.09, SD = .55). The representation of females (n = 2,546, 49%) and males (n = 2,687, 51%) was approximately equal in the sample. Minority students (n = 2,337, 45%), including African American students (n = 672, 13%), Hispanic students (n = 1,193, 23%), Asian students (n = 204, 4%), American Indian/Alaska Native students (n = 51, 1%), Native Hawaiian/Other Pacific Islander students (n = 40, 1%), and students of more than one race (n = 177, 3%) were underrepresented compared to the number of nonminority, white, students (n = 2,896, 55%). The representation of students' SES, measured by the PISA ESCS index, was approximately equal. The number of students with a low SES (i.e., the students whose ESCS index fell below the median ESCS z-score of 501), n = 2,601 (50%), was slightly less than the number of students with a high SES (i.e., the students whose ESCS index was greater than 501), n = 2,632 (50%).

In terms of the school level characteristics, the number of students who reported having less English class time (i.e., 250 minutes or less of class time per week), n = 2,841 (54%), was higher than the number of students who reported having more English class time (i.e., more than 250 minutes of class time per week), n = 2,392 (46%). Also, the number of students who reported being a member of a small class (i.e., a class of 20 students or less), n = 2,469 (47%), was lower than the number of students who reported being a member of a large class (i.e., a class of 21 or more students), n = 2,764 (53%).

Data Measures

Observed and latent data measures were utilized in this study. The observed measures included the reading comprehension achievement variable as well as the student and school level control variables. The latent data measures included the reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes variables.

The measure of high school students' reading comprehension achievement was obtained from the PISA 2009 cognitive test of reading literacy achievement. The measures for the five student (i.e., gender, minority status, and SES) and school (i.e., class time and class size) level control variables were obtained from the PISA 2009 U.S. Student Questionnaire, one of the two supplemental student questionnaires included in the PISA 2009 assessment. The PISA 2009 U.S. Student Questionnaire consisted of 46 questions that surveyed student's home backgrounds, their approaches to learning, their learning environments, and their familiarity with computers (OECD, 2010a). The other supplemental student questionnaire, the PISA 2009 Reading for School Questionnaire, was a two-question supplemental survey of students' school-based reading practices. The two items on the PISA 2009 Reading for School Questionnaire were designed to elicit responses regarding the types of texts and reading tasks required by students' reading related classwork and homework (OECD, 2012). All 19 items hypothesized to measure the latent factors (reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes) were obtained from these two questionnaires. The complete PISA 2009 U.S. Student Questionnaire and PISA 2009 Reading for School Questionnaire are provided in Appendix B and Appendix C, respectively.

Observed variables. The observed variables in this study included three student level control variables, two school level control variables, and the reading comprehension achievement outcome variable.

Student level control variables. The three student level control variables selected for this study were gender, minority status, and socio-economic status.

Gender. The values used to measure this variable were obtained from PISA 2009 U.S. Student Questionnaire question 4, "Are you female or male?" (p 4). Responses were coded, 1 = female and 2 = male. A dummy variable for gender, where 0 = female and 1 = male, was created to maintain consistency with the other dichotomous variables.

Minority status. The data used to measure this variable were obtained from the race/ethnicity composite variable. This composite variable was a combination of student responses from questions 5 and 6 on the PISA 2009 U.S. Student Questionnaire. Question 5 asked students to indicate if they were Hispanic or Latino. And question 6, "Which of these categories best describes your race?" (p.5), asked students to select one of the following race categories with which they most closely identified: white, African American, Hispanic, Asian, American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, or More Than One Race. The various race categories were collapsed into a dichotomous variable, minority status, which was coded 0 = minority and 1 = nonminority. The minority category included African American, Hispanic, Asian, American Indian/Alaska Native, and Native Hawaiian/Other Pacific Islander students and students of more than one race and the nonminority category included white students.

Socio-economic status (SES). The values designated to measure students' SES were chosen from the Index of Educational, Social, and Cultural Status (ESCS) in the PISA 2009 (OECD, 2012). According to Hopstock and Pelczar (2011), the ESCS index consisted of the following three components: the higher parental occupation, the higher parental education (expressed as years of schooling), and the index of home possessions (including the number of books in the students' home and all the items on the family wealth possessions, cultural possessions, and home educational resources indices). Each component was a compilation of different questions from the student survey. A sample item from this scale asked the students to indicate their mother's main job. The Cronbach's alpha internal consistency reliability estimate for this scale was, $\alpha = 0.69$. A dummy variable for SES was created to maintain consistency with the other dichotomous variables. The minimum and maximum ESCS values ($X_{min} = 110.92$, $X_{max} = 799.29$) and the median ESCS z-score (Mdn = 501) were used to assign values to the binary SES categories. The low SES category, labeled 0, included students with ESCS values ranging from 110-500.99, and the high SES category, labeled 1, included students with values ranging from 501-800.

School level control variables. Class time and class size were the two school level control variables selected for this study. They were chosen due to the statistically significant positive effects of additional class time (Huebener et al., 2016) and small class size (Krassel & Heinesen, 2014) on overall academic achievement.

Class time. The values of the learning time for English class variable, LMINS, presented in the PISA 2009 U.S. student codebook were used as the values for the class time variable in this study (National Center for Education Statistics [NCES], 2011b). The LMINS variable was a product of two questions on the PISA 2009 U.S. Student Questionnaire (Hopstock & Pelczar, 2011). The two questions from PISA 2009 U.S. Student Questionnaire that comprised the LMINS variable for English class included: question 32, "How many minutes, on average, are there in a class period for the following subjects? Minutes in a class period in English" (p. 21) and question 33, "How many class periods per week do you typically have for the following subjects? Number of class periods per week in English" (p. 21). Thus, the LMINS variable represented the total number of minutes per week student had of English class time. The minimum and maximum class time values ($X_{min} = 0$, $X_{max} = 1000$) and the median class time value (Mdn = 250) were used to assign values to the binary class time categories (0 = less class time, 1 = more class time). The less class time category, labeled 0, included class times of 0-250 minutes per week, and the more class time category, labeled 1, included class times of 250.01-1000 minutes per week.

Class size. The concept of class size is defined in this study as the number of students assigned to a class that one teacher is responsible for teaching (Finn, 2002). It is important to distinguish between the terms class size and student-teacher ratio because they are often conflated or confused. In comparison to class size, student-teacher ratio is the number of students in a class divided by the number of teachers. For example, the student-teacher ratio of a large class of 35 students is small if there are also three teaching aids in the classroom (Finn, 2002).

Data for this variable were obtained from question 39 on the PISA 2009 U.S. Student Questionnaire, “On average, about how many students attend your English class?” (p. 26). A dummy class size variable for was created to maintain consistency with the other dichotomous variables. Values for the binary class size categories (0 = small class size, 1 = large class size) were assigned using the minimum and maximum class size values ($X_{min} = 1$, $X_{max} = 90$) and the classification of small class size as classes with 20 students or less. Classes consisting of one through 20 students were recoded to the small class size category, and those with 21-90 students were recoded to the large class size category. The decision to code class sizes of 20 students or less as small classes was based on Finn’s (2002) recommendation for small class sizes.

Reading comprehension achievement. In this study, reading comprehension achievement is a dependent variable measured by data from the reading literacy achievement portion of the PISA 2009 U.S. cognitive test. The reading literacy achievement portion of the test included 102 questions targeting three aspects of reading: accessing and retrieving information, integrating and interpreting information, and reflecting on and evaluating information (Fleischman et al., 2010). Approximately 50% of the questions on the paper-and-pencil test were multiple-choice, 30% were short answer constructed response items, and 20% were closed constructed response items (Hopstock & Pelczar, 2011). A sample item for this construct included, “One purpose of the speech in Part B [of the Democracy in Athens passage] was to honor soldiers who fell in the first year of the Peloponnesian War. What was another purpose of this speech?" (OECD, 2010a, p. 58). Responses to the multiple-choice and closed constructed response items were coded either correct or incorrect with no partial credit awarded. Whereas, the short answer constructed response items, like the sample item above, were either awarded no credit, partial credit, or full credit (OECD, 2010a).

Plausible values. In addition to the 102 reading literacy items, the PISA 2009 assessment also included 35 mathematics literacy items and 52 science literacy items. Every student booklet contained all the reading literacy items, but the math and science literacy items were distributed differently across the various versions of the test booklet (NCES, 2011b). Because students were not administered every test item is was not possible to estimate scores for individual students (NCES, 2011b). Instead, the dataset included five reading achievement plausible values for each student. Plausible values represent the distribution of potential scores for all students in the population with similar characteristics and identical item response patterns (Fleischman et al., 2010). The PISA 2009 plausible values were randomly selected from the total distribution of scores that could be reasonably assigned to each individual student (OECD, 2012).

Two common approaches for using plausible values to represent student achievement include averaging all the values or selecting one value, typically the first, to work with (Perry, Wiederhold, & Ackerman-Piek, 2014). According to Perry et al., averaging the plausible values generally provides unbiased estimates of students' skills, however, the corresponding standard errors are underestimated. While selecting only one plausible value also causes the standard errors to be underestimated, it is usually to a lesser extent. Perry et al. cautioned that using only one plausible value can result in differing estimates depending on which values is used. Using the PISA 2009 dataset, Jerrim and Micklewright (2014) tested the concern of using one plausible value to represent the achievement variable. First, they estimated five separate models for each plausible value and averaged the resulting parameter estimates, then they estimated a model using only one plausible value. When they compared the results of two procedures they only found a negligible difference in the student achievement estimates. Based on the support provided by Perry et al. and Jerrim and Micklewright for selecting only one plausible value, the first plausible value from the PISA dataset was selected to represent the reading achievement variable throughout this study.

Latent variables. The three reading related latent variables (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes) included in this study were measured by a total of 19 observable indicators.

Reading comprehension strategy use. This scale consisted of seven items from the PISA 2009 Reading for School Questionnaire that measured students reading comprehension strategy use. In this study reading comprehension strategy use is broadly defined as the intentional application of a reading strategy to repair or improve comprehension. The seven items selected to measure the reading strategy use latent factor were obtained from question 2 of the PISA 2009 Reading for School Questionnaire which stated, “During the last month, how often did you have to do the following kinds of tasks for school (in the classroom or for homework)?” (p.1). The following items of question 2 on the PISA 2009 Reading for School Questionnaire were selected to define the reading strategy use construct: “Find information from a graph, diagram, or table”, “Explain the cause of events in a text”, “Explain the way characters behaved in a text”, “Learn about the life of a writer”, “Explain the purpose of a text”, “Learn about the place of a text in the history of literature”, and “Explain the connection between different parts of a text” (p. 1). Options for student responses ranged from 1 to 4, 1 = Many Times, 2 = Two or Three Times, 3 = Once, or 4 = Not at All, where the higher values indicated less reading strategy use. All items were reverse coded so the higher values indicate higher reading strategy use. The Cronbach’s and LVM reliability of the strategy use scale were equivalent (α and ρ_{xx} = 0.76). Item means ranged from 2.34 to 3.21, item variances ranged from 0.87 to 1.15, and the corrected item-total correlations ranged from 0.33 to 0.56.

Reading comprehension strategy instruction. This scale included five items from the PISA 2009 U.S. Student Questionnaire to measure reading strategy instruction. Reading comprehension strategy instruction is defined as the explicit instruction of reading comprehension strategies including teaching students what comprehension strategies are, why they are important, as well as how, when, and why to apply specific strategies. Explicit instruction, as defined in this study, also entails frequent teacher modeling of appropriate strategy use, ongoing guided practice, fostering of students’ independent comprehension strategy use, progress monitoring, and feedback (Ness, 2009). The five items selected to measure the reading strategy instruction latent construct were obtained from question 41 of the PISA 2009 U.S. Student Questionnaire which stated, “In your English classes, how often does the following occur?” (p. 27). The following items of question 41 from the PISA 2009 U.S. Student Questionnaire were selected to explain the reading strategy instruction latent factor: “The teacher asks questions that challenge students to get a better understanding of a text”, “The teacher gives students enough time to think about their answers”, “The teacher recommends a book or author to read’, “The teacher encourages students to express their opinion about a text”, and “The teacher shows students how the information in texts builds on what they already know” (p. 27). Student response options ranged from 1 to 4, 1 = Never of Hardly Ever, 2 = In Some Classes, 3 = In Most Classes, or 4 = In All Classes, with the higher values indicating more effective reading strategy instruction. The Cronbach’s and LVM reliability were the same (α and ρ_{xx} = 0.83). Item means ranged from 2.39 to 2.97, item variances ranged from 0.65 to 0.91, and the corrected item-total correlations ranged from 0.51 to 0.70.

Reading attitudes. This scale included seven items from the PISA 2009 U.S. Student Questionnaire that measured students’ attitudes toward reading. Reading attitudes was defined in this study as the spectrum of positive to negative feelings about reading that cause a reader to pursue or avoid reading opportunities and reading related activities (Alexander & Filler, 1976). The seven items selected to measure the reading attitudes latent construct were obtained from question 28 of the PISA 2009 U.S. Student Questionnaire which stated, “How much do you agree or disagree with these statements about reading?” (p. 17). The following indicators of question 28 from the PISA 2009 U.S. Student Questionnaire were hypothesized to define the reading attitudes construct:

“I read only if I have to”*, “Reading is one of my favorite hobbies”, “I find it hard to finish books”*, “I feel happy if I receive a book as a present”, “For me, reading is a waste of time”*, “I enjoy going to a bookstore or a library”, and “I read only to get information that I need”* (p. 17). Options for student responses ranged from 1 to 4, 1 = Strongly Disagree, 2 = Disagree, 3 = Agree, or 4 = Strongly Agree, with the higher values indicating a more positive attitude toward reading. The asterisks after four of the indicators denote negatively phrased items whose response categories were reverse coded to maintain a uniform direction (negative to positive) among the responses. The Cronbach’s and LVM reliability of the scale were equal (α and ρ_{xx} = 0.89). Item means ranged from 2.11 to 2.97, item variances ranged from 0.78 to 0.95, and the corrected item-total correlations ranged from 0.50 to 0.76.

Statistical Data Analysis Procedures

Various preliminary and structural equation modeling (SEM) statistical analysis procedures were employed to examine the research questions posed in this study. The research questions and corresponding data analytic method used to investigate each question are presented in Table 2.

Preliminary analyses. Preliminary analyses were conducted prior to investigating the three main research questions. The preliminary analyses included an examination of descriptive statistics and an internal consistency reliability analysis of the three latent constructs (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes).

Table 2

Research Questions and Statistical Data Analysis Procedures

Research Question	Statistical Data Analysis Method
RQ1 – Do the observable indicators selected to measure the three latent constructs (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes) appropriately define them?	Confirmatory Factor Analysis (Under SEM)
RQ2 – Are there direct and indirect effects among the three latent variables, the five student and school level variables, and high school students' reading comprehension achievement? RQ2a – What are the direct and indirect effects of the three latent constructs and the five student (i.e., gender, minority status, and SES) and school (i.e., class time and class size) level variables on high school students' reading comprehension achievement? RQ2b – What are the direct effects among the three latent constructs (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes)? RQ2c – What are the direct effects of the five student and school level variables on the three latent constructs?	Structural Regression Analysis (Under SEM)
RQ3 – Are there significant group differences in high school students' reading comprehension achievement, reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes across the five student (i.e., gender, minority status, and SES) and school (i.e., class time and class size) level variables?	Multiple Indicator, Multiple Cause Group-Code Modeling (Under SEM)

Descriptive statistics. A variety of descriptive statistics were obtained using SPSS. The frequencies and percentages of the gender, race/ethnicity, and minority status variables as well as the means, standard deviations, minimum values, and maximum values of the age and grade variables were obtained to describe the study sample presented above. The means, medians, minimum values, and maximum values of the SES, class time, and class size variables were obtained to determine the appropriate recoding values. Additionally, the frequencies and percentages of the student and school level variables were obtained to describe their composition. Finally, the means, standard deviations, and variances of the 19 items that measured the three latent variables were obtained to provide a description of their distributions.

Reliability. It was necessary to investigate the reliability of the three latent factors (reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes) to demonstrate internal consistency reliability prior to conducting any analysis involving the three latent constructs. It was important to first establish the reliability of the three latent constructs because any interpretations and generalizations based on them were only valid if the constructs are reliable. In general, reliability refers to the accuracy, consistency, and replicability of a measure (Dimitrov, 2012). Internal consistency reliability, one specific type of reliability, estimates how consistently items intended to measure the same construct produce similar scores (Kirk & Vigeland, 2014).

Internal consistency reliability estimates (i.e. Cronbach's Alpha coefficient), total item correlations (i.e., Corrected Item Values), and Cronbach's Alpha if Deleted for the, three latent constructs (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes) were obtained using SPSS. Nunnally's (1978) widely-accepted social science cutoff $\alpha = 0.70$ was used to assess the reliability of the three scales. The corrected item total correlations were inspected to determine the amount of correlation each item had with the scale. Nunnally and Bernstein (1994) recommended that items with a correlation of $r \leq .30$ to their scale should be retained. Cronbach's Alpha if Deleted values were then analyzed to determine other potential problems by examining if any of the values exceeded the alpha reliability estimate for its factor. If the Cronbach's Alpha if Deleted exceeds its factor, it signals that the reliability would increase if the item were removed. Any items flagged by the Cronbach's Alpha if Deleted were examined, and if it made sense substantively the item was removed.

Structural equation modeling. Once the reliability of the three latent constructs was assessed, structural equation modeling (SEM) was used to test the hypothetical relationships among the reading related variables against the empirical data. SEM was selected for this study because it is a comprehensive statistical technique used to quantify and test substantive theories (Raykov & Marcoulides, 2006). Specifically, SEM was selected because it can be used to explain how sets of indicators define latent constructs, how latent constructs are related to each other, and the extent to which the hypothetical models are supported by the sample data (Schumacker & Lomax, 2010). Further, the use of SEM is preferable to traditional regression analysis because it accounts for the measurement error in both the dependent and independent observed variables in a model. Whereas, as Raykov and Marcoulides (2006) explained, potential measurement error in the independent variables is overlooked in traditional regression analysis, which can cause the results to be incorrect.

Structural equation models consist of two components: a measurement model and a structural model. The measurement model is a multivariate regression model that describes the relationship between a set of latent variables and their indicators (Muthén &

Muthén, 2010). Latent variables are hypothetical constructs that are measured by a set of observable indicators because they cannot be measured directly (MacCallum & Austin, 2000). The use of latent variables is another advantage of SEM because it allows for the estimation of relationships among latent constructs that are free of the effects of measurement unreliability (Raykov, Tomer, & Nesselroade, 1991). Once the relationships between the latent constructs and their indicators are defined, the structural models are tested. According to Muthén and Muthén (2010), structural models have several uses. They can be used to describe relationships among latent constructs, relationships among observed variables, and/or relationships between latent constructs and observed variables that are not factors indicators (Muthén & Muthén, 2010).

Implementation of SEM procedures in this study began with the review of the relevant theoretical and empirical literature to refine and gather support for the hypothetical model. It is important to note that SEM analysis often involves the examination of alternative models since more than one model can fit a given set of data. However, after consideration, it was determined inappropriate to examine alternative models in this study. This decision was made because the hypothetical model was developed based on specific theoretical and research evidence regarding the relationships among the variables. Further, the purpose of this study required the simultaneous examination of the relationships among the selected variables and their relationships with reading comprehension achievement that were indicated in the hypothetical model.

Confirmatory factor analysis was then used to examine the SEM measurement model to determine whether the observed indicators appropriately defined the latent factors. Finally, after the measurement model was assessed, two specific SEM procedures (i.e., structural regression analysis and MIMIC analysis) were used to test the hypothesized relationships among the constructs, and their relationships with the observed variables in the model that were not factor indicators.

SEM measurement model – Confirmatory factor analysis. Confirmatory factor analysis (CFA) was employed to investigate the first research question, 'Do the observable indicators selected to measure the three latent constructs (reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes) appropriately define them?'. CFA is considered the measurement model in SEM (Muthén & Muthén, 2010). CFA models include latent factors and the observable indicators used to measure them. CFA, under the SEM framework, was used to test the relationships between the latent constructs and their observed indicators. The specific procedures required to conduct the CFA analysis, described below, included: model specification, model identification, examination of the CFA assumptions, testing for model fit, and evaluating the model parameters. A latent variable modeling reliability analysis was also conducted due to the finding of correlated errors in the CFA model.

The first step in CFA, under the SEM framework, was to specify a hypothetical CFA model (Dimitrov, 2012). The hypothetical CFA model specification consisted of determining which factors to include in the model, which observable variables defined the factors, which factors were expected to correlate, which errors were expected to correlate, and which factor loadings should be held equal. The CFA model specifications used in this investigation were supported by the relevant theoretical and empirical research presented in the literature review.

The next step after specifying the hypothetical confirmatory factor analysis (CFA) model was to determine its identification status. A model is considered identified when it is possible to compute a unique set of parameter estimates for all unknown model parameters. According to Dimitrov (2012), this can only happen when the number of known elements in the model is greater than or equal to the number of unknown model parameters. Models can be classified as underidentified, just-identified, or overidentified depending on the ratio of their known to unknown parameters. The scientifically desirable model is the overidentified model. In the overidentified model the number of known elements is greater than the number of unknown elements, which means its $df > 0$ (Dimitrov, 2012). Overidentified models are viewed as scientifically desirable because they can be used to test for model fit to the data as well as to compare alternative models.

Determining the identification status of the hypothetical CFA model (underidentified, just-identified, or overidentified) required first calculating the number of known elements and degrees of freedom in the model. *Equation 2*, was used to calculate the number of known elements in the model, p^*, where p was the number of indicators in the model (Dimitrov, 2012).

$$p^* = p(p+1)/2. \tag{2}$$

The degrees of freedom, df, were then computed by subtracting the number of unknown parameters in the model from the number of known parameters. The number of unknown parameters (i.e., free parameters) was obtained using Mplus Version 6.1 (hereafter referred to as Mplus). The identification status of the model was then determined based on whether the *df* of the model were negative (underidentified), equal to zero (just-identified), or positive (overidentified).

The final task prior to testing the fit of the hypothesized CFA was to examine the multivariate normality of the observed data in the model to ensure the assumption of multivariate normality was met. The usual maximum likelihood method of estimate for confirmatory factor analysis (CFA) in the framework of structural equation modeling can only be employed when the observed data in the model (i.e., the indicators) meet the assumption of multivariate normality (Raykov & Marcoulides, 2006). If the assumption of multivariate normality is not met, the chi-square statistic and goodness-of-fit indices can be inaccurate which inflates the Type I error in chi-square based hypotheses testing (Dimitrov, 2012). The following tests were conducted with the observed data from the hypothesized CFA model to determine whether the assumption of multivariate normality was met: univariate normality, outliers, and multivariate normality.

The purpose of the first test was to examine the univariate normality of the data for each observed variable. The assumption of univariate normality was met if the data for the observed variable has a normal distribution. As stated by Raykov and Marcoulides (2006), examining the skewness and kurtosis of the observed variables is the easiest way to test for univariate normality. Therefore, the skewness and kurtosis of the data for each observed variable in the CFA model were assessed. The guidelines below were used to determine if the assumption of univariate normality was met. Skewness is an index of the asymmetry of a univariate distribution. The guidelines for evaluating skewness suggested that variables with a skewness value less than the absolute value of 2 meet the assumption of univariate normality (Dimitrov, 2012). Likewise, a kurtosis value less than the absolute value of 7 meets the assumption of univariate normality (Dimitrov, 2012). According to Raykov and Marcoulides (2006), kurtosis refers to how flat or peaked the univariate distribution is in relation to the normal distribution.

The observed data was then evaluated for the presence of outliers. Outliers are extreme or irregular data values that can cause the results of the statistical analysis to be inaccurate (Dimitrov, 2012). A box plot was created for each observed variable to examine the data for outliers. A box plot is a standardized way of displaying a distribution of data (Dawson, 2011). The data in a box plot are separated into four sections, two inner fences and two outer fences. Both mild and extreme outliers can be identified using the fences in a box plot. Mild outliers are data points that fall outside the inner fences but inside the outer fences, and extreme outliers are data points that fall outside the outer fences (Dawson, 2011). Since outliers can distort the value of the means, standard deviations, and correlation coefficients, if the box plot analyses indicate the presence of an outlier(s), it is necessary to either explain, remove, or accommodate them using robust statistical tests (Schumacker & Lomax, 2010).

Multivariate normality of the observed variables was tested following the examination of univariate normality, and outliers. The first step in testing for multivariate normality was to obtain the standardized predicted values by conducting a second simple linear regression test of the hypothesized CFA model. A scatterplot was then developed to plot the standardized residuals of the model against the standardized predicted values. The absence of a pattern in the scatterplot among the standardized residuals and predicted values was evidence supporting the assumption of multivariate normality. All confirmatory factor analysis tests thus far were conducted using SPSS.

Following the model specification, model identification, and examination of the CFA assumptions, the first research question was examined by testing the fit of the model to the data using Mplus. As illustrated by the CFA model in Figure 2, it was hypothesized that the reading comprehension strategy use construct was defined by seven items, the reading comprehension strategy instruction construct was defined by five items, and the reading attitudes construct was defined by seven items. It was also hypothesized that the three latent variables were positively correlated with the two other latent variables (i.e., strategy use was positively correlated with strategy instruction, strategy use was positively correlated with reading attitudes, and reading attitudes were positively correlated with instruction).

Model fit testing usually involves the examination of an inferential goodness-of-fit index along with other descriptive indices (Raykov & Marcoulides, 2006). Dimitrov (2012) recommended using the following five goodness-of-fit indices to jointly examine the data fit of the CFA model: chi-square fit statistic, comparative fit index (CFI), Tucker-Lewis Index (TLI), standardized root-mean-square residual (SRMR), and root mean-square error of approximation (RMSEA).

The chi-square test of model fit is a nonparametric test that compares the observed sample data with the expected probability distribution to determine if the observed values are consistent or significantly different from the expected value. In other words, it tests how well the theoretical distribution fits the empirical distribution. The chi-squares value is an inferential index that represents the *T* test statistic of the goodness-of-fit model (Raykov & Marcoulides, 2006), and a nonsignificant chi-square value is evidence of good model fit (Dimitrov, 2012). Basing the evaluation of model fit solely on the chi-square test is cautioned because it is extremely sensitive to sample size. Therefore, it was necessary to also evaluate the fit of the measurement and structural models based on the other four goodness-of-fit indices.

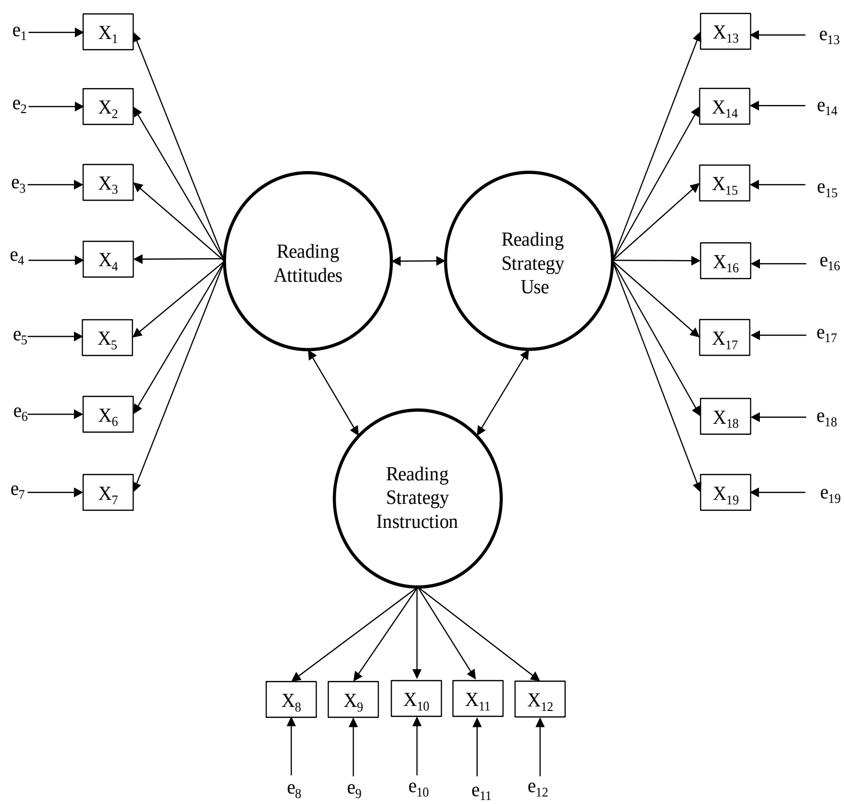


Figure 2. Hypothetical confirmatory factor analysis model.

The second goodness-of-fit index used to evaluate the CFA model fit was the comparative fit index (CFI). CFI is the ratio of improvement between the null CFA model and the hypothesized CFA model (Raykov & Marcoulides, 2006). According to Hu and Bentler (1999), a CFI greater than .95 (CFI > .95) is considered as evidence of a good model fit. The Tucker-Lewis index (TLI) was the third goodness-of-fit index used. TLI is similar to, but more stringent than, the CFI. Like the CFI, the TLI is a comparative fit index. Yet, it is more stringent than the CFI in that it corrects for model complexity imposing a penalty for any freely estimated parameters that do not largely improve the fit of the model (Dimitrov, 2012). Hu and Bentler (1999) stated that evidence of good model fit is supported by a TLI value greater than .95 (TLI > .95).

The next goodness-of-fit index used to evaluate the fit of the CFA model was the standardized root mean square residual (SRMR). The SRMR is an index that represents the difference between the observed and predicted correlation. Although a perfect fit is indicated by no difference between the observed and expected correlations (i.e., SRMS = 0), values less than .08 (SRMR < .08) are still indicative of a good model fit.

The root mean-square error of approximation (RMSEA) was the fifth, and final, goodness-of-fit index used in this investigation. The RMSEA index represents an approximate fit of the data calculated using the chi-square value, sample size, and degrees of freedom for the model. A RMSEA of .05 or less (RMSEA < .05) is evidence of a good model fit (Brown & Cudeck, 1993). Further evidence of model fit was obtained by evaluating the limits of the 90% confidence interval (CI) for the RMSEA. Good model fit was further supported when value of the lower limit was close to or equal to zero and the value of the upper limit was smaller than .08. MacCallum and Austin (2000) encouraged using RMSEA to evaluate model fit because it is sensitive to model misspecification, the guidelines for judging model quality seem accurate, and information about the precision of the fit estimation is provided by the confidence intervals.

The Initial CFA Model did not indicate a good fit between the model and the observed data. Consequently, the model results were examined for the presence of mis-specified, or non-significant, parameters to determine if any modifications could be made to improve the model fit. The mis-specified parameters were identified using the model modification indices (MIs) provided in the CFA results. The large number and high values of modification indices (MI) for this model suggested that modifying the model would certainly increase its fit to the observed data.

Dimitrov (2012) recommended that the MI with the highest value should be modified first, assuming the modification makes sense substantively. In this model, the highest MI values corresponded to the correlated error variances among the items. When item error variances are correlated, it suggests that the items measure something in common other than the latent factors represented in the model. After reviewing the item pairs, it made sense substantively to estimate their correlated error variances in to try to improve the fit of the data to the model. Stepwise procedures were used to individually estimate the correlated errors in descending order beginning with the MI with the highest value to the next highest, and so on. The goodness-of-fit indices were evaluated after each modification, and additional correlated errors were estimated until the model demonstrated good fit to the data. The model that resulted after estimating all the correlated item error variances was referred to as the Final CFA Model. Following the verification of a good model fit, the standardized residual variances in the Final CFA Model were examined. According to Raykov and Marcoulides (2006), standardized residuals represent inconsistencies between the model and the data. Typically, standardized residuals above 2.0 or below -2.0 indicate the model under- or over-explains the relationship between two variables, respectively. Further model modifications should be considered if the value of any standardized residual falls above or below this range.

After evaluating the goodness-of-fit indices and standardized residual variances, good model fit was further verified by determining whether the parameter estimates made sense both statistically and substantively. First, the statistical viability of the model was evaluated by examining the factor correlations, factor error variances, and item error variances. Statistical support for the model fit is provided by factor correlations that do not exceed an absolute value of 1, positive factor error variances, and positive item error variances (Dimitrov, 2012).

Substantive support for the model was then evaluated by examining whether the direction, magnitude, and statistical significance of the parameter estimates were consistent with the model specifications. The magnitudes of the standardized factor loadings were evaluated to identify the relative practical importance of each item in defining its construct. Finally, the squared values of the standardized factor loadings (R^2) were calculated to determine the proportion of variance in each item that was accounted for by its respective factor.

The confirmation of good fit of the Final CFA Model was followed by an examination of the Pearson internal consistency reliability estimates. In the preliminary data analyses, Cronbach's alpha test of internal consistency reliability was conducted to assess the reliability of the three latent factors (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes). Cronbach's alpha is based on the essentially tau-equivalent measurement model, which requires that certain assumptions are met to ensure the accuracy of the reliability estimate (Graham, 2006). The correlated errors revealed by the CFA analysis are a violation of the assumptions of essentially tau-equivalent measures. Therefore, to ensure the accuracy of the reliability estimates, the estimation of internal consistency reliability was reanalyzed for each factor using the latent variable modeling (LVM) approach. The Pearson internal consistency reliability estimates (ρ_{xx}), standard error of measurement (*SEM*) estimates, and their 90%, 95%, and 99% CIs were obtained for each factor using the bootstrap procedure in Mplus.

SEM structural model – Structural regression and MIMIC analyses. Once the relationships between the latent constructs and their indicators were defined, the SEM structural models were tested. According to Muthén and Muthén (2010), structural models have multiple uses. They can be used to describe the relationships among latent constructs, among observed variables, and/or between latent constructs and observed variables that are not factors indicators (Muthén & Muthén, 2010).

After the measurement model was assessed, the structural models were tested using structural regression and MIMIC analyses. The structural regression model was used to explore the direct and indirect effects of the hypothesized relationships among the constructs, and between the constructs and observed variables that were not the factor indicators. The MIMIC model was used to examine mean gender, minority status, SES, class time, and class size group differences in reading comprehension strategy use, reading strategy instruction, reading attitudes, and reading comprehension achievement.

Structural regression analysis. The second research question, 'Are there direct and indirect effects among the five student and school level variables, the three latent variables, and high school students' reading comprehension achievement?', was explored following the confirmation of a good model fit of the CFA measurement model. An analysis of the hypothesized direct and indirect effects of the independent variables on the dependent variables in the structural equation model was conducted using structural regression procedures in Mplus. The following four procedures, described below, were used to conduct the structural regression analysis: model specification, model identification, testing for model fit, and evaluating the parameter estimates in the model (i.e., path coefficients/partial regression coefficients of the direct and indirect effects).

Structural regression analysis was used to examine the explanatory relationships among latent variables, and the explanatory relationships between latent constructs and observed variables. Structural regression models are similar to both CFA models and path diagrams, yet they have distinctly important differences. Like CFA models, structural regression models contain latent constructs measured by observable variables. However, while CFA provides information regarding the correlation between the latent constructs, it does not offer information about the specific directional relationships among the constructs. Structural regression analysis differs from CFA in that the specific directional relationships among the constructs are tested by regressing the latent constructs on the other latent constructs (Raykov & Marcoulides, 2006). The main similarity between structural regression analysis and path analysis are the paths used in their models to signify the explanatory (direct and indirect) relationships among the variables. The main difference between the two types of analysis is the types of variables that can be included in their models. Raykov and Marcoulides (2006) explained that while path analysis is strictly intended to examine the relationships between observed variables, structural regression analysis can be thought of as an extension of path analysis in that it also allows for the examination of relationships between latent variables.

In addition to providing information concerning the specific directional relationships among the constructs, structural regression analysis also offers information regarding indirect effects in the causal relationships among the variables. This information is obtained by decomposing the sources of the correlations between the independent and dependent variable into the direct effects, indirect effects (total indirect and specific indirect effects), and total effects. The purpose of decomposing these effects is to discover precisely what is correlated with, or predicts, the dependent variable.

According to Raykov and Marcoulides (2006), direct effects are the effects of an independent variable on a dependent variable that are unmediated by any other variable in the model. Whereas, an indirect effect is the effect of an independent variable on a dependent variable that is mediated by at least one intervening variable. When more than two variables are included in a statistical test, it is important to measure the indirect effects in addition to the direct effects between the two main variables of interest. It is important because if indirect effects exist among the variables but are not measured, the resulting relationship between two variables of interest will not be fully explained.

Information regarding the total and specific indirect effects among the variables is also provided because indirect effects can be mediated by more than one intervening variable. Total indirect effects are all the paths from the independent variable to the dependent variable that are mediated by one additional variable, and a specific indirect effect is one single effect of the independent variable on the dependent variable that is mediated by a third variable (Bollen, 1987). Finally, the total effect of an independent variable on a dependent variable is the combination of all the direct and indirect effects.

The hypothetical structural regression model for this study, illustrated in Figure 3, was designed to examine the three subquestions that comprised the second research question. In order to address the first aspect of the second research question (also referred to as research subquestion 2a), 'What are the direct and indirect effects of the three latent constructs and the five student (i.e., gender, minority status, and SES) and school (i.e., class time and class size) level variables on high school students' reading comprehension achievement?', the structural regression model was used to examine the predictive power of the student and school factors as well as the three latent variables on high school students' reading comprehension achievement by simultaneously regressing reading comprehension achievement on all eight predictor variables.

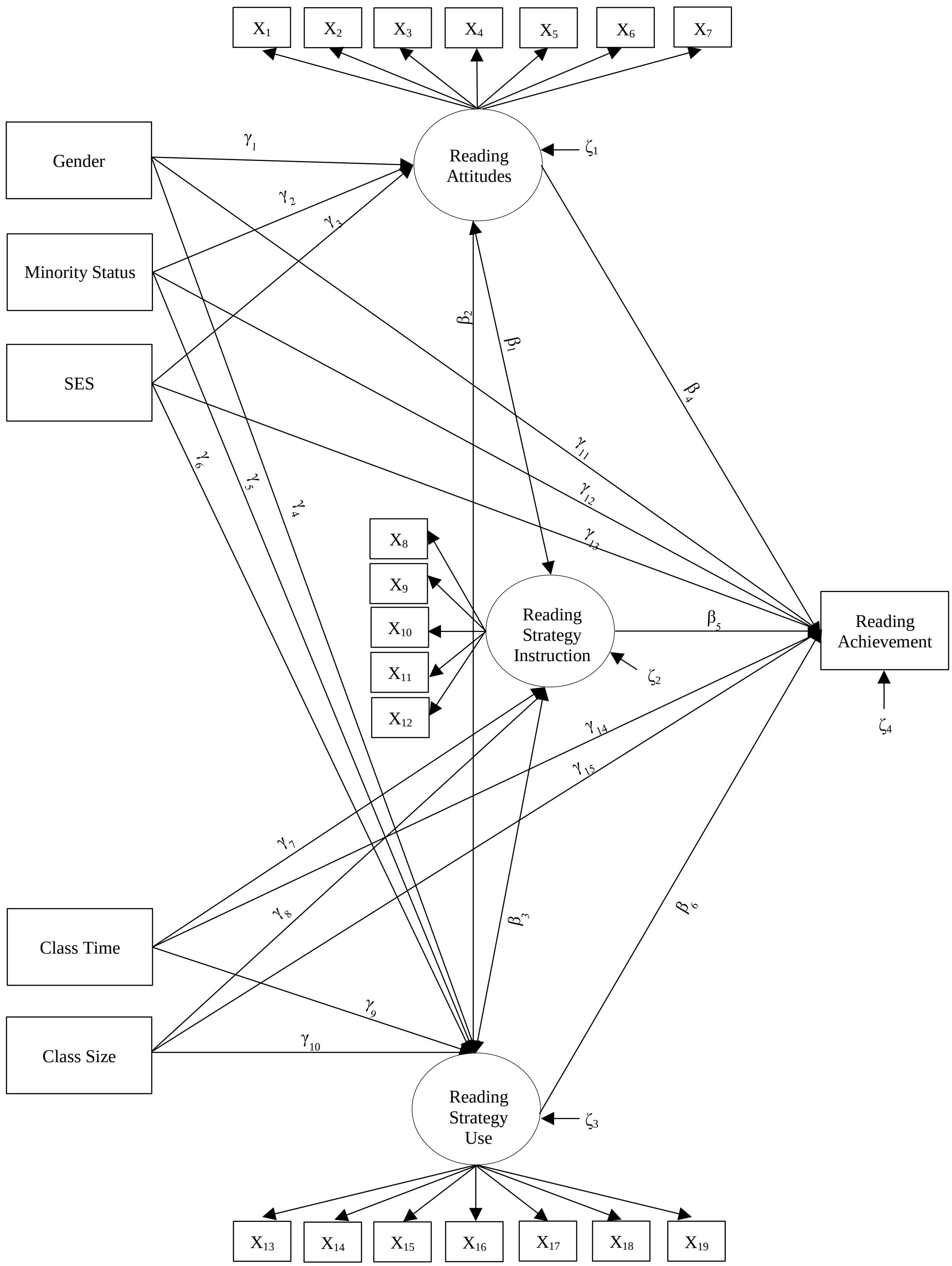


Figure 3. Hypothetical structural regression model

To explore the next aspect of the second research question (also referred to as research subquestion 2b), 'What are the direct effects among the three latent constructs (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes)?', the structural regression model was employed to analyze whether any of the latent variables were predicted by either of the other two latent variables by simultaneously regressing them on each other.

Finally, to investigate the last component of the second research question (also referred to as research subquestion 2c), 'What are the direct effects of the five student and school level variables on the three latent constructs?', the structural regression model was utilized to determine the predictive power of the student and school level variables on the latent constructs by simultaneously regressing the latent variables on the five control variables. Specification of the hypothetical model was the first step in this multi-step analysis process. Specifying the hypothetical structural regression model required identifying the expected effects of the relationships among all the variables in the model. The specifications of this model were based on a combination of the theoretical and empirical evidence presented in the literature review.

After specifying the hypothetical model, the identification status of the model needed to be determined. The number of known elements and degrees of freedom in the model needed to be calculated to determine the identification status of the hypothetical model. *Equation 2* was used to calculate the number of known elements in the model (Dimitrov, 2012). Then the degrees of freedom, *df*, were determined by subtracting the unknown parameters from the known parameters. The number of unknown parameters (i.e., free parameters) was obtained using Mplus. Finally, the identification status of the model was determined based on whether the *df* of the model were negative (underidentified), equal to zero (just-identified), or positive (overidentified).

Following the model specification and identification, and prior to evaluating the direct and indirect effects in the model, its data fit was examined using the guidelines for assessing the CFA model fit: $\chi^2 - p > .05$, CFI > .95, TLI > .95, SRMR < .08, RMSEA ≤ .05 (90% CI with a LL close to or including zero and an UL < .08), and standardized residuals (-2 < SR < +2). Once the goodness-of-fit of the structural regression model was confirmed, the standardized partial regression coefficients (i.e., path coefficients) for the direct and indirect effects were examined.

Path coefficients reflect the predictive power of the independent variable on the dependent variable. The magnitude of the statistically significant standardized estimates of the direct and indirect path coefficients were interpreted as the relative importance of the independent variable in predicting the amount of variation in the dependent variable (Raykov & Marcoulides, 2006). In path analysis, where all variables are observed variables, the positive and negative signs of the path coefficients can be used to interpret group differences. However, in structural regression analysis, where models include both observed and latent variables, subsequent analysis (i.e., MIMIC modeling analysis) is needed to meaningfully interpret group differences. The MIMIC group mean analysis procedures are described in following subsection of this chapter.

Additional analyses utilizing the structural regression results were then conducted separately on each of the four dependent variables in this model to determine the amount of variation in the dependent variable that was explained by the combination of all the independent variables together. In other words, analysis was conducted to identify the percentage of student differences in reading comprehension strategy use, reading comprehension strategy instruction, and reading comprehension achievement that was jointly explained by all their respective hypothesized predictors.

The first step in this analysis was to calculate the R^2 value (i.e., coefficient of multiple determination) for each dependent variable in the model (i.e., reading comprehension strategy use, reading comprehension strategy instruction, reading attitudes, and reading comprehension achievement). *Equation 3*, $R^2 = 1 – \mathrm{VAR}(\zeta)$, where $\mathrm{VAR}(\zeta)$ represents the residual variance in a dependent variable, was used to calculate the R^2 values (Dimitrov, 2012). The standardized residual variances of the four dependent variables needed to calculate each R^2 value were obtained from the Mplus results. The residual in the dependent variable, ζ, was the portion of the dependent variable that was not explained by the independent variable(s). Hence, the residual variance in the dependent variable was the variance of the portion of the dependent variable that was not explained by the independent variable(s). The residuals of the four dependent variables in the structural regression model are illustrated in Figure 3. Ultimately, after the four R^2 values were calculated, they were then converted to percentages. These percentages represented the amount of reading comprehension achievement, reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes that were explained by their hypothesized predictors.

Next, an F test to assess the statistical significance of the R^2 value for each dependent variable was conducted. A statistically significant F-statistic provides evidence that the variance in the dependent variable explained by the predictors is not equal to zero for the population (Dimitrov, 2009). *Equation 4*, was used to calculate the F-statistic: $F = (R^2/p)/[(1 - R^2)/(n - p - 1)]$, where p is the number of independent predictors and n is the sample size (Schumacker & Lomax, 2010). An F-statistic is statistically significant when its value is larger than the F-critical value for df_1, df_2. A table of the critical values of the F-distribution ($\alpha = .05$) was used to determine the F-critical value for df_1, df_2 (Schumacker & Lomax, 2010, Table A-5). Finally, the F-statistic was compared to the F-critical value to judge its statistical significance.

It is recommended that along with reporting the p-values of statistically significant results, researchers should also report the effect size (i.e., practical significance) of the results (Cohen, 1992). An effect size is the magnitude of the statistical significance of a result that meaningfully communicates the size, or strength, of the difference between groups (Cohen, 1992). Reporting both the statistical significance and the practical significance (i.e., effect size) is important because it allows for a more accurate interpretation of the study's findings. Although a statistically significant result indicates that a difference due to chance does exist, the p-value does not convey the size of the difference. Only reporting the p-value of a statistically significant result can be problematic in studies with a large sample size, such as this study. When the sample size is extremely large, the results of statistical tests are frequently significant (Sullivan & Feinn, 2012). However, these statistically significant differences typically lack practical significance. Without knowledge of the effect size, the practical meaning of a statistically significant result is easily misinterpreted as meaningful.

There are a variety of indexes used to measure effect size. According to Cohen (1992), there is a different effect size index for every type of statistical test. The effect size index, f^2, is the index associated with the F test used to determine the statistical significance of the squared multiple correlations (R^2). Thus, the effect size of the significant R^2 values of the four dependent variables in the structural regression model were calculated using *Equation 5*, $\text{ES}(f^2) = R^2 - [p/(N\text{-}1)]$, where p is the number of independent predictors of the dependent variable, and N is the total sample size (Schumacker & Lomax, 2010). The following guidelines, recommended by Cohen (1992), were used to interpret the magnitude of the f^2 effect sizes: small = .02, medium/moderate = .15, and large = .35.

Multiple Indicator, Multiple Cause (MIMIC) group code modeling. The third research question in this study was, ‘Are there significant group differences in high school students’ reading comprehension achievement, reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes across the five student (i.e., gender, minority status, and SES) and school (i.e., class time and class size) level variables?’. In order to address this question, it was necessary to test several hypotheses regarding mean differences of the gender, minority status, SES, class time, and class size groups in reading comprehension achievement, reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes. Multiple

Indicator, Multiple Cause (MIMIC) group code modeling procedures were employed because they allow for the testing of group mean differences.

MIMIC analysis is a special case of structural equation modeling that consists of latent and outcome variables predicted by observed variables (Schumacker & Lomax, 2010). The purpose of the MIMIC modeling approach was to further examine the direct effects of the observed variables on the latent and outcome variables established through the structural regression analysis, by comparing group mean differences. When comparing group mean differences using MIMIC modeling, the data from both groups were not separated, thus, the validity of the comparisons depended on the assumption that the construct being measured meant the same thing in both groups (MacCallum & Austin, 2000). This assumption is referred to as factorial invariance, or the invariance of factor loadings, item uniquenesses, and factor variances/covariances (Dimitrov, 2012). Thus, factorial invariance had to be tested and confirmed prior to conducting MIMIC analysis test of group differences. The following procedures, described below, were used to conduct the MIMIC analysis: specifying and identifying the MIMIC model, testing for factorial invariance, testing for model fit, and evaluating the model parameters.

The initial step of the MIMIC analysis was to specify the hypothetical MIMIC model. MIMIC model specifications consisted of hypothesizing which control variables were expected to demonstrate mean group differences on the three latent factors and on the reading achievement outcome variable. Any expected correlations in the model also needed to be specified. The specifications for the MIMIC model were supported by the theoretical and empirical research discussed in the review of the literature.

The identification status the hypothetical MIMIC model needed to be determined once it was specified. The first step to determine the identification status of the model was to calculate the number of known elements in the model using *Equation 2* (Dimitrov, 2012). Next, the degrees of freedom, *df*, were calculated by subtracting the number of unknown parameters from the number of known parameters. The number of unknown parameters (i.e., free parameters) was obtained using Mplus. Last, the identification status of the model was determined based on whether the *df* of the model were negative (underidentified), equal to zero (just-identified), or positive (overidentified).

After the MIMIC model was specified and identified it needed to be tested for factorial invariance. Factorial invariance is defined as the invariance of a factor model across different groups including the invariance of factor loadings, item intercepts, item residual variances/covariances, and factor variances/covariances (Dimitrov, 2012). Factorial invariance indicates that the construct(s) have the same meaning for each group (i.e., the relationships specified in the CFA model, between the latent factors and their indicators, are equivalent across groups). The combined testing of configural invariance, measurement invariance, and structural invariance was conducted using Mplus to obtain evidence that confirms the assumption of factorial invariance was met.

The goal of the third research question was to examine group differences across five population groups (gender, minority status, SES, class time, and class size). Thus, the following procedures for testing configural, measurement, and structural invariance were performed to assess the factorial invariance of each population group.

The first step in factorial invariance testing is to test for configural invariance because configural invariance is a prerequisite for measurement and structural invariance (Wu, Li, & Zumbo, 2007). Configural, or form, invariance is the invariance of the factor model specification (i.e., the pattern of free and fixed model parameters) across groups (Dimitrov, 2010). The purpose of testing for configural invariance is to demonstrate that the model specifications do not vary across the population groups.

Testing for configural invariance first required the identification of a baseline model. According to Byrne (2004), the baseline model is the model that best fits the data, while maintaining parsimony and substantive meaning. The fit of the baseline model to the data is estimated separately with each group separately. Configural invariance of the baseline model is supported when the goodness-of-fit indices for each group separately indicate very good model fit. A lack of configural invariance means that different constructs were measured across groups (Wu et al., 2007) therefore it would not be appropriate to test for group differences using the MIMIC analysis.

The three-factor Final CFA Model, with 19 items, was used as the baseline model to test for configural invariance. Data fit of the baseline model was estimated twice for each of the five population groups using Mplus. In other words, for each of the population groups, the fit of the baseline model was tested once with each portion of the sample (i.e., males and females, minority and nonminority, high SES and low SES, more class time and less class time, large class size and small class size). The goodness-of-fit indices of the three models (both sample groups together and each sample group separate) were examined for good model fit. After configural invariance across the population groups was confirmed, it was appropriate to proceed with the testing for measurement and structural invariance.

The purpose of examining measurement and structural invariance was to determine that the causal relationships between variables worked the same across the two groups for each observed variable. A model is considered to demonstrate measurement invariance when the scores of its' latent constructs hold the same meaning across groups (Dimitrov, 2006). Establishing measurement invariance prior to analyzing mean group differences was necessary, otherwise analyzing and interpreting mean differences as genuine group differences would have been misleading. Structural invariance is defined as the invariance of factor variances/covariances (Dimitrov, 2012). In their discussion of factorial invariance, Wu et al. (2007) stated that structural invariance is a useful, but not necessary, condition for comparing group mean differences. They explained that the structural model defines the causal relationships among the latent variables, thus structural invariance (the equality of factor variances/covariances) is not necessary to define the relationships between the items and the factors (Wu et al., 2007). However, they pointed out that support for structural invariance may suggest that the two groups belong to the same population. Therefore, as potentially additional support for factorial invariance, the structural invariance of all population groups was tested.

The step-up constraints method was used in this investigation to conduct the tests of measurement and structural invariance. Under this approach to testing factorial invariance, measurement invariance was examined at three increasingly constrained levels including: weak measurement invariance (i.e., invariant factor loadings), strong measurement invariance (i.e., invariant intercepts), and strict measurement invariance (i.e., invariant item variances/covariances). Finally, Dimitrov (2012) stated, once measurement invariance is established, the final step in testing for factorial invariance is to test for structural invariance (i.e., invariant factor variances/covariances).

The step-up approach to testing for factorial invariance is based on a chi-square difference test ($\Delta\chi^2$) between a constrained model that is nested within an unconstrained model ($\Delta\chi^2 = \chi^2_{\text{constrained}} - \chi^2_{\text{unconstrained}}$). In this study, the chi-square difference ($\Delta\chi^2$) is the difference of the usual maximum-likelihood chi-square (ML χ^2) values of the two nested models because, as previously discussed, the assumption of multivariate normality for each observed variable was met (Dimitrov, 2012).

The first step in determining the invariance of the nested models was to calculate the chi-square difference ($\Delta\chi^2 = \chi^2_{\text{constrained}} - \chi^2_{\text{unconstrained}}$) and the change in degrees of freedom ($\Delta df = df_{\text{constrained}} - df_{\text{unconstrained}}$) between the constrained and unconstrained models by hand. Next, a table of the critical values of the chi-square distribution (α = .05) was used to determine the critical value of the change in degrees of freedom between the constrained and unconstrained models (Dimitrov, 2009, Table A-3). Finally, the statistical significance of the chi-square difference was compared to the critical value of the chi-square distribution to determine the change in degrees of freedom between the constrained and unconstrained models. A $\Delta\chi^2$ is statistically significant when it is larger than the critical value for the Δdf. In terms of invariance testing, a non-significant chi-square difference is considered evidence of invariance. Therefore, invariance of the nested models was supported when the $\Delta\chi^2$ was smaller than the critical value of the Δdf.

In instances when the $\Delta\chi^2$ of the nested models was statistically significant, signaling a lack of invariance, partial invariance of the models was tested. Partial invariance testing involves removing the restrictions on some of the parameters, identified through an examination of the model modification indices, while continuing to hold the others invariant. When full invariance is not possible, if at least one parameter (aside from the marker indicator) remains invariant, evidence of partial invariance is sufficient evidence to continue with group comparisons (Milfont & Fischer, 2010).

While the $\Delta\chi^2$ test was used as the main determinant of measurement and structural invariance, the root mean square error of approximation (RMSEA) was used as additional support for good model fit. The change in the comparative fit index (ΔCFI) between constrained and unconstrained models ($\Delta CFI = CFI_{constrained} - CFI_{unconstrained}$) was used as additional evidence of invariance (Dimitrov, 2012). Due to the sensitivity of the chi-square statistic to sample size, the ΔCFI was also used to interpret invariance because it provides stable and dependable evidence that is not sensitive to sample size. Under the assumption of multivariate normality, Cheung and Rensvold's (2002) guidelines were used to determine invariance based on the ΔCFI. Their guidelines state that invariance is supported when the value of the decrease in the comparative fit index between the constrained and unconstrained models is smaller than or equal to -.01 ($\Delta CFI \leq -.01$), equal to zero (indicating no change), or positive (Cheung & Rensvold, 2002).

A study conducted by Vandenberg and Lance (2000) contains a good example of the proper use of the guidelines to interpret the ΔCFI as a supplemental determinant of invariance. Although the guidelines they used were obtained from a 1999 conference publication of Cheung and Rensvold, they are the same those previously described. The purpose of their research was to discuss the importance of determining measurement invariance across groups prior to testing population group differences. Vandenberg & Lance (2000) conducted and presented a sample data analysis as part of their study in order to illustrate this importance, and to subsequently recommend procedures for carrying out tests of measurement equivalence. They used existing data from a new employee survey to test the measurement invariance in order to determine whether it would be appropriate to analyze and interpret differences among the employee groups. The data they used corresponded to three organizational commitment constructs (i.e., compliance, internalization, and identification commitment).

Vandenberg & Lance's (2000) explanation of the results of the sample data analysis provided a clear illustration of how Cheung and Rensvold's guidelines are used to interpret the ΔCFI, specifically in the section of the results regarding the internalization commitment construct. After discussing the poor fit to the data of Model 0 and the acceptable fit of Model 1, Vandenberg & Lance (2000) used the ΔCFI values as additional evidence to support to the $\Delta\chi^2$ significance test of the nested models in their evaluation of measurement invariance (i.e., metric and scalar invariance). Metric invariance was confirmed by the results of the comparison of Model 1 and Model 2, $\Delta\chi^2(8) = 5.33$, $p > .01$; ΔCFI = .00, as evidenced by the nonsignificant chi-square difference and the change in CFI equal to zero. However, based on the results of the comparison between Models 2 and 3, $\Delta\chi^2(8) = 41.65$, $p < .01$; ΔCFI = –.04, scalar invariance was not supported, as evidenced by the statistically significant chi-square difference and the change in CFI larger than -.01 (Vandenberg & Lance, 2000).

Testing for measurement invariance began by testing the least constrained model, Model 0 (with no invariance assumed). The constraints were gradually increased in each subsequent model; Model 1 (weak measurement invariance/metric invariance), Model 2 (strong measurement invariance/scalar invariance), and Model 3 (strict measurement invariance/invariance of item uniquenesses). The first level of testing for measurement invariance, weak measurement invariance, examined the metric invariance of the model. Metric invariance of the model is supported by equal factor loadings across the population groups. Two models (Model 0 and Model 1) were created to test for metric invariance. Model 0 was the baseline model where all parameters were freely estimated. The constraints in Model 1 were stepped up from Model 0 by fixing all factor loadings to be invariant across the groups. Model 1 was therefore nested in Model 0. A nonsignificant chi-square difference between the two models, $\Delta\chi^2_{M1\text{-}M0}$, would indicate weak measurement invariance or invariant of factor loadings across the two groups.

The next level of measurement invariance testing was testing for strong measurement invariance. In addition to metric invariance, scalar invariance was also required as support for strong measurement invariance. Scalar invariance is the equality of item intercepts, or item means, across the population groups (Dimitrov, 2012). Model 2 was developed to test for strong measurement invariance by stepping up the constraints of Model 1. In addition to the existing constraint of equal factor loadings across the two groups, in Model 2 the item intercepts were also constrained to be invariant across groups. Strong measurement invariance, the invariance of factor loadings and item intercepts, was supported by a nonsignificant chi-square difference between Model 2 and Model 1, $\Delta\chi^2_{M2\text{-}M1}$. Once strong measurement invariance was confirmed, the comparison of factor means across groups was permissible because scalar invariance indicates that the origin of the scale does not differ across groups (Dimitrov, 2012).

Strict measurement invariance was the third, and final, level of testing for measurement invariance. Strict measurement invariance is the most constrained type of invariance requiring a combination of metric invariance, scalar invariance, and invariance of item uniquenesses. Invariance of item uniquenesses is defined as the equivalence of residual item variances/covariances across groups (Dimitrov, 2012). To test for strict measurement invariance Model 3 was stepped up by imposing the additional constraint of equal residual item variances/covariances to Model 2, while continuing to hold factor loadings and item intercepts invariant across groups. A nonsignificant chi-square difference between Model 3 and Model 2, $\Delta\chi^2_{M3\text{-}M2}$, offered evidence to support strict measurement invariance. Strict measurement invariance indicates that the items in both groups were measured with the same amount of precision. Thus, evidence of strict measurement invariance is used to support the assumption that group differences on any item are only attributable to group differences on the common factors (Dimitrov, 2010).

Once the tests of measurement invariance (which confirmed the weak, strong, and strict measurement invariance of the model) were finished, the test for structural invariance was conducted. This test required creating Model 4, the structural model, by stepping up the constraints of Model 2. In addition to constraining factor loadings and item intercepts, all factor variances/covariance were constrained to be equal across the two groups. As Model 4 was nested in Model 2, the nonsignificant chi-square difference between the two models, $\Delta\chi^2_{M4\text{-}M2}$, provided evidence of structural invariance.

After the tests of configural, measurement, and structural invariance confirmed the assumptions of completely (or partially) invariant factor loadings, invariant item intercepts, invariant item residual variances/covariances, and factor variances/covariances across the population groups were met, it was permissible to examine the mean group differences using MIMIC analysis. MIMIC analysis procedures were used to test the hypothetical model, illustrated in Figure 4, because they are a special case of structural equation modeling that allows for the testing of group mean differences.

First, a joint examination of the fit indices (χ^2, CFI, TLI, SRMR, and RMSEA,) and the standardized residuals was conducted using the guidelines for assessing the quality of the model fit to the data (χ^2 - p > .05, CFI > .95, TLI > .95, SRMR < .08, RMSEA ≤ .05 (90% CI LL close to or including zero and UL < .08), and standardized residuals (-2 < SR < +2). After a good model fit was established, the significance and the sign (+/-) of the structural regression coefficients were examined. A statistically significant structural regression coefficient indicates population heterogeneity, or group differences (Dimitrov, 2012). Specifically, a significant positive structural coefficient indicates that the group coded 1 scored higher on the construct compared to the reference group, coded 0. Conversely, a significant negative structural coefficient indicates that the group coded 1 scored lower on the construct compared to the reference group, coded 0.

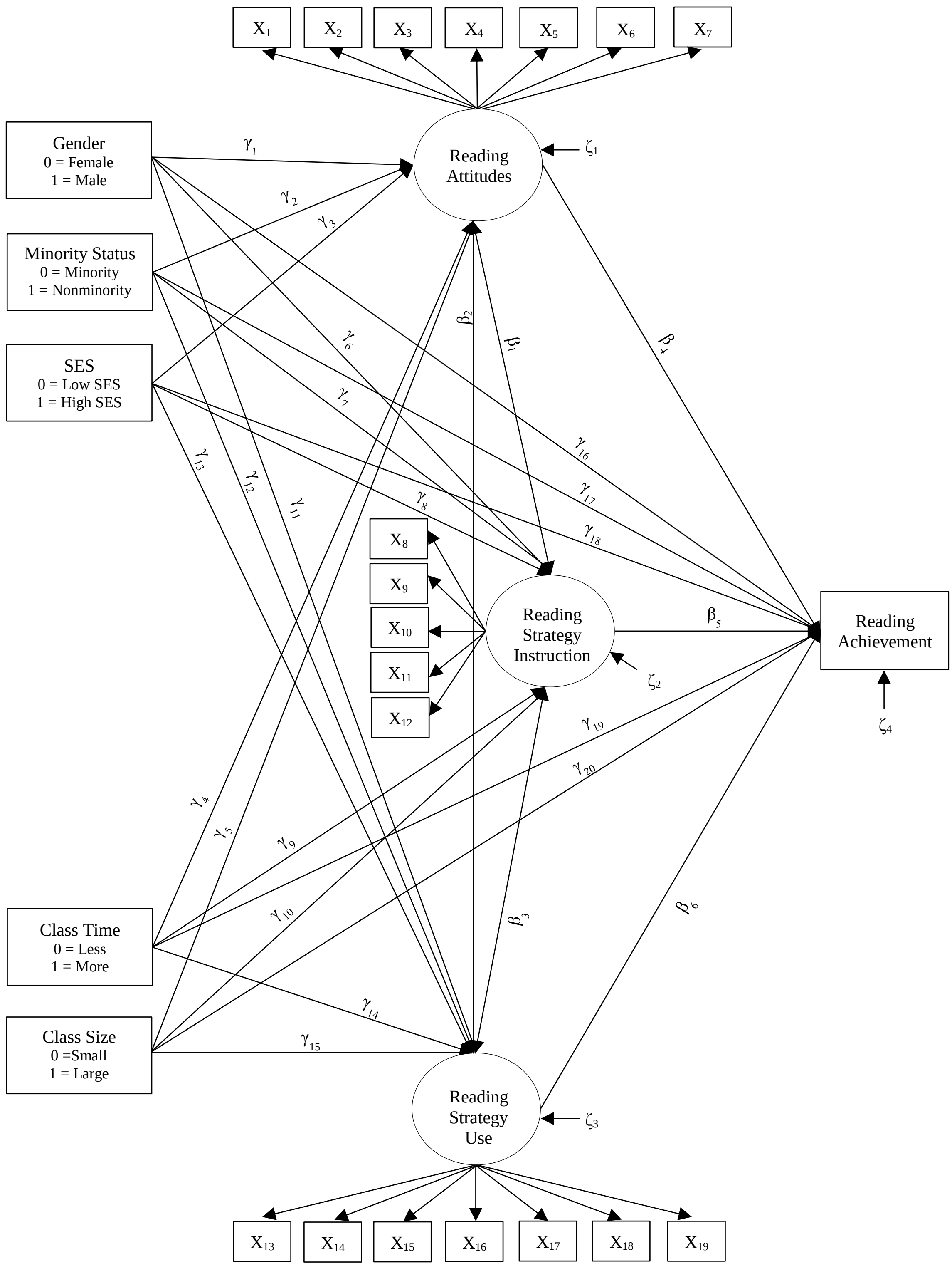


Figure 4. Hypothetical MIMIC model.

As previously mentioned, there are different effect size indexes for every statistical test. According to Cohen (1992), the *d* effect size index is used with the test of group mean differences. *Equation 6*, $\hat{d} = |\hat{\gamma}|/\sqrt{VAR(\zeta)}$, where $\hat{\gamma}$ was the unstandardized estimate of the structural coefficient and $VAR(\zeta)$ was the unstandardized factor residual variance, was used to calculate the effect sizes of the statistically significant group differences (Dimitrov, 2012). Cohen's (1992) guidelines for the *d* effect size index (small = .20, medium/moderate = .50, and large = .80) were used to interpret the magnitudes of the group mean difference effect sizes.

Chapter Four

In this chapter, the results of the data analyses employed to investigate the three research questions in this study are presented. The results are organized by the statistical techniques used to analyze the data including the preliminary data analyses, confirmatory factor analysis (CFA), reliability analysis, structural regression analysis, and MIMIC analysis. Internal consistency reliability of the three latent factors (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes) was analyzed prior to exploring the three research questions.

The first research question, 'Do the observable indicators selected to measure the three latent constructs appropriately define the three factors (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes)?', was examined using CFA. Next, structural regression analysis was used to address the second research question, 'Are there direct and indirect effects among the five student and school level variables, the three latent variables, and reading comprehension achievement?'. Finally, MIMIC analysis was used to answer the third research question, 'Are there significant group differences in high school students' reading comprehension achievement, reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes across the five student level (i.e., gender, minority status, and SES) and school level (i.e., class time and class size) variables?'.

Preliminary Analyses

The preliminary analyses included an analysis of descriptive statistics obtained to describe the composition of the five control variables. It also included an internal consistency reliability analysis of the three latent constructs (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes).

Descriptive statistics. The descriptive statistics obtained to describe the composition of the student and school level control variables are presented in Table 3. The frequencies and percentages of the two gender, minority status, SES, class time, and class size groups were included.

Table 3

Frequencies and Percentages of the Control Variables

Variable	Frequency	Percentage
Student Level Characteristics		
Gender		
Female	2,546	48.7
Male	2,687	51.3
Minority Status		
Minority	2,337	44.7
Non-Minority	2,896	55.3
Socio-Economic Status (SES)		
Low SES	2,601	49.7
High SES	2,632	50.3
School Level Characteristics		
Class Time		
Less Class Time	2,841	54.3
More Class Time	2,392	45.7
Class Size		
Small Class Size	2,469	47.2
Large Class Size	2,764	52.8

Likewise, the descriptive statistics obtained to describe the distributions of the latent variable indicators are presented in Table 4. The table includes the means, standard deviations, and variances of the 19 individual items hypothesized to define the latent constructs. The item means ranged from $M = 2.11$ to $M = 3.21$, the standard deviations ranged from $s = 0.81$ to $s = 1.07$, and the item variances ranged from $s^2 = 0.65$ to $s^2 = 1.15$.

Table 4

Distributions for the 19 Indicators of the Three Reading Constructs

Reading Constructs	Indicators	Mean	Standard Deviation	Variance
Reading Attitudes				
Item 1	X1 – Have To	2.51	0.95	0.91
Item 2	X2 – Favorite Hobby	2.11	0.93	0.86
Item 3	X3 – Finish Books	2.86	0.88	0.78
Item 4	X4 – Book as a Present	2.16	0.92	0.85
Item 5	X5 – Waste of Time	2.97	0.94	0.89
Item 6	X6 – Bookstore/Library	2.51	0.97	0.95
Item 7	X7 – Read for Information	2.56	0.89	0.80
Reading Strategy Instruction				
Item 1	X8 – Asks Questions	2.97	0.81	0.65
Item 2	X9 – Wait Time	2.91	0.83	0.69
Item 3	X10 – Books/Authors	2.39	0.96	0.91
Item 4	X11 – Encourages Opinions	2.85	0.89	0.79
Item 5	X12 – Prior Knowledge	2.73	0.89	0.79
Reading Strategy Use				
Item 1	X13 – Use Graphs	3.21	0.93	0.87
Item 2	X14 – Cause and Effect	3.05	0.96	0.92
Item 3	X15 – Character Actions	3.07	0.99	0.97
Item 4	X16 – Author Background	2.34	1.07	1.15
Item 5	X17 – Explain Text Purpose	3.06	0.99	0.98
Item 6	X18 – Historical Context	2.63	1.04	1.08
Item 7	X19 – Connect Text Parts	2.72	1.06	1.12

Note. $N = 5{,}233$.

Reliability. The internal consistency reliability of the three latent constructs was tested using SPSS. The internal consistency reliability estimates (i.e., Cronbach's Alpha coefficient), total item correlations (i.e., Corrected Item Values), and Cronbach's Alpha if Deleted for each of the three hypothesized factors are presented in Table 5.

Table 5

Internal Consistency Reliability Estimates for the Three Reading Constructs

Construct/Indicator	Reliability Estimates (α)	Mean	Corrected Item Total Correlation	Cronbach's Alpha if Item Deleted
Reading Attitudes	0.89			
X1 – Have To		2.51	.76	0.87
X2 – Favorite Hobby		2.11	.75	0.87
X3 – Finish Books		2.86	.50	0.90
X4 – Book as a Present		2.16	.70	0.87
X5 – Waste of Time		2.97	.72	0.87
X6 – Bookstore/Library		2.51	.71	0.87
X7 – Read for Information		2.56	.67	0.88
Reading Strategy Instruction	0.83			
X8 – Asks Questions		2.97	.61	0.80
X9 – Wait Time		2.91	.61	0.80
X10 – Books/Authors		2.39	.56	0.82
X11 – Encourages Opinions		2.85	.70	0.77
X12 – Prior Knowledge		2.73	.66	0.79
Reading Strategy Use	0.76			
X13 – Use Graphs		3.21	.33	0.76
X14 – Cause and Effect		3.05	.55	0.71
X15 – Character Actions		3.07	.56	0.71
X16 – Author Background		2.34	.43	0.74
X17 – Explain Text Purpose		3.06	.55	0.71
X18 – Historical Context		2.63	.46	0.73
X19 – Connect Text Parts		2.72	.46	0.73

Note. α = Cronbach's coefficient alpha internal consistency reliability estimate.

The internal consistency test of reliability for the first factor, reading comprehension strategy use, $\alpha = 0.76$, was slightly lower than the other two factors, but it still met the established social sciences cutoff value of $\alpha = 0.70$. The internal consistency reliability estimate of the reading comprehension strategy instruction factor, $\alpha = 0.83$, indicated that the data of this scale demonstrated a high level of reliability. No items on the reading comprehension strategy use or the reading comprehension strategy instruction scales were flagged by the Cronbach's Alpha if Deleted test, thus, all items were retained.

The internal consistency reliability estimate for the reading attitudes scale, signified that the data demonstrated a high level of reliability, $\alpha = 0.89$. According to the Cronbach's Alpha if Item Deleted test, the Finish Books (Cronbach's α if Item Deleted = 0.896) item was flagged for deletion. This meant that the unrounded Cronbach's Alpha reliability estimate for this scale, $\alpha = 0.890$, would increase by .06 if this item was removed from the scale. Since the reliability of the scale was already sufficiently high, the substantive value of retaining the item in the scale was more important than the amount of increased reliability that would be gained if it was removed. Consequently, this item was not removed from the reading attitudes scale.

The corrected item total correlations did not suggest the need to remove any items because the correlation of all 19 items with their respective scales was greater than the recommended cutoff value, $r \leq .30$ (Nunnally & Bernstein, 1994). Overall, the internal consistency reliability for each of the three scales is relatively high (reading strategy use, $\alpha = 0.76$, reading comprehension strategy instruction, $\alpha = 0.83$, and reading attitudes, $\alpha = 0.89$), thereby indicating the stability of the three latent reading constructs.

Structural Equation Modeling

The results of the structural equation modeling (SEM) analysis are presented in the following sections, including the results of the CFA used to examine the measurement model and the results of the structural regression analysis and MIMIC analysis used to examine the structural model.

SEM measurement model – Confirmatory factor analysis. The first research question, ‘Do the observable indicators selected to measure the three latent constructs appropriately define the three factors (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes)?’, was examined using confirmatory factor analysis (CFA). The results of the following CFA procedures are presented in the sections below: model specification, model identification, examination of the CFA assumptions, test of model fit, estimation of model parameters, and latent variable modeling analysis of internal consistency reliability.

CFA model specification. The initial procedure, under the SEM framework, was to specify a hypothetical CFA model (Dimitrov, 2012). The hypothetical CFA model specified in this study consisted of three latent constructs measured by 19 observable indicators from the PISA 2009 U.S. Student Questionnaire and the PISA 2009 Reading for School Questionnaire. A description of each item is presented in Table 6. It was hypothesized that the reading attitudes construct was defined by Items 1-7, the reading comprehension strategy instruction construct was defined by Items 8-12, and the reading strategy use construct was defined by Items 13-19. It was also hypothesized that each latent variable was positively correlated with the two other latent variables. More specifically, it was hypothesized that reading comprehension strategy use was positively correlated with reading strategy instruction, reading strategy use was positively correlated with reading attitudes, and reading attitudes were positively correlated with reading strategy instruction. Item 1 was selected as the reference indicator (i.e., the factor loading is fixed to 1.00) for the reading attitudes construct. By fixing the factor loading to 1.00, the scale of Item 2 became the scale for the reading attitudes construct. Similarly, Items 8 and 13 were the reference indicators for the reading comprehension strategy instruction and reading strategy use constructs, respectively.

CFA model identification. Once the hypothetical CFA model was specified, the identification status of the model needed to be determined. First, the number of known elements in the model was calculated using *Equation 2* (Dimitrov, 2012). The hypothetical CFA model had 25 indicators, $p = 25$. Thus, the number of known elements in the model was 190, 19(19+1)/2 = 190. The number of unknown model parameters, 60, was then obtained from Mplus. Next, the *df* of the hypothetical CFA model were computed by subtracting the number of unknown elements, 60, from the number of known elements, 190, $df = 130$. The number of known elements was greater than the number of unknown elements, therefore the *df* of this model were positive. Thus, this CFA model was classified as overidentified, which as previously mentioned, is scientifically desirable for testing model fit.

Table 6

19 Indicators for the Reading Constructs in the Hypothetical CFA Model

Construct/Indicator	Survey Item
Reading Attitudes	How much do you agree or disagree with these statements about reading?
Item 1	I read only when I have to.*
Item 2	Reading is one of my favorite hobbies.
Item 3	I find it hard to finish books.*
Item 4	I feel happy if I receive a book as a present.
Item 5	For me, reading is a waste of time.*
Item 6	I enjoy going to a bookstore or library.
Item 7	I only read to get information that I need.*
Reading Strategy Instruction	In your English classes, how often does the following occur?
Item 8	The teacher asks questions that challenge students to get a better understanding of the text.
Item 9	The teacher gives students enough time to think about answers.
Item 10	The teacher recommends a book or author to read.
Item 11	The teacher encourages students to express opinions about text.
Item 12	The teacher shows students how the information in texts builds on what they already know.
Reading Strategy Use	During the last month, how often did you have to do the following kinds of tasks for school?
Item 13	Find information from a graph, diagram, or table.
Item 14	Explain the cause of events in a text.
Item 15	Explain the way characters behave in a text.
Item 16	Learn about the life of a writer.
Item 17	Explain the purpose of a text.
Item 18	Learn about the place of a text in the history of literature.
Item 19	Explain the connection between different parts of a text.

Note. * Indicates a negatively worded item whose response categories were reverse coded.

CFA assumptions. Prior to testing for CFA model fit, tests of univariate normality, outliers, and multivariate normality of the observed data were conducted using SPSS to determine whether the assumption of multivariate normality was met. First, the skewness and kurtosis of the data for each observed latent variable was examined to determine if the assumption of univariate normality was met. The skewness and kurtosis of all 19 observed indicators are presented in Table 7. The skewness values for all the indicators were less than recommended cutoff of 2.0, ranging from -1.00 to 0.49. Additionally, the kurtosis values for all the indicators were less than recommended cutoff of 7.0, ranging from -1.24 to 0.05. Therefore, the skewness and kurtosis values illustrated that the data of all the observed indicators demonstrated univariate normality.

Table 7

Skewness and Kurtosis of the 19 Observed Latent Variable Indicators

Construct/Indicator	Skewness	Kurtosis
Reading Attitudes		
X1 – Have To	0.01	- 0.93
X2 – Favorite Hobby	0.49	- 0.61
X3 – Finish Books	- 0.41	- 0.54
X4 – Book as a Present	0.22	- 0.94
X5 – Waste of Time	- 0.65	- 0.46
X6 – Bookstore/Library	- 0.10	- 0.98
X7 – Read for Information	- 0.03	- 0.75
Reading Strategy Instruction		
X8 – Asks Questions	- 0.36	- 0.50
X9 – Wait Time	- 0.31	- 0.58
X10 – Books/Authors	0.18	- 0.89
X11 – Encourages Opinions	- 0.28	- 0.76
X12 – Prior Knowledge	- 0.13	- 0.79
Reading Strategy Use		
X13 – Use Graphs	- 1.00	0.05
X14 – Cause and Effect	- 0.71	- 0.50
X15 – Character Actions	- 0.79	- 0.45
X16 – Author Background	0.17	- 1.24
X17 – Explain Text Purpose	- 0.73	- 0.57
X18 – Historical Context	- 0.17	- 1.14
X19 – Connect Text Parts	- 0.31	- 1.13

The outliers in the data of the observed latent variable indicators were then examined. The box plots created to examine the data are provided in Appendix D. The results of the box plot analysis illustrated a total of seven cases with extreme outliers. As a total of seven out of the 5,233 cases in the sample was a negligible number of outliers, the decision to not remove the cases with outliers was made.

Finally, to examine the multivariate normality of the observed data in the CFA model, a second simple linear regression test was conducted to obtain the standardized predicted values. The scatterplot in Figure 5 was used to plot the standardized residuals of the model against the standardized predicted values. The lack of pattern among the residuals and predicted values in the scatterplot provided evidence to support the assumption of multivariate normality was met.

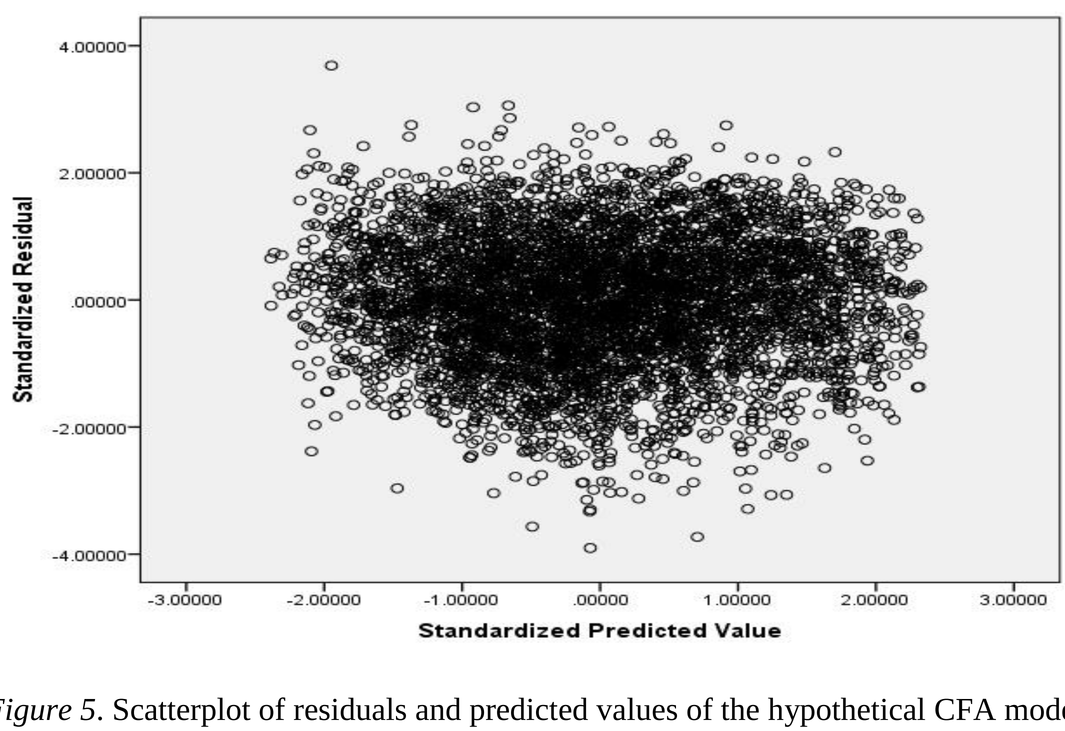


Figure 5. Scatterplot of residuals and predicted values of the hypothetical CFA model.

CFA model testing and parameter estimation. Following the specification, identification, and verification of the assumption of multivariate normality, the first research question was examined by testing the model fit of the hypothesized CFA model using Mplus. The adequacy of the initial CFA model fit, illustrated in Figure 6, was examined to test the measurement portion of the structural equation model. In addition to the expectation of a non-significant chi-square statistic, the following recommended cutoff values for the goodness-of-fit statistics were used to determine the adequacy of the CFA model fit: a CFI and TFI greater than .95 (Hu & Bentler, 1999), SRMR less than .08, and RMSEA less than .05 where the lower limit of the 90% CI is close to or includes zero, and the upper limit is less than .08 (Browne & Cudeck, 1993).

The results of the initial test of CFA model fit, $\chi^2(149) = 1{,}829.25$, $p < .001$, CFI = .95, TLI = .946, SRMR = .03, and RMSEA = .05, with 90% CI (.045, .048), indicated this model did not demonstrate adequate model fit. Basing the evaluation of model fit solely on the chi-square test is cautioned because it is extremely sensitive to sample size. Since the sample size in this study was large ($N = 5{,}233$), the result of the chi-square test was significant. In other words, the result of the chi-square test was biased because of the large sample size. Therefore, it was necessary to also evaluate the fit of the measurement and structural models based on the other four goodness-of-fit indices. The CFI, SRMR, and RMSEA met the criteria for model fit. However, the TLI (.946) was not above the .95 cutoff.

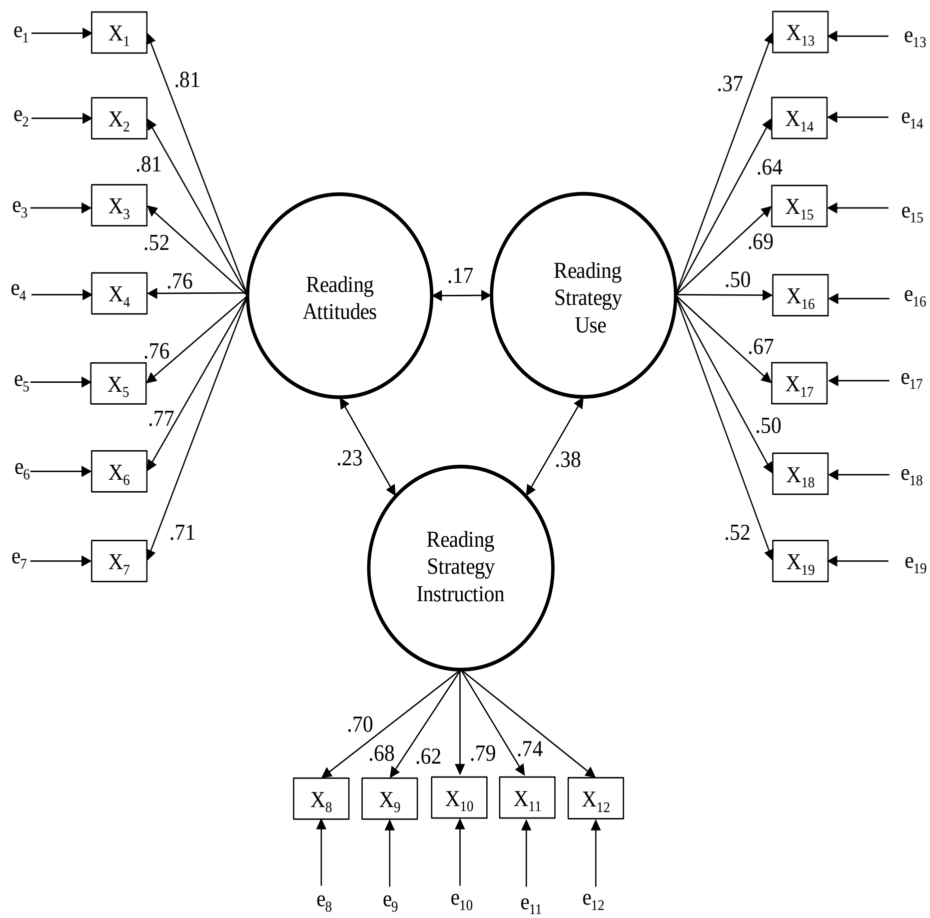


Figure 6. Initial confirmatory factor analysis model.

Modification indices were then examined to determine if any modifications could be made to improve its fit to the data. The large number, and high values, of MIs for this model suggested several modifications that would improve its fit to the data. In this model, the 11 largest MIs corresponded to the following correlated item error variances:

- Item 4 (I feel happy if I receive a book as a present) with Item 2 (Reading is one of my favorite hobbies)
- Item 5 (For me, reading is a waste of time) with Item 3 (I find it hard to finish books)
- Item 6 (I enjoy going to a bookstore or library) with Item 2 (Reading is one of my favorite hobbies)
- Item 6 (I enjoy going to a bookstore or library) with Item 4 (I feel happy if I receive a book as a present)
- Item 7 (I read to get information that I need) with Item 1 (I only read if I have to)
- Item 9 (The teacher gives students enough time to think about their answers) with Item 8 (The teacher asks questions that challenge students to get a better understanding of a text)
- Item 14 (In the last month I had to explain the cause of events in a text) with Item 13 (In the last month I had to find information on a graph, diagram, or table)
- Item 15 (In the last month I had to explain the way characters behave in a text) with Item 14 (In the last month I had to explain the cause of events in a text)
- Item 16 (In the last month I had to learn about the life of a writer) with Item 15 (In the last month I had to explain the way characters behave in a text)
- Item 17 (In the last month I had to explain the purpose of a text) with Item 15 (In the last month I had to explain the way characters behave in a text)
- Item 18 (In the last month I had to learn about the place of a text in the history of literature) with Item 16 (In the last month I had to learn about the life of a writer)

The decision to estimate correlated item error variances was made once each item pair was reviewed to ensure they were not duplicate questions. Stepwise procedures were used to estimate one correlated error at a time, in descending order starting with the largest MI. Eleven modifications of the model were made, each estimating an additional correlated item error variance and improving the overall model fit. The standardized estimates of the correlated error variances are included in Figure 7.

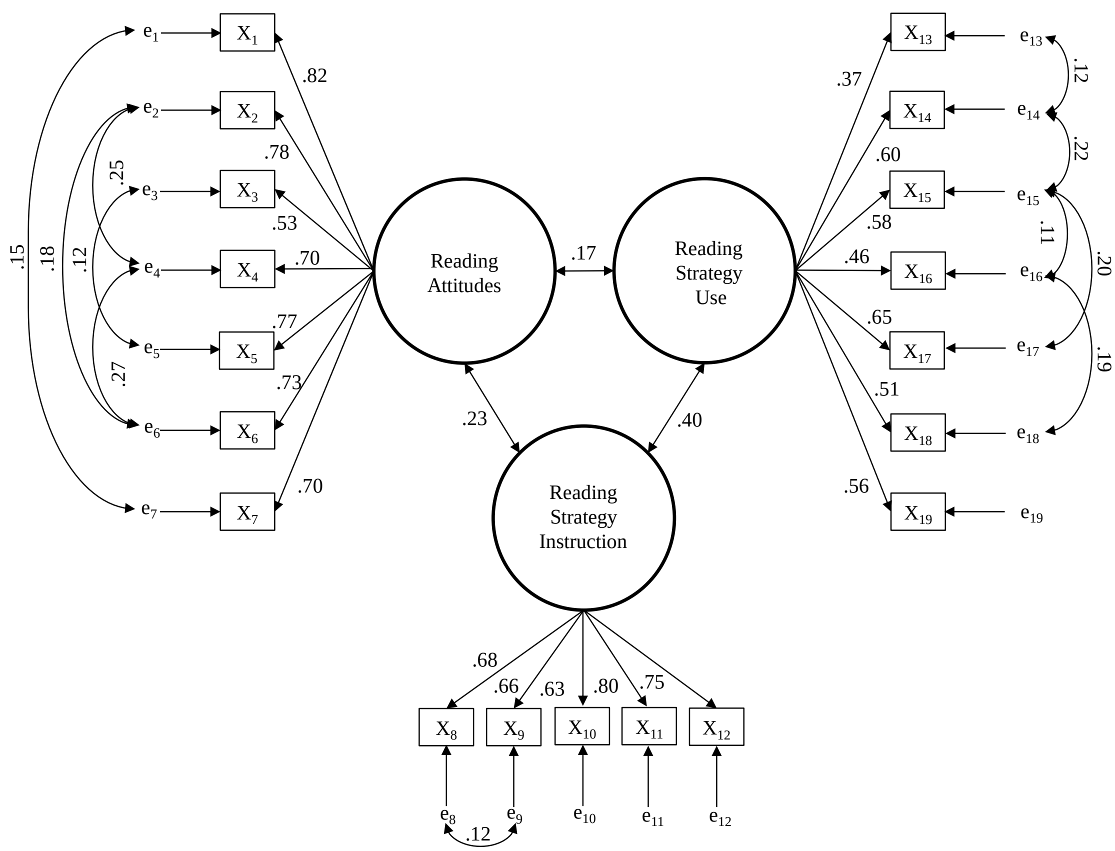


Figure 7. Final confirmatory factor analysis model with correlated errors.

The goodness-of-fit indices of the Initial and Final CFA Models are presented in Table 8. Clearly, the Final CFA Model, provided a better fit to the data. The statistically significant chi-square estimate of the Final Model, χ^2 (138) = 713.21, p < .001, was not problematic because this estimate was inflated by the large sample size. Given the sample of 5,233 students in this study, a significant chi-square estimate was expected. Despite the statistically significant χ^2, the joint examination of the other goodness-of-fit indices signified an improvement in all the values from the initial CFA model, CFI = .98, TLI = .98, SRMR = .02, and RMSEA = .03, with 90% CI (.026, .030). Specifically, the TLI (.98) met the cutoff criteria (TLI < .95) recommended by Hu and Bentler (1999). In addition to the increased adequacy of the goodness-of-fit indices, an improved model fit of the Final CFA Model was also supported by smaller and fewer modification indices than the Initial CFA Model.

Table 8

Goodness-of-Fit Indices for the Initial and Final CFA Models

CFA Model	χ^2	*df*	CFI	TLI	SRMR	RMSEA	90% CI LL	90% CI UL
Initial Model	1,829.25*	149	.95	.95	.03	.05	.045	.048
Final Model	713.21*	138	.98	.98	.02	.03	.026	.030

Note. CFA = confirmatory factor analysis; CFI = comparative fit index; TLI = Tucker–Lewis Index; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; CI = confidence interval; LL = lower limit; UL = upper limit.
[a]N = 5,233.
*p < .001.

The standardized residual variances in the Final CFA Model, presented in Table 9, were evaluated to further examine model fit. The standardized residual variances ranged from 0.33 to 0.86. The values were sufficiently small and within the recommended range of -2 to 2, which indicated no model-to-data inconsistencies and provided additional evidence in support of good model fit to the data.

Good fit of the Final CFA Model was further examined by determining whether the parameter estimates were both statistically and substantively viable. First, the statistical viability of the model was evaluated by examining the factor correlations, factor error variances, and item error variances. Statistical support for the model fit is provided by factor correlations that do not exceed an absolute value of 1, positive factor error variances, and positive item error variances (Dimitrov, 2012). In Figure 7 above, the statistically significant standardized coefficients associated with the two-way arrows connecting the three latent variables constructs represent positive correlations among the constructs, ($r = .23$, $p < .001$, $r = .17$, $p < .001$, and $r = .40$, $p < .001$, respectively). The positive factor error variances and positive item error variances, ranging from 0.33 to 0.86, offer statistical support for the good fit of the Final CFA Model. The three positive correlations in the model provide additional support of its good fit to the data because none of the unstandardized correlation coefficients exceeded an absolute value of 1.

The consistency between the model specifications and the direction, magnitude, and statistical significance of the parameter estimates were then examined to obtain substantive support for the model. As previously mentioned, the correlations between the factors were all positive and statistically significant. Further, the factor loadings for the indicators were all positive, moderate to large in magnitude ranging from .37 to .82, and statistically significant ($p < .001$). As these results are consistent with the model specifications, they provide substantive support for the good fit of the Final CFA Model.

Table 9

Standardized Factor Loadings, Standard Errors, R^2 Values, and Residual Variances of the Final CFA Model

Construct/Item	Factor Loading	Standard Error	R^2 Value	Residual Variance
Reading Attitudes				
X1 – Have To	.82	.01	.67	.33
X2 – Favorite Hobby	.78	.01	.61	.39
X3 – Finish Books	.53	.01	.28	.72
X4 – Book as a Present	.70	.01	.49	.51
X5 – Waste of Time	.77	.01	.60	.40
X6 – Bookstore/Library	.73	.01	.53	.47
X7 – Read for Information	.70	.01	.49	.51
Reading Strategy Instruction				
X8 – Asks Questions	.68	.01	.46	.54
X9 – Wait Time	.66	.01	.43	.57
X10 – Books/Authors	.63	.01	.39	.61
X11 – Encourages Opinions	.80	.01	.64	.36
X12 – Prior Knowledge	.75	.01	.56	.44
Reading Strategy Use				
X13 – Use Graphs	.37	.02	.14	.86
X14 – Cause and Effect	.60	.01	.36	.64
X15 – Character Actions	.58	.01	.33	.67
X16 – Author Background	.46	.01	.21	.79
X17 – Explain Text Purpose	.65	.01	.43	.57
X18 – Historical Context	.51	.01	.26	.74
X19 – Connect Text Parts	.56	.01	.31	.69

Note. All values are statistically significant ($p < .001$). Critical ratios ranged from 24.86 to 124.79.

Next, the values of the standardized factor loadings, presented in Table 9 above, were examined to determine the relative practical importance of each item used to define its respective latent construct. The standardized factor loadings for all the items included in the Final CFA model, depicted in Figure 7, are displayed on the one-way arrows from the latent variables to their indicators.

Overall, the factor loadings of the seven items used to define the reading comprehension strategy use factor, ranging from .37 to .65, were smaller than the factor loadings of the items that defined the reading comprehension strategy instruction and the reading attitudes constructs described below. Nevertheless, as mentioned, all seven factor loadings were statistically significant. Item 17, ‘In the last month I had to explain the purpose of a text’, had the largest practical importance in defining the factor (.65). Item 14, ‘In the last month I had to explain the cause of events in a text’ (.60), had the second largest practical importance in defining the reading comprehension strategy use construct, followed by Item 15, ‘In the last month I had to explain the way characters behaved in a text’ (.58), and Item 19, ‘In the last month I had to explain the connection between different parts of a text’ (.56). The three items with the smallest practical importance in defining the construct were Item 18, ‘In the last month I had to learn about the place of a text in the history of literature’ (.51), Item 16, ‘In the last month I had to learn about the life of a writer’ (.46), and finally Item 13, ‘In the last month I had find information from a graph, diagram, or table’ (.37).

The factor loadings of the five reading comprehension strategy instruction indicators ranged from .63 to .80. Item 11, ‘The teacher encourages students to express opinions about a text', was the item with the greatest relative importance in defining this construct (.80). The item with the next largest relative importance in the strategy instruction construct was Item 12, 'The teacher shows students how the information in texts builds on what they already know' (.75). The third largest item of practical importance in this construct was Item 8, 'The teacher asks questions that challenge students to get a better understanding of a text' (.68). The relative importance of Item 8 was followed closely by that of Item 9, 'The teacher gives students enough time to think about their answers' (.66). Lastly, the item with the smallest relative importance in explaining the reading comprehension strategy instruction factor was Item 10, 'The teacher recommends a book or author to read' (.63).

The factor loadings for the seven items that defined the third latent factor, reading attitudes, ranged from .53 to .82. The factor loading of Item 1, 'I do not only read if I have to' (.82), was the largest of the seven items, which signified that this item had the largest practical importance in defining the reading attitudes construct. The next two items in order of practical importance were Item 2, 'Reading is one of my favorite hobbies' and Item 5, 'For me, reading is not a waste of time' (.78 and .77, respectively). The next largest item of relative importance was Item 6, 'I enjoy going to a bookstore or library' (.73). The relative importance of Item 6 was slightly larger than that of Item 4, 'I feel happy if I receive a book as a present' and Item 7, 'I do not only read to get information that I need', that had the same factor loading (.70). Finally, Item 3, 'I do not find it hard to finish books', had the smallest factor loading (.53), thus the smallest relative importance among the seven items used to define the reading attitudes construct.

Finally, the squared value of each item's standardized factor loading, R^2, was calculated to determine the amount of variance in the item that was explained (i.e., accounted for) by its respective factor. The results indicated that reading comprehension strategy use accounted for 13.91% of the difference in the students' answers to how often they had to find information on a graph, diagram, or table in the last month, 35.64% of the variability in the responses of students regarding how often in the last month they had to explain the cause of events in a text, 31.02% of the differences in students' ratings of the frequency that they had to explain the way characters behaved in a text in the past month, 21.44% of the variability in the responses of students about how often in the last month they had to learn about the life of a writer, 42.77% of the differences in how students responded to the question asking how often in the last month they had to explain the purpose of a text, 25.91% of the differences in students' answers concerning how often in the last month they had to learn about the place of a text in the history of literature, and 31.36% of the variability in how students rated the frequency that they had to explain the connection between different parts of a text in the last month.

Likewise, it was determined that reading comprehension strategy instruction accounted for 45.70% of the variability in students' answers to how often their teachers asked questions that challenge them to get a better understanding of the text, 42.90% of the differences in how students rated the frequency that their teachers gave them enough time to think about their answers, 39.18% of the variability in how students responded to the question asking how often their teachers recommended books and authors for them to read, 63.68% of the differences in the responses of students about how often their teachers encouraged them to express their opinions about a text, and 55.65% of the variability in students' ratings of the frequency that their teachers showed them how the information in texts builds on what they already know.

Lastly, the results revealed that high school students' attitudes toward reading accounted for 67.08% of the differences in how strongly students rated their agreement with the statement that they do not only read when they have to, 60.53% of the variability in students' ratings of how much they agreed that they considered reading to be one of their favorite hobbies, 27.89% of the differences in the student answers concerning how much agreed or disagreed that they do not find it hard to finish reading books, 48.86% of the variability in how students rated the degree to which they agreed with feeling happy if they receive a book as a present, 59.75% of the differences in how students responded to the question asking how strongly they agreed or disagreed with the statement that they do not believe reading is a waste of time, 52.70% of the variability in the responses of students about how much they agreed that they enjoy going to a bookstore or library, and 49.42% of the differences in student answers regarding how strongly they agreed that they do not only read to get the specific information they need.

Latent variable modeling reliability. The correlated errors among the factor items identified through the CFA made it necessary to re-estimate the internal consistency reliability. Re-estimation was necessary because correlated errors violate the assumption of essentially tau-equivalent measures that must be met to ensure the Cronbach's alpha reliability estimates were accurate. Thus, the estimation of internal consistency reliability was conducted again for each factor using the latent variable modeling (LVM) approach in Mplus. Table 10 contains the Pearson internal consistency reliability estimates, standard error of measurement estimates, and corresponding confidence intervals (90%, 95%, and 99%) for the three reading constructs (reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes).

Table 10

Latent Variable Modeling Internal Consistency Reliability and Standard Error Estimates for the Three Reading Constructs

Construct	Lower Confidence Interval				Upper Confidence Interval		
	99%	95%	90%	Estimate	90%	95%	99%
Reading Strategy Use							
Reliability	0.745	0.749	0.750	0.760	0.768	0.770	0.773
SEM	2.171	2.177	2.180	2.198	2.216	2.219	2.225
Reading Strategy Instruction							
Reliability	0.820	0.822	0.824	0.831	0.838	0.839	0.842
SEM	1.361	1.367	1.370	1.386	1.402	1.405	1.411
Reading Attitudes							
Reliability	0.887	0.889	0.889	0.893	0.897	0.898	0.899
SEM	1.622	1.627	1.631	1.647	1.663	1.667	1.673

Note. CI = confidence interval; *SEM* = standard error of measurement.

The LVM estimate of reliability for the reading comprehension strategy use construct was, ρ_{xx} = 0.76, with a 95% CI (0.749, 0.770). The standard error of measurement for the reading strategy use construct was *SEM* = 2.20 with a 95% CI (2.177, 2.219). The LVM reliability of the strategy use scale was slightly lower than the reliability of the reading comprehension strategy instruction and the reading attitudes scaled, but it still demonstrated acceptable internal consistency reliability as it meets the social sciences recommended reliability cutoff of α = 0.70.

Likewise, the LVM reliability for the reading comprehension strategy instruction factor was estimated as, ρ_{xx} = 0.83, with a 95% CI (0.822, 0.839). This estimate indicated that the strategy instruction construct demonstrated high internal consistency reliability. The standard error of measurement for the reading strategy instruction construct was *SEM* = 1.39 with a 95% CI (1.367, 1.405).

Finally, for the reading attitudes latent construct, the LVM estimate of reliability was, ρ_{xx} = 0.89, with a 90% CI (0.889, 0.897), a 95% CI (0.889, 0.898), and a 99% CI (0.887, 0.899). This finding suggested that the internal consistency reliability for the reading attitudes construct was high. Further, the estimate of standard error of measurement for the reading attitudes construct was *SEM* = 1.65, with a 95% CI (1.627, 1.667). The comparison of the Cronbach's alpha and Latent Variable Modeling internal consistency reliability estimates for each of the three latent variables are presented in Table 11. This comparison reveals that both types of reliability analyses produced identical reliability estimates when used to measure the three constructs, thus confirming that the initial Cronbach's alpha reliability estimates were accurate.

Table 11

Comparison of the Reliability Estimates for the Three Reading Constructs

Construct	Reliability Estimates	
	Cronbach's Alpha (α)	LVM Estimate (ρ_{xx})
Reading Strategy Use	.76	.76
Reading Strategy Instruction	.83	.83
Reading Attitudes	.89	.89

Note. α = Cronbach's coefficient alpha internal consistency reliability estimate; ρ_{xx} = Pearson internal consistency reliability estimate.

SEM structural model – Structural regression and MIMIC analyses. Following the confirmation of good model fit of the CFA measurement model, the structural models were tested using structural regression analysis to investigate the direct and indirect effects among the variables in the model, and MIMIC analysis to examine group mean differences.

Structural regression analysis. Structural regression analysis was conducted in Mplus to evaluate the second research question, 'Are there direct and indirect effects among the five student and school level variables, the three latent variables, and high school students' reading comprehension achievement?'. The process of structural regression analysis included specifying and identifying of the hypothetical model, testing the fit of model to the data, and estimation and evaluation of the model parameters.

Structural regression model specification. Specifying the hypothetical model was the first step in the structural regression analysis. The hypothetical structural regression model, illustrated in Figure 3, posited a variety of direct effects among the control, latent, and dependent variables. It was hypothesized that reading attitudes, class time, and class size each had a direct effect on high school students' reading comprehension strategy use. It was further hypothesized that reading comprehension strategy instruction, gender, minority status, and SES all had direct effects on both reading strategy use and reading attitudes. The direct effects of both class time and class size on reading comprehension strategy instruction were also hypothesized. Furthermore, it was hypothesized that each of the five control variables (gender, minority status, SES, class time, and class size) and the three latent variables (reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes) had a direct effect on high school students' reading comprehension achievement. No specific indirect effects were hypothesized, however, for exploratory purposes the mediating effects of all five control variables and three latent variables were included in the structural equation model.

Structural regression model identification. Following the specification of the hypothetical structural regression model, the identification status of the model needed determined. The first step of determining the identification status of the model was to calculate the number of known elements in the model using *Equation 2* (Dimitrov, 2012). With 25 indicators in the model, p = 25, the number of known elements was 325, 25(25+1)/2 = 325. Next, the number of unknown model parameters, 91, was obtained from Mplus. The *df* of the hypothetical CFA model were then computed by subtracting the number of unknown elements, 91, from the number of known elements, 325, *df* = 234. Because the number of known elements was greater than the number of unknown elements, the *df* of this model were positive. Therefore, the hypothetical structural regression model was overidentified.

Structural model testing and parameter estimation. In order to examine research subquestion 2a, 'What are the direct and indirect effects of the three latent constructs and the five student (i.e., gender, minority status, and SES) and school (i.e., class time and class size) level variables on high school students' reading comprehension achievement?', the direct and indirect effects of the three latent variables and the student and school characteristics on the dependent variable (reading comprehension achievement) were assessed by simultaneously regressing the dependent variable on all eight independent variables. Research subquestion 2b, ‘What are the direct effects among the three latent constructs (reading strategy use, reading strategy instruction, and reading attitudes)?’, was examined next by simultaneously regressing the three latent variables on each other. Finally, in response to research subquestion 2c, ‘What are the direct effects of the five student and school level variables on the three latent constructs?’, the direct effects of the five control variables on the three latent variables were tested by simultaneously regressing the latent variables on the five control variables.

A joint examination of the goodness-of-fit indices for the structural regression model, illustrated in Figure 8, revealed a good model fit to the data, χ^2 (239) = 1502.43, $p < .001$, CFI = 0.96, TLI = .95, SRMR = .03, RMSEA = .03 (90% CI .030-.033). The chi-square estimate was statistically significant, however as previously mentioned, this is not considered an issue because the χ^2 estimate was expected to be inflated by large sample size in this study (N = 5,233). Further, additional support for good fit of the structural regression model to the data was offered by the standardized residuals, ranging from .34 to 1.00, which were well within the recommended range.

Following the confirmation of a good fit of the structural regression model, the standardized path coefficients (i.e., partial regression coefficients) associated with the direct and indirect effects in the model were examined. The statistically significant and nonsignificant standardized path coefficients for the direct effects in the model are depicted in Figure 8. Additionally, the decomposition of the direct effects, indirect effects (total indirect effect and specific indirect effects), and total effects corresponding to the paths associated with research subquestion 2a are presented in Table 12.

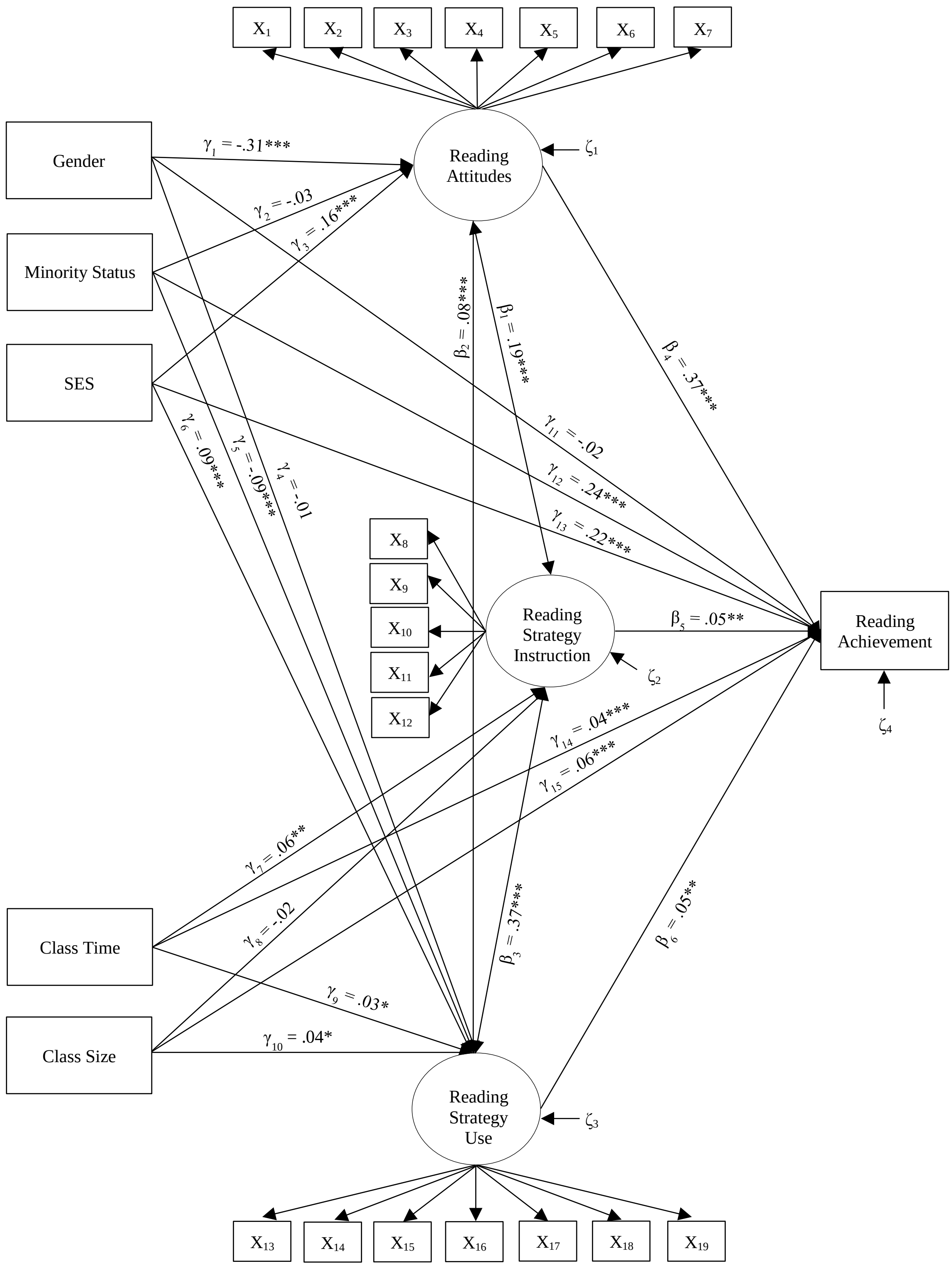


Figure 8. Final structural regression model.
**p* < .05. ***p* < .01. ****p* < .001.

Table 12

Standardized Effects of the Control Variables and the Three Reading Constructs on Reading Comprehension Achievement

Direct and Indirect Paths	Direct Effects	Specific Indirect Effects	Total Indirect Effects	Total Effects
Reading Strategy Use → Reading Achievement	.05**			.05**
Reading Strategy Instruction → Reading Achievement	.05**		.089***	.13***
RSI → RA → Reading Achievement		.070***		
RSI → RSU → Reading Achievement		.018**		
RSI → RA → RSU → Reading Achievement		.001*		
Reading Attitudes → Reading Achievement	.37***		.004**	.37***
RA → RSU → Reading Achievement		.004**		
Gender → Reading Achievement	-.02		-.116***	-.13***
Gender → RA → Reading Achievement		-.114***		
Gender → RSU → Reading Achievement		.000		
Gender → RA → RSU → Reading Achievement		-.001*		
Minority Status → Reading Achievement	.24***		-.014*	.23***
Minority Status → RA → Reading Achievement		-.010		
Minority Status → RSU → Reading Achievement		-.004*		
Minority Status → RA → RSU → Reading Achievement		.000		
Socio-Economic Status → Reading Achievement	.22***		.063***	.29***
SES → RA → Reading Achievement		.058***		
SES → RSU → Reading Achievement		.005*		
SES → RA → RSU → Reading Achievement		.001*		
Class Time → Reading Achievement	.04**		.01***	.05***
Class Time → RSI → Reading Achievement		.002*		
Class Time → RSU → Reading Achievement		.002		
Class Time → RSI → RA → Reading Achievement		.004**		
Class Time → RSI → RSU → Reading Achievement		.001*		
Class Size → Reading Achievement	.06***		-.001	.06***
Class Size → RSI → Reading Achievement		-.001		
Class Size → RSU → Reading Achievement		.002		
Class Size → RSI → RA → Reading Achievement		-.001		
Class Size → RSI → RSU → Reading Achievement		.000		

Note. RSI = Reading Strategy Instruction; RA = Reading Attitudes; RSU = Reading Strategy Use; SES = socio-economic status.
*$p < .05$. **$p < .01$. ***$p < .001$.

Among the significant direct effects of the three latent variables on high school students' reading comprehension achievement, reading attitudes had the greatest direct effect on their reading achievement (β_4 = .37, p < .001). The statistically significant specific indirect effect of students' reading attitudes on reading achievement via reading comprehension strategy use was markedly lower than the direct effect of reading attitudes (β_2-β_6 = .004, p = .03). Further, the significant and equivalent direct effects of strategy instruction and strategy use on reading achievement were both notably smaller than the direct effect of reading attitudes (β_5 = .05, p = .01 and β_6 = .05, p = .01, respectively). Despite having equivalent direct effects of reading achievement, the total effects of strategy instruction on achievement was larger (TE = .13, p < .001) than the total effect of reading strategy use on reading achievement (TE = .05, p < .01).

The total effect of reading comprehension strategy instruction on reading achievement was also larger than its direct effect. The total effects were larger than the direct effect because the statistically significant total indirect effects of strategy instruction on reading achievement (TIE = .09, p < .001) were included in the estimation of the total effects. These total indirect effects consisted of the specific indirect effects of reading attitudes (β_1-β_4 = .07, p < .001), strategy use (β_3-β_6 = .02, p = .01), and reading attitudes via strategy use (β_1-β_2-β_6 = .001, p = .03) on the relationship between reading strategy instruction and reading achievement.

Although the direct effect of gender on high school students' reading comprehension achievement was not significant in the structural regression model, the total effects (TE = -.13, p < .001) and the total indirect effects (TIE = -.17, p < .001) were both statistically significant. This indicates an indirect relationship between gender and reading comprehension achievement. Based on the structural regression results, this relationship was mediated by students' attitudes toward reading (γ_1-β_4 = -.12, p < .001) and by a combination of reading attitudes and reading comprehension strategy use (γ_1-β_2-β_6 = -.12, p = .04). Finally, the mediating effect of reading comprehension strategy use, also indicated in the results, was not statistically significant (γ_4-β_6 = .001, p = .61).

Among the other four student and school level variables (i.e., minority status, SES, class time, and class size), high school students' minority status and their SES had the greatest direct effects on reading comprehension achievement (γ_{12} = .24, p < .001 and γ_{13} = .22, p < .001, respectively). While the total value of the indirect effects in the relationship between students' minority status and their reading comprehension achievement was statistically significant (TIE = -.01, p = .02), only one of the three specific indirect effects expressed in the findings was significant. The results of the structural regression analysis revealed the significant mediating effect of reading comprehension strategy use on the relationship between students' minority status and their reading comprehension achievement (γ_5-β_6 = -.004, p = .02). Whereas, neither the mediating effect of reading attitudes nor the combined mediating effect of reading attitudes and reading comprehension strategy use on the relationship between students' minority status and their reading comprehension achievement were statistically significant (γ_2-β_4 = -.01, p = .08 and γ_2-β_2-β_6 = .00, p = .18, respectively).

The three specific indirect effects on the relationship between SES and student's reading comprehension achievement were all statistically significant. According to the structural regression results, the relationship between SES and reading comprehension achievement was mediated by students' reading attitudes (γ_3-β_4 = .06, $p < .001$) and reading comprehension strategy use (γ_6-β_6 = .01, $p = .02$). This relationship was also mediated by the statistically significant indirect effect of reading attitudes via reading comprehension strategy use (γ_3-β_2-β_6 = .001, $p = .04$).

Despite the statistical significance of these indirect effects, it is important to point out that the total indirect effect (TIE = .06 ($p < .001$) was markedly smaller than the total effect of students' SES on their reading comprehension achievement (TE = .29 ($p <$ .001). This meant that the combined effect of the mediating role that students' reading attitudes and reading comprehension strategy use play in the relationship between SES and reading comprehension achievement had little relative importance in the total prediction of reading comprehension achievement accounted for by students' SES.

Compared to the relative importance students' minority status and SES had in predicting reading comprehension achievement, the relative importance of class time and class size was considerably smaller. Further, the statistically significant direct effect of class size on students' reading achievement (γ_{15} = .06, $p < .001$), was slightly larger than the direct effect of class time (γ_{14} = .04, $p < .001$). The four specific indirect effects of class size on reading comprehension achievement via reading strategy instruction, via reading strategy use, via strategy instruction and reading attitudes, and via reading strategy instruction and reading strategy use, were not significant (γ_8-β_5 = -.001, $p = .29$; γ_{10}-β_6 = .002, $p = .10$; γ_8-β_1-β_4 = -.001, $p = .25$; and γ_8-β_3-β_6 = .00, $p = .29$, respectively). The nonsignificant indirect effects indicated that students' reading comprehension strategy use, reading strategy instruction, and students' attitudes toward reading did not mediate the relationship between class size and reading comprehension achievement.

Whereas, all three of these latent variables significantly mediated the relationship between class time and reading comprehension achievement. The mediating effect of reading comprehension strategy instruction via reading attitudes on the relationship between class time and reading comprehension achievement (γ_7-β_1-β_4 = .004, p = .001), was slightly larger than the mediating effect of reading comprehension strategy instruction (γ_7-β_5 = .002, p = .03), and the mediating effect of reading comprehension strategy instruction via reading strategy use (γ_7-β_3-β_6 = .001, p = .04). The only nonsignificant indirect effect of class time on reading achievement was the indirect effect via reading comprehension strategy (γ_8-β_6 = .002, p = .12).

According to the results of the structural regression analysis for research subquestion 2b, 'What are the direct effects among the three latent constructs?', reading comprehension strategy instruction and high school students' attitudes toward reading had statistically significant direct effects on reading comprehension strategy use. The statistically significant regression coefficient for the direct path from reading strategy instruction to reading strategy use (β_3 = .37, p < .001) was larger than the statistically significant regression coefficient for the direct path from reading attitudes to reading strategy use (β_2 = .08, p < .001), suggesting that strategy instruction was a better predictor of comprehension strategy use in this model than reading attitudes. The results of research subquestion 2b also revealed a statistically significant positive relationship between high school students' attitudes toward reading and reading comprehension strategy instruction ($\beta_1 = .19$, $p < .001$). Table 13 presents the standardized partial regression coefficients for the direct effects among high school student' reading comprehension strategy use, reading strategy instruction, and reading attitudes.

Table 13

Standardized Direct Effects Among the Three Reading Constructs

Direct Paths	Direct Effects
Reading Strategy Use	
Reading Strategy Instruction → Reading Strategy Use	.37***
Reading Attitudes → Reading Strategy Use	.08***
Reading Attitudes	
Reading Strategy Instruction → Reading Attitudes	.19***

*$p < .05$. **$p < .01$. ***$p < .001$.

The results of the structural regression analysis employed to examine research subquestion 2c, 'What are the direct effects of the five student and school level variables on the three latent constructs?', indicated first, that minority status, SES, class time, and class size were all statistically significant student/school level predictors of high school students' reading strategy use. Based on these findings, high school students' minority status and SES both have a statistically significant impact on their reported reading comprehension strategy use. The absolute value of the partial regression coefficient for the direct path from minority status to strategy use ($\gamma_5 = -.09$, $p < .001$) was equal to the statistically significant partial regression coefficient for the direct path from SES to strategy use ($\gamma_6 = .09$, $p < .001$). This signified that high school students' minority status and SES equally predicted their reading strategy use.

The structural regression results also intimated statistically significant and positive effects of class time and class size on high school students reading comprehension strategy use (γ_9 = .03, p = .05 and γ_{10} = .04, p = .04, respectively). Further, the results intimated that the hypothesized direct effect of students' gender on their reported use of reading comprehension strategies was nonsignificant (γ_4 = -.01, p = .60). Table 14 presents the standardized path coefficients for the direct effects of the student and school level variables on high school students' reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes.

Table 14

Standardized Direct Effects of the Control Variables on the Three Reading Constructs

Direct Paths	Direct Effects
Reading Strategy Use	
Gender → Reading Strategy Use	-.01
Minority Status → Reading Strategy Use	-.09***
SES → Reading Strategy Use	.09***
Class Time → Reading Strategy Use	.03*
Class Size → Reading Strategy Use	.04*
Reading Strategy Instruction	
Class Time → Reading Strategy Instruction	.06**
Class Size → Reading Strategy Instruction	-.02
Reading Attitudes	
Gender → Reading Attitudes	-.31***
Minority Status → Reading Attitudes	-.03
SES → Reading Attitudes	.16***

Note. SES = socio-economic status.
*p < .05. **p < .01. ***p < .001.

Next, the findings of the structural regression analyses showed that two of the three hypothesized student/school level predictor variables had statistically significant direct effects on high school students' attitudes toward reading. The path coefficient for the direct effect of gender ($\gamma_1 = -.31$, $p < .001$) on reading attitudes showed that it had the greatest relative importance in predicting the variability of students' reading attitudes. The statistically significant path coefficient for direct effect of students' SES on their reading attitudes suggested that it had about half of the relative important in predicting students' reading attitudes as their gender ($\gamma_3 = .16$, $p < .001$). Minority status, the last of the three hypothesized student/school level predictor variables did not have a statistically significant direct effect on students' attitudes toward reading ($\gamma_2 = -.03$, $p = .08$).

Finally, only one of the two hypothesized student/school level predictor variables had a statistically significant direct effect on high school reading strategy instruction. The results revealed a significant positive direct effect of high school students' reported amount reading class time on their reading comprehension strategy instruction, ($\gamma_7 = .06$, $p = .001$). However, effect of high school students' class size on their reading comprehension strategy instruction was not statistically significant ($\gamma_8 = -.02$, $p = .25$).

R^2 values and f^2 effect sizes. The amount of variation in each dependent variable jointly explained by all its predictors together (i.e., R^2 values) was examined after determining the relative importance of each predictor variable in explaining the amount of variation in reading comprehension achievement, reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes. The statistical significances and effect sizes of the R^2 values were also examined to enhance the

meaningful interpretation of results. The standardized residual variance, R^2 value, number of independent predictors, sample size used to calculate the R^2 value, and the effect size for each dependent variable are presented in Table 15.

Table 15

Residual Variances, R^2 Values, Number of Predictors, Sample Sizes, and Effect Sizes for Reading Achievement and the Three Reading Constructs

Dependent Variable	Residual Variance	R^2 Value	Number of Predictors	Sample Size	Effect Size
Reading Achievement	.667	.333	8	5,233	.33
Reading Strategy Use	.997	.172	7	5,233	.17
Reading Strategy Instruction	.828	.003	2	5,233	.00
Reading Attitudes	.848	.152	4	5,233	.15

Note. All residual variances and R^2 values are statistically significant ($p < .05$).

The R^2 value results in this section are part of the results of the second research question. However, this section was separated from the results of research subquestions 2a, 2b, and 2c because the R^2 value results include findings related to both subquestions 2b and 2c that have to be presented together. For example, it was hypothesized that reading comprehension strategy instruction, reading attitudes, gender, minority status, SES, class time, and class size each had a direct effect on reading comprehension strategy use. The hypothesized relationships between instruction and strategy use addressed research subquestion 2b, whereas the hypothesized relationship between gender and strategy use is related to research subquestion 2c. The structural regression model estimated the hypothesized relationships related to the reading comprehension strategy use construct simultaneously. Thus, the R^2 value results were associated with both research subquestions 2b and 2c, inextricably linking the two subquestions. Therefore, the R^2 value results needed to be presented together.

The coefficient of multiple determination for the prediction of high school students' reading comprehension achievement, $R^2 = .333$, denoted that 33.30% of the student differences in students' reading comprehension achievement were jointly accounted for by all the eight of the hypothesized predictors (i.e., the five student/school level variables and three latent variables) together. The order of strength of the predictors of students' reading achievement, ranging from largest to smallest, was the following: reading attitudes ($\beta_4 = .37$, $p < .001$), minority status ($\gamma_{12} = .24$, $p < .001$), SES ($\gamma_{13} = .22$, $p < .001$) , class size ($\gamma_{15} = .06$, $p < .001$), reading comprehension strategy instruction and reading comprehension strategy use ($\beta_5 = .05$, $p = .01$ and $\beta_6 = .05$, $p = .01$, respectively), class time ($\gamma_{14} = .04$, $p < .001$), and finally gender ($\gamma_{11} = -.02$, $p = .19$).

The R^2 value for the prediction of high school students' reading comprehension achievement was statistically significant, $p < .05$, based on the results of the *F*-test. The *F*-statistic, $F = 420.00$, was statistically significant because it exceeded the F-critical value, F-CV = 2.37, for $df(8, 5{,}224)$. The following guidelines were used for interpreting the magnitude of an f^2 effect size index: small effect = .02, medium/moderate effect = .15, and large effect = .35 (Cohen, 1992). Therefore, the effect size of the statistically significant R^2 value, $f^2 = .33$, indicated a large practical significance for the importance of the eight independent variables in predicting reading achievement.

The second R^2 value, for the prediction of high school students' reading strategy use, R^2 = .172, meant that 17.20% of the variance in reading comprehension strategy use was accounted for by all seven of the hypothesized predictors (i.e., reading attitudes, reading comprehension strategy instruction, gender, minority status, SES, class time, and class size) combined. The following list indicates the strength of the seven predictors of high school students' reading strategy use in descending order: reading strategy instruction ($\beta_3 = .37$, $p < .001$), minority status and SES ($\gamma_5 = -.09$, $p < .001$ and $\gamma_6 = .09$, $p < .001$, respectively), reading attitudes ($\beta_2 = .08$, $p < .001$), class size ($\gamma_{10} = .04$, $p = .04$), class time ($\gamma_9 = .03$, $p = .05$), and gender ($\gamma_4 = -.01$, $p = .60$).

Further, the results of the F-test to assess the statistical significance of the R^2 value for the prediction of students' reading comprehension strategy use, indicated that the F-statistic, $F = 125.00$, was statistically significant at the .05 level of significance. It was deemed statistically significant because it exceeded the F-critical value, F-CV = 2.01, for $df(7, 5{,}225)$. Based on the statistically significant F-test, the R^2 value for the prediction of high school students' reading comprehension strategy use was statistically significant, $p < .05$. Considering the effect size of the statistically significant R^2 value, $f^2 = .17$, the importance of these seven independent variables in predicting high school students' strategy use was considered to have a moderate practical significance.

The coefficient of multiple determination for the prediction of high school students' attitudes toward reading, R^2 = .152, suggested that 15.20% of the student differences in reading attitudes were jointly explained by the four hypothesized predictors (i.e., reading strategy instruction, gender, minority status, and SES) together. The order, ranging from largest to smallest, of strength of the predictors of high school students' reading attitudes was gender ($\gamma_1 = -.31$, $p < .001$), reading comprehension instruction ($\beta_1 = .19$, $p < .001$), SES ($\gamma_3 = .16$, $p < .001$), and minority status ($\gamma_2 = -.03$, $p = .08$).

Further, the results of the *F*-test to assess the statistical significance of the R^2 value for the prediction of high school students' reading attitudes, intimated that the *F*-statistic, $F = 237.50$, was statistically significant at the .05 level of significance. It was determined to be statistically significant because it exceeded the F-critical value, *F*-CV = 2.37, for $df(4, 5{,}228)$. The significant *F*-test provided evidence that the variance in high school students' reading attitudes explained by the four predictors was not equal to zero for the population (i.e., the R^2 value was statistically significant). The effect size of the significant R^2 value, $f^2 = .15$, signified a moderate practical significance for the importance of the four predictors in predicting high school students' reading attitudes.

The coefficient of multiple determination for the prediction of high school reading comprehension strategy instruction, $R^2 = .003$, meant that only 0.30% of the student differences in reading comprehension strategy instruction were jointly accounted for by the two hypothesized predictors, class time ($\gamma_7 = .06$, $p = .001$) and class size ($\gamma_8 = -.02$, $p = .25$), together. Despite this extremely small R^2 value, its significance was supported by the results of the *F*-test. The *F*-statistic, $F = 7.50$, was statistically significant at the .05 level of significance because it exceeded the *F*-critical value, *F*-CV = 3.00, for $df(2, 5{,}230)$. However, the effect size of the statistically significant R^2 value, $f^2 = .003$, signified that the lack of importance of class time and class size in predicting high school reading comprehension instruction had no practical significance.

With a few exceptions, the results of the structural regression analyses confirmed the hypothesized model. The overall results of the structural regression analyses revealed that 17 of the 21 hypothesized direct effects were supported by their statistically significant path coefficients. Most importantly, the results of structural regression analysis showed high school students' reading comprehension achievement is directly affected by seven of the eight student/school level and latent variables. In addition to the 17 statistically significant direct effects, the structural regression results revealed 13 statistically significant indirect effects.

Multiple Indicator, Multiple Cause (MIMIC) group code modeling. The third research question, 'Are there significant group differences in students' reading comprehension achievement, reading comprehensions strategy use, reading comprehension strategy instruction, and reading attitudes across the five student (i.e., gender, minority status, and SES) and school (i.e., class time and class size) level variables?', was investigated using the Multiple Indicator, Multiple Cause (MIMIC) modeling approach. The MIMIC analysis procedures (i.e., model specification, model identification, examination of factorial invariance, test of model fit, and estimation of the model parameters) were similar to the initial CFA and structural regression procedures. Mplus was used to conduct the MIMIC analysis.

MIMIC model specification. The first step in the MIMIC analysis was to specify the hypothetical model. The MIMIC model presented in Figure 4, illustrates the hypothesized group differences in reading comprehension achievement, reading strategy use, reading strategy instruction, and reading attitudes across the five student (gender, minority status, and SES) and school (class time and class size) level control variables.

MIMIC model identification. The identification status of the model was determined once the hypothetical MIMIC model was specified. First, *Equation 2*, was used to calculate the number of known elements in the model (Dimitrov, 2012). The model had 25 indicators, $p = 25$. Therefore, the number of known elements was 325, $25(25+1)/2 = 325$. Next, number of unknown model parameters, 96, was obtained from Mplus. Then the *df* of the model were computed by subtracting the number of unknown elements, 96, from the number of known elements, 325, $df = 229$. Since the number of known elements was greater than the number of unknown elements, the *df* of this model were positive. Thus, this CFA model was classified as overidentified.

Testing for factorial invariance. Factorial invariance across groups is a major assumption of MIMIC modeling. Therefore, testing for group differences in reading achievement and the three reading constructs began with confirming the assumption of factorial invariance was met for each control variable.

The first step of factorial invariance testing was to establish the configural invariance of each control variable. Thus, the configural invariance of the five variables was tested. This testing began with the selection of a baseline model. The final CFA model was chosen because it had already demonstrated good model fit while maintaining parsimony and substantive meaning. Next, baseline model fit was tested separately for the two groups in all five control variables. Configural invariance was supported when the fit indices for both groups of the variable individually indicated very good model fit. Table 16 contains sample sizes and fit indices from the configural invariance tests.

Table 16

Sample Size and Goodness-of-Fit Indices for Configural Invariance Tests

Variable/ Group Model	*n*	χ^2	*df*	CFI	TLI	SRMR	RMSEA	90% CI LL	90% CI UL
Gender									
Female	2,546	453.51*	138	.98	.98	.03	.03	.027	.033
Male	2,687	389.29*	138	.99	.98	.02	.03	.023	.029
Minority Status									
Minority	2,337	403.56*	138	.98	.98	.03	.03	.025	.032
Non-Minority	2,896	483.65*	138	.98	.98	.03	.03	.027	.032
SES									
Low SES	2,601	410.11*	138	.98	.98	.03	.03	.024	.031
High SES	2,632	452.75*	138	.98	.98	.03	.03	.026	.032
Class Time									
Less Time	2,841	423.79*	138	.99	.98	.03	.03	.024	.030
More Time	2,392	451.42*	138	.98	.98	.03	.03	.028	.034
Class Size									
Small Class	2,469	402.15*	138	.99	.98	.03	.03	.028	.031
Large Class	2,764	452.57*	138	.98	.98	.03	.03	.029	.032

Note. *n* = sample size; χ^2 = chi-square; *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker–Lewis Index; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; CI = confidence interval; LL = lower limit; UL = upper limit.
*$p < .001$.

The goodness-of-fit indices for the female gender group [χ^2 (138) = 453.52, $p < .001$, CFI = .98, TLI = .98, SRMR = .03, and RMSEA = .03 (90% CI = .027 – .033)] and the male gender group [χ^2 (138) = 389.29, $p < .001$, CFI = .99, TLI = .98, SRMR = .02, and RMSEA = .03 (90% CI = .023 – .029)] demonstrated very good fit to the data when the baseline model was tested separately with each model. Likewise, the goodness-of-fit indices, presented in Table 16 above, indicated good model fit when the baseline model was tested with the two minority groups, the two SES groups, the two class size groups, and the two class time groups. The good model fit of all 10 models provide evidence of the configural invariance of the five control variables. After it was confirmed that configural invariance was in place across the two groups for each control variable, the measurement and structural invariance tests for invariance of the factor loadings, intercepts, item uniqueness, and factor variances/covariances were conducted.

Under the step-up invariance constraints method, testing for measurement and structural invariance of the gender, minority status, SES, class size, and class time population groups was organized in a logical sequence of increasingly restrictive nested models. The statistical significance of the chi-square difference was examined at each step to assess whether the $\Delta\chi^2$ provided evidence to support the invariance of factor loadings, invariance of item intercepts, invariance of residual item variances/covariances, and the invariance factor variances/covariances.

The first tests of measurement and structural invariance were conducted with the two gender groups (0 = female, 1 = male). No invariance constraints were imposed on Model 0, or the baseline model, that were used to begin the testing for measurement invariance. The model fit to the data of Model 0, with all parameters freely estimated, was initially tested with both gender groups together, χ^2 (276) = 842.79, $p < .001$; CFI = .984, RMSEA = .03 (please see Table 17).

Table 17

Tests of Measurement and Structural Invariance Across Gender Groups

Model	χ^2	*df*	Model Comparison	$\Delta\chi^2$	Δdf	CV of Δdf	CFI	ΔCFI	RMSEA
0	842.79*	276					.984		.03
1	969.75*	292	M1 – M0	126.96*	16	26.30	.980	-.004	.03
1P	856.71*	282	M1P – M0	13.92	8	15.51	.983	-.001	.03
2	993.84*	298	M2 – M1P	137.13*	16	26.30	.980	-.003	.03
2P	872.99*	293	M2P – M1P	16.28	11	19.68	.983	.000	.03
3	1164.11*	312	M3 – M2P	291.12*	19	30.14	.975	-.008	.03
3P	881.93*	306	M3P – M2P	8.95	13	22.36	.983	.000	.03
4	888.26*	299	M4 – M2P	15.27*	6	12.59	.983	.000	.03
4P	878.45*	298	M4P – M2P	5.46	5	11.07	.983	.000	.03

Note. χ^2 = chi-square; *df* = degrees of freedom; $\Delta\chi^2$ = chi-square difference; Δdf = difference in degrees of freedom; CV = critical value of the chi-square distribution; CFI = comparative fit index; ΔCFI = difference in comparative fit index; RMSEA = root mean square error of approximation; M0 = Model 0; M1 = Model 1; M2 = Model 2; M3 = Model 3; M4 = Model 4; P = partially invariant.
*$p < .001$.

For the first test of measurement invariance, the test of weak measurement invariance (i.e., metric invariance) Model 1 was obtained from Model 0 by constraining all factor loadings equal across the two gender groups. The results of the chi-square difference test between Model 1 and Model 0, $\Delta\chi^2_{M1\text{-}M0}$, was statistically significant, $\Delta\chi^2$ (16) = 126.96, $p < .001$. Being that invariance was supported by a non-significant chi-square difference, the statistically significant $\Delta\chi^2$ between Model 0 and Model 1 indicated the factor loadings were not completely invariant across the two groups. Due to the statistically significant chi-square difference of two models, the modification indices were examined to determine how to best proceed with testing for partial invariance. Model 1 was modified by freely estimating the factor loadings of Item 5 and Item 8 one at a time. After freeing the factor loading of Item 8, the chi-square value for the new model (Model 1P) dropped from 969.75 to 856.71, which made the chi-square difference between Model 1P and Model 0, $\Delta\chi^2_{M1P\text{-}M0}$, no longer statistically significant, $\Delta\chi^2$ (8) = 13.92, $p > .05$. Judging by the guidelines, a root mean square error of approximation (RMSEA) $\leq .05$ indicated a good model fit (Dimitrov, 2012). Thus, the RMSEA = .03 of Model 1P was evidence of a good fit of the model to the data. Because Model 1P was nested within Model 0, the nonsignificant chi-square difference between these two models provided evidence that the factor loadings across the two gender groups were partially invariant (i.e., weak measurement invariance/partial metric invariance). The partially invariant factor loadings of Model 1P were also supported by the difference in the comparative fit index between Model 1P and Model 0, $\Delta CFI_{M1P\text{-}M0} = -.001$. The new model developed to test for the partial invariance of factor loadings, Model 1P, was used in the subsequent chi-square difference test with Model 2.

The next test of measurement invariance, the test of strong measurement invariance (i.e., metric and scalar invariance) entailed stepping up the constraints of Model 1P to create Model 2, in which all item intercepts were constrained as equal across the two gender groups. Since Model 2 was obtained from Model 1P, the factor loadings continued to be constrained as equal across the two gender groups. The chi-square difference between Model 2 and Model 1P, $\Delta\chi^2_{M2\text{-}M1P}$, was statistically significant, $\Delta\chi^2$ (16) = 137.13, $p < .001$, thus it was concluded the item intercepts across the two gender groups were not completely invariant. Modification indices were then examined to determine the modifications that should be made to the model to test for partial invariance. Beginning with the modification index with the largest value one item at a time was freed to vary across the two gender groups. The first modification of Model 2 involved creating Model 2P in which the intercept of Item 3was freely estimated. The resulting $\Delta\chi^2$ between Model 2P and Model 1P was statistically significant. Three additional items were freed to vary before for the $\Delta\chi^2$ between Model 2P and Model 1P, $\Delta\chi^2_{M2P\text{-}M1P}$, was non-significant, $\Delta\chi^2$ (11) = 16.28, $p > .05$. The RMSEA of .03 provided evidence of a good fit of model 2P to the data. The nonsignificant chi-square difference between the Model 2P and Model 1P indicated that in addition to partially invariant factor loadings, the item intercepts across gender groups were partially invariant (i.e., partial scalar invariance/strong measurement invariance). The comparative fit index difference between Model 2P and Model 1P, $\Delta CFI_{M2P\text{-}M1P}$ = .000, offered additional evidence to support the partially invariant factor loadings and partially invariant item intercepts of Model 2P. The new model developed to test for the partial invariance of item intercepts, Model 2P, was used in the following chi-square difference tests.

The third and final test of measurement invariance was the test of strict measurement invariance (i.e., metric invariance, scalar invariance, and uniqueness invariance). To conduct this test, Model 3 was obtained from Model 2P by constraining all residual item variances/covariances to be equal across the two gender groups. Model 3 was nested in model 2P, therefore, the factor loadings and item intercepts continued to be invariant across the gender groups. The $\Delta\chi^2$ difference between Model 3 and Model 2P, $\Delta\chi^2_{M3\text{-}M2P}$, was statistically significant, $\Delta\chi^2$ (19) = 291.12, $p < .001$, which signified that the residual item variances/covariances were not completely invariant across the two gender groups. In order to test for partial invariance of the residual item variances/

covariances, the modification indices were examined. The residual variance of Item 5 was allowed to vary across the two gender groups, which resulted in a statistically significant $\Delta\chi^2$ between Model 3P and Model 2P. After freeing four additional items, the chi-square value for the new model (Model 3P) dropped from 1164.11 to 881.93, and the chi-square difference between Model 3P and Model 2P, $\Delta\chi^2_{M3P\text{-}M2P}$, was no longer statistically significant, $\Delta\chi^2$ (13) = 8.95, $p > .05$. The root mean square error of approximation goodness-of-fit index of Model 2P (RMSEA = .03) provided evidence of a good fit of the model to the data. Further, evidence was provided by the nonsignificant chi-square difference between the Model 3P and Model 2P that in addition to partially invariant factor loadings and partially invariant item intercepts, the residual item variances/covariances across the two gender groups were partially invariant (i.e., partial invariance of item uniquenesses/strict measurement invariance). The partial invariance of factor loadings, partial invariance of item intercepts, and partial invariance of residual item variances/covariances of Model 3P were also supported by the difference in the comparative fit index between Model 3P and Model 2P, $\Delta CFI_{M3P\text{-}M2P} = .000$.

In order to test for structural invariance (i.e., invariance of factor loadings, item intercepts, and factor variances/covariances), Model 4 was developed from Model 2P by constraining all factor variances/covariances equal across the two gender groups. Like Model 3, Model 4 was nested in model 2P, therefore, factor loadings and item intercepts continued to be held equal across the two groups. The significant $\Delta\chi^2$ between Model 4 and Model 2P, $\Delta\chi^2_{M4\text{-}M2P}$, $\Delta\chi^2$ (6) = 15.27, $p < .001$, indicated that the factor variances/

covariances were not completely invariant across the gender groups. Due to the significant difference in the chi-square of the two models, partial invariance was tested.

As a result of the examination of the model modification indices, the variances/covariances of the reading attitudes factor was freely estimated in Model 4P. By freeing the variances/covariances of the reading attitudes factor, there was a drop in the chi-square value from 888.26 to 878.45, and the chi-square difference between Model 4P and Model 2P, $\Delta\chi^2_{M4P\text{-}M2P}$, was then non-significant, $\Delta\chi^2$ (5) = 5.46, $p > .05$. As with Models 1P, 2P, and 3P, support of a good model fit to the data was offered by the RMSEA = .03 of Model 4P. Being that Model 4P was nested within Model 2P, the nonsignificant chi-square difference between these two models offered support for the partial invariance of factor loadings, partial invariance of item intercepts, and partial invariance of factor variances/covariances across the gender groups (i.e., partial structural invariance). The partially invariant factor variances/covariances of Model 4P were also supported by the difference in the comparative fit index between Model 4P and Model 2P, $\Delta CFI_{M4P\text{-}M2P} = .000$. The results of the first set of tests indicated partial measurement invariance (partially invariant factor loadings, partially invariant intercepts, and partially invariant residual item variances/covariances) and partial structural invariance (invariant factor variances/covariances) across the two gender groups.

The next set of measurement and structural invariance tests were conducted with the two minority status groups (0 = minority, 1 = nonminority). Again, no invariance constraints were imposed on Model 0, or the baseline model, that were used to initiate the testing for measurement invariance. The model which was fit to the data of Model 0, with all parameters freely estimated, was tested first with both minority status groups together, χ^2 (276) = 887.20, p < .001; CFI = .983, RMSEA = .03 (please see Table 18). For the first test of measurement invariance, the test of weak measurement invariance (i.e., metric invariance) Model 1 was obtained from Model 0 by constraining all factor loadings equal across the two minority status groups.

Table 18

Tests of Measurement and Structural Invariance Across Minority Groups

Model	χ^2	*df*	Model Comparison	$\Delta\chi^2$	Δdf	CV of Δdf	CFI	ΔCFI	RMSEA
0	887.20*	276					.983		.03
1	920.55*	292	M1 – M0	33.35*	16	26.30	.983	.000	.03
1P	901.32*	286	M1P – M0	14.11	10	18.31	.983	.000	.03
2	1120.84*	302	M2 – M1P	219.52*	16	26.30	.977	-.006	.03
2P	916.01*	295	M2P – M1P	4.70	9	16.92	.983	.000	.03
3	1079.48*	314	M3 – M2P	163.47*	19	30.14	.979	-.004	.03
3P	920.85*	308	M3P – M2P	4.83	13	22.36	.983	.000	.03
4	938.69*	301	M4 – M2P	22.68*	6	12.59	.982	-.001	.03
4P	925.96*	300	M4P – M2P	9.95	5	11.07	.983	.000	.03

Note. χ^2 = chi-square; *df* = degrees of freedom; $\Delta\chi^2$ = chi-square difference; Δdf = difference in degrees of freedom; CV = critical value of the chi-square distribution; CFI = comparative fit index; ΔCFI = difference in comparative fit index; RMSEA = root mean square error of approximation; M0 = Model 0; M1 = Model 1; M2 = Model 2; M3 = Model 3; M4 = Model 4; P = partially invariant.
**p* < .001.

The results of the chi-square difference test between Model 1 and Model 0, $\Delta\chi^2_{M1-M0}$, was statistically significant, $\Delta\chi^2$ (16) = 33.35, p < .001. Since invariance is supported by a non-significant chi-square difference, the statistically significant $\Delta\chi^2$ between Model 0 and Model 1 indicated the factor loadings were not completely invariant across the two groups. Because of the statistically significant chi-square difference of two models, the modification indices were examined to determine how to best proceed with testing for partial invariance. Model 1 was modified by freely estimating the factor loading of Item 2. After freeing Item 2, the chi-square value for the new model (Model 1P) dropped from 920.55 to 901.32, which made the chi-square difference between Model 1P and Model 0, $\Delta\chi^2_{M1P\text{-}M0}$, no longer statistically significant, $\Delta\chi^2$ (10) = 14.11, $p > .05$.

As stated in the guidelines, a root mean square error of approximation (RMSEA) $\leq .05$ indicates a good model fit (Dimitrov, 2012). Thus, the RMSEA = .03 of Model 1P was evidence of a good fit of the model to the data. Since Model 1P was nested within Model 0, the nonsignificant chi-square difference between these two models provided evidence that the factor loadings across the two minority status groups were partially invariant (i.e., weak measurement invariance/partial metric invariance). Further, the partially invariant factor loadings of Model 1P were also supported by the difference in the comparative fit index between Model 1P and Model 0, $\Delta CFI_{M1P\text{-}M0} = .000$. Model 1P, which was developed to test for the partial invariance of factor loadings, was used in the following chi-square difference test with Model 2.

The second measurement invariance test, the test of strong measurement invariance, involved stepping up Model 1P constraints to create Model 2, where all item intercepts were constrained as equal across the minority status groups. Because Model 2 was developed from Model 1P, the factor loadings were still constrained as equal across the two minority status groups. The chi-square difference between Model 2 and Model 1P, $\Delta\chi^2_{M2\text{-}M1P}$, was significant, $\Delta\chi^2$ (16) = 219.52, $p < .001$, thus it was concluded the item intercepts across the two minority status groups were not completely invariant.

Modification indices were then examined to determine the modifications that should be made to the model in order to test for partial invariance. Beginning with the modification index with the largest value one item at a time was freed to vary across the two minority status groups. The first modification of Model 2 involved creating Model 2P in which the intercept of Item 5 was freely estimated. The resulting $\Delta\chi^2$ between Model 2P and Model 1P was statistically significant. Five additional items were freed to vary before for the $\Delta\chi^2$ between Model 2P and Model 1P, $\Delta\chi^2_{M2P\text{-}M1P}$, was non-significant, $\Delta\chi^2$ (9) = 4.70, $p > .05$. The RMSEA of .03 provided evidence of a good fit of model 2P to the data. The nonsignificant chi-square difference between the Model 2P and Model 1P indicates that in addition to partially invariant factor loadings, the item intercepts across the two minority status groups were partially invariant (i.e., partial scalar invariance/strong measurement invariance). The difference in the comparative fit index between Model 2P and Model 1P, $\Delta CFI_{M2P\text{-}M1P} = .000$ offers additional evidence in support of the partially invariant factor loadings and partially invariant item intercepts of Model 2P. The model created to test the partial invariance of item intercepts, Model 2P, was also utilized in the chi-square difference tests for Model 3 and Model 4.

The test of strict measurement invariance was the last test of measurement invariance. This test required using Model 2P to develop Model 3, which was accomplished by constraining all residual item variances/covariances as equal across the two minority status groups. The factor loadings and item intercepts were still invariant across the minority status groups since Model 3 was nested in model 2P. The $\Delta\chi^2$ difference between Model 3 and Model 2P, $\Delta\chi^2_{M3\text{-}M2P}$, was statistically significant, $\Delta\chi^2$

(19) = 163.47, $p < .001$, which signified that the residual item variances/covariances were not completely invariant across the two minority status groups.

In order to test for partial invariance of the residual item variances/covariances, the modification indices were examined. The residual variance of Item 5 was allowed to vary across the two minority status groups, which resulted in a statistically significant $\Delta\chi^2$ between Model 3P and Model 2P. After freeing four additional items, the chi-square value for the new model (Model 3P) dropped from 1079.48 to 920.85, and the chi-square difference between Model 3P and Model 2P, $\Delta\chi^2_{M3P\text{-}M2P}$, was no longer statistically significant, $\Delta\chi^2$ (13) = 4.83, $p > .05$. The root mean square error of approximation goodness-of-fit index of Model 3P (RMSEA = .03) provided evidence of a good fit of the model to the data. Further, evidence was provided by the nonsignificant chi-square difference between the Model 3P and Model 2P, that in addition to partially invariant factor loadings and partially invariant item intercepts, the residual item variances/ covariances across the two minority status groups were partially invariant (i.e., partial invariance of item uniquenesses/strict measurement invariance). The partial invariance of factor loadings, partial invariance of item intercepts, and partial invariance of residual item variances/covariances of Model 3P were also supported by the difference in the comparative fit index between Model 3P and Model 2P, $\Delta CFI_{M3P\text{-}M2P} = .000$.

Next, Model 4 was created from Model 2P by constraining all factor variances/ covariances equal across the two groups to test for structural invariance. Model 4 was nested in model 2P, like Model 3, thus, factor loadings and item intercepts needed to be held equal across the two minority status groups. The significant $\Delta\chi^2$ between Model 4 and Model 2P, $\Delta\chi^2_{M4\text{-}M2P}$, $\Delta\chi^2$ (6) = 22.68, $p < .001$, indicated that the factor variances/ covariances were not completely invariant across minority status groups. Because of the significant difference in the chi-square of the two models, partial invariance was tested.

As a result of the examination of the model modification indices, the variances/ covariances of the reading attitudes factor was freely estimated in Model 4P. By freeing the variances/covariances of the reading attitudes factor, there was a drop in the chi-square value from 938.69 to 925.96, and the chi-square difference between Model 4P and Model 2P, $\Delta\chi^2_{M4P\text{-}M2P}$, was then non-significant, $\Delta\chi^2$ (5) = 9.95, $p > .05$. As with Models 1P, 2P, and 3P, support of a good model fit to the data was offered by the RMSEA = .03 of Model 4P. Being that Model 4P was nested within Model 2P, the nonsignificant chi-square difference between these two models offered support for the partial invariance of factor loadings, partial invariance of item intercepts, and partial invariance of factor variances/covariances across the two groups (i.e., partial structural invariance). The partially invariant factor variances/covariances of Model 4P were also supported by the comparative fit index difference between Model 4P and Model 2P, $\Delta CFI_{M4P\text{-}M2P} = .000$.

Evidence to support partially invariant factor loadings, partially invariant intercepts, partially invariant residual item variances/covariances, and partially invariant factor variances/covariances across the two minority status groups was obtained from the results of the second group of measurement and structural invariance tests.

The third group of tests of measurement and structural invariance were conducted with the two SES groups (0 = low SES, 1 = high SES). In the initial test of measurement invariance, no invariance constraints were imposed on Model 0, or the baseline model.

The model fit to the data of Model 0, with all parameters freely estimated, was initially tested with both SES groups together, χ^2 (276) = 862.86, $p < .001$; CFI = .983, RMSEA = .03 (please see Table 19). For the first test of measurement invariance, the test of weak measurement invariance (i.e., metric invariance) Model 1 was obtained from Model 0 by constraining all factor loadings equal across the two SES groups. The results of the chi-square difference test between Model 1 and Model 0, $\Delta\chi^2_{M1\text{-}M0}$, was significant, $\Delta\chi^2$ (16) = 28.36, $p < .001$. Because invariance is supported by a non-significant chi-square difference, the significant $\Delta\chi^2$ between Model 0 and Model 1 indicated the factor loadings were not completely invariant across the two groups.

Table 19

Tests of Measurement and Structural Invariance Across SES Groups

Model	χ^2	*df*	Model Comparison	$\Delta\chi^2$	Δdf	CV of Δdf	CFI	ΔCFI	RMSEA
0	862.86*	276					.983		.03
1	891.22*	292	M1 – M0	28.36*	16	26.30	.983	.000	.03
1P	866.15*	291	M1P – M0	3.29	15	25.00	.984	.001	.03
2	962.55*	307	M2 – M1P	96.40*	16	26.30	.981	-.003	.03
2P	872.70*	302	M2P – M1P	6.55	11	19.68	.984	.000	.03
3	1030.32*	321	M3 – M2P	157.62*	19	30.14	.980	-.004	.03
3P	886.13*	315	M3P – M2P	13.43	13	22.36	.984	.000	.03
4	892.74*	308	M4 – M2P	20.04*	6	12.59	.983	-.001	.03

Note. χ^2 = chi-square; *df* = degrees of freedom; $\Delta\chi^2$ = chi-square difference; Δdf = difference in degrees of freedom; CV = critical value of the chi-square distribution; CFI = comparative fit index; ΔCFI = difference in comparative fit index; RMSEA = root mean square error of approximation; M0 = Model 0; M1 = Model 1; M2 = Model 2; M3 = Model 3; M4 = Model 4; P = partially invariant.
*$p < .001$.

Due to the statistically significant chi-square difference of two models, the modification indices were examined to determine how to best proceed with testing for partial invariance. Model 1 was modified by freely estimating the factor loading of Item 12. After freeing the factor loading of Item 12, the chi-square value for the new model (Model 1P) dropped from 891.22 to 866.15, which made the chi-square difference between Model 1P and Model 0, $\Delta\chi^2_{M1P\text{-}M0}$, no longer statistically significant, $\Delta\chi^2$ (15) = 3.29, $p > .05$. Based on the guidelines, a root mean square error of approximation (RMSEA) ≤ .05 indicates a good model fit (Dimitrov, 2012). Thus, the RMSEA = .03 of Model 1P was evidence of a good fit of the model to the data. Because Model 1P was nested within Model 0, the nonsignificant chi-square difference between these two models provided evidence that the factor loadings across the two SES groups were partially invariant (i.e., weak measurement invariance/partial metric invariance). The partially invariant factor loadings of Model 1P were also supported by the difference in the comparative fit index between Model 1P and Model 0, $\Delta CFI_{M1P\text{-}M0} = .001$. The new Model 1P was used in the subsequent chi-square difference test with Model 2.

The next test of measurement invariance, the test of strong measurement invariance, required stepping up the constraints of Model 1P to develop Model 2. All item intercepts in Model 2 were constrained as equal across the SES groups. Because Model 1P was used to create from Model 2, the factor loadings were still constrained as equal across the two SES groups. The chi-square difference between Model 2 and Model 1P, $\Delta\chi^2_{M2\text{-}M1P}$, was statistically significant, $\Delta\chi^2$ (16) = 96.40, $p < .001$, thus it was concluded the item intercepts across the two SES groups were not completely invariant.

Modification indices were then examined to determine the modifications that should be made to the model in order to test for partial invariance. Beginning with the modification index with the largest value one item at a time was freed to vary across the two SES groups. The first modification of Model 2 involved creating Model 2P in which the intercept of Item 10was freely estimated. The resulting $\Delta\chi^2$ between Model 2P and Model 1P was statistically significant. Three additional items were freed to vary before for the $\Delta\chi^2$ between Model 2P and Model 1P, $\Delta\chi^2_{M2P\text{-}M1P}$, was non-significant, $\Delta\chi^2$ (11) = 6.55, $p > .05$. The RMSEA of .03 provided evidence of a good fit of model 2P to the data.

The nonsignificant chi-square difference between the Model 2P and Model 1P indicated that in addition to partially invariant factor loadings, the item intercepts across the two SES groups were partially invariant (i.e., partial scalar invariance/strong measurement invariance). The difference in the comparative fit index between Model 2P and Model 1P, $\Delta CFI_{M2P\text{-}M1P}$ = .000 offers additional evidence in support of the partially invariant factor loadings and partially invariant item intercepts of Model 2P. The new model designed to test for the partial invariance of item intercepts, Model 2P, was utilized in the subsequent chi-square difference tests for Models 3 and 4.

The third measurement invariance test was the test of strict measurement invariance. For this test, Model 2P was used to created Model 3. This was done by constraining all residual item variances/covariances to be equal across the SES groups. Because Model 3 was nested in model 2P, the factor loadings and item intercepts remained invariant across the SES groups. The $\Delta\chi^2$ difference between Model 3 and

Model 2P, $\Delta\chi^2_{M3\text{-}M2P}$, was significant, $\Delta\chi^2$ (19) = 157.62, $p < .001$, which signified that the residual item variances/covariances were not completely invariant across the groups.

Next, the modification indices we examined in order to test for partial invariance of the residual item variances/covariances. The residual variance of Item 5 was allowed to vary across the two SES groups, which resulted in a statistically significant $\Delta\chi^2$ between Model 3P and Model 2P. After freeing four additional items, the chi-square value for the new model (Model 3P) dropped from 1030.32 to 886.13, and the chi-square difference between Model 3P and Model 2P, $\Delta\chi^2_{M3P\text{-}M2P}$, was no longer statistically significant, $\Delta\chi^2$ (13) = 13.43, $p > .05$. The root mean square error of approximation goodness-of-fit index of Model 3P (RMSEA = .03) provided evidence of a good fit of the model to the data. Further, evidence was provided by the nonsignificant chi-square difference between the Model 3P and Model 2P that in addition to partially invariant factor loadings and partially invariant item intercepts, the residual item variances/ covariances across the two SES groups were partially invariant (i.e., partial invariance of item uniquenesses/strict measurement invariance). The partial invariance of factor loadings, partial invariance of item intercepts, and partial invariance of residual item variances/covariances of Model 3P were also supported by the difference in the comparative fit index between Model 3P and Model 2P, $\Delta CFI_{M3P\text{-}M2P} = .000$.

The test for structural invariance required using Model 2P to develop Model 4 by constraining all factor variances/covariances to be equal across the two SES groups. As with Model 3, Model 4 was nested in model 2P, consequently, factor loadings and item intercepts remained equal across the two SES groups. The statistically significant $\Delta\chi^2$

between Model 4 and Model 2P, $\Delta\chi^2_{M4\text{-}M2P}$, $\Delta\chi^2$ (6) = 20.04, $p < .001$, indicated that the factor variances/covariances were not completely invariant across the two SES groups. Because of the significant difference in the chi-square of two models, partial invariance was attempted. The modification indices for Model 4 did not include factor variances/ covariances that could be freely estimated. Thus, it was not possible to test for partial structural invariance across groups, and structural invariance could not be confirmed.

The third set of tests, which examined the measurement and structural invariance across the two SES groups, offered evidence that supported partially invariant factor loadings, partially invariant intercepts, and partially invariant residual item variances/covariances across the two groups. Neither invariant nor partially invariant factor variances/covariances were supported by the evidence, thus structural invariance across the two SES groups was not confirmed. The lack of structural invariance is not a concern. As previously mentioned structural invariance is a useful, but not required, condition for comparing group mean differences because the equality of factor variances/covariances (structural invariance) is not necessary to define the relationships between the items and the factors (Wu et al., 2007).

The fourth set of measurement and structural invariance tests were conducted with the class time groups (0 = less class time, 1 = more class time). Invariance constraints were not imposed on Model 0, or the baseline model, used to initiate this set of tests. The model fit to the data of Model 0, with all parameters freely estimated, was initially tested with both class time groups together, χ^2 (276) = 875.21, $p < .001$; CFI = .983, RMSEA = .03 (please see Table 20). For the first test of measurement invariance, the test of weak measurement invariance (i.e., metric invariance), Model 1 was obtained from Model 0 by constraining all factor loadings equal across groups.

Table 20

Tests of Measurement and Structural Invariance Across Class Time Groups

Model	χ^2	*df*	Model Comparison	$\Delta\chi^2$	Δdf	CV of Δdf	CFI	ΔCFI	RMSEA
0	875.21*	276					.983		.03
1	897.52*	292	M1 – M0	22.31	16	26.30	.983	.000	.03
2	945.37*	308	M2 – M1	47.86*	16	26.30	.982	-.001	.03
2P	914.05*	306	M2P – M1	16.54	14	23.69	.983	.000	.03
3	961.36*	325	M3 – M2P	47.30*	19	30.14	.982	-.001	.03
3P	925.01*	323	M3P – M2P	10.96	17	27.59	.983	.000	.03
4	925.25*	312	M4 – M2P	11.19	6	12.59	.983	.000	.03

Note. χ^2 = chi-square; *df* = degrees of freedom; $\Delta\chi^2$ = chi-square difference; Δdf = difference in degrees of freedom; CV = critical value of the chi-square distribution; CFI = comparative fit index; ΔCFI = difference in comparative fit index; RMSEA = root mean square error of approximation; M0 = Model 0; M1 = Model 1; M2 = Model 2; M3 = Model 3; M4 = Model 4; P = partially invariant.
$^*p < .001$.

The results of the chi-square difference test between Model 1 and Model 0, $\Delta\chi^2_{M1-M0}$, was not significant, $\Delta\chi^2$ (16) = 22.31, $p > .05$. The RMSEA = .03 of Model 1 was further evidence of a good fit of the model to the data. Because Model 1 was nested within Model 0, the nonsignificant chi-square difference between the models provided evidence that the factor loadings across groups were invariant (i.e., weak measurement invariance/metric invariance). The comparative fit index difference between Model 1 and Model 0, $\Delta\text{CFI}_{M1-M0} = .000$, also supported the invariant factor loadings of Model 1.

The test of strong measurement invariance was the second test of measurement invariance. This test involved stepping up the constraints of Model 1P to create Model 2, where all item intercepts were constrained as equal across the two class time groups. Since Model 2 was developing using Model 1P, the factor loadings continued to be constrained as equal across the two class time groups. The chi-square difference between Model 2 and Model 1, $\Delta\chi^2_{M2-M1}$, was statistically significant, $\Delta\chi^2$ (16) = 47.86, $p < .001$, thus it was concluded the item intercepts across the two class time groups were not completely invariant. Subsequently, the modification indices were examined to determine the modifications that should be made in order to test for partial invariance. One item at a time, the intercepts for Item 18 and Item 8, respectively, were set free to vary across the two class time groups.

The first modification of Model 2 involved creating Model 2P in which the intercept of Item 18 was freely estimated. The resulting $\Delta\chi^2$ between Model 2P and Model 1P was significant. After freeing the intercept for Item 8, the chi-square value for Model 2P dropped from 945.37 to 914.05 and the chi-square difference between Model 2P and Model 1P, $\Delta\chi^2_{M2P-M1P}$, was no longer significant, $\Delta\chi^2$ (14) = 16.54, $p > .05$. Further, the RMSEA of .03 provided evidence of a good fit of model 2P to the data.

The nonsignificant chi-square difference between the Model 2P and Model 1P indicated that in addition to invariant factor loadings, the item intercepts across the two class time groups were partially invariant (i.e., partial scalar invariance/strong measurement invariance). The difference in the comparative fit index between Model 2P and Model 1P, $\Delta CFI_{M2P-M1P}$ = .000 offers additional evidence in support of the invariant factor loadings and partially invariant item intercepts of Model 2P. Model 2P, the model developed to test for the partial invariance of item intercepts, was later used in the chi-square difference tests for Models 3 and 4.

The test of strict measurement invariance was the final measurement invariance test. To carry out this third test, Model 3 was developed from Model 2P by constraining all residual item variances/covariances to be equal across groups. The factor loadings and item intercepts stayed invariant across the class time groups since Model 3 was nested in model 2P. The $\Delta\chi^2$ difference between Model 3 and Model 2P, $\Delta\chi^2_{M3\text{-}M2P}$, was statistically significant, $\Delta\chi^2$ (19) = 47.30, $p < .001$, which signified that the residual item variances/covariances were not completely invariant across the two class time groups.

In order to test for partial invariance of the modification indices were examined and Item 19 was allowed to vary across the two class time groups. This caused the chi-square value for Model 3P to drop from 961.36 to 925.01, and the chi-square difference between Model 3P and Model 2P, $\Delta\chi^2_{M3P\text{-}M2P}$, was no longer statistically significant, $\Delta\chi^2$ (17) = 10.96, $p > .05$. The RMSEA = .03 of Model 3P provided evidence of a good fit of the model to the data. Further, evidence was provided by the nonsignificant chi-square difference between the Model 3P and Model 2P, that in addition to invariant factor loadings and partially invariant item intercepts, the residual item variances/covariances across the two class time groups were partially invariant (i.e., partial invariance of item uniquenesses/strict measurement invariance). The invariance of factor loadings, partial invariance of item intercepts, and partial invariance of residual item variances/ covariances of Model 3P were also supported by the difference in the comparative fit index between Model 3P and Model 2P, $\Delta CFI_{M3P\text{-}M2P} = .000$.

By constraining all the factor variances/covariances in Model 2P to be equal across the two class time groups, Model 4 was created to conduct the test of structural invariance. Model 4 was nested in model 2P, like Model 3, as such, factor loadings and item intercepts continued to be held equal across the two class time groups. The non-significant $\Delta\chi^2$ between Model 4 and Model 2P, $\Delta\chi^2_{M4\text{-}M2P}$, $\Delta\chi^2$ (6) = 11.19, $p > .05$, indicated that the factor variances/covariances were invariant across the two class time groups. As with Models 1P, 2P, and 3P, support of a good model fit to the data was offered by the RMSEA = .03 of Model 4. Being that Model 4 was nested within Model 2P, the nonsignificant chi-square difference between the two models offered support for the invariance of factor loadings, partial invariance of item intercepts, and invariance of factor variances/covariances across the class time groups (i.e., structural invariance). The invariant factor variances/covariances of Model 4 were also supported by the difference in the comparative fit index between Model 4 and Model 2P, $\Delta CFI_{M4\text{-}M2P} = .000$. In conclusion, evidence to support invariant factor loadings, partially invariant intercepts, partially invariant residual item variances/covariances, and invariant factor variances/covariances across the two class time groups was provided by the results of the last group of measurement and structural invariance tests.

Finally, the last set of measurement and structural invariance tests were conducted with the two class size groups (0 = small class size, 1 = large class size). Like the four previous tests, no invariance constraints were imposed on Model 0, or the baseline model, that were used in the test of measurement invariance. The model fit to the data of Model 0, with all parameters freely estimated, was initially tested with both class size groups together, χ^2 (276) = 854.72, $p < .001$; CFI = .984, RMSEA = .03 (please see Table 21). For the first test of measurement invariance, the test of weak measurement invariance (i.e., metric invariance) Model 1 was obtained from Model 0 by constraining all factor loadings equal across the two class size groups. The results of the chi-square difference test between Model 1 and Model 0, $\Delta\chi^2_{M1\text{-}M0}$, was statistically significant, $\Delta\chi^2$ (16) = 49.84, $p < .001$. Since invariance is supported by a non-significant chi-square difference, the statistically significant $\Delta\chi^2$ between Model 0 and Model 1 indicated the factor loadings were not completely invariant across the two groups.

Table 21

Tests of Measurement and Structural Invariance Across Class Size Groups

Model	χ^2	*df*	Model Comparison	$\Delta\chi^2$	Δdf	CV of Δdf	CFI	ΔCFI	RMSEA
0	854.72*	276					.984		.03
1	904.56*	292	M1 – M0	49.84*	16	26.30	.983	-.001	.03
1P	865.03*	290	M1P – M0	10.31	14	23.69	.984	.000	.03
2	926.62*	306	M2 – M1P	61.59*	16	26.30	.983	-.001	.03
2P	872.81*	303	M2P – M1P	7.78	13	22.36	.984	.000	.03
3	904.94*	322	M3 – M2P	32.13*	19	30.14	.984	.000	.03
3P	894.05*	321	M3P – M2P	21.25	18	28.87	.984	.000	.03
4	905.58*	309	M4 – M2P	32.78*	6	12.59	.983	-.001	.03
4P	882.72*	308	M4P – M2P	9.91	5	11.07	.984	.000	.03

Note. χ^2 = chi-square; *df* = degrees of freedom; $\Delta\chi^2$ = chi-square difference; Δdf = difference in degrees of freedom; CV = critical value of the chi-square distribution; CFI = comparative fit index; ΔCFI = difference in comparative fit index; RMSEA = root mean square error of approximation; M0 = Model 0; M1 = Model 1; M2 = Model 2; M3 = Model 3; M4 = Model 4; P = partially invariant. *$p < .001$.

Due to the statistically significant chi-square difference of two models, the modification indices were examined to determine how to best proceed with testing for partial invariance. Model 1 was modified by freely estimating the factor loadings of Item 12 and Item 9 one at a time. After freeing the factor loading of Item 8, the chi-square value for the new model (Model 1P) dropped from 904.56 to 865.03, which made the chi-square difference between Model 1P and Model 0, $\Delta\chi^2_{\text{M1P-M0}}$, no longer statistically significant, $\Delta\chi^2$ (14) = 10.31, $p > .05$. Based on the guidelines, a root mean square error of approximation (RMSEA) $\leq .05$ indicates a good model fit (Dimitrov, 2012). Thus, the RMSEA = .03 of Model 1P was evidence of a good fit of the model to the data. Because Model 1P was nested within Model 0, the nonsignificant chi-square difference between these two models provided evidence that the factor loadings across the two class size groups were partially invariant (i.e., weak measurement invariance/partial metric invariance). The partially invariant factor loadings of Model 1P were also supported by the difference in the comparative fit index between Model 1P and Model 0, $\Delta\text{CFI}_{\text{M1P-M0}}$ = .000. Model 1P, the new model developed to test for the partial invariance of factor loadings, was used in the following chi-square difference test with Model 2.

The next test of measurement invariance, the test of strong measurement invariance consisted of stepping up the constraints of Model 1P to create Model 2, in which all item intercepts were constrained as equal across the two class size groups. Because Model 2 was created from Model 1P, the factor loadings were still constrained as equal across the two class size groups. The chi-square difference between Model 2 and Model 1P, $\Delta\chi^2_{\text{M2-M1P}}$, was significant, $\Delta\chi^2$ (16) = 61.59, $p < .001$, which indicated that the item intercepts across the two class size groups were not completely invariant.

The modifications that should be made to the model to test for partial invariance were then determined by an examination of the model modification indices. Beginning with the modification index with the largest value one item at a time was freed to vary across the two class size groups. The first modification of Model 2 involved creating Model 2P in which the intercept of Item 5 was freely estimated. The resulting $\Delta\chi^2$ between Model 2P and Model 1P was statistically significant. One additional item was freed to vary before for the $\Delta\chi^2$ between Model 2P and Model 1P, $\Delta\chi^2_{M2P\text{-}M1P}$, was non-significant, $\Delta\chi^2$ (13) = 7.78, $p > .05$. The RMSEA of .03 provided evidence of a good fit of model 2P to the data. This nonsignificant chi-square difference indicated that the item intercepts across the two class size groups were partially invariant (i.e., partial scalar invariance/strong measurement invariance). The difference in the comparative fit index between Model 2P and Model 1P, $\Delta CFI_{M2P\text{-}M1P}$ = .000 offers additional evidence in support of the partially invariant factor loadings and partially invariant item intercepts of Model 2P. Model 2P, which was created to test the partial invariance of item intercepts, was also used for chi-square difference tests of Model 3 and Model 4.

The last test of measurement invariance was the test of strict measurement invariance. The first step in this test required using Model 2P to create Model 3, which was achieved by constraining all residual item variances/covariances as equal across the groups. Model 3 was nested in model 2P, thus, the factor loadings and item intercepts stayed invariant across the class size groups. The $\Delta\chi^2$ difference between Model 3 and Model 2P, $\Delta\chi^2_{M3\text{-}M2P}$, was significant, $\Delta\chi^2$ (19) = 32.13, $p < .001$, which signifying that the residual item variances/covariances were not completely invariant across groups.

In order to test for partial invariance of the residual item variances/covariances, the modification indices were examined. Model 3 was modified by freely estimating the residual variance of Item 5. After freeing this residual variance, the chi-square value for the new model (Model 3P) dropped from 904.94 to 894.05, which made the chi-square difference between Model 3P and Model 2P, $\Delta\chi^2_{M3P\text{-}M2P}$, no longer significant, $\Delta\chi^2$ (18) = 21.25, $p > .05$. The root mean square error of approximation goodness-of-fit index of Model 3P (RMSEA = .03) provided evidence of a good fit of the model to the data.

Further, evidence was provided by the nonsignificant chi-square difference between the Model 3P and Model 2P, that in addition to partially invariant factor loadings and partially invariant item intercepts, the residual item variances/covariances across the two class size groups were partially invariant (i.e., partial invariance of item uniquenesses/strict measurement invariance). The partial invariance of factor loadings, partial invariance of item intercepts, and partial invariance of residual item variances/covariances of Model 3P were also supported by the difference in the comparative fit index between Model 3P and Model 2P, $\Delta CFI_{M3P\text{-}M2P}$ = .000.

To conduct the test for structural invariance, Model 2P was used to develop Model 4 by constraining all factor variances/covariances equal across the groups. Like Model 3, Model 4 was nested in model 2P, thus, factor loadings and item intercepts were required to be held equal across the two class size groups. The significant $\Delta\chi^2$ between Model 4 and Model 2P, $\Delta\chi^2_{M4\text{-}M2P}$, $\Delta\chi^2$ (6) = 32.78, $p < .001$, indicated that the factor variances/covariances were not completely invariant across the two class size groups.

Because of the statistically significant difference in the chi-square of the two models, partial invariance was tested. Because of the examination of the model modification indices, the reading comprehension strategy use factor was freely estimated in Model 4P. By freeing this factor, there was a drop in the chi-square value from 905.58 to 882.72, and the chi-square difference between Model 4P and Model 2P, $\Delta\chi^2_{M4P-M2P}$, was then non-significant, $\Delta\chi^2$ (5) = 9.91, $p > .05$. As with Models 1P, 2P, and 3P, support of a good model fit to the data was offered by the RMSEA = .03 of Model 4P. Since Model 4P was nested within Model 2P, the nonsignificant chi-square difference between the two models supported the partial invariance of factor loadings, partial invariance of item intercepts, and partial invariance of factor variances/covariances across the two class size groups (i.e., partial structural invariance). The partially invariant factor variances/covariances of Model 4P were also supported by the difference in the comparative fit index between Model 4P and Model 2P, $\Delta CFI_{M4P-M2P} = .000$. The results of the fifth set of tests for measurement and structural invariance provided evidence to support the partially invariant factor loadings, partially invariant intercepts, partially invariant residual item variances/covariances, and partially invariant factor variances/covariances across the two class size groups.

In conclusion, the results of testing for factorial invariance showed invariance of form (configural invariance) and structural invariance (partially invariant factor loadings, partially invariant intercepts, partially invariant residual item variances/covariances, and partially invariant factor variances/covariances) across the gender, minority status, and class size groups. Additionally, although the factorial invariance test results did not support the structural invariance of the SES groups, configural invariance, partial invariance of factor loadings, partial invariance of item intercepts, and partially invariance of residual item variances/covariances were supported. Finally, the results of testing for factorial invariance across class time groups provided evidence of configural and structural invariance (invariant factor loadings, partially invariant intercepts, partially invariant residual item variances/covariances, and invariant factor variances/covariances).

The results of testing for configural, measurement, and structural invariance, of the gender, minority status, SES, class size, and class time variables confirmed that the MIMIC assumption of factorial invariance (invariant factor loadings, invariant intercepts, invariant item residual variances and covariance, and invariant factor variances/ covariances) or partial factorial invariance was met for each variable. Thus, it was permissible to test for mean group differences in response to the third research question, 'Are there significant group differences in high school students' reading comprehension achievement, reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes across the five student level (i.e., gender, minority status, and SES) and school level (i.e., class time and class size) variables?'.

MIMIC model testing and parameter estimation. The Multiple Indicator, Multiple Cause (MIMIC) group code modeling was employed to test the mean differences between each of the two gender, minority status, SES, class time, and class size groups in reading comprehension achievement and the three reading constructs. The results of the MIMIC group differences tests are presented in Figure 9.

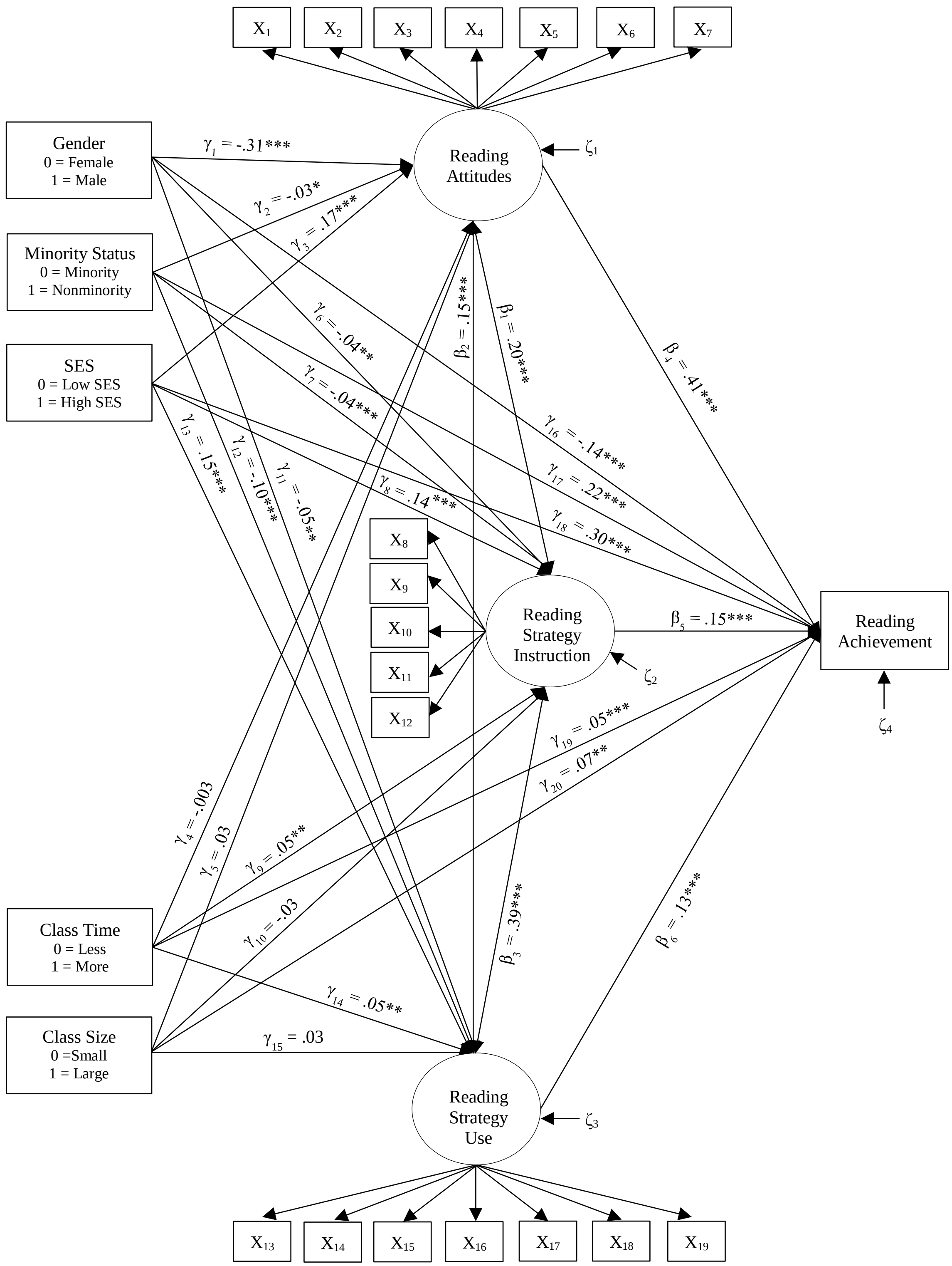


Figure 9. Final MIMIC model.
**p* < .05. ***p* < .01. ****p* < .001.

A joint examination of the goodness-of-fit indices suggested a good model fit to the data, CFI = .96, TLI = .95, SRMR = .03, RMSEA = .03 (90% CI = .030-.033), aside from the significant chi-square, $\chi^2(234) = 1426.79$, $p < .001$, which was overlooked because of the large sample size in this study. Additional support for good fit of the MIMIC model was offered by the standardized residuals, ranging from 0.34 to 0.98, which were well within the recommended range.

All structural coefficients in the MIMIC model, presented in Figure 9, are standardized. Thus, the two-way paths (i.e., two-way arrows) signify correlations between the variable they connect. The structural coefficients corresponding to the two-way paths are the correlation estimates.

The following positively correlated relationships are depicted in Figure 9: the relationship between reading comprehension achievement and reading comprehension strategy use ($\beta_6 = .13$, $p < .001$), the relationship between reading achievement and reading comprehension strategy instruction ($\beta_5 = .15$, $p < .001$), the relationship between reading achievement and reading attitudes ($\beta_4 = .41$, $p < .001$), the relationship between reading strategy use and reading strategy instruction ($\beta_3 = .39$, $p < .001$), the relationship between reading comprehension strategy use and reading attitudes ($\beta_2 = .15$, $p < .001$), and the relationship between reading attitudes and reading comprehension strategy instruction ($\beta_1 = .20$, $p < .001$). The magnitude of the positive correlations among the three latent factors are consistent with the results of the confirmatory factor analysis.

The findings of the MIMIC analyses of group differences in high school students' reading comprehension achievement indicated statistically significant mean differences for all the student and school level control variables in the model (gender, minority status, SES, class size, and class time). Gender differences in high school students' reading comprehension achievement were suggested by the significant structural path coefficient for the path from gender to achievement ($\gamma_{16} = -.14$, $p < .001$). Based on the negative sign of the path coefficient and the designated group coding (0 = female, 1 = male), the MIMIC analysis results meant that female high school students demonstrated higher reading comprehension achievement than their male peers. Based on Cohen's (1992) guidelines for interpreting the magnitude of a *d* effect size index (small effect = .20, moderate effect = .50, and large effect = .80), the effect size, $d = .31$, signified a small to moderate magnitude of the gender differences in reading comprehension achievement. Table 22 presents the absolute values of the statistically significant unstandardized structural coefficients and factor residual variances.

The statistically significant structural coefficient for the path from minority status to reading achievement showed that there were also differences in high school students' reading achievement based on their minority status ($\gamma_{17} = .22$, $p < .001$). Moreover, based on the ascribed group coding (0 = minority, 1 = nonminority), the positive path coefficient signified that the reading comprehension achievement of the nonminority students was higher than the minority students. Considering the effect size, $d = .50$, the magnitude of the achievement difference between the groups of students was moderate.

The standardized structural coefficient for the path from SES to reading achievement was also statistically significant, which indicated differences in achievement between the two SES groups ($\gamma_{18} = .30$, $p < .001$). Based on the positive path coefficient and the coding for the two SES groups (0 = low SES, 1 = high SES), the reading comprehension achievement of the high SES group was greater than the achievement of the low SES group. The magnitude of the effect size difference, *d* = .67, in reading achievement between the low and high SES groups was moderate to large.

Table 22

Structural Coefficients, Residual Variances, and Effect Sizes for the Group Mean Differences in Reading Achievement and the Three Reading Constructs

Construct/ Groups	Structural Coefficient	Residual Variance	Effect Size
Reading Achievement			
Gender	27.70	8117.89	.31
Minority Status	45.21	8117.89	.50
SES	60.76	8117.89	.67
Class Time	9.05	8117.89	.10
Class Size	13.49	8117.89	.15
Reading Strategy Use			
Gender	.05	.28	.10
Minority Status	.05	.28	.21
SES	.15	.28	.30
Class Time	.06	.28	.11
Reading Strategy Instruction			
Gender	.03	.12	.09
Minority Status	.07	.12	.09
SES	.10	.12	.28
Class Time	.04	.12	.11
Reading Attitudes			
Gender	.49	.53	.67
Minority Status	.05	.53	.07
SES	.27	.53	.38

Note. The unstandardized structural coefficients are shown as absolute values.

According to the statistically significant structural coefficient for the path from class time to reading comprehension achievement (γ_{19} = .05, *p* = .001). Given the group coding for the two groups, (0 = less class time, 1 = more class time), the positive sign of this coefficient intimated that the reading comprehension achievement of students who reported having more class time was higher than the achievement of students who reported having less class time. The mean difference effect size, *d* = .10, meant that the magnitude of the reading achievement class time group difference was small.

Finally, the results of the MIMIC group mean analysis revealed differences in high school students' reading achievement based on the size of their reading classes. Considering the coding for the two class size groups (0 = small class size, 1 = large class size), the positive sign of the path from class size to reading comprehension achievement, (γ_{20} = .07, *p* < .001), signified that high school students in large classes outperformed students in small classes. The effect size, *d* = .15, indicated a small magnitude in the difference between the reading achievement of the students in the small and large classes.

The tests of group difference in reading comprehension strategy use showed significant differences in strategy across the gender, minority status, SES, and class time groups. The structural coefficient for the path from gender to reading comprehension strategy use was significant, which suggested that the two gender groups differed on their reported use of reading comprehension strategies (γ_{11} = -.05, *p* = .01). Based on the group coding (0 = female, 1 = male), the negative coefficient indicated that female students reported using strategies more frequently than male students. Like the gender group differences in reading strategy instruction, the effect size magnitude of the gender group difference in the reading comprehension strategy use, $d = .10$, was also small.

The structural path coefficient for the path from minority status to reading strategy use ($\gamma_{12} = -.10$, $p < .001$) was also statistically significant. This means that minority and nonminority students also differed in terms of their reported use of reading strategies. Based on the coding for the two groups (0 = minority, 1 = nonminority), the finding that minority students reported using reading strategies more frequently than nonminority group was supported by the negative sign of the structural path coefficient. The effect size of this difference, $d = .21$, signified that the differing frequency of reading strategy use between minority and nonminority students a small.

According to the statistically significant path coefficient for the path from SES to reading comprehension strategy use ($\gamma_{13} = .15$, $p < .001$), there were also SES differences in the frequency of reading comprehension strategy use reported by high school students. The positive sign of this path coefficient along with the SES group coding (0 = low SES, 1 = high SES) suggesting that students with a high SES reported using reading strategies more often than students with a low SES. The mean difference effect size, $d = .30$, meant that the magnitude of the reported difference in strategy use between the two groups of students was small to moderate.

While there were no significant class size group differences in reading strategy use based on the results of the MIMIC analysis, ($\gamma_{15} = .03$, $p = .11$), the significant standardized structural coefficient for the path from class time to reading strategy use showed that there were class time group differences in strategy use ($\gamma_{14} = .05$, $p = .003$).

Given the group coding for the two groups, (0 = less class time, 1 = more class time), the positive sign of this coefficient indicated that students who reported having more reading comprehension strategy instructional class time also reported using reading strategies more frequently than the students with less class time. The effect size, $d = .11$, indicated that the practical significance of the difference in reading strategy use according the students reported amount of class time was small.

Like the group difference in high school students' reading comprehension strategy use, the results of the MIMIC analyses of group differences in reading comprehension strategy instruction revealed statistically significant differences in the gender, minority status, SES, and class time groups. The statistically significant standardized structural coefficient for the path from gender to reading comprehension strategy instruction denoted gender difference in high school reading strategy instruction ($\gamma_6 = -.04$, $p = .01$).

Based on the coding for the two gender groups (0 = female, 1 = male) and the negative sign of the path coefficient, female high school students reportedly received more reading comprehension strategy instruction than males. Although the results were statistically significant, the effect size, $d = .09$, signified that the magnitude of the difference in reading strategy instruction reported by female and male students was small.

Similarly, there were differences in the reading strategy instruction reported by minority and nonminority high school students. The standardized structural coefficient for the path from minority status to reading comprehension strategy instruction was statistically significant, signifying minority status groups difference in reading strategy instruction ($\gamma_7 = -.04$, $p = .01$). Minority students reported that they received more strategy instruction than nonminority students, judging by the negative sign of the standardized structural coefficient and the group coding (0 = minority, 1 = nonminority). The magnitude of the effect size, *d* = .09, indicated that the difference in strategy instruction reported by the nonminority and minority groups was small.

Further, SES group difference in reading comprehension strategy instruction were suggested by the significant structural coefficient for the path from SES to reading strategy instruction (γ_8 = .14, p < .001). Given the group coding for the two SES groups, (0 = low SES, 1 = high SES), the positive sign of this coefficient revealed that students with a high SES reportedly receive more reading comprehension strategy instruction than students in the low SES group. The effect size, *d* = .28, meant that a small to moderate magnitude in the difference strategy instruction reported by the high and low SES groups.

According to the significant coefficient for the path from class time to reading comprehension strategy instruction (γ_9 = .05, p = .001), there were also differences in the reported amount of reading strategy instruction between the two class time groups. The designated group coding for the two groups (0 = less class time, 1 = more class time) and the positive sign of this path coefficient intimated that the students who reported having more class time also reported receiving more reading strategy instruction compared to the students who reported having less class time. The mean difference effect size, *d* = .11, indicated that the magnitude of the difference in strategy instruction based on the amount of class time reported by students was small. Further, the one control variable for which there were no significant group differences in high school reading comprehension strategy instruction was the school level variable, class size (γ_{10} = -.03, p = .10).

Finally, the MIMIC analysis results related to the students' reading attitudes revealed statistically significant gender, minority status, and SES group differences. The standardized structural coefficient for the path from gender to reading attitudes was statistically significant, which showed that the two gender groups differed on reading attitudes ($\gamma_1 = -.31$, $p < .001$). Given the coding for the two gender groups, (0 = female, 1 = male), the negative sign of the path coefficient signified that female students reported more positive reading attitudes than male students. The effect size of the reading attitude gender group differences was moderate to large, $d = .67$.

The significant standardized structural coefficient for the path from minority status to reading attitudes revealed differences in high school students' attitudes toward reading based on their race/ethnicity ($\gamma_2 = -.03$, $p = .04$). Considering the assigned group coding (0 = minority, 1 = nonminority), the negative sign of the path coefficient indicated that minority students reported more positive attitudes toward reading than nonminority students. The mean difference effect size, $d = .07$, indicated that the magnitude of the difference in reading attitudes between minority and nonminority students was small.

Further, the statistically significant path coefficient for the path from SES to reading attitudes ($\gamma_3 = .18$, $p < .001$), signified differences in reading attitudes between the students with a high SES and the students with a low SES. The positive sign of this path coefficient along with the SES group coding (0 = low SES, 1 = high SES) meant that the students with a high SES reported more positive attitudes toward reading than the students with a low SES. The magnitude of the effect size, $d = .38$, of the difference between the reading attitudes of the low and high SES groups was small to moderate.

Lastly, the results of the MIMIC group mean analysis of class time and class size group differences in high school students' attitudes toward reading were not statistically significant ($\gamma_{4} =$ -.003, p = .90 and $\gamma_{5} =$ .03, p = .06, respectively). This suggests that there were no significant differences in high school students' reported attitudes toward reading related to reading comprehension strategy instructional class time or class sizes.

In conclusion, 16 of the 20 hypothesized group differences in high school students' reading comprehension achievement, reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes across the five student and school level variables were confirmed by the results of the MIMIC analysis. All four of the hypothesized gender, minority status, and SES group differences were statistically significant. The hypothesized class time group differences in students' reading comprehension strategy use, strategy instruction, and reading comprehension achievement were also statistically significant. However, there were no class time group differences in students' attitudes toward reading. The only significant class size group differences were observed for differences in high school students' reading comprehension achievement. Class size group differences in reading comprehension strategy use, strategy instruction, and reading attitudes were not statistically significant.

Chapter Five

Most educational reform efforts in the US are intended to increase student achievement, in order to minimize academic achievement gaps in reading and mathematics (Lang et al., 2009). Despite the ongoing national attention and ardent goals designed to improve reading proficiency, on average, more than 60% of high school seniors in the US continue to struggle to read (NCES, 2017). Further, over 1.5 million college freshmen are unprepared for the challenging requirements of advanced college reading (NCES, 2014). As a result of this prevalent reading failure, the overriding purpose of this study was to investigate various relationships among reading related variables to potentially discover new avenues of exploration that may conceivably lead to improvements in the reading comprehension proficiency of U.S. high school students.

The main purpose of this study was to examine the relationships among U.S. high school students' reading comprehension strategy use, reading comprehension strategy instruction, attitudes toward reading, and their reading comprehension achievement (as measured by PISA 2009 U.S. Reading Literacy achievement scores). Additionally, because evidence from prior empirical research posited that gender, minority status, SES, class time, and class size each have a significant relationship either with one of the latent constructs and/or reading comprehension achievement, they were also included as the student and school level control variables in the SEM model. Finally, because of the inconsistent and/or incomplete evidence in the reading research literature defining these relationships, group differences in high school students' reading comprehension achievement, reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes across the five student and school level variables (gender, minority status, SES, class size, and class time) were investigated with the intent of developing a more robust understanding of these relationships.

Due to the extensive number of results obtained in this study, a summary of the findings is provided first in this chapter. The summary is followed by a more elaborate discussion of results as well as the various implications of the study findings. Finally, the limitations of this study and suggested areas of future research are presented.

Discussion

As indicated by the preliminary reliability analyses, the reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes scales all demonstrated moderate to high levels of reliability according to the Cronbach's alpha and latent variable modeling tests of internal consistency reliability. The results related to the first research question, 'Do the observable indicators selected to measure the three latent constructs appropriately define them?', revealed statistically significant relationships between each indicator and its respective latent factor. As hypothesized, the results indicated that reading strategy use was accurately defined by seven indicators, reading comprehension strategy instruction was accurately defined by five indicators, and reading attitudes was accurately defined by seven indicators. The results of the confirmatory factor analysis, employed to examine this question, also revealed that the three latent constructs were all positively correlated.

Structural regression analysis was used to examine the second research question, ‘Are there direct and indirect effects among the five student and school level variables, the three latent variables, and high school students’ reading comprehension achievement?’. The second research question consisted of three subquestions. The results of research subquestion 2a, ‘What are the direct and indirect effects of the three latent constructs and the five student (i.e., gender, minority status, and SES) and school (i.e., class time and class size) level variables on high school students’ reading comprehension achievement?’, obtained via structural regression analysis, indicated that except for gender, all of the other variables in the model were statistically significant predictors of high school student’ reading comprehension achievement.

The structural regression analysis of indirect effects on high school students’ reading comprehension achievement revealed that in this model many of the relationships were mediated by reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes. Specifically, the following indirect relationships were found to be significant: the relationship between reading comprehension strategy instruction and reading achievement was mediated by both reading comprehension strategy use and reading attitudes, the relationship between gender and reading achievement was mediated by reading strategy use and reading attitudes, the relationship between minority status and reading achievement was mediated by strategy use, the relationship between SES and reading achievement was mediated by reading attitudes and strategy use, and finally, the relationship between class time and reading achievement was mediated by reading strategy instruction and reading attitudes.

The results of research subquestion 2b, ‘What are the direct effects among the three latent constructs (i.e., reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes)?’, suggested that strategy instruction and reading attitudes had a direct effect on reading strategy use, reading strategy instruction had a direct effect on reading attitudes, and as anticipated, neither strategy use nor reading attitudes had a direct effect on strategy instruction.

Finally, the results of research subquestion 2c, ‘What are the direct effects of the five student and school level variables on the three latent constructs?’, revealed that minority status, SES, class time, and class size had statistically significant positive direct effects on reading comprehension strategy use. The results also revealed that the effect of gender on reading comprehension strategy use was not statistically significant. While class time had a statistically significant direct effect on reading comprehension strategy instruction, the effect of class size was insignificant. Finally, gender and SES had statistically significant direct effects on students’ reading attitudes. However, the effects of students’ minority status on their attitudes toward reading were not significant.

The R^2 value results discussed in this section are part of the results of the second research question. This section was separated from the discussion of the subquestions of the second research question because the R^2 value results were associated with both subquestions 2b and 2c, and need to be discussed together. For example, it was hypothesized that reading comprehension strategy instruction, gender, minority status, and SES each had a direct effect on reading attitudes. While research subquestion 2b involved the hypothesized relationship between instruction and attitudes, the hypothesized relationship between gender and attitudes was related to research subquestion 2c. Since the structural regression model estimated the hypothesized relationships related to the reading comprehension strategy instruction construct simultaneously, the R^2 value results are related to both research subquestions 2b and 2c. Consequently, these results needed to be presented, interpreted, and discussed together.

First, the R^2 value results revealed that 33.30% of the variance in reading comprehension achievement was jointly accounted for by all eight predictor variables. The following list indicates the strength of the eight hypothesized predictors of reading comprehension achievement in descending order: reading attitudes, minority status, SES, class size, reading comprehension strategy instruction and reading comprehension strategy use (equivalent), class time, and gender. The large practical significance of the f^2 effect size of the R^2 value for the prediction of students' reading achievement, $f^2 = .33$, confirmed the importance of these independent variables in predicting achievement. The importance of these findings will be discussed in the 'Implications' section below.

Based on the R^2 value results 17.20% of the student differences in high school students' reading comprehension strategy use were explained by the combination of their attitudes toward reading, reading comprehension strategy instruction, gender, minority status, SES, class time, and class size. The order of strength of the predictors of strategy use, ranging from largest to smallest, was indicated by the results as the following: reading comprehension strategy instruction, minority status and SES (equivalent), reading attitudes, class size, class time, and gender. The importance of these seven variables in predicting students' reading strategy use was confirmed by the moderate f^2 effect size of the R^2 value for the prediction of strategy use, $f^2 = .17$.

The results also revealed that 15.20% of the variance in high school students' reading attitudes was jointly explained by all four hypothesized predictor variables together. This signified that reading comprehension strategy instruction, gender, minority status, and SES were all important predictors of students' attitudes toward reading. In the following list, the four predictors of students' reading attitudes are ordered from the largest predictor to the smallest: gender, reading strategy instruction, SES, and minority status. The moderate f^2 effect size of the R^2 value for the prediction of reading attitudes, $f^2 = .15$, supported the importance that these four variables have in predicting high school students' attitudes toward reading.

According to the structural regression results, only 0.30% of the student differences in high school reading comprehension strategy instruction were jointly accounted for by the two hypothesized predictors, class time and class size. The strength of the prediction from class size was larger than the prediction from class time. The f^2 effect size of the significant R^2 value in predicting strategy instruction, $f^2 = .003$, meant that the lack of importance of class time and class size in predicting reading strategy instruction was not practically significant.

The results related to the third research question, 'Are there significant group differences in students' reading comprehension strategy use, reading comprehension strategy instruction, reading attitudes, and reading comprehension achievement across the five student level (i.e., gender, minority status, and SES) and school level (i.e., class time and class size) variables?', were obtained using MIMIC analysis. The findings revealed statistically significant differences across all five groups on reading comprehension achievement. There were also statistically significant differences in gender, minority status, SES, and class time for reading comprehension strategy use. Whereas, the class size group differences in reading comprehension strategy use were not significant. The class time group differences in reading comprehension strategy instruction were statistically significant, however, the class size group differences were insignificant. Finally, statistically significant gender, minority status, and SES groups differences in reading attitudes were found. On the other hand, difference in students reading attitudes based on class time and class size groups were not statistically significant.

The expectation that the significant group differences in the four variables would consistently favor girls, nonminority students, students with a high SES, more class time, and smaller class sizes was not entirely accurate. The results suggested that the students in the minority group reported more positive attitudes toward reading, a higher rate of reading strategy instruction, and a higher frequency of reading comprehension strategy use. Further, the students in the large class size group performed higher on the reading comprehension assessment than the students in the small class size group.

Reading comprehension achievement. The hypothesis that reading comprehension strategy use had a direct effect on reading comprehension achievement was confirmed by the results of this study. The statistically significant, positive relationship between high school students' reading comprehension strategy use and reading comprehension achievement found in this study is consistent with previous research conducted by Denton et al. (2015), Hong-Nam et al. (2014), and Shera (2014). The large number of statistically significant indirect effects related to high school students' reading comprehension strategy use provide additional support for the important role it plays in high school students' reading comprehension achievement. The findings of this study indicated that reading comprehension strategy use mediated the relationships between reading comprehension achievement and reading attitudes, reading comprehension achievement and reading comprehension strategy instruction, reading comprehension achievement and gender, reading comprehension achievement and minority status, and reading comprehension achievement and SES.

The SEM results indicated that high school students' reading comprehension strategy use and high school reading comprehension instruction equally predicted reading comprehension achievement. The finding that reading comprehension strategy instruction had a direct effect on high school students' reading comprehension achievement is in alignment with a 2015 study conducted by Sari, that reported a statistically significant, positive, relationship between reading strategy instruction and reading achievement. The results also suggested that reading comprehension strategy instruction mediated the relationship between reading comprehension achievement and class time and between reading comprehension achievement and class size.

It was also hypothesized in this study that high school students' attitudes toward reading had a direct effect on their reading comprehension achievement. The findings in this study also confirmed this hypothesis. According to the results of the model, high school students' attitudes toward reading was the greatest predictor of their reading comprehension achievement. Although the results of this study are comparable to studies identified in the literature review, the magnitude of the relationship was different than both previously identified studies. The magnitude of the relationship between reading attitudes and high school students' reading comprehension achievement found in this study, $\beta = .37$, was lower than the results obtained in Kasapoglu's 2014 study, $\beta = .57$, but higher than the results obtained Sari's 2015 study, $\gamma = 0.16$. The inconsistent findings among these three multivariate studies is extremely noteworthy because all three studies were conducted using the results of the PISA 2009. While the sample in this study was composed of students in the United States, the other two studies utilized the results of students in Turkey. This inconsistency may also be related to the fact that the variables used to examine high school students' reading comprehension achievement differed across the three studies. Further support for the important impact of high school students' attitudes toward reading on their comprehension achievement is offered by the results of the study indicating several statistically significant indirect effects related to reading attitudes. The findings suggested that reading attitudes mediated the relationships between reading comprehension achievement and reading strategy instruction, reading comprehension achievement and gender, reading comprehension achievement and SES, and reading comprehension achievement and class time.

Gender, minority status, and socio-economic status. Although the results of this study indicated that high school students' gender did not have a direct effect on reading achievement, there were statistically significant gender differences in high school students' reading comprehension achievement, favoring girls. The effect size of this difference, $d = .31$, indicated a small to moderate practical significance. This finding is concordant with the results of multiple existing empirical studies that reported significant gender differences in reading, in favor of female students (e.g., Chui & McBride-Chang, 2006; Huang, 2015; Kasapoglu, 2014; Lietz, 2006; Shera, 2014; Singh, 2008). As stated by Watson and Kehler (2012) the lower literacy achievement typically demonstrated by male students, compared to female students, has been an area of major concern since the 1990s. There are a host of possible explanations for this gender disparity, including the differing socio-cultural experiences, neurobiological functioning, cognitive processing, and reading behaviors of male and female students (Gurian, 2010; Singh, 2008).

The hypothesis that high school students' minority status had a direct effect on their reading comprehension achievement and that there were statistically significant differences in high school students' reading comprehension achievement based on their minority status, was also confirmed by the results of the study. According to the test of minority group differences, nonminority students scored higher than minority students. The effect size, $d = .50$, indicates a moderate practical significance of this difference in reading achievement. The findings indicating that high school students' minority status had a significant direct effect on their reading comprehension achievement and that there were statistically significant achievement differences according to students' minority status, were consistent with the findings of Brown-Jeffy's 2006 study of the race gap in high school students' reading achievement. The result of this study, indicating a significant reading achievement gap between minority and nonminority students in the

US favored nonminority students, is also consistent with Brown-Jeffy's (2006) findings. Minority students have been consistently identified as at risk for academic failure. The reading achievement disparity between minority and nonminority students has been attributed to different factors including student SES, school SES, peer influence, home environment, number of parents in the home, parental support, highest level of parental education (Johnson, Gupta, Rosen, & Rosen, 2013; Sirin, 2005).

Students with a low SES have also been consistently identified as at risk for academic failure. Similar to the differences in reading achievement based on students' minority status, the SES reading achievement gap has been attributed to a multitude of risk factors including inadequate food or shelter, violence in the home or community, parents suffering from mental health issues, parents who are addicted to alcohol and/or drugs, poor school/parent relationship (i.e., parents are alienated from school or feel unwelcome at their child's school), parental unemployment (Johnson et al., 2013). It was hypothesized in this study that students' SES had a direct effect on their reading comprehension achievement and that there were significant differences in high school students' reading comprehension achievement based on their SES status. The results of this study supported these hypotheses. Students' SES had a statistically significant direct effect on reading comprehension achievement. This finding was consistent the findings of Chui and McBride-Chang (2006), Huang (2015), Özdemir and Gelbal (2014), Shera (2014), and Singh (2008), all of which indicated that the positive relationship between high school students' SES and their reading achievement was statistically significant.

Further, group differences in students reading achievement among the SES groups, in favor of students with a high SES, were also statistically significant. These results were commensurate with the results reported in Shera's 2014 multilevel analysis of school effects, gender, and SES differences in high school students' reading achievement. The results were also in agreement with the results of Huang's 2015 investigation of the SES achievement gap in the United States. The finding that students with a high SES scored higher than students with a low SES had moderate to large practical significance, $d = .67$. Based on the definition of SES adopted for this study, the practical interpretation of these findings is that high school students whose parents have more years of schooling, good paying careers, and whose families have more financial, cultural, and home education resources (specifically, the number of books in the home), demonstrate higher levels of reading comprehension proficiency.

Class time and class size. As hypothesized, class time also had a direct effect on the reading comprehension achievement of high school students. The finding of this study, indicating that increased reading instruction time significantly improves high school students' reading achievement, are consistent with the findings of Cattaneo et al. (2016), Huang (2015), Kasapoglu (2014), and Lavy (2015). Furthermore, there were statistically significant class time group differences in high school students' reading comprehension achievement. The tests of class time group differences revealed that students who reported more class time per week demonstrated higher reading achievement than high school students who reported less class time. This finding had a small practical significance, $d = .10$. The results of this study, indicating that increased reading instruction class time positively influenced high school students reading achievement, were corroborated by Huebener et al.'s (2016) research that examined the relationship between high school reading class time and reading achievement. Specifically, Huebener et al. (2016) reported that an increase of two hours of instructional time per week improved high school students' reading comprehension achievement by an average of 6% of an international standard deviation.

The hypothesized direct effect of high school students' class size on their reading comprehension achievement was confirmed by the findings in this study. As stated in the literature review, no studies specifically investigating the relationship between high school class sizes and high school students' reading comprehension achievement were found in the literature. However, in a more general study that investigated the effects of class size on high school students' overall academic achievement, Krassel and Heinesen (2014) found a statistically significant positive effect of class size on achievement. Other studies conducted by Chatterji (2006), Finn and Achilles (1999), and Molnar et al. (1999) examined the effects of class size on the reading comprehension achievement of elementary students. The significant influence of class size on reading achievement reported in the results of each of these related studies adds credence to the results of this study. However, because they examined elementary versus high school aged students, they don't entirely corroborate the results.

The findings of the current study also revealed statistically significant differences in high school students' reading comprehension achievement based on their reported class sizes. The MIMIC test of class size group differences indicated that high school students in the larger classes outperformed students in the smaller classes in reading comprehension achievement. The effect size, $d = .15$, indicated a small magnitude in the difference between the reading achievement of the students in the small and large class size groups. No studies of high school students reading comprehension achievement were found to compare these results, however, the results of two similar, but not entirely compatible, studies were consistent with this finding. For example, a study conducted by Johnson (2000) that investigated class size group differences in the reading achievement of eighth grade students also reported that the reading performance of students in larger classes was higher than students in smaller classes. Likewise, the results of Rice's (1999) study that examined class size group differences in high school students' academic achievement reported that high school students in smaller classes did not perform as well as those in larger classes. However, findings in this study related to the class size group differences conflicted with the majority of results in the studies that have examined class size group differences in academic achievement. The findings in most of the class size studies discussed in the literature review indicated the achievement of students in small classes was higher than the students in large classes (e.g., Chatterji, 2006; Finn & Achilles, 1999; Glass & Smith, 1979; Krassel & Heinesen, 2014; Molnar et al., 1999). The class size group differences revealed in this study were also incompatible with the results of a study conducted by Leuven et al. (2008) that found no statistically significant class size group differences in achievement.

Initially, the finding that students in the large classes outperformed students in the small classes was surprising. However, upon further reflection, this unanticipated result is likely due to the cutoff value used in this study to differentiate the small and large class size groups for the MIMIC group differences test. Prior to identifying the cutoff value, the class size research and other reputable academic resources were examined to identify the size of high school classes most beneficial to student learning. The only suggestions found were for elementary and middle school classes. Approximately 20 students or less were typically recommended as the desirable elementary class size (e.g., Finn, 2002). Consequently, the cutoff of 20 students was selected for the small class size group. It is plausible that the optimal size of high school classes is larger than elementary classes considering that older students are typically more independent and require less individualized attention. If so, this may explain the unexpected class size results.

Reading comprehension strategy use. The CFA results supported the hypothesis that three latent constructs (i.e., reading comprehension strategy use, reading strategy instruction, and reading attitudes) were all positively correlated with each other. Also, as anticipated, the structural regression analysis confirmed that reading comprehension strategy use did not have a statistically significant direct effect on either reading comprehension strategy instruction or reading attitudes. Additional findings regarding the direction of the relationships among the latent constructs, which allows for more meaningful interpretation, will be discussed in the relevant sections below.

Gender, minority status, and socio-economic status. The hypothesis that gender had a direct effect on reading comprehension strategy use was not supported by the structural regression model results. The insignificant direct effect of gender on strategy use contradicts the findings in the study of adolescents' reading comprehension strategy use conducted by Denton et al. in 2015. Although there was not a statistically significant direct relationship between gender and reading comprehension strategy use, the results indicated statistically significant differences in reading comprehension strategy use across the two gender groups, in favor of girls. The effect size of this difference, $d = .10$, indicated a small practical significance. This outcome is consistent with existing studies that reported high school girls use reading strategies to support their comprehension more frequently than boys (e.g., Denton et al., 2015; Lim et al., 2015; Shera, 2014).

In addition to the direct effect of gender on students' reading comprehension strategy use, a direct effect of the high school students' minority status was hypothesized. As opposed to the hypothesized direct effect of gender, this hypothesized direct effect was confirmed by the results of this study. The results also confirmed the hypothesized minority status differences in high school students' reading comprehension strategy use. The statistically significant group differences indicating that minority students reported using reading comprehension strategies more than nonminority student had a small practical significance, $d = .21$. No prior studies that examined the relationship between high school students' minority status and their reading comprehension strategy use were found in which to compare these results. However, this finding seems inconsistent with several other findings in this study that indicated minority students demonstrated lower reading comprehension achievement than nonminority students, that reading comprehension strategy use has a positive direct effect on reading comprehension achievement, and that the relationship between minority status and reading comprehension achievement was mediated by reading comprehension strategy use. It is possible that other mediating factors in the relationship between minority status and reading comprehension achievement not included in the SEM model have a greater impact on high school students' reading comprehension achievement.

It was also hypothesized that reading comprehension strategy use was directly impacted by students' SES. The results indicating that high school students' SES had a significant direct effect on their reading strategy use confirmed this hypothesis. This positive direct effect of SES on high school students' strategy use is consistent with prior research findings presented in the study conducted by Lim et al. (2015). Further, the statistically significant results of the MIMIC group mean analysis employed to explore SES differences in high school students' strategy use revealed that students with a high SES reported using reading strategies more frequently than students with a low SES. The effect size, d = .30, indicated a small to medium practical significance of this finding. Similar to the minority status differences in students' strategy use, no prior studies that examined the differences in high school students reading strategy use related to their SES were found in the literature in which to compared these results.

Class time and class size. Class time had a significant and positive effect on high school students' reading comprehension strategy use, as anticipated. The tests of class time group differences indicated that students in the more class time group reported using reading strategies more than the students with less class time. The small practical significance of this finding was supported by the effect size, d = .11. Also, as hypothesized, the results of this study indicated that class size had a statistically significant direct effect on high school students' reading comprehension strategy use, however, no significant class size group differences in reading comprehension strategy use were found. The lack of statistically significant class size group differences in strategy use may also be related to the issue mentioned above, regarding the designated class size cut off value used in this study. Only prior research that investigated the relationship between high school students' reading comprehension achievement and class time was available in the literature. Unfortunately, no existing research specifically focused on the relationship between reading comprehension strategy use and class time or class size was located to compare to the findings of this study.

Reading comprehension strategy instruction. The confirmatory factor analysis results indicated a statistically significant, positive, correlation between reading comprehension strategy instruction and high school students' use of reading comprehension strategies. The direction of the relationship provided by structural regression results indicated that strategy instruction had a direct effect on reading comprehension strategy use. In other words, effective reading comprehension strategy instruction increases the frequency of high school students' reading strategy use. This finding implies that students who receive explicit reading comprehension strategy instruction are more likely to use the strategies to aid their comprehension of text. This result is concordant with the results Lim et al. (2015) obtained in their SEM study of high school students' reading attitudes and reading behaviors.

The hypothesized positive correlation between reading comprehension strategy instruction and high school students' attitudes toward reading was also confirmed by the CFA. According to the structural regression results, the hypothesized direct effect of reading comprehension strategy instruction on high school students' reading attitudes was statistically significant. This finding is commensurate with the findings of existing studies that have examined the relationship between high school reading comprehension instruction and high school students reading attitudes (e.g., Lim et al., 2015; Jhang, 2014). The direct effect of high school reading comprehension strategy instruction on students' reading attitudes implies that students who are taught strategies that support their comprehension of texts are more likely to pursue reading opportunities and reading related activities due to more positive attitudes toward reading.

Gender, minority status, and socio-economic status. The results of the MIMIC model indicated that female students reportedly received more reading comprehension strategy instruction. The small effect size, $d = .09$, indicates that this finding has little practical significance. Minority students also reported receiving more reading comprehension strategy instruction, $d = .09$. Considering this finding, the positive impact that high school reading comprehension strategy instruction has on students' reading comprehension strategy use and on their attitudes toward reading found in this study may explain why minority students reported using reading strategies more frequently than nonminority students, and why they reported more positive attitudes toward reading than their peers. It is possible that minority students receive more instruction regarding reading comprehension strategies because teachers are aware that as a group minority students historically demonstrate lower levels of reading achievement. If this conclusion is accurate, and teachers provide additional attention to minority students for this reason, the results indicating that students with a low SES reportedly received less reading comprehension strategy instruction, reported a lower frequency of reading comprehension strategy use, and reported less positive attitudes toward reading suggests teachers are either unaware that students with a low SES also underperform in reading or identifying students with a low SES is less apparent than identifying minority students.

Class time and class size. The structural regression results revealed that the amount of class time had a statistically significant, albeit small, positive impact on high school reading comprehension instruction. The results of the MIMIC analysis also indicated significant class time group differences in reading comprehension strategy instruction. This result implies that students who reported more class time per week also reported receiving more reading comprehension strategy instruction. The small practical significance of this finding was supported by the effect size, $d = .11$. Unfortunately, no specific literature in the reading or class time research related to these results were found.

The hypothesis that class size had a direct effect on reading comprehension strategy instruction and that there were significant class size group differences in high school students' reading comprehension strategy instruction were not supported by the results of this study. These results are incompatible with existing qualitative research that reported the positive impact of smaller class sizes on the teaching and learning processes in high school (e.g., Blatchford et al., 2011; Harfitt & Tsui, 2015). It is likely that the small class size cutoff of 20 students, explained above, also impacted these findings.

Reading attitudes. The third hypothesized correlation among the three latent constructs, the correlation of high school students' attitudes toward reading and their reading comprehension strategy use, was also supported by the results of the CFA. The structural regression analysis results provided further information regarding the direction of this relationship. Specifically, they indicated that student's attitudes toward reading had a statistically significant direct effect on their reading comprehension strategy use. The previously mentioned study conducted by Lim et al. (2015) also examined the relationship between high school students' use of reading comprehension strategies and their reading attitudes. Similarly, they found that students' attitudes toward reading positively predicted their use of reading comprehension strategies.

Gender, minority status, and socio-economic status. It was hypothesized that the gender of high school students had a direct effect on their attitudes toward reading. The statistically significant direct effect of students' gender on their reading attitudes revealed in this study confirm this hypothesis. This finding is consistent with the results of existing research (e.g., Jhang, 2014; Lim et al., 2015; OECD, 2001; OECD, 2004). Further, results of the MIMIC group mean analysis confirmed the statistically significant differences in high school students' reading attitudes that were also hypothesized in this study. According to this test of group differences, female high school students have more positive attitudes toward reading than males. The effect size, d = .67 indicated the moderate to large practical significance of this finding. These results in this study are consistent with the results of prior studies that have investigated gender difference in high school students' reading attitudes (e.g., Bussert-Webb & Zhang, 2016; Gökhan, 2012; Lim et al., 2015). This finding may be related to the commonly held belief that females are more successful readers than males. Male students who are conditioned to believe that they will not perform as well as their female peers may consequently have lower expectations for themselves, be less interested in making an effort to increase their reading proficiency, and/or develop negative attitudes toward reading.

The results of the study did not support the hypothesis that high school students' minority status had a statistically significant direct effect on their reading attitudes. However, the hypothesized minority status differences in high school students' attitudes toward reading were supported. The statistically significant group differences in high school students' reading attitudes across the two minority status groups, indicating that minority students had more positive attitudes than nonminority student had a small practical significance, $d = .07$. This finding has promising implications for the reading achievement of minority students. Given the lower reading comprehension achievement demonstrated by minority students and the significant positive relationship between reading achievement and reading attitudes described above, focusing on increasing the positive reading attitudes of minority could be an important key to narrowing the high school reading achievement gap. As mentioned in the introduction, no studies were found that investigated the specific relationship between the minority status of high school students' and their reading attitudes, but in the 1995 study conducted by McKenna et al., this relationship was examined with elementary aged students in first through sixth grade. The significant relationship between minority status and students reading attitudes found in each of the six grades are consistent with the results of this study.

It was also hypothesized that high school students' attitudes toward reading were positively predicted by their SES. This hypothesis was confirmed by the results of the study. The statistically significant positive relationship between students' SES and their reading attitudes is consistent with existing research. In 2012, Gökhan conducted a study to investigate effects of multiple variables on the reading attitudes of ninth and 12th grade students, including SES. The results of the ANOVA analysis employed in that study indicated that reading attitudes of high school students were significantly and positively related to their SES level, $F(2,423) = 542.777$, $p < .05$. Additionally, significant SES differences in high school students' reading attitudes were hypothesized. This hypothesis was also supported by the results of this study. The tests of SES group differences indicated that students with a high SES have more positive attitudes toward reading than students with a low SES. The moderate practical significance of this finding was supported by the effect size, $d = .38$. Gökhan (2012) also found significant SES differences in high school students' attitudes toward reading. It is possible that students with a higher SES have more positive attitudes toward reading because, as discussed above, these high school students also demonstrate greater success in reading achievement. Successful academic reading experiences are likely to increase their enjoyment of reading, their motivation to read, and ultimately their reading attitudes.

Class time and class size. The results of the MIMIC analysis did not confirm the hypothesized class time or class size group differences in high school students' attitudes toward reading. Unfortunately, there is currently no empirical research regarding the relationships between high school students' reading attitudes and class time or class size to corroborate or dispute these findings.

Conclusion. In addition to the findings of no class time or class size group differences in high school students' reading attitudes, the following results of this study add to the existing high school reading literature base: high school class size has a direct effect on students' reading comprehension achievement, there are high school class size group differences in reading achievement favoring large classes, high school students' gender did not have a direct effect on their reading comprehension strategy use, minority status has a direct effect on high school students' reading comprehension strategy use, there are minority status and SES group differences in high school students' reading comprehension strategy use, class time and class size have a direct effect on high school students' reading comprehension strategy use, there are class time and class size group differences in high school students' reading comprehension strategy use, class time has a direct effect on high school students' reported reading comprehension strategy instruction, there are class time group differences in high school students' reported reading comprehension strategy instruction, class size does not have a direct effect on high school students' reported reading comprehension strategy instruction, there are no class size group differences in high school students' reported reading strategy instruction, minority status has a direct effect on high school students' reading attitudes, and there are minority status group differences in high school students' reading attitudes.

Implications

The detrimental effects of low reading comprehension achievement and ongoing concerns about the ability of U.S. schools to teach students to read (Therrien, 2004) keep the reading achievement gap in the forefront of priorities for all stakeholders. Teachers, parents, administrators, policymakers, and researchers all share the same goal of providing reading instruction that maximizes the reading achievement of all students

(IRA, 2002). The results of this study offer several important implications stakeholders can utilize to increase the reading comprehension achievement of high school students.

Implications for teachers. The low level of U.S. high school students' reading proficiency, in conjunction with the impact that positive reading attitudes have on reading achievement and the fact that students' reading attitudes decline with age, make it exceedingly critical for educators to focus more attention on increasing and/or reinforcing these students' positive attitudes toward reading. Indeed, out of all the variables in the SEM model, the results of this study indicated that high school students' attitudes toward reading had the most positive impact on their reading comprehension achievement. As McKenna et al. (2012) stated, effective reading instruction requires the awareness of students' reading attitudes. This awareness allows teachers to appropriately design instruction aimed at increasing and/or sustaining students' positive reading attitudes (Bokhorst-Heng & Pereira, 2008). For example, teachers could incorporate specific instructional techniques into their plans that have been found to positively influence students' reading attitudes such as positive reinforcement, whole class discussions, and modeling of positive reading attitudes (Kush & Watkins, 1996).

In order to appropriately design reading instruction focused on reading attitudes, it is necessary for teachers to assess students' reading attitudes. Teachers often believe they can accurately judge their students' attitudes toward reading, however, teacher's ratings of student attitudes toward reading do not always match their actual attitudes (Kush & Watkins, 1996). Accurate knowledge of high school students' reading attitudes obtained from a reading attitude assessment is also important for teachers because students reading attitudes tend to become more negative over time (McKenna et al., 2012) and negative reading attitudes hinder students' learning of reading comprehension strategies (Petscher, 2010). This makes it especially important for high school teachers to be aware of students' reading attitudes because they tend to become more negative over time (McKenna et al., 2012). Further, the results of the MIMIC test of gender differences indicated that boys had more negative attitudes toward reading than girls, suggesting that while teachers should not overlook girls' reading attitudes, they need to pay special attention to increasing boys' reading attitudes.

Often, struggling high school readers have experienced failure over such a considerable period of time that they become frustrated and develop negative attitudes towards reading (Woolley, 2011). Providing these students with the instructional support necessary to help them overcome their frustration with reading is a significant challenge made even more difficult by the fact that this frustration has caused many of these students to disengage from any type of academic reading. Positive attitudes toward reading are associated with increased reading engagement (McKenna et al., 1995). Thus, reading instruction that focuses on developing and/or increasing students' reading attitudes can help struggling readers reengage in academic reading. In addition to a focus on improving students' reading attitudes, encouraging struggling readers to engage in reading activities should be a top priority for teachers because according to Guthrie et al. (2004), reading engagement also leads to increased reading achievement.

The results of this study suggest that students' attitudes toward reading is prime area for reading comprehension intervention. Unlike other interventions that can take a long period of time to produce results, changes in students' reading attitudes can be quite notable over a short period of time (Kazelskis et al., 2005). For example, social cognitive learning theory suggests that students' self-beliefs are easily manipulated by teachers and other influential adults (Bandura, 1997; Schunk & Pajares, 2001). More specifically, teachers' attitudes towards reading influence their students' reading attitudes (Applegate & Applegate, 2004; Lim et al., 2015). Therefore, something as simple as a teacher consistently modeling positive attitudes toward reading could positively influence students' reading attitudes and subsequently their reading comprehension achievement. Another rather easy way to positively influence students' attitudes toward reading is to allow them to select the books they read in class. The autonomy provided by self-selecting texts makes the reading experience more positive and enjoyable for students (Petscher, 2010). Enjoyment of reading increases students' desire to read more, and once students become more motivated they will begin reading more frequently (Williams, 2014). Which ultimately increases students' reading attitudes as well as their reading comprehension achievement (Petscher, 2010).

Implications for teachers related to gender, minority status, and SES. Over the past 20 years, student diversity has significantly increased in U.S. schools. The wide array of background characteristics represented in today's classrooms make it extremely difficult for teachers to meet the individual needs of every student (Klein, 2008). In line with this notion, Williams (2014) added that the increasingly diverse student population in high schools has made it ever more challenging to improve students' reading comprehension achievement. In a broad sense, Luke, Woods, and Dooley (2011), stated that content area reading comprehension strategy instruction is one important approach that high school teachers can use to effectively increase the reading achievement of students from diverse backgrounds. They explained that these strategies allow diverse students to become familiar with important common background knowledge and become more competent with the various content area reading demands (Luke et al., 2011).

Specifically, the results of this study suggest that it is important for teachers to be aware that differences in gender, minority status, and SES significantly impact students' reading comprehension abilities and performance. As such, teachers should differentiate their reading comprehension instruction to appropriately address known student differences. Brozo et al. (2014) offered teachers several research based suggestions to address the gender differences in reading that disproportionally disadvantage boys. First, they recommend that teachers consciously make boys reading needs a priority. Next, they tell teachers to focus on increasing boys reading engagement and reading frequency. They also suggest that teachers assign boys texts related to their individual interests. Further, they encourage teachers to allow boys to read using electronic texts and alternative forms of media. Finally, they recommend that teachers develop reading programs for boys that involve their fathers and/or other influential adult males. Of course, these approaches are also beneficial to girls, thus, teachers should not focus on improving boys' reading comprehension achievement to the exclusion of girls'.

High school teachers also need to be cognizant of the impact that social and cultural factors have on students' reading comprehension achievement (Myrberg & Rosén, 2009). Considering the positive relationship between parental involvement and student achievement (Huang, 2015), coupled with the lower parental involvement associated with minority students and students with a low SES, these students are likely to have less exposure to or support for reading at home, which negatively affects their overall academic achievement (Sirin, 2005). Further, the results of the minority and SES tests of group differences indicated that minority students and students with low SES demonstrated significantly lower reading comprehension achievement. Thus, it is inappropriate for high school teachers to make broad assumptions regarding all students' reading comprehension proficiency. They need to understand that their students likely have dissimilar reading experience in the home and unequal reading comprehension competencies. Therefore, although it may seem like a challenging task, all high school content area teachers need to be aware of students' family background in order to effectively accommodate the significant influences of students' minority status and SES on their reading comprehension achievement (Jeynes, 2002; Sirin, 2005).

Last, and arguably the most important implication for high school teachers is simply that they need to explicitly teach students general cognitive, metacognitive, and content area specific reading comprehension strategies. The results of this study confirm the findings of prior research. Even though research evidence has consistently confirmed that explicit reading strategy instruction can effectively increase students' reading comprehension achievement, there is a disconnect between the research and current high school teaching practices. Despite the poor reading comprehension proficiency of a majority of U.S high school students and the positive relationship between reading comprehension proficiency and content area achievement, little to no time is spent teaching reading comprehension in high school content area classrooms (Ness, 2009). Many high school teachers think that teaching reading is not part of their instructional responsibilities. They believe that reading comprehension instruction is complete before high school, despite expert recommendations to continue teaching students' how to read as long as they need it (Ness, 2009). The fact that 63% of 12th grade students (NCES, 2017) cannot read proficiently suggests that a large number of high school students did not acquire adequate reading comprehension skills prior to high school and are in desperate need of continued reading comprehension instruction.

All classroom teachers at every level of schooling need to be able, ready, and willing to help students improve their reading comprehension proficiency. Especially considering that classroom teachers, not reading specialists or special educators, are ultimately responsible for their students' learning (Invernizzi, Landrun, Howell, & Warley, 2005). The low to moderate values of the factor loadings for the seven items used to define the reading comprehension strategy use latent construct offer a starting point that high school teachers can use to plan their reading comprehension strategy instruction. For example, the lowest scoring item on this scale addressed students' use of graphs, diagrams, and tables to locate important information in a text. Knowledge of, and the ability to accurately use this strategy is extremely important for high school students' comprehension of non-fiction content area texts. Similar to the above suggestion for improving students' reading attitudes, teachers need to assess students' use of reading comprehension strategies. Comparing the specific reading strategies that students report using against the strategies they expect students to be using would enable teachers to more effectively plan their reading instruction to meet the needs of their students.

Implications for higher education. The main implication of the results of this study for institutes of higher education is directly related to the previous implication that high school teachers need to teach their students content area reading comprehension strategies. Preservice secondary education teacher preparation programs need to expand their programs of study to include sufficient training in content area reading instruction. They need to ensure that high school teachers can effectively teach content area reading comprehension strategies and integrate explicit reading comprehension strategy instruction into their content area instruction.

According to Kissau and Hiller (2013), one of the most common reasons high school teachers give to explain why they allot minimal, if any, instructional time to teaching reading comprehension strategies is that they never learned how to teach reading comprehension strategies, much less how to incorporate reading comprehension strategy instruction into their instruction. Likewise, as Ness (2009) reported in the findings of her qualitative investigation of middle and high school content area teachers' reading comprehension instruction practices, some teachers stated that they did not feel qualified to teach reading comprehension and others surprisingly revealed that they didn't even know the meaning of comprehension instruction (Ness, 2009). It is not entirely surprising however, that many high school teachers do not feel prepared to teach reading comprehension strategies. Typically, high school teachers are only prepared to teach their subject area content (Irvin et al., 2007), and they have not been taught how to teach content area reading comprehension strategies or how to incorporate reading comprehension strategy instruction into their content classes.

Most preservice secondary teacher preparation programs only require one content area reading course (Sturtevant, 2003). Various statements of high school teachers reported throughout the literature point to an overall lack of competence in terms of reading comprehension strategy instruction and suggest that, in general, the amount of preservice training they received was insufficient. The probability that high school students learn content area reading comprehension strategies would increase if institutions of higher education developed and implemented a reading curriculum that directly instructs preservice high school teachers how to effectively teach the specific reading strategies and skills requisite of their content area. Well prepared high school teachers who can competently teach reading comprehension strategies are essential to eliminating the reading achievement gap.

Implications for administrators. Although there are a wide variety of strategies that aid students' reading comprehension, it is difficult for secondary content area teachers to know when and how to use these strategies with their students. This dearth in pedagogical knowledge can be attributed to the lack of a content area reading comprehension strategy instruction curriculum in pre-service teacher preparation programs discussed above. It can also be attributed to the lack of focus on professional development opportunities for in-service high school teachers (Vaughn et al., 2013).

The lack of pre-service training combined with the significant effects of reading comprehension strategy instruction and reading strategy use on high school students' reading achievement implies that education administrators play an extremely vital role in helping to reduce the abysmal reading achievement gap that has plagued students in the US for at least the last 30 years. An important first step is to reverse the current trend of limited professional development designed to improve the reading instruction competency of high school teachers (Shanahan & Shanahan, 2012).

District level administrators need to develop and school level administrators need to implement required content area specific reading comprehension strategy professional development with all high school content area teachers. To increase the probability that reading comprehension strategies are effectively taught to and used by high school students, all content area teachers need to be aware of the strategies, understand their importance, know how to effectively teach them to students, and be given extensive opportunities to practice using the strategies specific to their area of expertise. Essentially, they need receive much of the same content area reading comprehension strategy instruction they will, in turn, need to implement with their students.

District and school level administrators also need to develop a way to ensure these reading comprehension strategies are implemented in all content area classrooms. For example, school level administrators should conduct content area reading comprehension strategy instruction observation of all content area teachers. District and school level accountability teams as well as accountability procedures could also be developed. The school accountability teams would be responsible for collecting school level data. A representative from the school team would also be a member of the district level team responsible for reporting their results. The district level team would then synthesize the results across the various high schools, identify and communicate areas of concern, and potentially develop additional professional development targeting ongoing areas of need.

Another area of need that can be easily addressed by school administrators is the underperformance in reading comprehension of students with a low SES, which was reaffirmed by the results of this study. SES was the third largest positive predictor (out of eight) of high school students' reading comprehension achievement in the SEM model used in this study. A summer reading instruction program and a summer book program are two relatively easy approaches administrators could employ to support the reading achievement of high school students with a low SES. Summer reading programs are especially beneficially for students with a low SES because, according to Allington et al. (2010), the reading achievement of students from low income families regresses during the summer break from instruction. In addition to the benefits to student learning, summer reading programs should be an attractive option for administrators because federal financial assistance to support additional learning opportunities schools provide for at risk students from low income families is allocated under Title I, Part A of the Elementary and Secondary Education Act of 1965.

Administrators could also support the reading comprehension achievement of students with a low SES by facilitating a summer book program for low income students. They could develop a very manageable program similar to the experimental program implemented by Allington et al. (2010). They developed their program based on the notion that without books at home, students from low income families rarely read over the summer break and subsequently demonstrate a loss in previously attained reading comprehension achievement. The experimental program entailed giving students from low income families books of their choosing to reading during the summer break. Results of the program indicated that not only were the average summer reading losses experienced by students from low income families eliminated, the program participants also demonstrated reading gains similar to those typically observed in students from middle income families (Allington & McGill-Franzen, 2013; Allington et al., 2010).

Implications for policymakers. There are several reasons why the effective remediation of widespread reading comprehension deficits has not been implemented in U.S. high schools. Sadly, the top reason is that policy makers typically do not understand the complexities of literacy development or the critical relationship between literacy (i.e., reading and writing) and high school students' content area learning (Irvin et al., 2007). Consequently, scarce resources are allocated to support high school students who struggle with the reading and/or writing skills required for academic success.

For example, many high schools do not provide remedial reading classes for students with reading difficulties (Shanahan & Shanahan, 2012) and less than 10% of high schools have a literacy specialists on site to work with students (Irvin et al., 2007). The lack formal reading comprehension instruction and reading experts in most high schools are major roadblocks to any current efforts of improving high school students' reading comprehension achievement. The direct and indirect effects of reading comprehension strategy instruction, reading comprehension strategy use, and reading attitudes on high school students' reading comprehension achievement revealed in this study support the need for education policy makers to allocate more funding and resources to improve the currently bleak state of reading instruction in U.S. high schools.

The results of this study also provide evidence in support of the current and widely debated education policy initiative of increasing instructional time. Although the variance explained by the class time school level variables in the SEM model was relatively small, it is still important to consider the policy implications of this statistically significant result. The finding is important because a global policy change to increase instructional time has the potential to positively impact the reading achievement of a larger number of students compared to changes made at the individual student level.

A second reason policy makers should consider increasing instruction time is that due to the increased comprehensiveness and rigor of the curriculum standards, teachers often feel rushed to cover all the required content. In Ness's (2009) qualitative study mentioned above, a common reason many high school teachers reported for not teaching content area reading comprehension strategies was a lack of instructional time. They felt like they didn't have enough time to teach all the required content area subject matter and they didn't want to waste time teaching reading strategies. According to Ness, the teachers explained that they believed their main responsibility was to teach students their content area curriculum. Further, they explained that with the pressure to sufficiently cover all the content area material prior to the end-of-year state testing, they felt like teaching reading comprehension strategies took up too much valuable class time.

Policymakers should also consider adding explicit content area reading comprehension strategies to the individual high school subject area curriculum standards (i.e., biology, chemistry, history, algebra, etc.). Including content area reading comprehension strategies in the state standards would serve multiple purposes. First, it would increase teachers' awareness of the need, and their responsibility, to incorporate reading strategies into their instruction. Explicitly communicating the appropriate strategies students need per subject area would inform teachers exactly of what they need to teach. Third, including the strategies into the standards would serve as an additional measure of accountability to help ensure teachers teach students' these necessary skills.

Implications for parents. The positive relationship between parental involvement and students' academic achievement is supported extensively throughout the research literature (Flowers, 2007). In fact, according to Gamoran and Long (2007), families have a larger impact on student achievement than schools. A student's family plays a key role in their school success including preparing them for school, placing a high value on education, conveying a belief in their academic competence, and promoting their language development and comprehension through reading (Ladd, 2012).

Quite possibly the most important action a parent can take to support their high school child's learning is to pay close attention to their reading proficiency. If they believe that their child is unable to read at an appropriate level and/or they think that their child is struggling to comprehend difficult content areas texts, they need to contact the relevant content area teacher. If they find out their child is not being taught critical content area reading strategies, they need to bring this concern to the attention of school administrators. If, in fact, the teacher is providing proper reading comprehension strategy instruction, the child's reading difficulties may be indicative of a more severe reading deficit or underlying learning issue that should be promptly examined by the high school.

Further, parents of students with a low SES can help their child overcome some of the disadvantages students experience due to their low SES by devoting time to working with their children on the strategies and skills they need to develop reading proficiency (Buckingham et al., 2014). Understandably, in some cases, some parents may not know how to effectively support their child's reading comprehension development at home. In other cases, some parents may not have the time necessary to help their children. In either case, parents need to contact their child's teacher and/or school to request guidance and learn out about available resources they can take advantage of to help support the development of their child's reading comprehension proficiency. Finally, one relatively easy action parents can take to help their child, regardless of the circumstances just described, is to model positive attitudes toward reading because parental attitudes toward reading are positively associated with their child's reading attitudes (Petscher, 2010).

Implications for researchers. One promising method to improving high school students' reading comprehension is content area reading comprehension strategy instruction. For students to receive the most effective instruction, teachers need to be knowledgeable of and understand how to implement empirically based reading comprehension strategy instructional approaches. According to Hagaman and Reid (2008), existing reading research has predominantly focused on the foundational reading skills of young students. While early intervention research is necessary, the abundance of research on young readers has left a void in the literature of studies that address the distinct difficulties of struggling high school readers (Deshler & Hock, 2007). As a result, there are relatively less resources for teachers about how to help high school students overcome general reading comprehension difficulties or about high school reading comprehension strategy instruction. Further, the positive effects on high school students' reading achievement demonstrated by all variables in this study, coupled with the limited amount of existing research regarding many of the relationships examined in this study indicates there are several gaps in the literature that need to be filled.

Although some reading comprehension strategy interventions developed for elementary age students have been successfully extended with high school age students, high school content area teachers need evidenced based interventions that are designed intentionally for the high school age students. They need interventions that specifically address reading comprehension strategies for their content area that can be implemented with a wide range of students, yet still permit sufficient time for content area instruction. While a few such interventions have been developed, most have not been rigorously investigated and their effectiveness have not been systematically demonstrated (Lang et al., 2009). Therefore, there is a need for practical, evidenced based content area reading comprehension strategy interventions that can be easily implemented in high school classrooms. Finally, the results of this study indicating that the positive relationship between high school students' reading comprehension strategy instruction and reading comprehension achievement is mediated by students' reading comprehension strategy use and their reading attitudes, suggest that specific reading comprehension strategy instruction interventions with a combined focus on content area reading comprehension strategies and high school students reading attitudes are also needed.

Limitations and Future Research

There are several important limitations identified in this study that need to be discussed. The first main limitation of this study was the use of secondary data. While the secondary PISA 2009 dataset provided an extensive amount of valuable information that would not otherwise be as easily accessible, it's use presented various limitations in this study. For example, the data in this dataset was not collected to address the specific research questions in this study. As a result, rather than having data from students of all high school ages/grade levels, analysis in this study was restricted to the sample of 15-year-old high school students. Further, the items used to define the reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes latent variables were limited to the questions posed in the PISA 2009. Future survey research conducted with ninth through 12th grade students utilizing a precisely designed survey instrument including enough items to fully target all relevant aspects of the reading comprehension strategy use, reading comprehension strategy instruction, and reading attitudes constructs would offer more meaningful findings that could be used to enhance high school students' reading comprehension proficiency.

A second major limitation of this study was the cutoff value selected to distinguish the small and large class sizes for the analysis of class size group differences. As explained in the discussion section above, the only class size recommendations found in the literature were for elementary and middle school classes. In lieu of an optimal high school class size recommendation, Finn’s (2002) recommended small class size of 20 or less students was adopted to designate the small class group in this study. The cutoff of 20 students for the small class size group is thought a limitation in this study because, considering the results of existing class size research, the results of all four class size group differences tests in this study were questionable.

More specifically, the only hypothesized class size group difference (of the four hypothesized) confirmed by the results of the MIMIC analysis was the class size group differences in students’ reading comprehension achievement. While this difference was hypothesized, it was not anticipated that the students in the larger classes would outperform students in the smaller classes. This finding is questionable mainly because it contradicts the results of most existing class size research. It is also questionable because it conflicts with classroom teachers’ common assumption that smaller class sizes positively influence academic achievement.

The fact that three of the four hypothesized class size difference were not confirmed by the data, coupled with the counterintuitive outcome of the one confirmed class size group difference, make it likely that the class size groups were poorly specified in this study. While it is possible that the unexpected findings were not related to the class size cutoff, future research is needed to determine the size of high school classes that is most beneficial to student learning. Those findings can then be used to reanalyze the class size group differences examined in this study to determine if the results were due to the incorrectly specifying the class size groups, or if the higher achievement of the students in the larger classes has practical implications for high school class sizes.

The third major limitation of this study is related to the decisions made in defining the minority status groups. When defining the minority status groups, it was decided to combine all six minority groups represented in the U.S. portion of the PISA 2009 dataset into one group (i.e., the minority group). In doing so, valuable information regarding the differences among the six race/ethnicity groups (i.e., African American, Hispanic, Asian, American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, or more than one race) in relation to the various reading variables examined in the study was lost. This decision seems especially limiting in light of the number of statistically significant findings related to high school students' minority status found in this study and the paucity of existing research in these areas. Future research to examine the relationships addressed in this study across the specific race/ethnicity groups is recommended.

Likewise, future research to cross-examine gender and SES with the significant relationships students' minority status had with the latent reading constructs and reading comprehension achievement is recommended. The findings of such research would provide a deeper understanding of how teachers should develop and plan instruction that may serve to reduce the minority status reading achievement gap. Indeed, this type of cross examination with assorted student and school level control variables beyond gender and SES, (e.g., age, grade level, English Language Learner status, disability status, immigration status, school SES, school type, etc.) would be valuable future research.

Researchers are currently faced with the crucial responsibility of expanding the knowledge base regarding all aspects of reading comprehension strategies that high school teachers can rely upon to secure students reading comprehension proficiency prior to the end of compulsory education. Substantially more research is needed to inform teachers about the most effective approaches to improving high school students' reading comprehension proficiency. Overall, additional research regarding each of the relationships examined in this study is recommended. In the review of the literature for this study alone, no research was found in three important areas of interest related to high school students reading comprehension strategy use, including its relationship with students' minority status, the amount of reading comprehension strategy instructional class time, and the size of the classes in which comprehension strategies are taught. The limited amount of prior empirical research addressing these relationships, specifically addressing the population of high school students, inhibited the discussion of findings in this study. Future research in these areas would provide meaningful information to allow for a complete and robust discussion of the significant findings in this study.

Further, the indirect effects examined in this study for exploratory purposes yielded a variety of interesting results that have either not yet been addressed or have been inadequately addresses in the high school reading literature. Future research focused on any of the following relationships would be valuable to the field: the relationship between reading comprehension strategy instruction and reading achievement mediated by reading strategy use and reading attitudes, the relationship between gender and reading achievement mediated by reading strategy use and reading attitudes, the relationship between minority status and reading achievement mediated by strategy use, the relationship between SES and reading achievement mediated by reading attitudes and strategy use, and finally, the relationship between class time and reading achievement mediated by reading strategy instruction and reading attitudes.

Neurobiological reading comprehension intervention research is one final, and particularly interesting area of reading research in need of greater attention. Advances in neuroplasticity research indicating that the brain can be taught to change has prompted research that examined the relationship between the brain and several different reading disabilities (e.g., Cutting, Materek, Cole, Levine, & Mahone, 2009; Shaywitz et al., 2003; Shaywitz et al., 2002). Effective brain based reading intervention research in areas such as phonological awareness (e.g., Shaywitz et al., 2004) and phonics (e.g., Simos et al., 2002) have also been conducted. Additionally, there have been some effective neurobiological reading comprehension intervention studies conducted with young students (e.g., Meyler, Keller, Cherkassky, Gabrieli, & Just, 2008), however, nothing was found in the literature investigating older students. Thus, future neurobiological reading comprehension intervention research targeting high school students is recommended. Further, considering the gender based reading achievement differences associated with brain development discussed in the literature review (e.g., Gurian, 2010; Prado & Plourde, 2011) and the gender based reading achievement differences favoring high school girls supported by the results of this study, future neurobiological reading comprehension intervention research, specifically focused on improving the reading comprehension proficiency of at-risk high school boys, is also recommended.

Conclusion

High school students' reading comprehension difficulties are commonly associated with student demographic variables such as gender, minority status, and SES. The No Child Left Behind Act of 2001, officially made U.S. schools accountable for closing the reading comprehension achievement gaps due to gender, minority status, and SES group differences (No Child Left Behind [NCLB], 2001). Unfortunately, the evidence presented in the literature review and the results of this study demonstrated persisting disparities in reading comprehension proficiency associated with high school students' gender, minority status, and SES. While the minority reading achievement gap is shrinking, and the gender achievement gap has remained relatively stable, the SES reading achievement gap is widening (Reardon et al., 2012). In today's economy, academic success is more widely recognized and rewarded than in the past. Considering the positive relationship between reading achievement and overall academic success, eliminating the SES reading achievement gap could potentially expand the opportunities for academic success of students with a low SES. Furthermore, expanded opportunities for academic success could boost the otherwise minimal chances of social mobility currently afforded to students with a low SES. The potential of greater social mobility for students with a low SES (Reardon et al., 2012) is one of many reasons improving high school students' reading comprehension proficiency should remain a national priority.

The first mention of reading as a national concern in the US dates back to World War I, when intelligence testing showed that approximately 25% of the 2.8 million drafted soldiers were unable to read a newspaper written in English (White, 1999).

Unfortunately, over 100 years later, reading remains a national concern in the US. According to the results of the NAEP 2015, among the 63% of 12th grade students who scored below proficient in reading comprehension, 35% of the students scored at the basic proficiency level, and 25% of the students scored below the basic level of reading proficiency (NCES, 2017). Further, from 1998 to the most recent NAEP administration in 2015, the reading scores of 12th grade students have declined (NCES, 2017).

Amongst the myriad of skills students must learn in school, reading comprehension is of utmost importance. Conceivably, in the future there will be new ways to obtain information that are as widespread and effective as text, but currently reading comprehension proficiency remains the key to students becoming fully informed and productive members of our society. Likewise, failure to develop proficient reading comprehension skills can have devastating long-term academic, psychological, social, behavioral, and economic effects on students as well as the society in which they live.

Reading failure does not just happen. Most likely, high school students who struggle to read were also middle school students who struggled to read, and were perhaps even elementary students who struggled to read. Reading failure is a result of students' instructional needs not being met somewhere along the continuum of their academic lives. After 30 years of high school students consistently failing to demonstrate proficient reading comprehension, teachers at every level of schooling from pre-school up and through adult continuing education programs must become wholly committed to ensuring the reading success of all students. It is also imperative that policymakers allocate the funding and resources necessary to support the reading achievement of high school students. Continued apathy on the part of teachers or policymakers toward addressing the reading needs of struggling U.S. students is indefensible.

There are currently millions of science, technology, and engineering jobs in the US that cannot be filled because of an overall skills deficit in the country. Unfortunately, a widespread deficit in the skills required by these highly complex careers is not surprising when considering the prevalence of reading failure in the US. To learn the advanced skills required for these in demand job, students must possess the reading comprehension strategies and skills so many are presently lacking (Reardon et al., 2012). The results of this study demonstrated that the explicit teaching of reading strategies can help high school students develop these fundamental skills, and various actions to begin reversing the pattern of reading failure in the United States have been recommended.

The ability to comprehend text is not innate. And unfortunately, for some students, comprehension will always be challenging. Yet, all students must have the opportunity to develop the reading skills necessary to pursue their goals in life. While the process of learning to read requires commitment and effort on the part of the student, it is incumbent on everyone in their lives to do everything within their means to help them achieve reading proficiency. Without this essential skill, their opportunities for success in life will be severely and unnecessarily limited. As Flippo, Holland, McCarthy, and Swinning (2009) eloquently opined, helping “students acquire and enjoy the lifelong habit of reading…is their right, and our responsibility” (p. 82). Educators, administrators, parents, policy makers, and researchers must work together and with students to help them to achieve their potential in life.

Appendix A

Institutional Review Board Letter of Exemption

GEORGE MASON UNIVERSITY

Office of Research Integrity and Assurance

Research Hall, 4400 University Drive, MS 6D5, Fairfax, Virginia 22030
Phone: 703-993-5445; Fax: 703-993-9590

DATE:	March 31, 2016
TO:	Jennifer Buxton
FROM:	George Mason University IRB
Project Title:	[841066-1] Examining the Relationships Among Reading Strategy Instruction, Reading Strategy Use, Reading Attitudes, and Reading Achievement
SUBMISSION TYPE:	New Project
ACTION:	DETERMINATION OF NOT HUMAN SUBJECT RESEARCH
DECISION DATE:	March 31, 2016

Thank you for your submission of New Project materials for this project. The Office of Research Integrity & Assurance (ORIA) has determined this project does not meet the definition of human subject research under the purview of the IRB according to federal regulations.

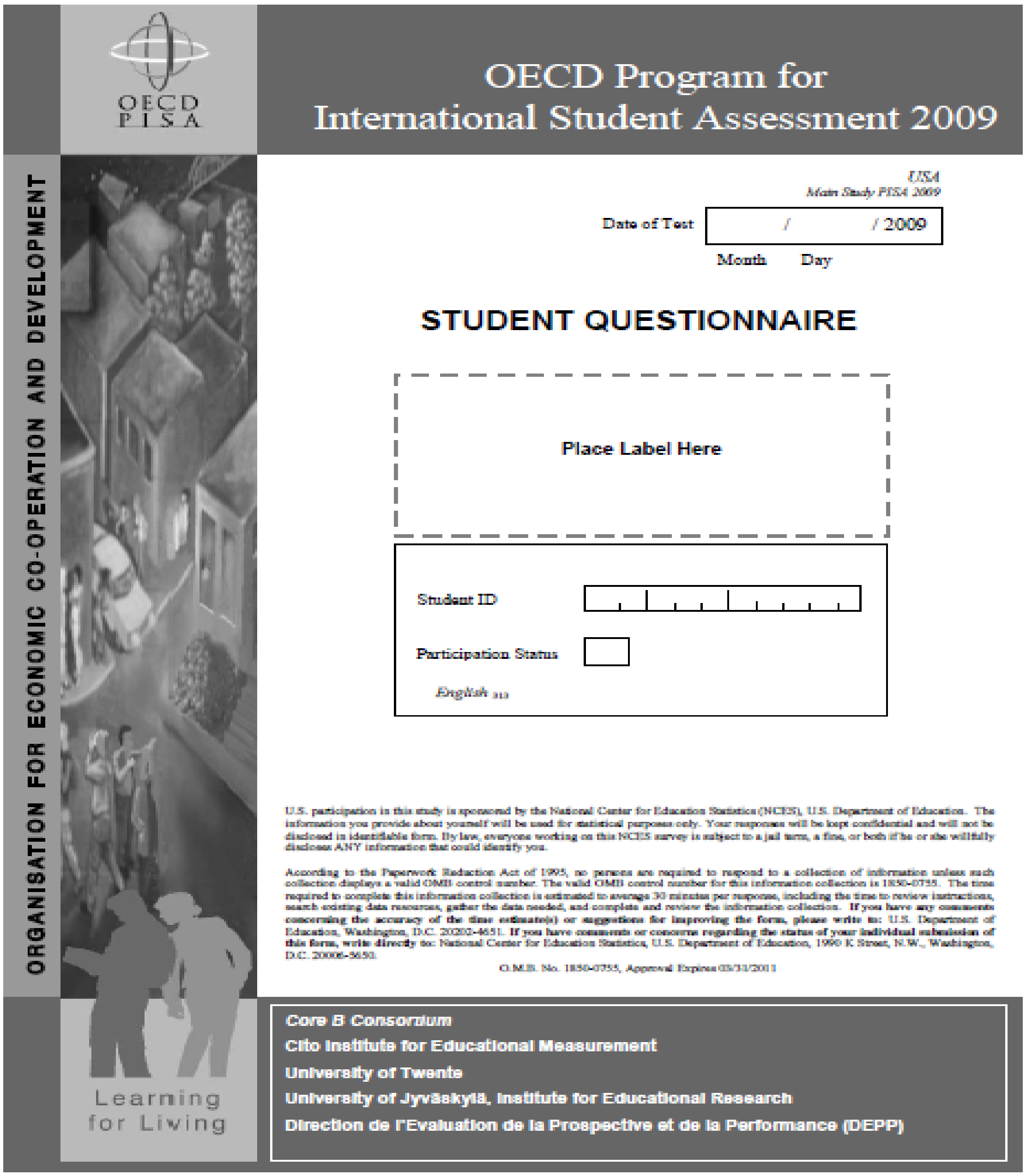
If you have any questions, please contact Karen Motsinger at 703-993-4208 or kmotsing@gmu.edu. Please include your project title and reference number in all correspondence with this committee.

Please remember that if you modify this project to include human subjects research activities, you are required to submit revisions to the ORIA prior to initiation.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within George Mason University IRB's records.

Appendix B

PISA 2009 U.S. Student Questionnaire



OECD PISA

OECD Program for International Student Assessment 2009

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

USA
Main Study PISA 2009

Date of Test / / 2009

Month Day

STUDENT QUESTIONNAIRE

Place Label Here

Student ID

Participation Status

English 313

U.S. participation in this study is sponsored by the National Center for Education Statistics (NCES), U.S. Department of Education. The information you provide about yourself will be used for statistical purposes only. Your responses will be kept confidential and will not be disclosed in identifiable form. By law, everyone working on this NCES survey is subject to a jail term, a fine, or both if he or she willfully discloses ANY information that could identify you.

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. The valid OMB control number for this information collection is 1850-0755. The time required to complete this information collection is estimated to average 30 minutes per response, including the time to review instructions, search existing data resources, gather the data needed, and complete and review the information collection. If you have any comments concerning the accuracy of the time estimate(s) or suggestions for improving the form, please write to: U.S. Department of Education, Washington, D.C. 20202-4651. If you have comments or concerns regarding the status of your individual submission of this form, write directly to: National Center for Education Statistics, U.S. Department of Education, 1990 K Street, N.W., Washington, D.C. 20006-5650.

O.M.B. No. 1850-0755, Approval Expires 03/31/2011

Learning for Living

Core B Consortium
Cito Institute for Educational Measurement
University of Twente
University of Jyväskylä, Institute for Educational Research
Direction de l'Evaluation de la Prospective et de la Performance (DEPP)

In this booklet you will find questions about:

- You
- Your family and your home
- Your reading activities
- Learning time
- Classroom and school climate
- Your English classes
- Libraries
- Your strategies in reading and understanding texts

In some of the questions you will be asked about *reading*. What we specifically mean by reading is the skill to understand, use and think about *written* texts. This skill is needed to reach one's goals, to develop one's knowledge and potential, and to take part in society.

Please read each question carefully and answer as accurately as you can. In the test you usually circled your answers. For this questionnaire, you will normally answer by darkening a circle. For a few questions you will need to write a short answer.

If you make a mistake when darkening a circle, erase your mistake and darken the correct circle. If you make a mistake when writing an answer, simply cross it out and write the correct answer next to it.

In this questionnaire, there are no right or wrong answers. Your answers should be the ones that are right for you.

You may ask for help if you do not understand something or are not sure how to answer a question.

Your answers will be combined with others to make totals and averages in which no individual can be identified. All your answers will be kept confidential.

SECTION 1: ABOUT YOU

Q1 **What grade are you in?**

____
grade

Q2 **How long have you been in this school?**

(Please darken only one circle.)

Less than one year ○[1]

One to two years ○[2]

Three to four years ○[3]

More than four years ○[4]

Q3 **When were you born?**

(Please write the month, day and year you were born)

____________ _______ *19*___
Month *Day* *Year*

Q4 **Are you female or male?**

Female *Male*

○[1] ○[2]

Q5 Which best describes you?

(Please darken only one circle.)

I am Hispanic or Latino.	$\bigcirc_1$
I am <u>not</u> Hispanic or Latino.	$\bigcirc_2$

Q6 Which of these categories best describes your race?

(Please darken one or more circles.)

White	$\bigcirc_1$
Black or African American	$\bigcirc_1$
Asian	$\bigcirc_1$
American Indian or Alaska Native	$\bigcirc_1$
Native Hawaiian or Other Pacific Islander	$\bigcirc_1$

Q7 Did you attend pre-school?

No	O 1
Yes, for one year or less	O 2
Yes, for more than one year	O 3

Q8 Did you attend kindergarten?

No	O 1
Yes	O 2

Q9 How old were you when you started first grade?

______ *years*

Q10 Have you ever repeated a grade?

(Please darken only one circle in each row)

	No, never	*Yes, once*	*Yes, twice or more*
a) In kindergarten	O 1	O 2	O 3
b) In grades 1-6	O 1	O 2	O 3
c) In grades 7-9	O 1	O 2	O 3
d) In grades 10-12	O 1	O 2	O 3

Q11 What is the highest grade or level of school you expect to complete?

(Please darken only one circle.)

Less than high school	O 1
High school	O 2
Vocational or technical certificate (such as cosmetology or auto mechanics)	O 3
Associate's degree (2-year degree from a community college)	O 4
Bachelor's degree (4-year college degree)	O 5
Master's degree	O 6
Doctoral or professional degree such as medicine or law	O 7

SECTION 2: YOUR FAMILY AND YOUR HOME

In this section you will be asked some questions about your family and your home.

Some of the following questions are about your mother and father or those persons who are like a mother or father to you — for example, guardians, step-parents, foster parents, etc.

If you share your time with more than one set of parents or guardians, please answer the following questions for those parents/guardians you spend the most time with.

Q12 Who usually lives at home with you?

(Please darken one circle in each row)

	Yes	No
a) Mother (including stepmother or foster mother)	O_1	O_2
b) Father (including stepfather or foster father)	O_1	O_2
c) Brother(s) (including stepbrothers)	O_1	O_2
d) Sister(s) (including stepsisters)	O_1	O_2
e) Grandparent(s)	O_1	O_2
f) Others (e.g., cousin)	O_1	O_2

Q13a What is your mother's main job?
(e.g., school teacher, cook, sales manager)

(If she is not working now, please tell us her last main job)

Please write in the job title. ______________________________

Q13b What does your mother do in her main job?
(e.g., teaches high school students, helps prepare meals in a restaurant, manages a sales team)

Please use a sentence to describe the kind of work she does or did in that job.

____

Q14 What is the highest level of schooling (not including college) completed by your mother?

If you are not sure which circle to choose, please ask the test administrator for help.

(Please darken only one circle)

She completed grade 12 (high school diploma or GED).	O 1
She completed grade 9.	O 2
She completed grade 6.	O 3
She did not complete grade 6.	O 4

Q15 Does your mother have any of the following degrees, certificates or diplomas?

If you are not sure how to answer this question, please ask the test administrator for help.

(Please darken one circle in each row)

	Yes	*No*
a) Master's, doctoral, or professional degree such as medicine or law	O₁	O₂
b) Bachelor's degree (4-year college degree)	O₁	O₂
c) Associate's degree (2-year degree from a community college)	O₁	O₂
d) Vocational or technical certificate/diploma after high school (such as cosmetology or auto mechanics)	O₁	O₂

Q16 What is your mother currently doing?

(Please darken only one circle)

Working full-time for pay	O₁
Working part-time for pay	O₂
Not working, but looking for a job	O₃
Other (e.g., home duties, retired)	O₄

Q17a What is your father's main job?
(e.g., school teacher, cook, sales manager)

(If he is not working now, please tell us his last main job)

Please write in the job title. ______________________________

Q17b What does your father do in his main job?
(e.g., teaches high school students, helps prepare meals in a restaurant, manages a sales team)

Please use a sentence to describe the kind of work he does or did in that job.

____

Q18 What is the highest level of schooling (not including college) completed by your father?

If you are not sure how to answer this question, please ask the test administrator for help.

(Please darken only one circle)

He completed grade 12 (high school diploma or GED).	O 1
He completed grade 9.	O 2
He completed grade 6.	O 3
He did not complete grade 6.	O 4

Q19 Does your father have any of the following degrees, certificates or diplomas?

If you are not sure which circle to choose, please ask the test administrator for help.

(Please darken one circle in each row)

	Yes	No
a) Master's, doctoral, or professional degree such as medicine or law	O 1	O 2
b) Bachelor's degree (4-year college degree)	O 1	O 2
c) Associate's degree (2-year degree from a community college)	O 1	O 2
d) Vocational or technical certificate/diploma after high school (such as cosmetology or auto mechanics)	O 1	O 2

Q20 What is your father currently doing?

(Please darken only one circle)

Working full-time for pay	O 1
Working part-time for pay	O 2
Not working, but looking for a job	O 3
Other (e.g., home duties, retired)	O 4

Q21 In what country were you and your parents born?

(Please darken one circle in each column)

	You	*Mother*	*Father*
United States*	○ 01	○ 01	○ 01
Other country	○ 02	○ 02	○ 02

*NOTE: the "United States" refers to the 50 states, District of Columbia, and U.S. military bases abroad.

Q22 If <u>you</u> were NOT born in the United States, how old were you when you arrived in the United States?

If you were less than 12 months old, please write zero (0).

If you were born in the United States please skip this question and go to Q23.

_________ *years*

Q23 What language do you speak at home most of the time?

(Please darken only one circle)

English	○ 303
Spanish	○ 156
Other language	○ 859

13

Q24 Which of the following are in your home?

(Please darken one circle in each row)

	Yes	*No*
a) A desk to study at	O1	O2
b) A room of your own	O1	O2
c) A quiet place to study	O1	O2
d) A computer you can use for school work	O1	O2
e) Educational software	O1	O2
f) A link to the Internet	O1	O2
g) Classic literature (e.g., Shakespeare)	O1	O2
h) Books of poetry	O1	O2
i) Works of art (e.g., paintings)	O1	O2
j) Books to help with your school work	O1	O2
k) Technical reference books or manuals	O1	O2
l) A dictionary	O1	O2
m) A dishwasher	O1	O2
n) A DVD player	O1	O2
o) A guest room	O1	O2
p) A high-speed Internet connection	O1	O2
q) A musical instrument	O1	O2

Q25 **<u>How many</u> of these are there at your home?**

(Please darken only one circle in each row)

	None	One	Two	Three or more
a) Cell phones	O	O	O	O4
b) Televisions	O	O	O	O4
c) Computers	O	O	O	O4
d) Cars	O	O	O	O4
e) Bathrooms with a bathtub or shower	O	O	O	O4

Q26 **How many books are there in your home?**

There are usually about 15 books per foot of shelving. Do not include magazines, newspapers, or your schoolbooks.

(Please darken only one circle)

0-10 books	O1
11-25 books	O2
26-100 books	O3
101-200 books	O4
201-500 books	O5
More than 500 books	O6

SECTION 3: YOUR READING ACTIVITIES

The questions in this section are mainly about your reading activities outside school.

Q27 About how much time do you usually spend reading for enjoyment?

(Please darken only one circle)

I do not read for enjoyment.	O 1
30 minutes or less a day	O 2
More than 30 minutes to less than 60 minutes a day	O 3
1 to 2 hours a day	O 4
More than 2 hours a day	O 5

Q28 How much do you agree or disagree with these statements about reading?

(Please darken only one circle in each row)

	Strongly disagree	*Disagree*	*Agree*	*Strongl agree*
a) I read only if I have to.	O 1	O 2	O 3	O 4
b) Reading is one of my favorite hobbies.	O 1	O 2	O 3	O 4
c) I like talking about books with other people.	O 1	O 2	O 3	O 4
d) I find it hard to finish books.	O 1	O 2	O 3	O 4
e) I feel happy if I receive a book as a present.	O 1	O 2	O 3	O 4
f) For me, reading is a waste of time.	O 1	O 2	O 3	O 4
g) I enjoy going to a bookstore or a library.	O 1	O 2	O 3	O 4
h) I read only to get information that I need.	O 1	O 2	O 3	O 4
i) I cannot sit still and read for more than a few minutes.	O 1	O 2	O 3	O 4
j) I like to express my opinions about books I have read.	O 1	O 2	O 3	O 4
k) I like to exchange books with my friends.	O 1	O 2	O 3	O 4

Q29 How often do you read these materials <u>because you want to</u>?

(Please darken only one circle in each row)

	Never or almost never	*A few times a year*	*About once a month*	*Several times a month*	*Several times a week*
a) Magazines	O 1	O 2	O 3	O 4	O 5
b) Comic books	O 1	O 2	O 3	O 4	O 5
c) Fiction (e.g., novels, narratives, stories)	O 1	O 2	O 3	O 4	O 5
d) Non-fiction books	O 1	O 2	O 3	O 4	O 5
e) Newspapers	O 1	O 2	O 3	O 4	O 5

Q30 How often are you involved in the following reading activities?

(Please darken only one circle in each row. If you don't know what the activity is, darken "I don't know what it is.")

	I don't know what it is	Never or almost never	Several times a month	Several times a week	Several times a day
a) Reading emails	O1	O2	O3	O4	O5
b) Chatting on line	O1	O2	O3	O4	O5
c) Reading online news	O1	O2	O3	O4	O5
d) Using an online dictionary or encyclopedia (e.g., Wikipedia®)	O1	O2	O3	O4	O5
e) Searching online information to learn about a particular topic	O1	O2	O3	O4	O5
f) Taking part in online group discussions or forums	O1	O2	O3	O4	O5
g) Searching for practical information on line (e.g., schedules, events, tips, recipes)	O1	O2	O3	O4	O5
h) Text-messaging	O1	O2	O3	O4	O5

Q31 When you are studying, how often do you do the following?

(Please darken only one circle in each row)

	Almost never	*Sometimes*	*Often*	*Almost always*
a) When I study, I try to memorize everything that is covered in the text.	O1	O2	O3	O4
b) When I study, I start by figuring out what exactly I need to learn.	O1	O2	O3	O4
c) When I study, I try to memorize as many details as possible.	O1	O2	O3	O4
d) When I study, I try to relate new information to prior knowledge acquired in other subjects.	O1	O2	O3	O4
e) When I study, I read the text so many times that I can recite it.	O1	O2	O3	O4
f) When I study, I check if I understand what I have read.	O1	O2	O3	O4
g) When I study, I read the text over and over again.	O1	O2	O3	O4
h) When I study, I figure out how the information might be useful outside school.	O1	O2	O3	O4
i) When I study, I try to figure out which concepts I still haven't really understood.	O1	O2	O3	O4
j) When I study, I try to understand the material better by relating it to my own experiences.	O1	O2	O3	O4
k) When I study, I make sure that I remember the most important points in the text.	O1	O2	O3	O4
l) When I study, I figure out how the text information fits in with what happens in real life.	O1	O2	O3	O4
m) When I study and I don't understand something, I look for additional information to clarify this.	O1	O2	O3	O4

SECTION 4: LEARNING TIME

Q32 How many minutes, on average, are there in a class period for the following subjects?

Minutes in a class period in English (English classes may include those in literature, creative writing, journalism, etc.):	________	Minutes
Minutes in a class period in mathematics:	________	Minutes
Minutes in a class period in science:	________	Minutes

Q33 How many class periods per week do you typically have for the following subjects?

Number of class periods per week in English:	________	class periods
Number of class periods per week in mathematics:	________	class periods
Number of class periods per week in science:	________	class periods

Q34 In a normal, full week at school, how many class periods do you have in total?

Number of ALL class periods:	________	class periods

Q35 What type of out-of-school-time lessons do you attend currently?

These are only lessons in subjects that you are also learning at school, that you spend extra time learning outside of normal school hours. The lessons may be given at your school, at your home or somewhere else.

(Please darken only one circle in each row)

	Yes	*No*
a) Enrichment lessons in English	O 1	O 2
b) Enrichment lessons in mathematics	O 1	O 2
c) Enrichment lessons in science	O 1	O 2
d) Enrichment lessons in other school subjects	O 1	O 2
e) Remedial lessons in English	O 1	O 2
f) Remedial lessons in mathematics	O 1	O 2
g) Remedial lessons in science	O 1	O 2
h) Remedial lessons in other school subjects	O 1	O 2
i) Lessons to improve your study skills	O 1	O 2

Q36 How many hours do you typically spend per week attending out-of-school-time lessons in the following subjects (at school, at home or somewhere else)?

These are only lessons in subjects that you are also learning at school, that you spend extra time learning outside of normal school hours. The lessons may be given at your school, at your home or somewhere else.

(Please darken one circle in each column)

	English	*Mathematics*	*Science*	*Other subjects*
I do not attend out-of-school-time lessons in these subjects.	O1	O1	O1	O1
Less than 2 hours a week	O2	O2	O2	O2
2 or more but less than 4 hours a week	O3	O3	O3	O3
4 or more but less than 6 hours a week	O4	O4	O4	O4
6 or more hours a week	O5	O5	O5	O5

SECTION 5: YOUR SCHOOL

Q37 *Thinking about what you have learned in school:* **To what extent do you agree or disagree with the following statements?**

(Please darken only one circle in each row)

	Strongly disagree	*Disagree*	*Agree*	*Strongly agree*
a) School has done little to prepare me for adult life when I leave school.	O 1	O 2	O 3	O 4
b) School has been a waste of time.	O 1	O 2	O 3	O 4
c) School has helped give me confidence to make decisions.	O 1	O 2	O 3	O 4
d) School has taught me things which could be useful in a job.	O 1	O 2	O 3	O 4

Q38 How much do you disagree or agree with each of the following statements about teachers at your school?

(Please darken only one circle in each row.)

	Strongly disagree	*Disagree*	*Agree*	*Strongly agree*
a) I get along well with most of my teachers.	O 1	O 2	O 3	O 4
b) Most of my teachers are interested in my well-being.	O 1	O 2	O 3	O 4
c) Most of my teachers really listen to what I have to say.	O 1	O 2	O 3	O 4
d) If I need extra help, I will receive it from my teachers.	O 1	O 2	O 3	O 4
e) Most of my teachers treat me fairly.	O 1	O 2	O 3	O 4

SECTION 6: YOUR ENGLISH CLASSES

Q39 On average, about how many students attend your English class(es)?

________ *students*

Q40 How often do these things happen in your English classes?

(Please darken only one circle in each row)

	Never or hardly ever	*In some classes*	*In most classes*	*In all classes*
a) Students don't listen to what the teacher says.	O 1	O 2	O 3	O 4
b) There is noise and disorder.	O 1	O 2	O 3	O 4
c) The teacher has to wait a long time for the students to quiet down.	O 1	O 2	O 3	O 4
d) Students cannot work well.	O 1	O 2	O 3	O 4
e) Students don't start working for a long time after the class begins.	O 1	O 2	O 3	O 4

Q41 In your English classes, how often does the following occur?

(Please darken only one circle in each row)

	Never or hardly ever	*In some classes*	*In most classes*	*In all classes*
a) The teacher asks students to explain the meaning of a text.	O 1	O 2	O 3	O 4
b) The teacher asks questions that challenge students to get a better understanding of a text.	O 1	O 2	O 3	O 4
c) The teacher gives students enough time to think about their answers.	O 1	O 2	O 3	O 4
d) The teacher recommends a book or author to read.	O 1	O 2	O 3	O 4
e) The teacher encourages students to express their opinions about a text.	O 1	O 2	O 3	O 4
f) The teacher helps students relate the stories they read to their lives.	O 1	O 2	O 3	O 4
g) The teacher shows students how the information in texts builds on what they already know.	O 1	O 2	O 3	O 4

Q42 In your English classes, how often does the following occur?

(Please darken only one circle in each row)

	Never or hardly ever	*In some classes*	*In most classes*	*In all classes*
a) The teacher explains beforehand what is expected of the students.	O1	O2	O3	O4
b) The teacher checks that students are concentrating while working on the reading assignment.	O1	O2	O3	O4
c) The teacher discusses students' work, after they have finished the reading assignment.	O1	O2	O3	O4
d) The teacher tells students in advance how their work is going to be judged.	O1	O2	O3	O4
e) The teacher asks whether every student has understood how to complete the reading assignment.	O1	O2	O3	O4
f) The teacher grades students' work.	O1	O2	O3	O4
g) The teacher gives students the chance to ask questions about the reading assignment.	O1	O2	O3	O4
h) The teacher poses questions that motivate students to participate actively.	O1	O2	O3	O4
i) The teacher tells students how well they did on the reading assignment immediately after.	O1	O2	O3	O4

SECTION 7: LIBRARIES

In this section you are asked questions about libraries. These may be in your school and/or outside your school.

Q43 How often do you visit a library for the following activities?

(Please darken only one circle in each row)

	Never	*A few times a year*	*About once a month*	*Several times a month*	*Several times a week*
a) Borrow books to read for pleasure	O 1	O 2	O 3	O 4	O 5
b) Borrow books for school work	O 1	O 2	O 3	O 4	O 5
c) Work on homework, course assignments or research papers	O 1	O 2	O 3	O 4	O 5
d) Read magazines or newspapers	O 1	O 2	O 3	O 4	O 5
e) Read books for fun	O 1	O 2	O 3	O 4	O 5
f) Learn about things that are not course-related, such as sports, hobbies, people or music	O 1	O 2	O 3	O 4	O 5
g) Use the Internet	O 1	O 2	O 3	O 4	O 5

Q44 Does your school have a school library?

No	O 1
Yes	O 2

SECTION 8: YOUR STRATEGIES IN READING AND UNDERSTANDING TEXTS

There are several approaches to studying and understanding texts. Some of them are more useful than others, depending on the kind of reading task. The next two questions present two reading tasks, followed by a list of these approaches or "strategies." We want to know your opinion about the usefulness of these strategies for the different reading tasks.

Both questions begin with a short description of a particular reading task. Then several possible reading strategies are listed. Think about the usefulness of each of the strategies in relation to the given reading task only. Some strategies may be useful for one reading task but not for another.

Give a score between 1 and 6 to each strategy. A score of 1 means you think it is not a useful strategy at all for this reading task. A score of 6 means you think it is a very useful strategy for this reading task.

You can use the same score more than once if you think two or more strategies are similarly useful, but please darken only one circle in each row.

Here is an example question that a student has completed. (This example is about playing table tennis, not reading.)

Example Question

Task: You want to improve at playing table tennis so you can win a local competition.
How do you rate the usefulness of the following strategies for improving at playing table tennis?

Possible strategy	**Score** *Not useful at all (1)*	*(2)*	*(3)*	*(4)*	*(5)*	*Very useful (6)*
a) I read a book about table tennis technique.	o	●	o	o	o	o
b) I practice playing table tennis against a friend as often as possible.	o	o	o	o	o	●
c) I do general fitness exercises every morning.	o	o	o	●	o	o
d) I watch expert players and try to figure out their techniques.	o	o	o	●	o	o

Q45 Reading task: You have to understand and remember the information in a text.

How do you rate the usefulness of the following strategies for understanding and memorizing the text?

Possible strategy	**Score** *Not useful at all* (1)	(2)	(3)	(4)	(5)	*Very useful* (6)
a) I concentrate on the parts of the text that are easy to understand.	O$_1$	O$_2$	O$_3$	O$_4$	O$_5$	O$_6$
b) I quickly read through the text twice.	O$_1$	O$_2$	O$_3$	O$_4$	O$_5$	O$_6$
c) After reading the text, I discuss its content with other people.	O$_1$	O$_2$	O$_3$	O$_4$	O$_5$	O$_6$
d) I underline important parts of the text.	O$_1$	O$_2$	O$_3$	O$_4$	O$_5$	O$_6$
e) I summarize the text in my own words.	O$_1$	O$_2$	O$_3$	O$_4$	O$_5$	O$_6$
f) I read the text aloud to another person.	O$_1$	O$_2$	O$_3$	O$_4$	O$_5$	O$_6$

Q46 Reading task: You have just read a long and rather difficult two-page text about fluctuations in the water level of a lake in Africa. You have to write a summary.

How do you rate the usefulness of the following strategies for writing a summary of this two-page text?

Possible strategy	Score *Not useful at all* *(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*Very useful* *(6)*
a) I write a summary. Then I check that each paragraph is covered in the summary, because the content of each paragraph should be included.	O 1	O 2	O 3	O 4	O 5	O 6
b) I try to copy out accurately as many sentences as possible.	O 1	O 2	O 3	O 4	O 5	O 6
c) Before writing the summary, I read the text as many times as possible.	O 1	O 2	O 3	O 4	O 5	O 6
d) I carefully check whether the most important facts in the text are represented in the summary.	O 1	O 2	O 3	O 4	O 5	O 6
e) I read through the text, underlining the most important sentences. Then I write them in my own words as a summary.	O 1	O 2	O 3	O 4	O 5	O 6

Thank you very much for your cooperation in completing this questionnaire!

Appendix C

PISA 2009 Reading for School Questionnaire

Questions about reading for school

We want to know about the kind of reading you typically do for school, either in the classroom or for homework.

Q1 During the last month, how often did you have to read the following types of texts <u>for school</u> (in the classroom or for homework)? *RFS1*

(Please tick only one box in each row.)

		Many times	*Two or three times*	*Once*	*Not at all*
a)	Information texts about writers or books	☐1	☐2	☐3	☐4
b)	Poetry	☐1	☐2	☐3	☐4
c)	Texts that include diagrams or maps	☐1	☐2	☐3	☐4
d)	Fiction (e.g., novels, short stories)	☐1	☐2	☐3	☐4
e)	Newspaper reports and magazine articles	☐1	☐2	☐3	☐4
f)	Instructions or manuals telling you how to make or do something (e.g., how a machine works)	☐1	☐2	☐3	☐4
g)	Texts that include tables or graphs	☐1	☐2	☐3	☐4
h)	Advertising material (e.g., advertisements in magazines, posters)	☐1	☐2	☐3	☐4

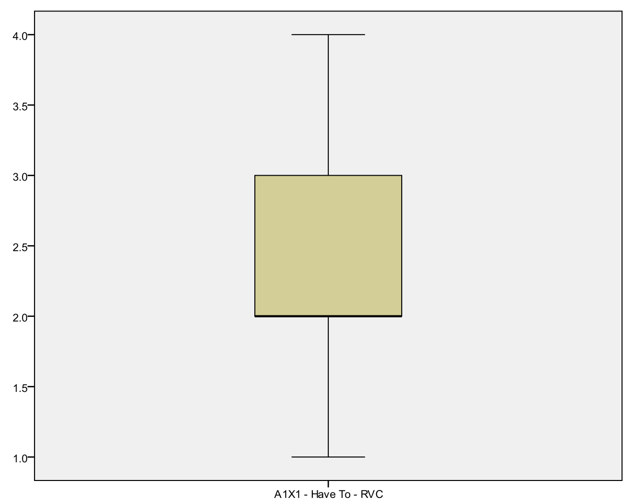
Q2 During the last month, how often did you have to do the following kinds of tasks <u>for school</u> (in the classroom or for homework)? *RFS2*

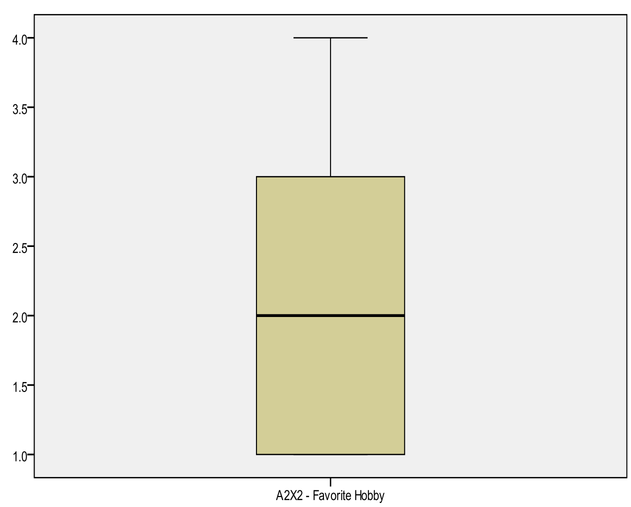
(Please tick only one box in each row.)

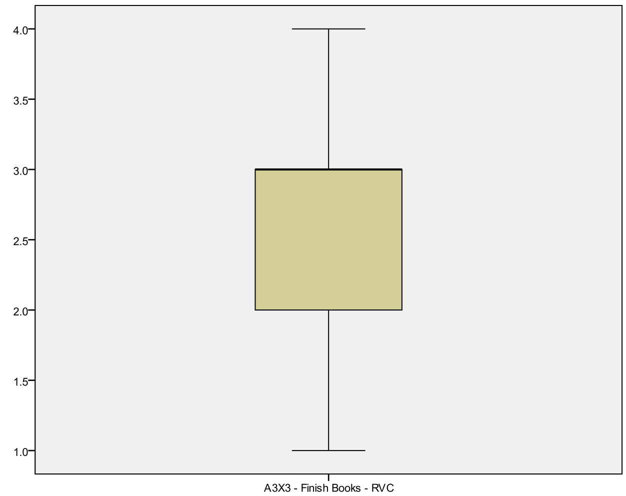
		Many times	*Two or three times*	*Once*	*Not at all*
a)	Find information from a graph, diagram or table	☐1	☐2	☐3	☐4
b)	Explain the cause of events in a text	☐1	☐2	☐3	☐4
c)	Explain the way characters behave in a text	☐1	☐2	☐3	☐4
d)	Learn about the life of the writer	☐1	☐2	☐3	☐4
e)	Explain the purpose of a text	☐1	☐2	☐3	☐4
f)	Memorise a text by heart (e.g., a poem or part of a play)	☐1	☐2	☐3	☐4
g)	Learn about the place of a text in the history of literature	☐1	☐2	☐3	☐4
h)	Describe the way the information in a table or graph is organised	☐1	☐2	☐3	☐4
i)	Explain the connection between different parts of a text (e.g., between a written part and a map)	☐1	☐2	☐3	☐4

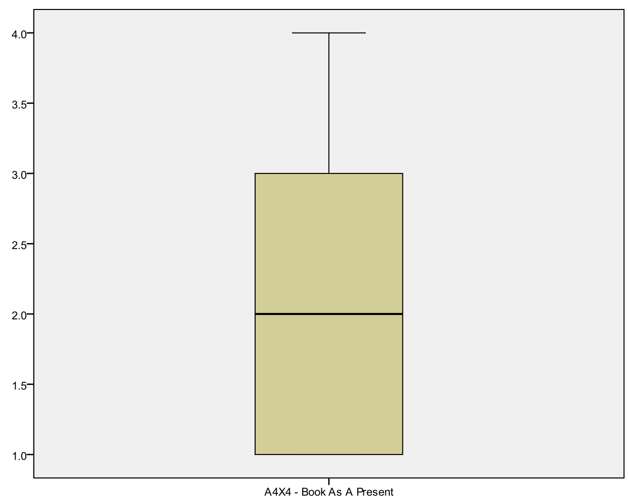
Appendix D

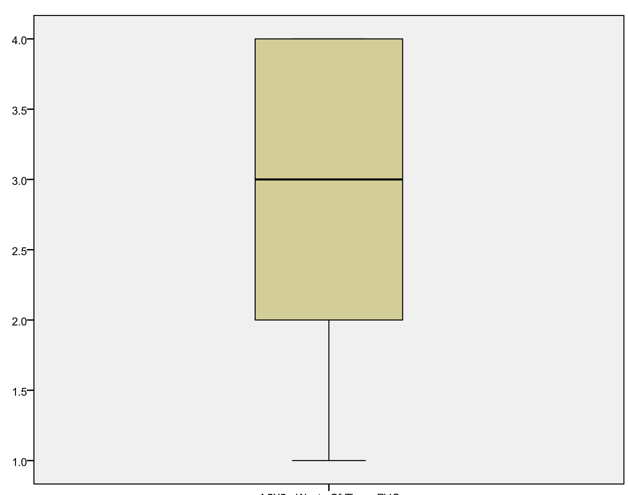
Box Plots of Observed Latent Variable Indicators

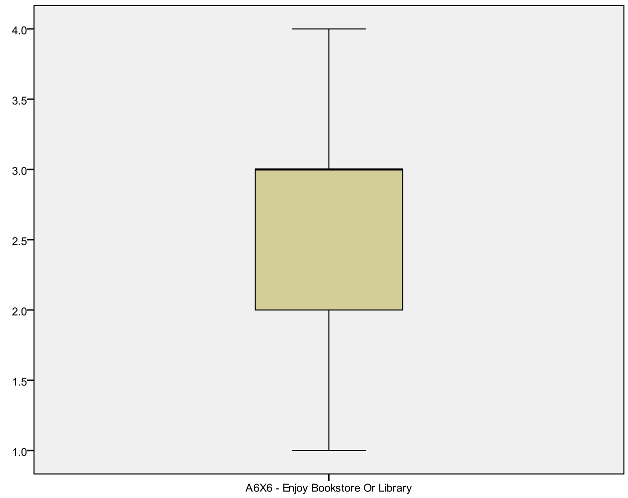


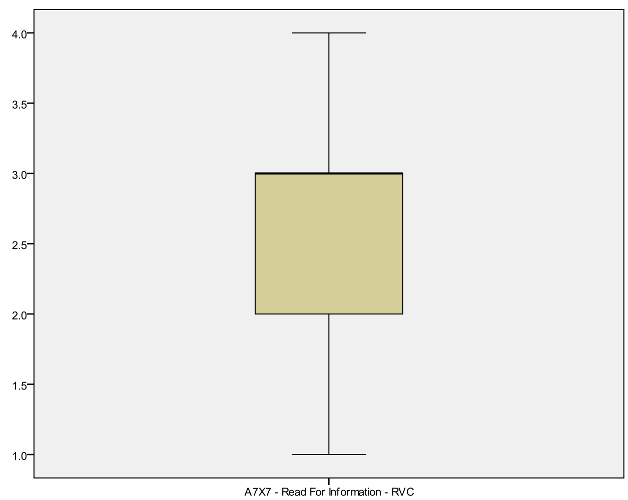


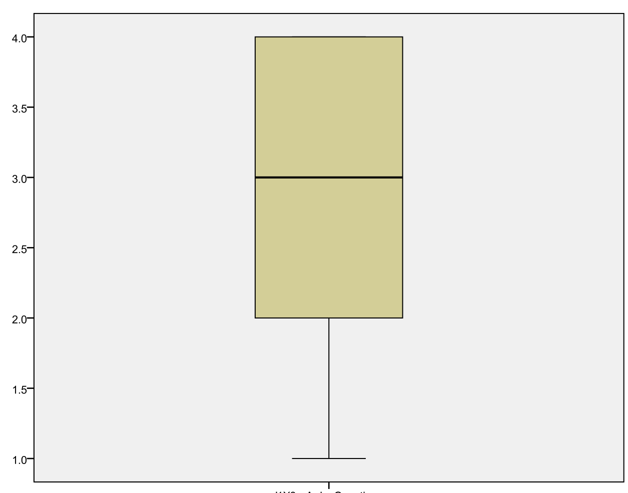

4.0
3.5
3.0
2.5
2.0
1.5
1.0
A3X3 - Finish Books - RVC

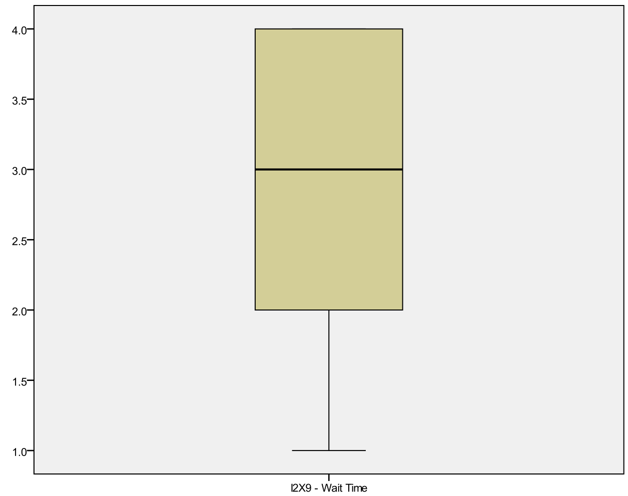

4.0
3.5
3.0
2.5
2.0
1.5
1.0
A4X4 - Book As A Present

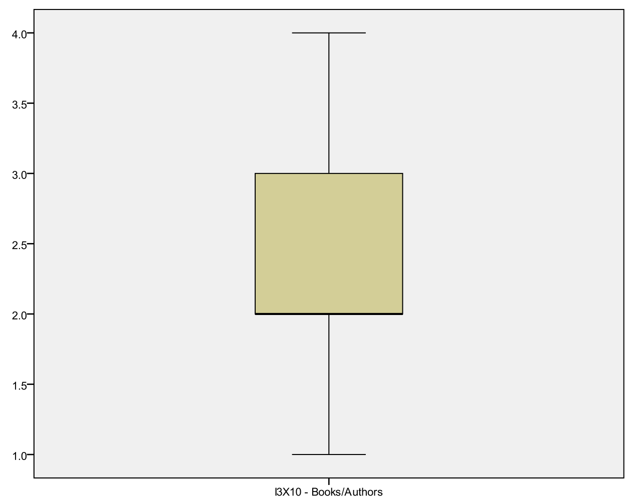

4.0
3.5
3.0
2.5
2.0
1.5
1.0
A5X5 - Waste Of Time - RVC

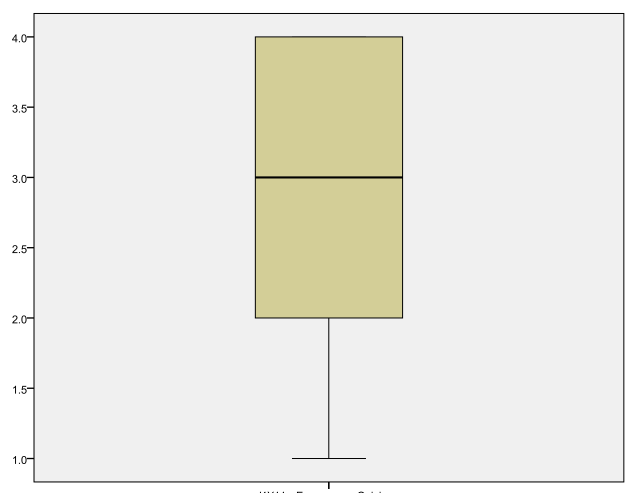

4.0
3.5
3.0
2.5
2.0
1.5
1.0
A6X6 - Enjoy Bookstore Or Library

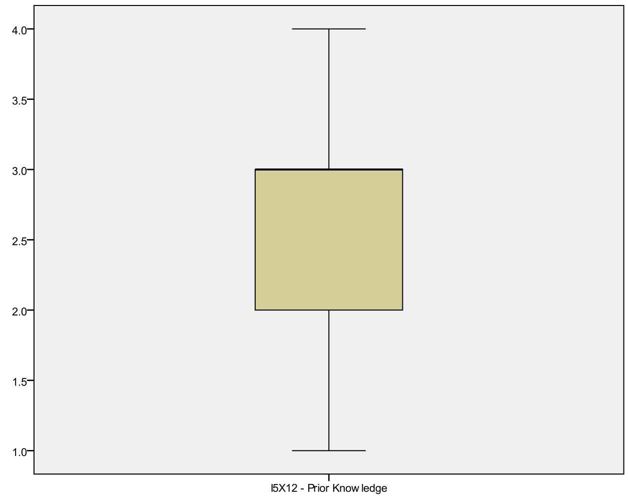

4.0
3.5
3.0
2.5
2.0
1.5
1.0
A7X7 - Read For Information - RVC

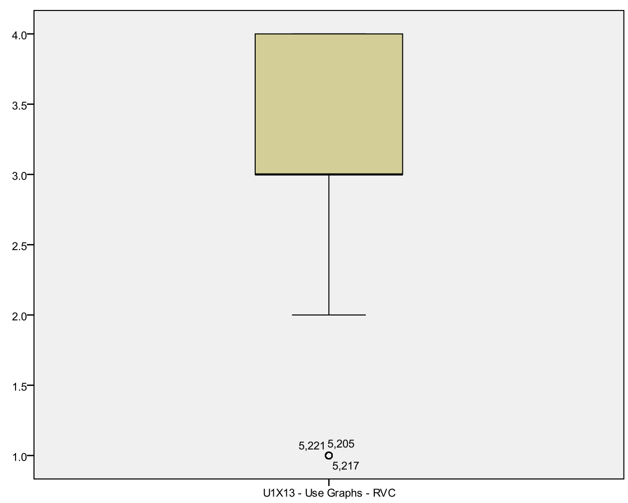

4.0
3.5
3.0
2.5
2.0
1.5
1.0
I1X8 - Asks Questions

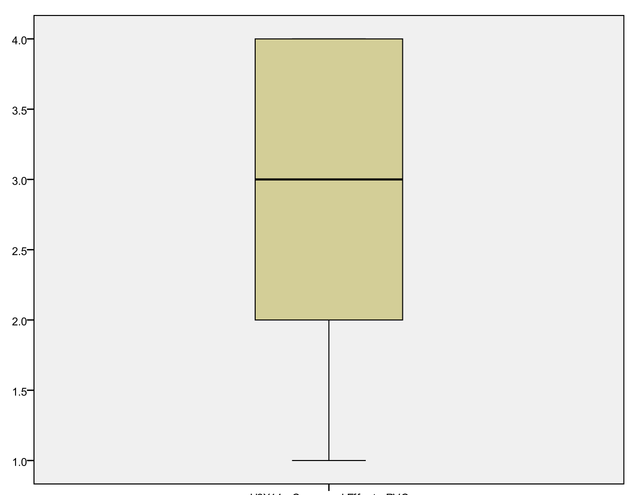

4.0
3.5
3.0
2.5
2.0
1.5
1.0
I2X9 - Wait Time

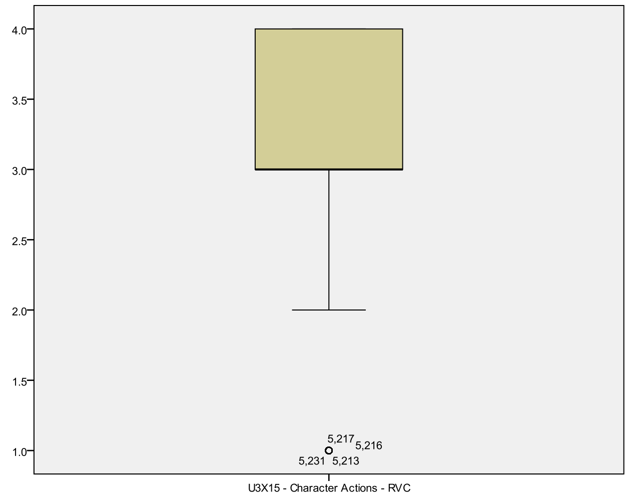

4.0
3.5
3.0
2.5
2.0
1.5
1.0
I3X10 - Books/Authors

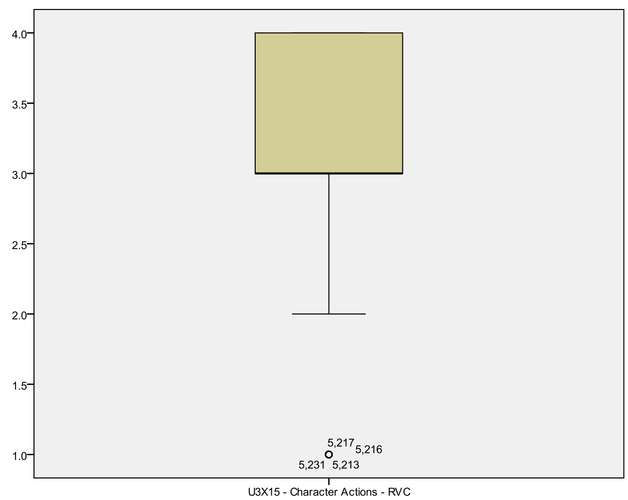

4.0
3.5
3.0
2.5
2.0
1.5
1.0
I4X11 - Encourages Opinions


4.0
3.5
3.0
2.5
2.0
1.5
1.0
I5X12 - Prior Knowledge

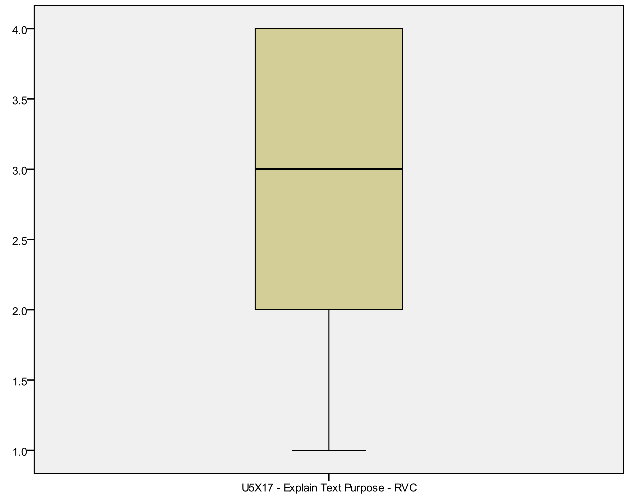

4.0
3.5
3.0
2.5
2.0
1.5
1.0
5,221
5,205
5,217
U1X13 - Use Graphs - RVC

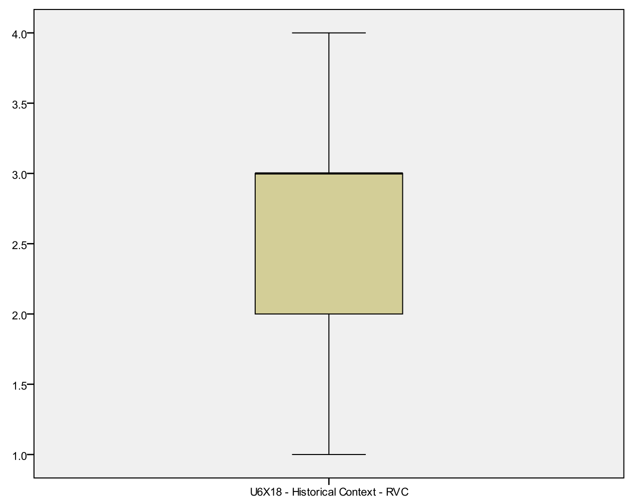

4.0
3.5
3.0
2.5
2.0
1.5
1.0
U2X14 - Cause and Effect - RVC

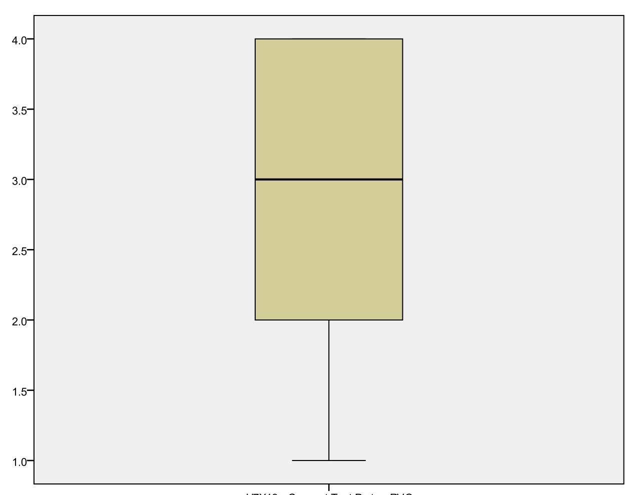

4.0
3.5
3.0
2.5
2.0
1.5
1.0
5,217
5,216
5,231
5,213
U3X15 - Character Actions - RVC


4.0
3.5
3.0
2.5
2.0
1.5
1.0
5,217
5,216
5,231
5,213
U3X15 - Character Actions - RVC


4.0
3.5
3.0
2.5
2.0
1.5
1.0
U4X16 - Author Background - RVC


4.0
3.5
3.0
2.5
2.0
1.5
1.0
U5X17 - Explain Text Purpose - RVC


4.0
3.5
3.0
2.5
2.0
1.5
1.0
U6X18 - Historical Context - RVC


4.0
3.5
3.0
2.5
2.0
1.5
1.0
U7X19 - Connect Text Parts - RVC

References

Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *Journal of Learning Disabilities, 39,* 414-431. doi:10.1177/00222194060390050401

Alexander, J. E., & Filler, R. C. (1976). *Attitudes and reading*. Newark, DE: International Reading Association.

Alfassi, M. (2004). Reading to learn: Effects of combined strategy instruction on high school students. *The Journal of Educational Research, 97*, 171-184. doi:10. 3200/JOER.97.4.171-185

Allington, R. L., & McGill-Franzen, A. (2013). Eliminating summer reading setback: How we can close the rich/poor reading achievement gap. *Reading Today,* 10-11.

Allington, R. L., McGill-Franzen, A., Camilli, G., Williams, L., Graff, J., Zeig, J.,…Nowak, R. (2010). Addressing summer reading setback among economically disadvantaged elementary students. *Reading Psychology, 31*, 411-427. doi:10. 1080/02702711.2010.505165

Allison, P. D. (2000). Multiple imputation for missing data: A cautionary tale. *Sociological Methods & Research, 28*(3), 301-309. doi:10.1177/ 0049124100028003003

Anastasiou, D., & Griva, E. (2009). Awareness of reading strategy use and reading comprehension among poor and good readers. *Elementary Education Online, 8,* 283-297.

Andreassen, R., & Bråten, I. (2010). Examining the prediction of reading comprehension on different multiple-choice tests. *Journal of Research in Reading, 33,* 263-283. doi:10.1111/j.1467-9817.2009.01413.x

Applegate, A. J., & Applegate, M. D. (2004). The Peter Effect: Reading habits and attitudes of preservice teachers. *The Reading Teacher, 57,* 554-563.

Artelt, C., Schiefele, U., & Schneider, W., (2001). Predictors of reading literacy. *European Journal of Psychology of Education, 16*, 363-383. doi:10.1007/bf03173188

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY: Freeman.

Bandura, A. (1999). Social cognitive theory: An agentic perspective. *Asian Journal of Social Psychology, 2*, 21-41. doi:10.1111/1467-839x.00024

Bastug, M. (2014). The structural relationship of reading attitude, reading comprehension, and academic achievement. *International Journal of Social Sciences & Education, 4*, 931-946.

Berkeley, S., Marshak, L., Mastropieri, M. A., & Scruggs, T. E. (2011). Improving student comprehension of social studies text: A self-questioning strategy for inclusive middle school classes. *Remedial and Special Education, 32*, 105-113. doi:10.1177/0741932510361261

Biancarosa, C., & Snow, C. E. (2006). *Reading next – A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.

Blatchford, P. (2003). A systematic observational study of teachers' and pupils' behaviour in large and small classes. *Learning and Instruction, 13*(6), 569-595. doi:10.1016/S0959-4752(02)00043-9

Blatchford, P., Bassett, P., & Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher-pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and Instruction, 21*(6), 715-730. doi:10.1016/j.learninstruc.2011.04.001

Blatchford, P., & Mortimore, P. (1994). The issue of class size for young children in schools: What can we learn from research? *Oxford Review of Education, 20*, 411-428. doi:10.1080/0305498940200402

Boardman, A. G., Vaughn, S. R., Buckley, P. Reutebuch, C., Roberts, G., & Klingner, J. K. (2016). Collaborative strategic reading for students with learning disabilities in upper elementary classrooms. *Exceptional Children, 82*, 409-427. doi:10.1177/0014402915625067

Bokhorst-Heng, W., & Pereira, D. (2008). Non-at-risk adolescents' attitudes towards reading in a Singapore secondary school. *Journal of Research in Reading, 31*, 285-301. doi:10.1111/j.1467-9817.2008.00369.x

Bollen, K. A. (1987). Total, direct, and indirect effects in structural equation models. In C. C. Clogg (Ed.), *Sociological methodology* (pp. 37-69). Washington, DC: American Sociological Association.

Bouchamma, Y., Poulin, V., & Ruel, C. (2014). Impact of reading strategy use on girls' and boys' achievement. *Reading Psychology, 35*, 312-331. doi:10.1080/02702711.2012.724043

Brown-Jeffy, S. (2006). The race gap in high school reading achievement: Why school racial composition still matters. *Race, Gender, & Class, 13*(3/4), 268-294.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Brozo, W. G., Sulkunen, S., Shiel, G., Garbe, C., Pandian, A., & Valtin, R. (2014). Reading, gender, and engagement: Lessons from five PISA countries. *Journal of Adolescent & Adult Literacy, 57*, 584-593. doi:10.1002/jaal.291

Bryant, D. P., Bryant, B. R., Flower, A., McKenna, J., Muething, C., Shin, M., & Kim, M. (2013). *Secondary special education observation and intervention study: Technical report*. Austin, TX: Meadows Center for Preventing Educational Risk and Texas Education Agency.

Buckingham, J., Beaman, R., & Wheldall, K. (2014). Why poor children are more likely to become poor readers: The early years. *Educational Review, 66*, 428-446. doi:10.1080/00131911.2013.795129

Bussert-Webb, K., & Zhang, Z. (2016). Reading attitudes of Texas high school students. *Reading Psychology, 37*, 424-448. doi:10.1080/02702711.2015.1059396

Byrne, B. M. (2004). Testing for multigroup invariance using AMOS graphics: A road less traveled. *Structural Equation Modeling: A Multidisciplinary Journal, 11*, 272-300. doi:10.1207/s15328007sem1102_8

Caccamise, D., Franzke, M., Eckhoff, A., Kintsch, E., & Kintsch, W. (2007). Guided practice in technology-based summary writing. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 375-396). New York, NY: Erlbaum.

Calhoon, M. B. (2005). Effects of a peer-mediated phonological skill and reading comprehension program on reading skill acquisition for middle school students with reading disabilities. *Journal of Learning Disabilities, 38,* 424-433. doi:10.1177/00222194050380050501

Cantrell, S. C., Almasi, J. F., Carter, J. C., Rintamaa, M., & Madden, A. (2010). The impact of a strategy-based intervention on the comprehension and strategy use of struggling adolescent readers. *Journal of Educational Psychology, 102,* 257-280. doi:10.1037/a0018212

Cantrell, S. C., & Carter, J. C. (2009). Relationships among learner characteristics and adolescents' perceptions about reading strategy use. *Reading Psychology, 30,* 195-224. doi:10.1080/02702710802275397

Caro, D. H., McDonald, J. T., & Willms, J. D. (2009). Socio-economic status and academic achievement trajectories from childhood to adolescence. *Canadian Journal of Education, 32,* 558-590.

Cattaneo, M. A., Oggenfuss, C., & Wolter, S. C. (2016). *The more, the better? The impact of instructional time on student performance* (IZA Discussion Paper No. 9797). Retrieved from The Institute for the Study of Labor (IZA) website: http://ftp.iza.org/dp9797.pdf

Chamot, A. U., & O'Malley, J. M. (1987). The cognitive academic language learning approach: A bridge to the mainstream. *TESOL Quarterly, 21,* 227-249. doi:10. 2307/3586733

Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators, of early achievement: Evidence from the Early Childhood Longitudinal Study (ECLS). *Journal of Educational Psychology, 98,* 489-507. doi:10.1037/0022-0663. 98.3.489

Cheema, J. R. (2012). *Handling missing data in education research using SPSS* (Doctoral dissertation). Retrieved from ProQuest Dissertation and Theses database. (UMI No. 3503006)

Cheema, J. R. (2013). Does it matter how you measure it? The case of self-efficacy in mathematics. *Issues in Educational Research, 23,* 346-356.

Cheung, G.W., & Rensvold, R. B. (1999, August). *What constitutes significant differences in evaluating measurement invariance?* Paper presented at the 1999 conference of the Academy of Management, Chicago, IL.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 233-255. doi:10.1207/S15328007SEM0902_5

Chiu, M. M., & McBride-Chang, C. (2006). Gender, context, and reading: A comparison of students in 43 countries. *Scientific Studies of Reading, 10,* 331-362. doi:10. 1207/s1532799xssr1004_1

Clark, S. K., Dugan, C., Moretine, T., Prior, J. O., Ray, J., Rosenburg, M., & Trischitta, A. (2004). *Successful strategies for reading in the content areas: For grades 3-5*. Huntington Beach, CA: Shell.

Cohen, J. (1992). A Power Primer. *Psychological Bulletin, 112*, 155-159. doi:10.10. 1037/0033-2909.112.1.155

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, A. D., York, R. L. (1966). *Equality of educational opportunity* (Report No. OE-38001). Washington, DC: U.S. Department of Health Education and Welfare, U.S. Office of Education, U.S. Government Printing Office.

Conley, M. W. (2008). Cognitive strategy instruction for adolescents: What we know about the promise, what we don't know about the potential. *Harvard Educational Review, 78,* 84-106. doi:10.17763/haer.78.1.j612282134673638

Conradi, K., Jang, B. G., & McKenna, M. C. (2014). Motivational terminology in reading research: A conceptual review. *Educational Psychology Review, 26*, 127-164. doi:10.1007/s10648-013-9245-z

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 99*, 311-325. doi:10.1037/0022-0663.99.2.311

Cromley, J. G., & Azevedo, R. (2011). Measuring strategy use in context with multiple-choice items. *Metacognition Learning, 6*, 155-177. doi:10.1007/s11409-011-9070-z

Cutting, L. E., Materek, A., Cole, C. A. S., Levine, T. M., & Mahone, E. M. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Annals of Dyslexia, 59*, 34-54. doi:10.1007/ s11881-009-0022-0

Dawson, R. (2011). How significant is a boxplot outlier? *Journal of Statistics Education, 19*(2), 1-13.

Denton, C. A., Wolters, C. A., York, M. J., Swanson, E., Kulesz, P. A., & Francis, D. J. (2015). Adolescents' use of reading comprehension strategies: Differences related to reading proficiency, grade level, and gender. *Learning and Individual Differences, 37*, 81-95. doi:10.1016/j.lindif.2014.11.016

Denton, D. R. (1999). *Reading reform in the SREB states: Early assessment*. Atlanta, GA: Southern Regional Education Board.

Deshler, D. & Hock, M. (2007). Adolescent literacy: Where we are, where we need to go. In M. Pressley, A. Billman, K. Perry, K. Reffitt, & J. Reynolds (Eds.), *Shaping literacy achievement: Research we have, research we need* (pp. 98-128). New York, NY: Guilford.

Dewitz, P., Jones, J., & Leahy, S. (2010). Comprehension strategy instruction in core reading programs. *Reading Research Quarterly, 44*, 102-126. doi:10.1598/RRQ.44.2.1

Dimitrov, D. M. (2006). Comparing groups on latent variables: A structural equation modeling approach. *Work: A Journal of Prevention, Assessment, & Rehabilitation*, *26*, 429-436.

Dimitrov, D. M. (2009). *Quantitative research in education: Intermediate & advanced methods.* Oceanside, NY: Whittier.

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct of validation. *Measurement and Evaluation in Counseling and Development, 43,* 121-149. doi:10.1177/0748175610373459

Dimitrov, D. M. (2012). *Statistical methods for validation of assessment scale data in counseling and related fields.* Alexandria, VA: American Counseling Association.

Duke, N. K., & Pearson, P. D. (2008). Effective practices for developing reading comprehension. *The Journal of Education, 189*(1/2), 107-122.

Duke, N. K., Pearson, P. D., Strachan, S. L., & Billman, A. K. (2011). Essential elements of fostering and teaching reading comprehension. In S. J. Samuels & A. E. Farstrup (Eds.), *What Research Has to Say About Reading Instruction* (4th ed., pp. 51-93). Newark, DE: International Reading Association.

Duncan, L. G., McGeown, S. P., Griffiths, Y. G., Stothard, S. E., & Dobai, A. (2016). Adolescent reading skill and engagement with digital and traditional literacies as predictors of reading comprehension. *British Journal of Psychology, 107*, 209-238. doi:10.1111/bjop.12134

Durkin, D. (1978). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly, 14*, 481-533. doi:10.1598.rrq.14.4.2

Edmonds, M. S., Vaughn, S. R., Wexler, J., Reutebuch, C., Cable, A., Klingler Tackett, K., Wick Schnakenberg, J. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research, 79*, 262-300. doi:10.3102/0034654308325998

Elementary and Secondary Education Act of 1965, Title I, Part A, 20 U.S.C.A. § 6301 *et seq.*

Falk-Ross, F. C. (2002). Toward the new literacy: Changes in college students' reading comprehension strategies following reading/writing projects. *Journal of Adolescent & Adult Literacy, 45*, 278-288.

Fan, F. A. (2012). Class size: Effects on students' academic achievements and some remedial measures. *Research in Education, 87*, 95-98. doi:10.7227/RIE.87.1.7

Finn, J. D. (2002). Small classes in American schools: Research, practice, and politics. *Phi Delta Kappan, 83*, 551-560. doi:10.1177/003172170208300717

Finn, J. D. & Achilles, C. M. (1999). Tennessee's class size study: Findings, implications, misconceptions. *Educational Evaluation and Policy Analysis, 21*, 97-109. doi:10.3102/01623737021002097

Fleischman, H. L., Hopstock, P. J., Pelczar, M. P., & Shelley, B. E. (2010). Highlights from PISA 2009: Performance of U.S. 15-year-old students in reading, mathematics, and science literacy in an international context (Report No. NCES 2011-004). Retrieved from United States Department of Education, Institute of Education Sciences, National Center for Education Statistics website: https://nces.ed.gov/pubs2011/2011004.pdf

Flippo, R. F. (2014). *Assessing readers: Qualitative diagnosis and instruction* (2nd ed.). New York, NY: Routledge.

Flippo, R. F., Holland, D-A. D., McCarthy, M. T., & Swinning, E. A. (2009). Asking the right questions: How to select an informal reading inventory. *The Reading Teacher, 63,* 79-83. doi:10.1598/RT.63.1.8

Flowers, L. A. (2007). Recommendations for research to improve reading achievement for African American students. *Reading Research Quarterly, 42*, 424-428. doi:10.1598/RRQ.42.3.8

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N.,…O'Connor, R. (2002). Exploring the importance of reading programs for kindergarteners with disabilities in mainstream classrooms. *Exceptional Children, 68*, 295-311. doi:10.1177/001440290206800301

Fuchs, D., Fuchs, L. S., Thompson, A., Svenson, E., Yen, L., Al Otaiba, S.,…Saenz, L. (2001). Peer-assisted learning strategies in reading: Extensions for kindergarten, first grade, and high school. *Remedial and Special Education, 22*, 15-21. doi:10.1177/074193250102200103

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Monitoring reading growth using student recalls: Effects of two teacher feedback systems. *The Journal of Educational Research, 83*, 103-110. doi:10.1080/00220671.1989.10885938

Fuchs, L. S., Fuchs, D., & Kazdan, S. (1999). Effects of peer-assisted learning strategies on high school students with serious reading problems. *Remedial and Special Education, 20*, 309-318. doi:10.1177/074193259902000507

Gajria, M., Jitendra, A. K., Sood, S., & Sacks, G. (2007). Improving comprehension of expository text in students with LD: A research synthesis. *Journal of Learning Disabilities, 40*, 210-225. doi:10.1177/00222194070400030301

Gamoran, A., & Long, D.A. (2007). Equality of educational opportunity: A 40-year retrospective. In R. Teese, S. Lamb, M. Duru-Bellat, S. Helme (Eds.), *International studies in educational inequality: Theory and policy* (Vol. 1, pp. 23-47). New York, NY: Springer.

Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research, 71*, 279-321. doi:10.3102/00346543071002279

Glass, G. V., & Smith, M. L. (1979). Meta-analysis of research on class size and achievement. *Educational Evaluation and Policy Analysis, 1*(1), 2-16. doi:10.2307/1164099

Gökhan, B. (2012). Reading attitudes of high school students: An analysis from different variables. *International Journal on New Trends in Education and Their Implications, 3*(2), 47-58.

Gömleksiz, M. N. (2004). Validity and reliability of an attitude scale towards reading habit. *Fırat University Journal of Social Sciences, 14*, 185-195.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7,* 6-10. doi:10.1177/074193258600700104

Graesser, A. C. (2007). An introduction to strategic reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 3-26). New York, NY: Erlbaum.

Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement, 66*, 930-944. doi:10.1177/0013164406288165

Graue, E., Hatch, K., Rao, K, & Oen, D. (2007). The wisdom of class-size reduction. *American Educational Research Journal, 44,* 670-700. doi:10.3102/ 0002831207306755

Griffiths, G., Sohlberg, M. M., & Biancarosa, G. (2011). *A review of models of reading comprehension with implications for adults with mTBI and the Campus Reader.* Retrieved from Campus Reader website: http://www.campusreader.org/dev/ resources/pdfs/CampusReadertheory_final0401.pdf

Griva, E., Alevriadou, A., & Semoglou, K. (2012). Reading preferences and strategies employed by primary school students: Gender, socio-cognitive and citizenship issues. *International Education Studies, 5*(2), 24-34. doi:10.5539/ies.v5n2p24

Gurian, M. (with Stevens, K.). (2010). *Boys and girls learn differently! A guide for teachers and parents* (2nd ed.). San Francisco, CA: Jossey-Bass.

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., . . . & Tonks, S. (2004). Increasing reading comprehension and engagement through concept oriented reading instruction. *Journal of Educational Psychology, 96,* 403-423. doi:10.1037/0022-0663.96.3.403

Hagaman, J. L., & Reid, R. (2008). The effects of the paraphrasing strategy on the reading comprehension of middle school students at risk for failure in reading. *Remedial and Special Education, 29*, 222-234. doi:10.1177/ 0741932507311638

Hahs-Vaughn, D. L. (2005). A primer for using and understanding weights with national datasets. *The Journal of Experimental Education, 73*, 221-248. doi:10.3200/ JEXE. 73.3.221-248

Hamilton, E. W. (2009). *Raising bookworms: Getting kids reading for pleasure and empowerment.* Sag Harbor, NY: Beech Tree Books.

Harfitt, G. J., & Tsui, A. B. M. (2015). An examination of class size reduction on teaching and learning processes: a theoretical perspective. *British Educational Research Journal, 41,* 845-865. doi:10.1002/berj.3165

Harris, T. L., & Hodges, R. E. (Eds.). (1995). *The literacy dictionary: The vocabulary of reading and writing*. Newark, DE: International Reading Association.

Hasselbring, T. S., & Goin, L. I. (2004). Literacy instruction for older struggling readers: What is the role of technology? *Reading & Writing Quarterly, 20*, 123-144. doi:10.1080/10573560490262073

Hitchcock, C. H., Prater, M. A., & Dowrick, P. W. (2004). Reading comprehension and fluency: Examining the effects of tutoring and video self-modeling on first-grade students with reading difficulties. *Learning Disability Quarterly, 27*, 89-103. doi:10.2307/1593644

Hong-Nam, K., Leavell, A. G., Maher, S. (2014). The relationships among reported strategy use, metacognitive awareness, and reading achievement of high school students. *Reading Psychology, 35*, 762-790. doi:10.1080/02702711.2013. 807900

Hopstock, P., & Pelczar, M. (2011). *Technical report and user's guide for the Program for International Student Assessment (PISA): 2009 data files and database with U.S. specific variables* (Report No. NCES 2011-025). Retrieved from United States Department of Education, Institute of Education Sciences, National Center for Education Statistics website: https://nces.ed.gov/surveys/pisa/pdf/2011025.pdf

Hougen, M. (2015). *Evidenced-based reading instruction for adolescents, grades 6-12.* (CEEDAR Document No. IC-13). Retrieved from University of Florida, Collaboration for Effective Educator, Development, Accountability, and Reform Center website: http://ceedar.education.ufl.edu/wp-content/uploads/2015/05/IC-13_FINAL_05-26-15.pdf

Houtveen, A. A. M., & van de Grift, W. J. C. M. (2006). Effects of metacognitive strategy instruction and instruction time on reading comprehension. *School Effectiveness and School Improvement, 18*, 173-190. doi:10.1080/09243450601058717

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *Quarterly Journal of Economics, 115*, 1239-1285. doi:10.1162/003355300555060

Hoxby, C. M., Murarka, S., & Kang, J. (2009). *How New York City's charter schools affect achievement*. Cambridge, MA: New York City Charter Schools Evaluation Project.

Hu, L. T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6,* 1-55. doi:10.1080/10705519909540118

Huang, H. (2015). Can students themselves narrow the socioeconomic-status-based achievement gap through their own persistence and learning time? *Education Policy Analysis Archives, 23*, 1-36. doi:10.14507/epaa.v23.1977

Hudley, C., Graham, S., & Taylor, A. (2007). Reducing aggressive behavior and increasing motivation in school: The evolution of an intervention to strengthen school adjustment. *Educational Psychologist, 42*, 251-260. doi:10.1080/ 00461520701621095

Huebener, M., Kuger, S., & Marcus, J. (2016). *Increased instruction hours and the widening gap in student performance* (DIW Berlin Discussion Paper No. 1561). Retrieved from DIW Berlin, German Institute for Economic Research website: https://papers.ssrn.com/sol3/paperscfm?abstract _id=2753493

Hulme, C., & Snowling, M. J. (2011). Children's reading comprehension difficulties: Nature, causes, and treatments. *Current Directions in Psychological Science, 20*, 139-142. doi:10.1177/0963721411408673

Hurst, B., Franklin, K., & Scales, K. B. (2010). How does a high school reading class influence students' reading levels, habits, and perceptions? *Journal of Reading Education, 35*(3), 10-18.

Hyde, J. S., & Durik, A. M. (2005). Gender, competence, and motivation. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 375-391). New York, NY: Guilford Press.

International Reading Association. (2002). *What is evidenced-based reading instruction? A position statement of the International Reading Association*. Retrieved from https://www.literacyworldwide.org/docs/default-source/where-we-stand/evidence-based-position-statement.pdf?sfvrsn=6

Invernizzi, M. A., Landrum, T. J., Howell, J. L., & Warley, H. P. (2005). Toward the peaceful coexistence of test developers, policymakers, and teachers in an era of accountability. *The Reading Teacher, 58*, 610-618. doi:10.1598/RT.58.7.2

Irvin, J. L., Meltzer, J., & Dukes, M. (2007). *Taking action on adolescent literacy: An implementation guide for school leaders*. Alexandria, VA: Association for Supervision and Curriculum Development.

Jacobs, V. A. (2008). Adolescent literacy: Putting the crisis in context. *Harvard Educational Review, 78*, 7-39. doi:10.17763/haer.78.1.c577751kq7803857

Jerrim, J., & Micklewright, J. (2014). Socioeconomic gradients in children's cognitive skills: Are cross-country comparisons robust to who reports family background? (IZA Discussion Paper No. 8392). Retrieved from The Institute for the Study of Labor (IZA) website: http://ftp.iza.org/dp8392.pdf

Jeynes, W. H. (2002). The challenge of controlling for SES in social science and education research. *Educational Psychology Review, 14*, 205-221. doi:10. 1023/A:101467882

Jhang, F. H. (2014). The influences of inductive instruction and resources on students' attitudes toward reading: Evidence from PISA 2009. *Educational Research and Evaluation, 20*, 386-410. doi:10.1080/13803611.2014.966114

Johnson, K. A. (2000). *Do small classes influence academic achievement? What the National Assessment of Educational Progress shows: A report of the Heritage Center for Data Analysis* (Report No. CDA00-07). Retrieved from The Heritage Foundation website: http://www.heritage.org/education/report/do-small-classes-influence-academic-achievement-what-the-national-assessment

Johnson, K. F., Gupta, A., Rosen, H., & Rosen, H. (2013). Improving reading comprehension through holistic intervening and tutoring during after-school with high risk minority elementary school students. *Mentoring & Tutoring: Partnership in Learning, 21*, 431-443. doi:10.1080/13611267.2013.855861

Kaestle, C. F., Damon-Moore, H., Stedman, L. C., Tinsley, K., & Trollinger, W. V., Jr. (1991). *Literacy in the United States: Readers and reading since 1880*. Binghamton, NY: Vail-Ballou.

Kamil, M. L. (2008a). *Elements of successful reading instruction*. Retrieved from the Center on Innovation & Improvement website: http://www.centerii.org/techassist/solutionfinding/resources/SuccReadInstruct.pdf

Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008b). *Improving adolescent literacy: Effective classroom and intervention practices: A Practice Guide* (Report No. NCEE 2008-4027). Retrieved from United States Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance website: https://ies.ed.gov/ncee/wwc/Docs/ PracticeGuide/adlit_pg_082608.pdf

Kasapoglu, K. (2014). A logistic regression analysis of Turkey's 15-year-old's scoring above the OECD average on the PISA '09 reading assessment. *Educational Sciences: Theory & Practice, 14,* 649-667. doi:10.12738/estp.2014.2.1744

Kazelskis, R., Thames, D., Reeves, C., Flynn, R., Taylor, L., Beard, L. A., & Turnbo, D. (2005). Reliability and stability of Elementary Reading Attitude Survey (ERAS) scores across gender, race, and grade level. *The Professional Educator, 27*(1/2), 29-37.

Keene, E. (2006). *Assessing comprehension thinking strategies.* Huntington Beach, CA: Shell.

Khezrlou, S. (2012). The relationship between cognitive and metacognitive strategies, age, and level of education. *The Reading Matrix, 12,* 50-61.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95,* 163-182. doi:10.1037/0033-295x.95.2.163

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

Kintsch, W. (2005). An overview of top-down and bottom-up effects in comprehension: The CI perspective. *Discourse Processes, 39,* 125-128. doi:10.1207/s15326950dp3902&3_2

Kirk, C., & Vigeland, L. (2014). A psychometric review of norm-referenced tests used to assess phonological error patterns. *Language, Speech, and Hearing Services in Schools, 45*, 365-377. doi:10.1044/2014_LSHSS-13-0053

Kirmizi, F. S. (2011). The relationship between reading comprehension strategies and reading attitudes. *Education 3-13, 39*, 289-303. doi:10.1080/ 03004270903514320

Kissau, S., & Hiller, F. (2013). Reading comprehension strategies: An international comparison of teacher preferences. *Research in Comparative and International Education, 8,* 437-454. doi:10.2304/rcie.2013.8.4.437

Kitsantas, A., Ware, H. W., & Cheema, J. R. (2010). Predicting mathematics achievement from mathematics efficacy: Does analytical method make a difference? *The International Journal of Educational and Psychological Assessment, 5,* 25-44.

Klein, A. F. (2008). *Providing differentiated reading instruction to meet the individual needs of students*. Retrieved from http://www.readingaz.com/updates/reading_az_white_paper.pdf

Klingner, J. K., & Vaughn, S. R. (1996). Reciprocal teaching of reading comprehension strategies for students with learning disabilities who use English as a second language. *Elementary School Journal, 96,* 275-293. doi:10.1086/461828

Kolbe, T., Partridge, M., & O'Reilly, F. (2012). *Time and learning in schools: A national profile*. Boston, MA: National Center on Time & Learning.

Konstantopoulos, S., & Traynor, A. (2014). Class size effects on reading achievement using PIRLS data: Evidence from Greece. *Teachers College Record, 116*(2), 1-29.

Krassel, K. F., & Heinesen, E. (2014). Class-size effects in secondary school. *Education Economics, 22*, 412-426. doi:10.1080/09645292.2014.902428

Kush, J. C., & Watkins, M. W. (1996). Long-term stability of children's attitudes toward reading. *The Journal of Educational Research, 89,* 315-319. doi:10.1080/00220671.1996.9941333

Kush, J. C., Watkins, M. W., & Brookhart, S. M. (2005). The temporal-interactive influence of reading achievement and reading attitude. *Educational Research and Evaluation, 11*, 29-44. doi:10.1080/13803610500110141

Ladd, H. (2012). Presidential address: Education and poverty: Confronting the evidence. *Journal of Policy Analysis and Management, 31*, 203-227. doi:10.1002/pam.21615

Lang, L., Torgesen, J., Vogel, W., Chanter, C., Lefsky, E., Petscher, Y. (2009). Exploring the relative effectiveness of reading interventions for high school students. *Journal of Research on Educational Effectiveness, 2*, 149-175. doi:10.1080/19345740802641535

Lavy, V. (2015). Do differences in school's instruction time explain international achievement gaps? Evidence from developed and developing countries. *The Economic Journal, 125,* F397-F424. doi:10.1111/ecoj.12233

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher, 31*(1), 3-12. doi:10.3102/0013189x031001003

Leuven, E., Oosterbeek, H., & Rønning, M. (2008). Quasi-experimental estimates of the effect of class size on achievement in Norway. *The Scandinavian Journal of Economics, 110,* 663-693. doi:10.1111/J.1467-9442.2008.00556.X

Liang, L. A., & Dole, J. A. (2006). Help with teaching reading comprehension: Comprehension instructional frameworks. *The Reading Teacher*, *59*, 742-753. doi:10.1598/RT.59.8.2

Lietz, P. (2006). A meta-analysis of gender differences in reading achievement at the secondary school level. *Studies in Educational Evaluation, 32*, 317-344. doi:10.1016/j.stueduc.2006.10.002

Lim, H. J., Bong, M., & Woo, Y.-K. (2015). Reading attitude as a mediator between contextual factors and reading behavior. *Teachers College Record, 117*(1), 1-36.

Logan, S. & Johnston, R. (2009). Gender differences in reading ability and attitudes: Examining where these differences lie. *Journal of Research in Reading*, *32*, 199-214. doi:10.1111/j.1467-9817.2008.01389.x

Logan, S. & Johnston, R. (2010). Investigating gender differences in reading. *Educational Review, 62*, 175-187. doi:10.1080/00131911003637006

Luke, A., Woods, A., & Dooley, K. (2011). Comprehension as social and intellectual practice: Rebuilding curriculum in low socioeconomic and cultural minority schools. *Theory Into Practice, 50*, 157-164. doi:10.1080/00405841.2011. 558445

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201-226. doi:10. 1146/annurev.psych.51.1.201

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2001). *Gates-MacGinitie Reading Tests, Level 7/9*, *Form S* (4th ed.). Itasca, IL: Riverside.

Maerten-Rivera, J., Myers, N., Lee, O., & Penfield, R. (2010). Students and school predictors of high-stakes assessment in science. *Science Education, 94*, 937-962. doi:10.1002/sce.20408

Manset-Williamson, G., & Nelson, J. (2005). Balanced, strategic reading instruction for upper elementary and middle school student with reading disabilities: A comparative study of two approaches. *Learning Disability Quarterly, 28*, 59-74. doi:10.2307/4126973

Martínez, R. S., Aricak, O. T., & Jewell, J. (2008). Influence of reading attitude on reading achievement: A test of the temporal-interaction model. *Psychology in the Schools, 45*, 1010-1022. doi:10.1002/pits.20348

Mastropieri, M. A., & Scruggs, T. E. (1997). Best practices in promoting reading comprehension in students with learning disabilities 1976 to 1996. *Remedial and Special Education, 18,* 197-213. doi:10.1177/074193259701800402

Mathes, P. G., Grek, M. L., Howard, J. K., Babyak, A. E., & Allen, S. H. (1999). Peer-assisted learning strategies for first-grade readers: A tool for preventing early reading failure. *Learning Disabilities Research & Practice, 14,* 50-60. doi:10.1207/sldrp1401_5

Mathes, P. G., Howard, J. K., Allen, S. H., & Fuchs, D. (1998). Peer-assisted learning strategies for first-grade readers: Responding to the needs of diverse learners. *Reading Research Quarterly, 3*3, 62-94. doi:10.1598/RRQ.33.1.4

Mathewson, G. C. (1994). Model of attitude influence upon reading and learning to read. In R. B. Ruddell & M. E. Ruddell (Eds.), *Theoretical models and processes of reading* (4th ed.). Newark, NJ: International Reading Association.

McGeown, S. P., Johnston, R. S., Walker, J., Howatson, K., Stockburn, A., & Dufton, P. (2015). The relationship between young children's enjoyment of learning to read, reading attitudes, confidence and attainment. *Educational Research, 57*, 389-402. doi:10.1080/00131881.2015.1091234

McKenna, M. C., Conradi, K., Lawrence, C., Jang, B. G., & Meyer, J. P. (2012). Reading attitudes of middle school students: Results of a U.S. survey. *Reading Research Quarterly, 47,* 283-306. doi:10.1002/RRQ.021

McKenna, M. C., & Kear, D. J. (1990). Measuring attitude toward reading: A new tool for teachers. *The Reading Teacher, 43,* 626-639. doi:10.1598/RT.43.8.3

McKenna, M. C., Kear, D. J., & Ellsworth, R. A. (1995). Children's attitudes toward reading: A national survey. *Reading Research Quarterly, 30,* 934-956. doi:10.2307/748205

McMaster, K. L., Kung, S.-H, Han, I., & Cao, M. (2008). Peer-assisted learning strategies: A "Tier 1" approach to promoting English learners' response to intervention. *Council for Exceptional Children, 74,* 194-214. doi:10.1177/001440290807400204

McNamara, D. S., Ozuru, Y., Best, R., & O'Reilly, T. (2007). The 4-Prong Comprehension Strategy Framework. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 465-491). New York, NY: Erlbaum.

Meyer, B. J. F., & Wijekumar, K. (2007). A web-based tutoring system for the structure strategy: Theoretical background, design, and findings. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 347-374). New York, NY: Erlbaum.

Meyler, A., Keller, T. A., Cherkassky, V. L., Gabrieli, J. D. E., & Just, M. A. (2008). Modifying the brain activation of poor readers during sentence comprehension with extended remedial instruction: A longitudinal study of neuroplasticity. *Neuropsychologia*, *46*, 2580-2592. doi:10.1016/j.neuropsychologia.2008.03.012

Milesi, C., & Gamoran, A. (2006). Effects of class size and instruction on kindergarten achievement. *Educational Evaluation and Policy Analysis, 28*, 287-313. doi:10.3102/01623737028004287

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research, 3*(1), 111-121. doi:10.21500/20112084.857

Moats, L. C. (1999). *Teaching reading is rocket science: What expert teachers of reading should know and be able to do.* Washington, DC: American Federation of Teachers.

Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology, 94*, 249-259. doi:10. 1037//0022-0663.94.2.249

Molnar, A., Smith, P., Zahorik, J., Palmer, A., Halbach, A., & Ehrle, K. (1999). Evaluating the SAGE Program: A pilot program in targeted pupil-teacher reduction in Wisconsin. *Educational Evaluation and Policy Analysis, 21*, 165-177. doi:10.3102/01623737021002165

Morgan, P. L., Young, C., & Fuchs, D. (2005). Peer-assisted learning strategies: An effective intervention for young readers. *Insights on Learning Disabilities, 3*(1), 23-42.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Kennedy, A. M. (2003). *PIRLS 2001 international report: IEA's study of reading literacy achievement in primary school in 35 countries.* Retrieved from International Association for the Evaluation of Educational Achievement, TIMMS & PIRLS International Study Center website: https://timss.bc.edu/pirls2001i/pdf/p1_IR_book.pdf

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006 international report: IEA's progress in international reading literacy study in primary schools in 40 countries.* Retrieved from International Association for the Evaluation of Educational Achievement, TIMSS & PIRLS International Study Center website: https://pirls.bc.edu/PDF/p06_international_report.pdf

Muthén, L. K., & Muthén, B. O. (2010). *Mplus statistical analysis with latent variables: Users guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Myrberg, E., & Rosén, M. (2009). Direct and indirect effects of parents' education on reading achievement among third graders in Sweden. *British Journal of Educational Psychology, 79*, 695-711. doi:10.1348/000709909x453031

Nation, K. (2001). Exposing hidden deficits (Spearman Medal Lecture). *The Psychologist, 14*, 238–242.

National Center for Education Statistics. (2011a). *Program for International Student Assessment (PISA) 2009: U.S. public-use data files and electronic codebook*. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2011038

National Center for Education Statistics. (2011b). *Quick guide to using the Program for International Student Assessment (PISA) 2009 electronic codebook (ECB) for U.S. data*. Retrieved from https://nces.ed.gov/surveys/pisa/pdf/PISA2009CB_quickguide.pdf

National Center for Education Statistics. (2011c). *The nation's report card: Reading 2011* (Report No. NCES 2012-457). Retrieved from https://nces.ed.gov/nationsreportcard/pubs/main2011/2012457.asp

National Center for Education Statistics. (2013). *The nation's report card: Trends in academic progress 2012* (Report No. NCES 2013-456). Retrieved from https://nces.ed.gov/nationsreportcard/subject/publications/main2012/pdf/2013456.pdf

National Center for Education Statistics. (2014). *2003-04, 2007-08, 2011-12 National Postsecondary Student Aid Study: Percentage of first-year undergraduate students who reported taking remedial education courses, by selected student and institution characteristics: 2003-04, 2007-08, and 2011–12* (Table 311.40). Retrieved from https://nces.ed.gov/programs/digest/d15/tables/dt15_311.40.asp

National Center for Education Statistics. (2017). *National Assessment of Education Progress (NAEP): NAEP data explorer.* Retrieved from https://nces.ed.gov/nationsreportcard/ naepdata/dataset.aspx

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core State Standards for Reading.* Retrieved from http://www.corestandards.org/Math/Content/K/introduction

National Institutes of Health. (1999). *Drug use and related adverse behavioral and social consequences*. Retrieved from http://grants.nih.gov/grants/guide/pa-files/PA-99-113.html

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.* Washington, DC: National Institute of Child Health and Human Development.

Negley, H. (1975). *Survey of social studies teaching* [monograph]. Indianapolis, IN: Indiana State Department of Instruction.

Ness, M. K. (2009). Reading comprehension strategy instruction in secondary content area classrooms: Teachers' use of and attitudes toward reading comprehension instruction. *Reading Horizons*, *49*, 143-166.

No Child Left Behind Act of 2001, 20 U.S.C.A. § 6301 *et seq.* (West, 2003).

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Oakley, G. (2011). The assessment of reading comprehension cognitive strategies: Practices and perceptions of Western Australian teachers. *Australian Journal of Language and Literacy, 34*, 279-292.

Organization for Economic Cooperation and Development. (2001). *Knowledge and skills for life: First results from the OECD Programme for International Student Assessment (PISA)2000*. Paris, France: Author. doi:10.1787/9789264195905-en

Organization for Economic Cooperation and Development. (2004). *Learning for tomorrow's world: First results from the OECD Programme for International Student Assessment (PISA) 2003*. Paris, France: Author. doi:10.1787/ 9789264006416-en

Organization for Economic Cooperation and Development. (2009). *PISA data analysis manual: SPSS* (2nd ed.). Paris, France: Author. doi:10.1787/9789264056275-en

Organization for Economic Cooperation and Development. (2010a). *PISA 2009 assessment framework: Key competencies in reading, mathematics and science*. Paris, France: Author. doi:10.1787/9789264062658-en

Organization for Economic Cooperation and Development. (2010b). *PISA 2009 results: Student engagement, strategies, and practices* (Vol. 3). Paris, France: Author. doi:10.1787/9789264083943-en

Organization for Economic Cooperation and Development. (2010c). *PISA 2009 results: Students online digital technologies & performance* (Vol. 6). Paris, France: Author. doi:10.1787/9789264112995-en

Organization for Economic Cooperation and Development. (2010d). *PISA 2009 results: What students know and can do. Student performance in reading, mathematics, and science* (Vol. 1). Paris, France: Author. doi:10.1787/9789264091450-en

Organization for Economic Cooperation and Development. (2011). *Girls' and boys' reading performance since 2000, in PISA 2009 at a Glance*. Paris, France: Author. doi:10.1787/9789264095298-en

Organization for Economic Cooperation and Development. (2012). *PISA 2009 technical report*. Paris, France: Author. doi:10.1787/9789264167872-en

Organization for Economic Cooperation and Development. (2015). *OCED Programme for International Student Assessment (PISA): Database PISA 2009*. Retrieved from http://www.oecd.org/pisa/pisaproducts/pisa2009database-downloadabledata.htm

Oxford, R. L., & Burry-Stock, J. A. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the Strategy Inventory for Language Learning (SILL). *System, 23*, 1-23. doi:10.1016/0346-251X(94) 00047-A

Özdemir, B., & Gelbal, S. (2014). Investigating factors that affect Turkish students' academic success with canonical commonality analysis according to PISA 2009 results. *Education and Science, 39*, 41-57. doi:10.15390/EB.2014.3025

Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension fostering and comprehension monitoring activities. *Cognition and Instruction, 1*, 117-175. doi:10.1207/s1532690xci0102_1

Palinscar, A. S., & Schultz, K. M. (2011). Reconnecting strategy instruction with its theoretical roots. *Theory into Practice, 50*, 85-92. doi:10.1080/00405841. 2011.558432

Paris, S., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8*, 293-316. doi:10.1016/0361-476x (83) 90018-8

Pečjak, S., Podlesek, A., & Pirc, T. (2011). Model of reading comprehension for 5th grade students. *Studia Psychologica,* 53, 53-67.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227-247). Malden, MA: Blackwell.

Perry, A., Wiederhold, S., & Ackerman-Piek, D. (2014). How can skill mismatch be measured? New approaches with PIAAC. *Methods, Data, Analyses, 8*(2), 137-174. doi:10.12758/mda.2014.006

Petscher, Y. (2010). A meta-analysis of the relationship between student attitudes towards reading and achievement in reading. *Journal of Research in Reading, 33*, 335-355. doi:10.1111/j.1467-9817.2009.01418.x

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research, 31,* 459-470. doi:10. 1016/S0883-0355(99)00015-4

Prado, L., & Plourde, L. A. (2011). Increasing reading comprehension through the explicit teaching of reading strategies: Is there a difference among the genders? *Reading Improvement, 48*, 32-43.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale NJ: Erlbaum.

Pressley, M., & Harris, K. R. (2008). Cognitive strategies instruction: From basic research to classroom instruction. *The Journal of Education, 189*, 77-97.

Pressley, M., Wharton-MacDonald, R., Mistretta-Hampton, J., & Echevarria, M. (1998). Literacy instruction in 10 fourth-grade classrooms in upstate New York. *Scientific Studies of Reading, 2,* 159-194. doi:10.1207/s1532799xssr0202_4

Quirk, T. J., Trismen, D. A., Nalin, K. B., & Weinberg, S. F. (1975). The classroom behavior of teachers during compensatory reading instruction. *The Journal of Educational Research, 68*, 185-192. doi:10.1080/00220671.1975.10884742

Randi, J., Newman, T., & Grigorenko, E. (2010). Teaching children with autism to read for meaning: Challenges and possibilities. *Journal of Autism and Developmental Disorders, 40*, 890-902. doi:10.1007/s10803-010-0938-6

Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling* (2nd ed.). Mahwah, NJ: Erlbaum.

Raykov, T., Tomer, A., & Nesselroade, J. R. (1991). Reporting structural equation modeling results in *Psychology and Aging*: Some proposed guidelines. *Psychology and Aging, 6*, 499-503. doi:10.1037/0882-7974.6.4.499

Reardon, S. F., Valentino, R. A., & Shores, K. A. (2012). Patterns of literacy among U.S. students. *The Future of Children, 22*, 17-37. doi:10.1353/foc.2012.0015

Reis, S. M., McCoach, D. B., Coyne, M., Schreiber, F. J., Eckert, R. D., & Gubbins, E. J. (2007). Using planned enrichment strategies with direct instruction to improve reading fluency, comprehension, and attitude toward reading: An evidence-based study. *The Elementary School Journal, 108*, 3-23. doi:10.1086/522383

Rice, J. K. (1999). The impact of class size on instructional strategies and the use of time in high school mathematics and science courses. *Educational Evaluation and Policy Analysis, 21*, 215-229. doi:10.2307/1164301

Rice, J. M. (1902). Educational research: A test in arithmetic. *The Forum, 34*, 281-297.

Rivkin, S. G., & Schiman, J. C. (2015). Instruction time, classroom quality, and academic achievement. *The Economic Journal*, *125*, F425-F448. doi:10.1111/ecoj.12315

Roberts, G., Torgesen, J. K., Boardman, A., & Scammacca, N. (2008). Evidenced-based strategies for reading instruction of older students with learning disabilities. *Learning Disabilities Research & Practice*, *23*, 63-69. doi:10.1111/j.1540-5826. 2008.00264.x

Rogoff, B. (2003). *The Cultural Nature of Human Development: Thinking with the tools and institutions of culture.* New York, NY: Oxford.

Ruddell, R. B., & Speaker, R. B., Jr. (1985). The interactive reading process: A model. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (3rd ed., pp. 751-793). Hillsdale, NJ: Erlbaum.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55,* 68-78. doi:10.1037//0003-066X.55.1.68

Şahbaz, N. K. (2012). Evaluation of reading attitudes of 8th grade students in primary education according to various variables. *Educational Research and Reviews, 7*, 571-576. doi:10.5897/ERR12.101

Sallabaş, M. E. (2008). Relationship between 8th grade secondary school students' reading attitudes and reading comprehension skills. *Journal of the Faculty of Education, 9*, 141-155.

Sari, A. A. (2015). Using structural equation modeling to investigation students' reading comprehension skills. *Elementary Education Online, 14*, 511-521. doi:10. 17051/io.2015.32986

Savolainen, H., Ahonen, T., Aro, M., Tolvanen, A., & Holopainen, L. (2008). Reading comprehension, word reading, and spelling as predictors of school achievement and choice of secondary education. *Learning and Instruction, 18*(2), 201-210. doi:10.1016/j.learninstruc.2007.09.017

Scammacca, N. K., Roberts, G. J., Cho, E., Williams, K. J., Roberts, G., Vaughn, S. R., & Carroll, M. (2016). A century of progress: Reading interventions for students in grades 4-12, 1914-2014. *Review of Educational Research, 86*, 756-800. doi:10. 3102/0034654316652942

Scammacca, N. K., Roberts, G. J., Vaughn, S. R., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. (2007). *Reading interventions for adolescent struggling readers: A meta-analysis with implications for practice.* Portsmouth, NH: RMC Research Corporation, Center for Instruction.

Schenck, E. A., Walker, D. R., Nagel, C. R., & Webb, L. C. (2005). *Analysis of State K-3 Reading Standards and Assessments: Final Report*. Retrieved from U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service website: https://www2.ed.gov/ rschstat/eval/other/reading/state-k3-reading.pdf

Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). New York, NY: Routledge.

Schunk, D. H. (2011). *Learning theories: An educational perspective* (6th ed.) Boston, MA: Pearson.

Schunk, D. H., & Pajares, F. (2001). The development of academic self-efficacy. In A. Wigfield & J. Eccles (Eds.), *Development of achievement motivation* (1st ed., pp. 15-31). San Diego, CA: Academic Press.

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review, 78*, 40-51. doi:10.17763/haer.78.1.v62444321p602101

Shanahan, T., & Shanahan, C. (2012). What is disciplinary literacy and why does it matter? *Topics in Language Disorders*, *32*, 7-18. doi:10.1097/TLD. 0b013e318244557a

Shang, H.-F. (2010). Reading strategy use, self-efficacy, and EFL Reading Comprehension. *Asian EFL Journal Quarterly, 12*(2), 18-42.

Shaywitz, B. A., Shaywitz, S. E., Blachman, B., Pugh, K. R., Fulbright, R. K., Skudlarski, P.,…Gore, J. C. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biological Psychiatry, 55*, 926-933. doi:10.1016/j.biopsych.2003.12.019

Shaywitz, S. E., Shaywitz, B. A., Fulbright, R. K., Skudlarski, P., Mencl, W. E., Constable, R. T.,…Gore, J. C. (2003). Neural systems for compensation and persistence: Young adult outcome of childhood reading disability. *Biological Psychiatry, 54*, 25-33. doi:10.1016/s0006-3223(02)01836-x

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P.,…Gore, J. C. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry, 52*, 101-110. doi:10.1016/s0006-3223(02)01365-3

Sheorey, R., & Mokhtari, K. (2001). Differences in the metacognitive awareness of reading strategies among native and non-native readers. *System, 29*, 431-449. doi:10.1016/S0346-251X(01)00039-2

Shera, P. (2014). School effects, gender, and socioeconomic differences in reading performance: A multilevel analysis. *International Education Studies, 7*(11), 28-39. doi:10.5539/ies.v7n11p28

Shin, I.-S., & Chung, J. Y. (2009). Class size and student achievement in the United Stated: A meta-analysis. *KEDI Journal of Educational Policy, 6*, 3-19.

Shook, A. C., Hazelkorn, M., & Lozano, E. R. (2011). Science vocabulary for all: Strategies to improve vocabulary in an inclusive biology class. *The Science Teacher, 78*(3), 45-49.

Silva, S. M., Verhoeven, L., & van Leeuwe, J. (2011). Socio-cultural variation in reading comprehension development among fifth graders in Peru. *Reading and Writing, 24*, 951-969. doi:10.1007/s11145-010-9242-2

Simos, P. G., Fletcher, J. M., Bergman, E., Breier, J. I., Foorman, B. R., Castillo, E. M.,…Papanicolaou, A. C. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology, 58,* 1203-1213. doi:10.1212/WNL.58.8.1203

Singh, M. (2008). Factors contributing to reading literacy differences between male and females. *International Journal of Learning, 15*(3), 337-344. doi:10.18848/1447-9494/cgp/v15i03/45695

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*, 417-453. doi:10.3102/00346543075003417

Smith, C. B. (1976). Teaching reading in the secondary school. *Educational Leadership, 33*, 509-512.

Smith, M. C. (1990). A longitudinal investigation of reading attitude development from childhood to adulthood. *The Journal of Educational Research, 83*, 215-219. doi:10.1080/00220671.1990.10885958

Snow, C. E. (2002). *Reading for understanding: Toward a research and development program in reading comprehension.* Santa Monica, CA: Rand.

Snow, C. E., & Biancarosa, G. (2003). *Adolescent literacy and the achievement gap: What do we know and where do we go from here?* New York, NY: Carnegie Corporation.

Somers, M.-A., Corrin, W., Sepanik, S., Salinger T., Levin, J., & Zmach, C. (2010). *The enhanced reading opportunities study final report: The impact of supplemental literacy courses for struggling ninth-grade readers* (Report No. NCEE 2010-4021). Retrieved from United States Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance website: https://ies.ed.gov/ncee/pubs/ 20104021/pdf/20104021.pdf

Sonnenschein, S., Stapleton, L. M., & Benson, A. (2010). The relation between the type and amount of instruction and growth in children’s reading competencies. *American Educational Research Journal, 47,* 358-389. doi:10.3102/0002831209349215

Spörer, N. & Brunstein, J. C. (2009). Fostering the reading comprehension of secondary school students through peer-assisted learning: Effects on strategy knowledge, strategy use, and task performance. *Contemporary Educational Psychology, 34,* 289-297. doi:10.1016/j.cedpsych.2009.06.004

Sturtevant, E. G. (2003). *The literacy coach: A key to improving teaching and learning in secondary schools.* Washington, DC: Alliance for Excellent Education

Sulkunen, S. (2013). Adolescent literacy in Europe: An urgent call for action. *European Journal of Education, 48*, 528-542. doi:10.1111/ejed.12052

Sullivan, G. M., & Feinn, R. (2012). Using effect size – or why the p-value is not enough. *Journal of Graduate Medical Education, 4*, 279-282. doi:10.4300/JGME-D-12-00156.1

Swalander, L., & Taube, K. (2007). Influences of family based prerequisites, reading attitude, and self-regulation on reading ability. *Contemporary Educational Psychology*, *32*, 206-230. doi:10.1016/j.cedpsych.2006.01.002

Swanson, E., Wanzek, J., McCulley, L., Stillman-Spisak, S., Vaughn, S. R., Simmons, D.,…Hairell, A. (2016). Literacy and text reading in middle and high school social studies and English language arts classrooms. *Reading & Writing Quarterly, 32*, 199-222. doi:10.1080/10573569.2014.910718

Swanson, H. L. (1999). Reading research for students with LD: A meta-analysis of intervention questions. *Journal of Learning Disabilities, 32,* 504-532. doi:10.1177/002221949903200605

Takala, M. (2006). The effects of reciprocal teaching of reading comprehension in mainstream and special (SLI) education. *Scandinavian Journal of Education Research*, *50*, 559-576. doi:10.1080/00313830600953824

The Melissa Institute for Violence Prevention and Treatment. (2010). *2010 report to the community: Preventing violence through literacy.* Retrieved from https://www.melissainstitute.org/documents/TMI_AnnualReport_2010.pdf

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. *Remedial and Special Education*, *25*, 252-261. doi:10.1177/07419325040250040801

Tullock-Rhody, R., & Alexander, J. E. (1980). A scale for assessing attitudes for reading in secondary schools. *Journal of Reading*, *23*, 609-614.

Türkyilmaz, M. (2015). The relationship between reading attitude, metacognitive awareness of reading strategies, personality, and self-regulation: A study of modeling. *Education*, *136*, 11-18.

Türkyilmaz, M. & Aydemir, Y. (2014). The study of the validity and reliability of the scale for secondary education students' attitude towards reading. *International Journal of Language Academy, 2-3,* 194-207.

United States Census Bureau. (2017). *Current population survey, 2016 and 2017 Annual Social and Economic Supplements: People in poverty by selected characteristics: 2015 and 2016* (Table 3). Retrieved from https://www.census.gov/data/tables/2017/demo/income-poverty/p60-259.html

Valencia, S. W., Pearson, P. D., & Wixson, K. K. (2011). Assessing and tracking progress in reading comprehension: The search for keystone elements in college and career readiness. Paper presented at the Invitational Research Symposium on Through-Course Summative Assessments, Atlanta, GA. Retrieved from http://www.ets.org/Media/Research/pdf/TCSA_Symposium_Final_Paper_Valencia_Pearson_Wixson.pdf

Vandenberg, R. J. & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for the organizational research. *Organizational Research Methods, 3,* 4-70. doi:10.1177/109442810031002

Vaughn, S. R., Klingner, J. K., & Bryant, D. P. (2001). Collaborative strategic reading as a means to enhance peer-mediated instruction for reading comprehension and content-area learning. *Remedial and Special Education, 22*, 66-74. doi:10.1177/074193250102200201

Vaughn, S. R., Klingner, J. K., Swanson, E. A., Boardman, A. G., Roberts, G., Mohammed, S. S., & Stillman-Spisak, S. J. (2011). Efficacy of Collaborative Strategic Reading with middle school students. *American Educational Research Journal, 48*, 938-964. doi:10.3102/0002831211410305

Vaughn, S. R., Roberts, G., Schnakenberg, J. B., Fall, A.-M., Vaughn, M. G., & Wexler, J. (2015a). Improving reading comprehension for high school students with disabilities: Effects for comprehension and school retention. *Exceptional Children, 82*, 117-131. doi:10.1177/0014402915585478

Vaughn, S. R., Roberts, G., Wexler, J., Vaughn, M. G., Fall, A.-M., & Schnakenberg, J. B. (2015b). High school students with reading comprehension difficulties: Results of a randomized control trial of a two-year reading intervention. *Journal of Learning Disabilities, 48*, 546-588. doi:10.1177/0022219413515511

Vaughn, S. R., Swanson, E. A., Roberts, G., Wanzek, J., Stillman-Spisak, S. J., Solis, M., & Simmons, D. (2013). Improving reading comprehension and social studies knowledge in middle school. *Reading Research Quarterly, 48*. doi:10.1002/rrq. 039

Wang, D. (2009). Factors affecting the comprehension of global and local main idea. *Journal of College Reading and Learning, 39*(2), 34-52. doi:10.1080/10790195. 2009.10850317

Wanzek, J., Vaughn, S. R., Scammacca, N. K., Metz, K., Murray, C. S., Roberts, G., & Danielson, L. (2013). Extensive reading interventions for students with reading disabilities after Grade 3. *Review of Educational Research, 83*, 163-195. doi:10. 3102/0034654313477212

Wanzek, J., Wexler, J., Vaughn, S. R., & Ciullo, S. (2010). Reading instruction for struggling readers in the upper elementary grades: A synthesis of 20 years of research. *Reading and Writing, 23*, 889-912. doi:10.1007/s11145-009-9179-5

Watson, A., & Kehler, M. (2012). Beyond the "Boy Problem": Raising questions, growing concerns and literacy reconsidered. *The New England Reading Association Journal, 48*(1), 43-55.

Wexler, J., Mitchell, M. A., Clancy, E. E., & Silverman, R. D. (2016). An investigation of literacy practices in high school science classrooms. *Reading & Writing Quarterly, 33*, 258-277. doi:10.1080/10573569.2016.1193832

White, B. (1999). War preparations and ethnic and racial relations in the United States. In M. F. Boemeke, R. Chickering, & S. Förster (Eds.), *Anticipating total war: The German and American experiences, 1871-1914* (pp. 97-124). New York, NY: Cambridge University Press.

White, B. (2007). Are girls better readers than boys? Which boys? Which girls? *Canadian Journal of Education, 30*, 554-581. doi:10.2307/20466650

Williams, E. (2014). Breaking the barriers to reading success in middle and high schools. *Reading Improvement, 51*, 233-236.

Williams, J. P. (2007). Literacy in the curriculum: Integrating text structure and content area instruction. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 199-219). New York, NY: Erlbaum.

Woodcock, R. W. (1997). *Woodcock Diagnostic Reading Battery.* Itasca, IL: Riverside.

Woolley, G. (2011). *Reading Comprehension: Assisting children with learning difficulties.* New York, NY: Springer.

Wu, A. D., Li, Z., Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group Confirmatory Factor Analysis: A demonstration with TIMSS data. *Practical Assessment, Research & Evaluation, 12*, 1-26.

Biography

Jennifer A. Buxton graduated in 1995 from Greensburg Central Catholic High School, in Greensburg, Pennsylvania. She received her Bachelor of Arts in Psychology and her Masters of Science in Elementary Education from Duquesne University in 2000 and 2001, respectively. Before moving to Virginia in 2006, she worked as a Behavioral Specialist Consultant and a Program Manager for students with special needs. Once in Virginia, she taught fifth grade for six years in Loudoun County Public Schools. In 2009, she began her Ph.D. in Education Research Methodology at George Mason University. During her doctoral program, she worked as an adjunct professor and a student teaching supervisor at Marymount University in Arlington, Virginia. She also worked as an adjunct professor and a graduate research assistant at George Mason University in Fairfax and Arlington, Virginia.