Differences in the Reading Performance of Texas Grade 4 Girls and Boys: A Multiyear, Statewide Investigation

LaTracy Harris

John R. Slate

George W. Moore

Frederick C. Lunenburg

Sam Houston State University
Texas, United States

Abstract

In this investigation, the degree to which differences were present in reading between Grade 4 Texas boys and girls was examined. Data, obtained from the Texas Education Agency Public Education Information Management System for all Grade 4 boys and girls in Texas who took the State of Texas Assessment of Academic Readiness Reading exam, were analyzed for the 2012-2013, 2013-2014, and 2014-2015 school years. In all three years analyzed, statistically significant differences were established in not only overall reading performance, but also in all three Reading Reporting categories. A clear stair-step effect was present; girls outperformed boys in all three Reading Reporting categories as well as had higher percentages who met the Level II Final Satisfactory Standard. Future research recommendations and implications for policy and practice are suggested.

Keywords: Gender, STAAR Reading, Texas, Grade 4, Level II Final Satisfactory Standard and Literacy.

Gender differences in reading have been studied for centuries (Ayers, 1909). With reading being essential for academic success, it is apparent why researchers seek to understand the reading proficiency levels of girls and boys. Not only has the topic of gender differences in reading been examined in the United States, but it has been investigated in numerous other countries as well. This concern is one that spans across the globe due to similar gender differences in reading performance. In a recent international study, Mullis, Martin, Foy, and Hooper (2017), established that, in reading, girls had higher average scores than did boys in 48 of the 50 countries that participated in the 2016 Progress in International Reading Literacy Study. Furthermore, boys did not have higher reading achievement scores than girls in any of the 50 countries (Mullis et al., 2017). Strong reading comprehension and critical thinking skills are paramount in competing for jobs in the 21st century. All students, regardless of gender, must acquire such literacy skills early in their education (McGown, 2016).

In an analysis of gender differences from elementary through high school, Klecker (2006) examined Grade 4, 8, and 12 students' National Assessment of Educational Progress test scores across the 1992, 1994, 1998, 2000, 2002, and 2003 school years. Grade 4 girls outperformed boys in reading in all six years. Klecker (2006) reported, similar to Grade 4 results, that Grade 8 girls had higher reading scores than Grade 8 boys. Grade 12 results were congruent with Grade 4 and Grade 8 results in that Grade 12 girls had higher reading scores than Grade 12 boys.

In a similar study, Below, Skinner, Fearrington, and Sorrell (2010) investigated the degree to which gender differences were present in early literacy of kindergarten through Grade 5 students. Girls scored higher than boys on all four pre-literacy skills (Below et al., 2010). As such, findings were in agreement with previous researchers (e.g., Stinnett, 2011) that girls enter school with more advanced literacy skills than boys. Specifically established by Stinnett (2011) was that girls have stronger reading skill development entering kindergarten than do boys. From Kindergarten to Grade 5, statistically significant differences exist in reading performance between girls and boys that favored girls.

In another international study, the reading achievement of Grade 4 girls and boys across participating G-20 countries was examined (National Center for Education Statistics, 2015). All G-20 countries had higher percentages of Grade 4 girls outscore Grade 4 boys in reading, with differences ranging from 8 percentage points in France to 25 percentage points in Saudi Arabia (National Center for Education Statistics, 2015). In the United States, the difference between girls and boys was 13 percentage points. Girls outscored boys in reading at every grade level and at every age analyzed (National Center for Education Statistics, 2015).

In a recent investigation directly related to this article, McGown (2016) analyzed the degree to which gender differences were present in the reading performance of Texas Grade 3 students. She examined the State of Texas Assessment of Academic Readiness (STAAR) Reading assessment for three years. Regarding Reading Reporting Category 1, 2, and 3, Grade 3 girls had statistically significantly higher test scores than Grade 3 boys in all three of the school years examined. Concerning the Level II Final Satisfactory Performance Standard for girls and boys, results for all three school years were statistically significant. Grade 3 girls had statistically significantly higher percentages who met the Level II Final Satisfactory Performance Standard than Grade 3 boys in all three school years. In her investigation, Grade 3 girls had statistically significantly better reading performance in all of the STAAR Reading measures and in all three school years of data she analyzed.

In another recent study conducted in Texas, Schleeter (2017) examined the degree to which differences were present between Grade 3 English Language Learner boys and girls in their reading achievement. Participants in this study were Grade 3 English Language Learner boys and girls who took the State of Texas Assessment of Academic Readiness Reading assessments in the 2012-2013, 2013-2014, and 2014-2015 school year. A total of three years of STAAR Reading data were examined for English Language Learner girls and boys who were enrolled in Grade 3. In each year, English Language Learner girls outperformed English Language Learner boys (Schleeter, 2017). According to Schleeter (2017), the gender performance gap in the met standard category (Phase-in 1, Phase-in 2, and Final Satisfactory) was 5.9%. In the Level III Advanced Performance category, English Language Learner girls outperformed English Language Learner boys by an average of 2.7% (Schleeter, 2017). English Language Learner girls outperformed English Language Learner boys in every category measured. Of particular importance are reading disparities in high school. As noted by Wright and Slate (2015), Texas high school girls continue to outperform boys on state-mandated reading exams.

"For the last 100 years, researchers have expressed concern over a male deficit in reading achievement" (Stinnett, 2011, p. 72). Similarly, Klecker (2006) noted the positive relationship between being female and having higher reading scores than male students. As early as 1909, Ayers communicated a concern regarding a deficit with boys in reading. Some researchers (e.g., Northwestern University, 2008) have argued that girls have superior language abilities to boys. Concerning gender inequality, researchers (Catsambis et al., 2012; Condron, 2007; Tach & Farkas, 2006) have postulated the understanding that differences in reading performance between girls and boys exists; however, the underpinnings of that existence needs more exploring. What is certain is that for all students, regardless of gender, to compete in the 21st century; they must acquire proficient reading skills before Grade 3.

Statement of the Problem

Educators around the world aim to provide students with a solid reading foundation, regardless of gender. However, only a third of children in the United States read at grade level (Sanchez, 2018). Although a strong emphasis in the No Child Left Behind Act was on improving reading performance in early elementary, disaggregation of data by gender was not mandated in this legislative act. Therefore, gaps in the academic performance of boys and girls were and continue to not be monitored as closely as are other achievement gaps (Klecker, 2006). Only by analyzing data by gender will educational leaders become fully cognizant of the disparity in reading performance between girls and boys and implement strategies to close the achievement gap. However, as a result of this exclusion, the newly passed, Every Student Succeeds Act (2015) maintains the requirements for data disaggregation for accountability purposes but has included gender as a subgroup. As identified by Sadker and Zittleman (2005), girls receive better grades on their report card, perform higher than boys on standardized assessments, and are less likely to exhibit behavior challenges. Accordingly, it is imperative to analyze gender differences to inform educators on how to address these types of disparities. For this reason, the focus of this study was on Grade 4 girls and boys and the degree to which gender is related to reading performance on the state-mandated reading assessment in Texas.

Purpose of the Study

The purpose of this article was to examine the extent to which boys and girls differed in their reading achievement of Texas Grade 4 students. In this article, the degree to which gender differences existed in the reading performance of Texas Grade 4 students was addressed. For each of these studies, archival data from the Texas Education Agency Public Education Information Management System were analyzed. An analysis of academic performance for the 2012-2013, 2013-2014, and 2014-2015 school years on the state-mandated reading assessments for Texas Grade 4 students was conducted to determine the degree to which trends were present in the reading performance of boys and girls.

Significance of the Study

As outlined earlier, gender differences in reading achievement have been explored in great detail in the United States (Moore et al., 2012; Sadker & Zittleman 2005; Sanchez, 2018). Specifically, the underperformance of boys, both nationally and internationally, has raised much concern. Despite the numerous research findings regarding inequities in gender performance in reading, little has been implemented to make substantial improvement. Because the focus of the No Child Left Behind Act (2001) was not centered around monitoring student performance by gender, few researchers (e.g., McGown, 2016; Schleeter, 2017) have examined the relationship between gender and reading achievement as measured by the State of Texas Assessment of Academic Readiness. Unfortunately, limited progress has been made in closing the gender gap in reading.

Additionally, the reading performance of Grade 4 students by gender as measured by the STAAR assessment has not been previously examined. Therefore, this study is relevant and of utmost importance. In an era where the stakes are high for standardized testing and the findings from multiple studies (e.g., Below et al., 2010; Klecker, 2006; McGown, 2016) are that girls outperform boys in reading both nationally and internationally, it is imperative that educators identify methods for ensuring the success of all students by closing the achievement gaps. Disparities have been documented for centuries (Ayers, 1909) which is why the time to address differences in gender performance is now. Educators must analyze the differences in the performance of girls and boys on standardized assessments and use the information obtained in equipping the schools and districts. Therefore, the findings of this study may be helpful to educational leaders and policymakers.

Research Questions

In this study, the following overarching research question was addressed: What is the difference in reading performance between Texas Grade 4 boys and girls? Specific subquestions under this overarching research question were: (a) What is the difference between Texas Grade 4 boys and girls in their understanding across genres?; (b) What is the difference between Texas Grade 4 boys and girls in their comprehension and analysis of literary texts?; (c) What is the difference between Texas Grade 4 boys and girls in their comprehension and analysis of informative texts?; (d) What is the difference between Texas Grade 4 boys and girls in their performance on the Level II Final Satisfactory standard?; and (e) What is the degree to which trends are present in reading for Texas Grade 4 boys and girls? The first four research subquestions were addressed for three school years, whereas the last research question involved a comparison of results across all three school years.

Method

Research Design

The research design in this study was a quantitative, causal comparative, non-experimental research design (Johnson & Christensen, 2012). Researchers use causal comparative designs to find relationships between independent and dependent variables after the action has already taken place (Johnson & Christensen, 2012). In this investigation, the action that has already taken place was the STAAR Reading test that was administered to students in the 2012-2013, 2013-2014, and 2014-2015 school years. The independent variable in this research study was gender and the dependent variables were the three reporting categories (i.e., Reporting Category 1, Reporting Category 2, Reporting Category 3) and the Level II Final Satisfactory Performance Standard from the 2012-2013, 2013-2014, and 2014-2015 STAAR exams.

Instrumentation and Procedures

The data that were utilized in this study were previously obtained from the Texas Education Agency Public Education Information Management System database for the 2012-2013, 2013-2014, and 2014-2015 school years. To obtain the data, a Public Information Request was submitted to and fulfilled by the Texas Education Agency. Datasets were requested for (a) Texas Grade 4 students, (b) gender, and (c) STAAR Reading Reporting Categories.

Assessed by the STAAR Reading test are three categories for performance. In Reporting Category 1: The student will demonstrate an ability to understand a variety of written texts across reading genres (Texas Education Agency Student Assessment Division Frequently Asked Questions, 2016, p. 2). Outlined in this category is the focus on the reading and vocabulary development of the student. Students are expected to understand new vocabulary and use it when reading and writing (Texas Education Agency Student Assessment Division Frequently Asked Questions, 2016, p. 2). In addition, students are expected to identify the meaning of common prefixes and suffixes and know how they change the meaning of roots words (Texas Education Agency Student Assessment Division Frequently Asked Questions, 2016, p. 2).

In Reporting Category 2: The student will demonstrate an ability to understand and analyze literary texts (Texas Education Agency Student Assessment Division Frequently Asked Questions, 2016, p. 3). Reporting Category 2 is centered around comprehension of a variety of texts drawing on reading strategies (Texas Education Agency Student Assessment Division Frequently Asked Questions, 2016, p. 3). Students are expected to ask applicable questions, seek clarification, discover facts and details about stories, and support answers with textual evidence (Texas Education Agency Student Assessment Division Frequently Asked Questions, 2016, p. 4). In addition, students are expected to make inferences and draw conclusions about theme and genre in different cultural, historical, and contemporary contexts (Texas Education Agency Student Assessment Division Frequently Asked Questions, 2016, p. 4). Reporting Category 2 also measures students' skills on drawing conclusions about the structure and elements of poetry (Texas Education Agency Student Assessment Division Frequently Asked Questions, 2016, p. 4).

According to The Texas Education Agency, in Reporting Category 3: The student will demonstrate an ability to understand and analyze informational texts (Texas Education Agency Student Assessment Division Frequently Asked Questions, 2016, p. 5). Students are expected to analyze, draw conclusions, and make inferences about the author's purpose in cultural, contemporary, and historical contexts (Texas Education Agency Student Assessment Division Frequently Asked Questions, 2016, p. 5). Similar to Reporting Categories 1 and 2, students are expected to provide evidence from the text to support their understanding.

Each reporting category encompasses Readiness and Supporting Standards (Texas Education Agency The New STAAR Report Card Presentation, 2017, pp. 1-2). The general characteristics of Readiness Standards includes skills that are essential for success in the current grade (Texas Education Agency Student Assessment Division Frequently Asked Questions, 2016, p. 4). These standards are designed to measure student preparedness for the next grade level. In addition, these standards support college and career readiness benchmarks and measures specific content and concepts. Unlike Readiness Standards, Supporting Standards are introduced in the current grade level but emphasizes subject matter in a subsequent year. This standard addresses more narrowly defined content and concepts. Reporting Category 1 includes five multiple choice questions from both the Readiness and Supporting Standards; Reporting Category 2 contains 15 multiple choice questions from both the Readiness and Supporting Standards; and Reporting Category 3 includes 14 multiple choice questions also from both the Readiness and Supporting Standards (Texas Education Agency Student Assessment Division Frequently Asked Questions, 2016, p. 4). Level II Satisfactory Academic Performance is the performance category for students who demonstrate some knowledge of course content but may have a few deficits regarding critical elements. Those critical elements may require a student to still need additional support to master the objectives. However, this category of performance constitutes a passing score with some remediation for the next school year (Texas Education Agency, The New STAAR Report Card Presentation, 2017, p. 10). Also, students are expected to exhibit "a flexible range of met cognitive reading skills in both assigned and independent reading to understand an author's message… as they become self-directed, critical readers" by being evaluated in their mastery of Figure 19, a TEKS process standard, across the three Reporting Categories (Texas Education Agency Student

Assessment Division Frequently Asked Questions, 2016). Readers are directed to http.//tea.texas.gov/ for more reliability and validity information regarding the STAAR test.

Results

Prior to conducting a multivariate analysis of variance (MANOVA), its underlying assumptions were checked. Specifically examined were data normality, Box's Test of Equality of Covariance and the Levene's Test of Equality of Error Variances. The majority of these assumptions were not met, however, the robustness of a MANOVA procedure made it appropriate to use in this study (Field, 2009). Results of statistical analyses for Grade 4 boys and girls in Texas who took the STAAR Reading test will be described by Reading Reporting Category. Results in this study will be discussed in chronological order from 2012-2013, 2013-2014, and 2014-2015.

Overall Results for the Three School Years

With respect to the 2012-2013 school year, the MANOVA revealed a statistically significant difference, Wilks' $\Lambda = .98$, $p < .001$, partial $\eta^2 = .02$, in overall reading performance between boys and girls. The effect size for this statistically significant difference was small (Cohen, 1998). Regarding the 2013-2014 school year, the MANOVA revealed a statistically significant difference, Wilks' $\Lambda = .99$, $p < .001$, partial $\eta^2 = .01$, in overall reading performance between boys and girls. Using Cohen's (1988) criteria, the effect size was small. Concerning the 2014-2015 school year, the MANOVA revealed a statistically significant difference, Wilks' $\Lambda = .99$, $p < .001$, partial $\eta^2 = .01$, in overall reading performance between boys and girls. Based on Cohen's (1988) criteria, this effect size was small. In all three school years, the effect sizes were small. Statistically significant differences were yielded in all three school years between boys and girls in their overall reading performance.

Reading Reporting Category 1 Results (Understanding Across Genres) Across All Three School Years

Following the overall results of the MANOVA, univariate follow-up Analysis of Variance (ANOVA) procedures were conducted for each of the three STAAR Reading Reporting Categories. For the 2012-2013 school year, a statistically significant difference was yielded between boys and girls in their Reading Reporting Category 1 performance, $F(1, 372796) = 1456.96$, $p < .001$, partial $\eta^2 = .01$, small effect size (Cohen, 1998). With respect to the 2013-2014 school year, a statistically significant difference was revealed between boys and girls in their STAAR Reading Reporting Category 1 scores, $F(1, 377768) = 10689.08$, $p < .001$, partial $\eta^2 = .01$, small effect size (Cohen, 1998). Concerning the 2014-2015 school year, a statistically significant difference was again yielded between boys and girls in their STAAR Reading Reporting Category 1 scores, $F(1, 388726) = 8894.84$, $p < .001$, partial $\eta^2 = .004$, a below small effect size (Cohen, 1998). On the STAAR Reading Reporting Category 1, the effect sizes were small for the first two school years and below small in the last school year.

Regarding the 2012-2013 school year, girls had a statistically significantly higher average raw score, 0.13 points higher, than boys. Concerning the 2013-2014 school year, girls also had a statistically significantly higher average raw score, 0.34 points higher, than boys. Consistent with the other two years, in the 2014-2015 school year, girls had a statistically significantly higher average raw score, 0.30 points higher, than boys. In all three school years, girls scored statistically significantly higher on the STAAR Reading Reporting Category 1 than boys. Readers are referred to Table 1 for the descriptive statistics for this analysis.

Table 1 Descriptive Statistics for the STAAR Grade 4 Reporting Category 1 Scores of Boys and Girls for the 2012-2013, 2013-2014, and 2014-2015 School Years

School Year and Gender	*n*	*M*	*SD*
2012-2013			
Boys	190,267	6.93	2.31
Girls	182,531	7.06	2.21
2013-2014			
Boys	192,652	6.43	3.93
Girls	185,118	6.77	3.83
2014-2015			
Boys	199,474	5.97	2.51
Girls	189,254	6.27	2.45

Reading Reporting Category 2 (Understanding Literary Texts) Results Across All Three School Years

Regarding the 2012-2013 school year, a statistically significant difference was yielded between boys and girls in their Reading Reporting Category 2 performance, $F(1, 372796) = 68991.29$, $p < .001$, partial $\eta^2 = .01$, small effect size (Cohen, 1998). Concerning the 2013-2014 school year, a statistically significant difference was revealed between boys and girls in their STAAR Reading Reporting Category 2 scores, $F(1, 377768) = 50069.92$, $p < .001$, partial $\eta^2 = .01$, small effect size (Cohen, 1998). With respect to the 2014-2015 school year, a statistically significant difference was again yielded between boys and girls in their STAAR Reading Reporting Category 2 scores, $F(1, 388726) = 80876.90$, $p < .001$, partial $\eta^2 = .01$, small effect size (Cohen, 1998). On the STAAR Reading Reporting Category 2, the effect sizes for the statistically significant differences in the reading performance of boys and girls were in the small range for all three school years.

Concerning the 2012-2013 school year, girls had a statistically significantly higher average raw score, 0.86 points higher, than boys. In reference to the 2013-2014 school year, girls also had a statistically significantly higher average raw score, 0.73 points higher, than boys. Consistent with the other two years, in the 2014-2015 school year, girls had a statistically significantly higher average raw score, 0.92 points higher, than boys. Girls scored statistically significantly higher on the Reading Reporting Category 2 than boys in all three school years analyzed. Table 2 contains the descriptive statistics for this analysis.

Table 2 Descriptive Statistics for the STAAR Grade 4 Reporting Category 2 Scores of Boys and Girls for the 2012-2013, 2013-2014, and 2014-2015 School Years

School Year and Gender	*n*	*M*	*SD*
2012-2013			
Boys	190,267	11.01	3.92
Girls	182,531	11.87	3.82
2013-2014			
Boys	192,652	11.17	3.94
Girls	185,118	11.90	3.83
2014-2015			
Boys	199,474	11.65	4.19
Girls	189,254	12.57	3.92

Reading Reporting Category 3 (Understanding Informational Texts) Results Across All Three School Years

With respect to the 2012-2013 school year, a statistically significant difference was yielded between boys and girls in their Reading Reporting Category 3 performance, $F(1, 372796) = 5041.91$, $p < .001$, partial $\eta^2 = .001$, a below small effect size (Cohen, 1998). Regarding the 2013-2014 school year, a statistically significant difference was revealed between boys and girls in their STAAR Reading Reporting Category 3 scores, $F(1, 377768) = 22553.01$, $p < .001$, partial $\eta^2 = .01$, small effect size (Cohen, 1998). Concerning the 2014-2015 school year, a statistically significant difference was again yielded between boys and girls in their STAAR Reading Reporting Category 3 scores, $F(1, 388726) = 33128.80$, $p < .001$, partial $\eta^2 = .01$, small effect size (Cohen, 1998). On the STAAR Reading Reporting Category 3, the effect size was below small in the first school year and small in the last two school years.

Concerning the 2012-2013 school year, girls had a statistically significantly higher average raw score, 0.23 points higher, than boys. Regarding the 2013-2014 school year, girls also had a statistically significantly higher average raw score, 0.49 points higher, than boys. Consistent with the other two years, in the 2014-2015 school year, girls had a statistically significantly higher average raw score, 0.58 points higher, than boys. Girls scored statistically significantly higher on the Reading Reporting Category 3 than did boys in all three school years analyzed. Descriptive statistics for this analysis are contained in Table 3.

Table 3 Descriptive Statistics for the STAAR Grade 4 Reporting Category 3 Scores of Boys and Girls for the 2012-2013, 2013-2014, and 2014-2015 School Years

School Year and Gender	*n*	*M*	*SD*
2012-2013			
Boys	190,267	9.92	3.82
Girls	182,531	10.15	3.76
2013-2014			
Boys	192,652	9.79	3.65
Girls	185,118	10.28	3.57
2014-2015			
Boys	199,474	9.65	3.89
Girls	189,254	10.23	3.77

Results for the Level II Final Satisfactory Performance Analyses Over Time

Student performance on the Level II Final Satisfactory standard was examined next through the use of Pearson chi-square procedures. This statistical procedure was the most appropriate statistical procedure to use because dichotomous data were present for the Level II Final Satisfactory Performance Standard (i.e., met or did not meet this standard) and for gender. As such, the chi-square is the preferred statistical procedure when both variables are categorical (Field, 2009). Because a large sample size was present, the assumptions for utilizing a chi-square were met.

Concerning the STAAR Level II Final Satisfactory Performance Standard, a statistically significant difference was present between boys and girls in the 2012-2013 school year, $\chi^2(1) = 602.77$, $p < .001$. The effect size revealed for this finding, Cramer's V, was below small, .04 (Cohen, 1988). A statistically significantly higher percentage of girls, 3.9%, met the Level II Final Satisfactory Performance Standard than boys. Table 4 contains the descriptive statistics for this analysis.

Table 4 Frequencies and Percentages for the Grade 4 STAAR Reading Level II Satisfactory Performance Standard of Boys and Girls for the 2012-2013, 2013-2014, and 2014-2015 School Years

	Met Standard		Did Not Meet Standard	
School Year	*n*	%	*n*	%
2012-2013				
Boys	68,748	35.6	124,250	64.4
Girls	72,624	39.5	111,259	60.5
2013-2014				
Boys	64,861	33.2	130,669	66.8
Girls	71,057	38.1	115,509	61.9
2014-2015				
Boys	71,292	36.3	125,110	63.7
Girls	81,134	43.2	106,552	56.8

With regard to the Level II Final Satisfactory Performance Standard of boys and girls, the result for the 2013-2014 school year was statistically significant, $\chi^2(1) = 1{,}006.29$, $p < .001$. The effect size revealed for this finding, Cramer's V, was below small, .05 (Cohen, 1988). A statistically significantly higher percentage of girls, 4.9%, met the Level II Final Satisfactory Performance Standard than boys. Delineated in Table 4 are the descriptive statistics for this analysis.

Regarding the Level II Final Satisfactory Performance Standard of boys and girls, a statistically significant difference was present in the 2014-2015 school year, $\chi^2(1) = 1{,}925.34$, $p < .001$. The effect size revealed for this finding, Cramer's V, was below small, .07 (Cohen, 1988). A statistically significantly higher percentage of girls, 6.9%, met the Level II Final Satisfactory Performance Standard than boys. Readers are referred to Table 4 for the descriptive statistics for this analysis.

In analyzing the reading performance of Grade 4 students in Texas across the three years of data that were analyzed, clear trends were present in the reading scores of boys and girls. In each of the three STAAR Reading Reporting Categories for all three years, girls demonstrated statistically significant higher reading test scores than boys. Moreover, statistically significantly higher percentages of Grade 4 girls met the STAAR Reading Level II Satisfactory Standard than did Grade 4 boys. These results are depicted in Figures 1, 2, 3, and 4.

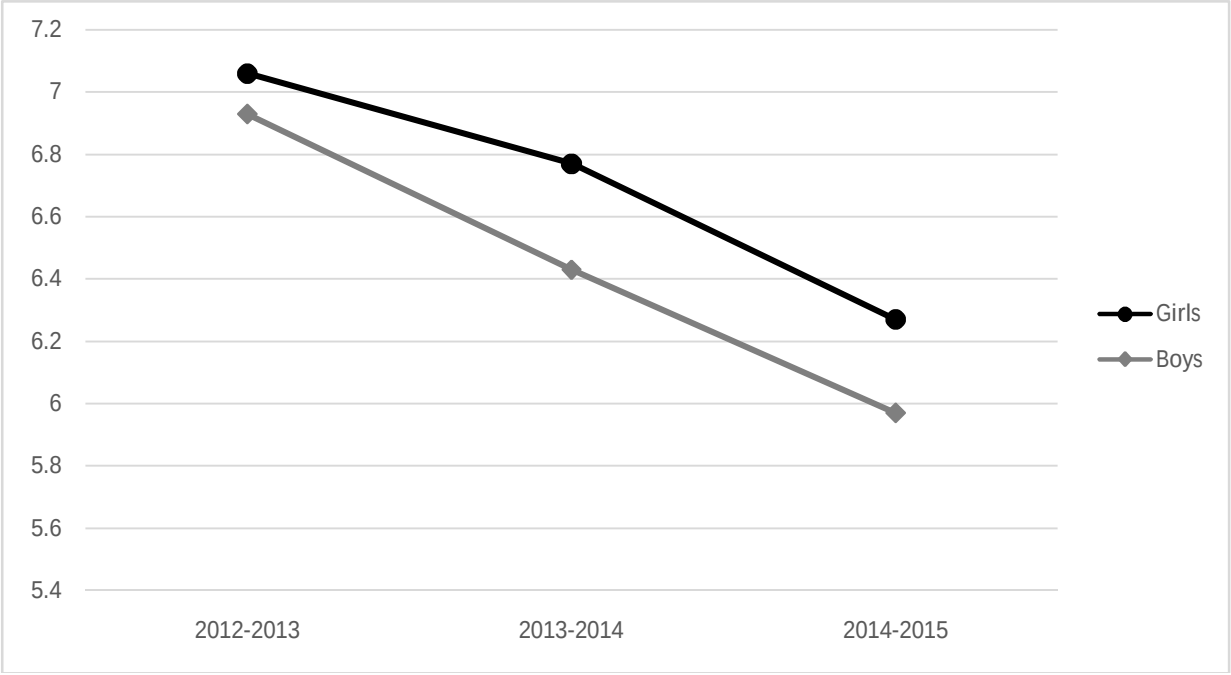

Figure 1. Average scores for boys and girls on the STAAR Grade 4 Reporting Category 1 for the 2012-2013, 2013-2014, and 2014-2015 school years.

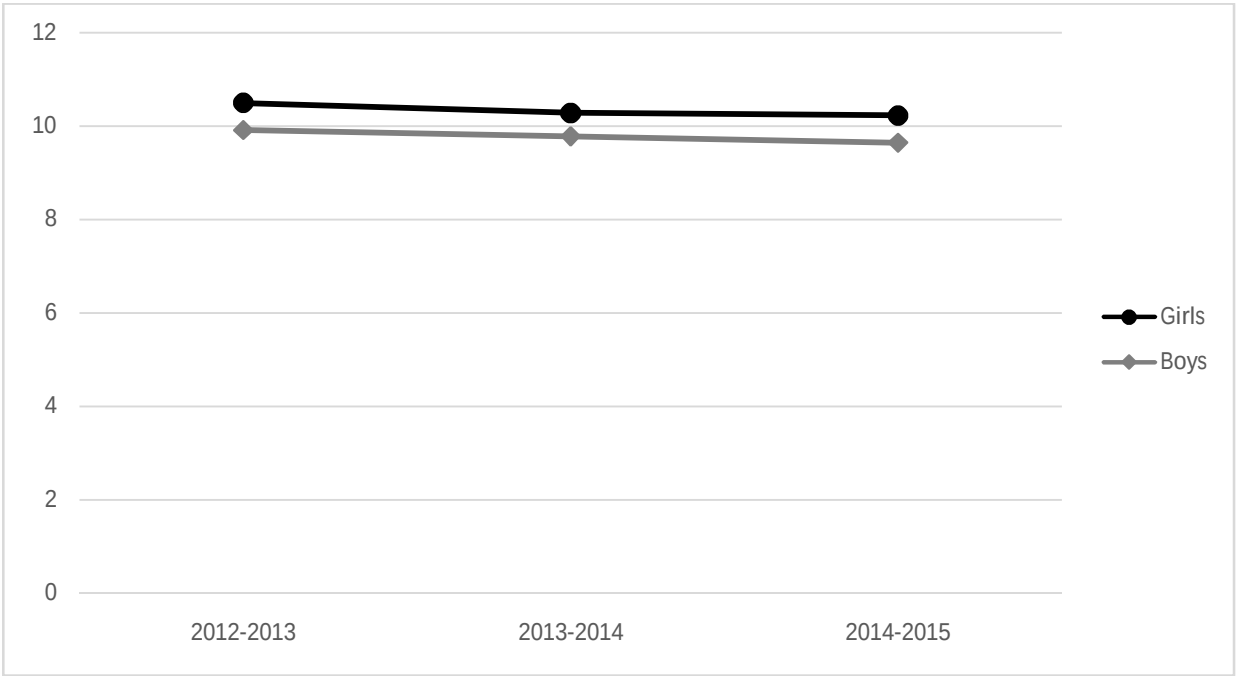

Figure 2. Average scores for boys and girls on the STAAR Grade 4 Reporting Category 2 for the 2012-2013, 2013-2014, and 2014-2015 school years.

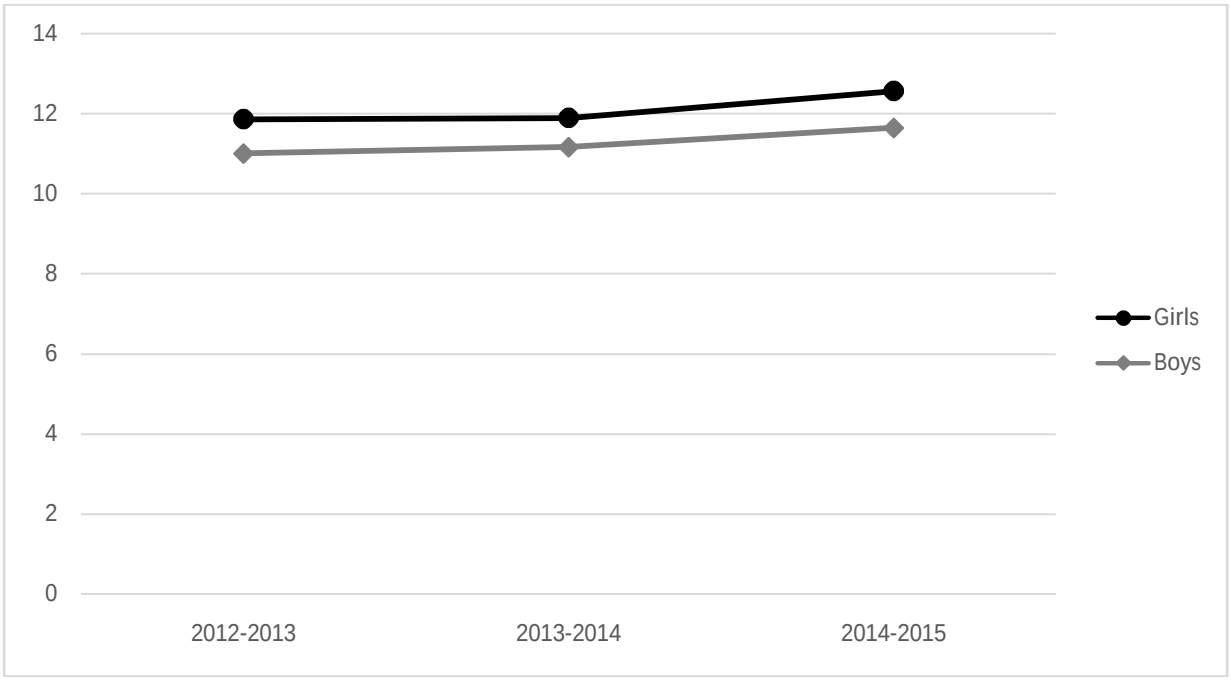

Figure 3. Average scores for boys and girls on the STAAR Grade 4 Reporting Category 3 for the 2012-2013, 2013-2014, and 2014-2015 school years.

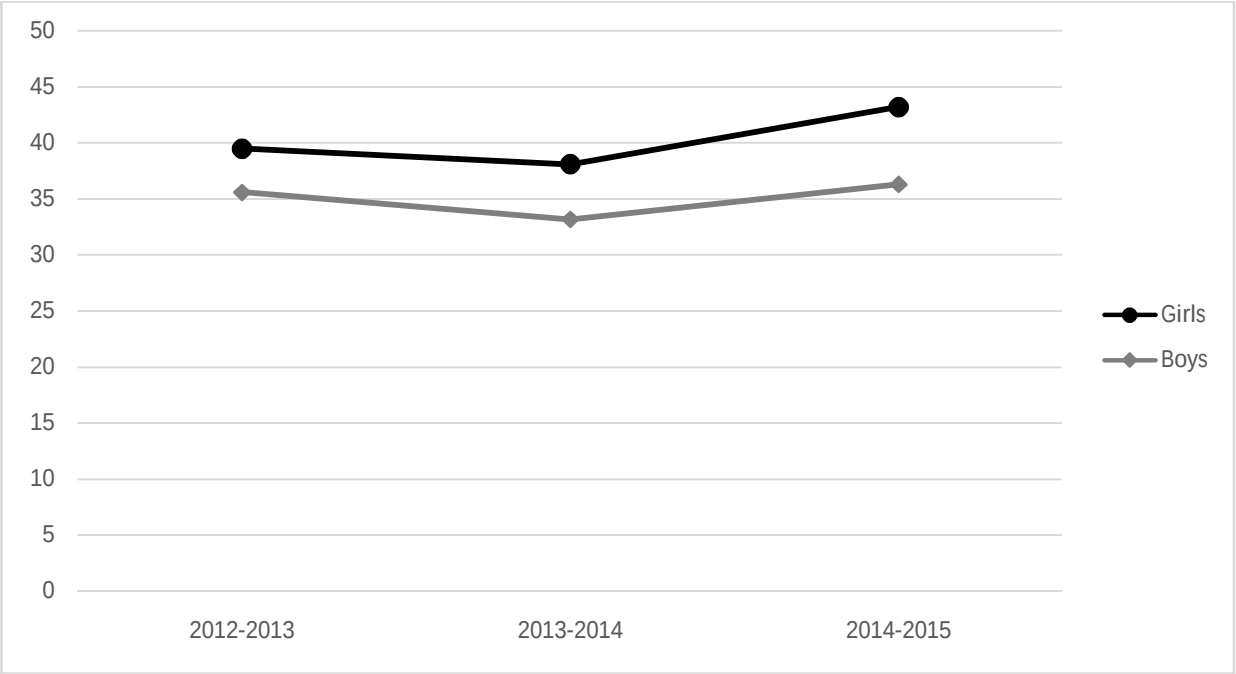

Figure 4. Grade 4 STAAR Reading Level II Satisfactory Performance Standard of boys and girls for the 2012-2013, 2013-2014, and 2014-2015 school years.

Discussion

Examined in this investigation was the degree to which differences were present between girls and boys in their reading performance on the Texas state-mandated assessment. Three years of statewide data on the three Grade 4 STAAR Reading Reporting Categories were analyzed for boys and girls. Inferential statistical analyses revealed the presence of statistically significant differences between boys and girls in their reading performance in all three school years. A summary of these results is contained in Table 5. Following these statistical analyses, the Level II Final Satisfactory Performance Standard by gender was addressed and determined to yield statistically significant gender differences.

Table 5 Summary of Reading Performance Results for the Grade 4 STAAR Reading Exam for Boys and Girls in the 2012-2013, 2013-2014, and 2014-2015 School Years

STAAR Reading Category	Statistically Significant	Effect Size	Lowest Performing Group
2012-2013			
Reporting Category 1	Yes	Small	Boys
Reporting Category 2	Yes	Small	Boys
Reporting Category 3	Yes	Below Small	Boys
2013-2014			
Reporting Category 1	Yes	Small	Boys
Reporting Category 2	Yes	Small	Boys
Reporting Category 3	Yes	Small	Boys
2014-2015			
Reporting Category 1	Yes	Below Small	Boys
Reporting Category 2	Yes	Small	Boys
Reporting Category 3	Yes	Small	Boys

Connections to Existing Literature

Concern over gender inequities in reading achievement has been present for generations (McGown, 2016). Education should be the great equalizer, however, gender literacy gaps have been extensively documented for decades. In a recent Texas, multiyear analysis, McGown (2016) examined the STAAR Reading test scores of Grade 3 boys and girls to ascertain the extent to which gender differences were present. In the three years of data she analyzed, girls had statistically significantly higher overall reading scores and higher STAAR Reading Reporting Category 1, 2, and 3 test scores. Moreover, higher percentages of girls met the passing standard on the STAAR Reading test than did boys.

As evidenced by the results of this study, Grade 4 girls in Texas have better reading scores than do boys. These findings are consistent with the existing literature regarding the disparity between girls and boys in their reading achievement. Although efforts have been made to close the achievement gaps in reading between ethnic and racial subgroups, limited success has occurred with regard to closing the gender gap in reading achievement (Klecker, 2006). When Klecker (2006) analyzed gender differences between students in elementary and high school, he determined that girls outperformed boys in all six years of data that were analyzed. Another researcher, Stinnett (2011), established that statistically significant differences were present between girls and boys in a study analyzing skill development in reading. Girls continue to outperform boys in school as evident by their report card grades and by the state-mandated standardized assessments given each year (Sadker&Zittleman, 2005).

In 2008, Northwestern University claimed that girls have a superior language ability to boys; and that girls enter school with more advanced literacy skills than boys (Stinnett, 2011). Not only are Grade 4 girls outperforming boys in the United States but around the world, scoring higher than boys in all G-20 countries (National Center for Education Statistics, 2015). As identified, the achievement gaps between girls and boys are seen at an early age. When those deficits are not addressed by Grade 3, the first year of state-mandated testing, girls will continue to perform higher than boys (McGown, 2016).

Implications for Policy and Practice

Based on the results of this multiyear statewide investigation in which STAAR reading scores were analyzed by gender, several implications for policy and practice can be recommended. First, additional funding should be provided to purchase reading material that interest boys. Librarians, Media Specialist, teachers, and Literacy Coaches should provide both girls and boys with an interest survey to determine which subjects, genres, and texts would be of interest. Boys should be allowed to read sports magazines, car magazines, or what reading material interest them the most. Educators should find a process for incorporating rigor as well as relevance in every reading lesson to increase engagement and excitement. Second, school districts in collaboration with state and federal agencies should provide professional development opportunities specifically designed to target the reading gaps between girls and boys. Teachers should be equipped to provide research-based strategies and techniques aimed at ensuring that boys acquire reading skills at the same rate of proficiency as girls. Third, districts should receive funding for full-day pre-kindergarten programs which would provide the early literacy foundation needed for all students to read at or above their reading level. Fourth, schools in conjunction with their district should provide parenting classes to shed light on the disparities between girls and boys and provide resources for parents to use at home. These efforts will assist the teachers, administrators, and coaches in closing the gaps between girls and boys in reading.

Recommendations for Future Research

Based on the results of this empirical multiyear investigation, several recommendations for future research can be made. First, researchers are encouraged to examine the relationship between reading performance and gender in other grade levels. The findings from this investigation are limited to Grade 4 boys and girls. As such, researchers should examine the reading performance of boys and girls in middle school and high school to determine if the gender differences delineated herein are also present at other grade levels. In this study, reading achievement was analyzed by gender only. Accordingly, a second recommendation would be for researchers to analyze other demographic characteristics such as economic status and ethnicity/race to ascertain whether relationships are present between those characteristics and student reading performance.

With only reading achievement being analyzed in this study, the third recommendation is to examine if performance differences by gender are present in other subjects such as mathematics, writing, and science. Fourth, researchers should examine reading performance by gender in other states. Only data on the students in Texas were examined in this study. The extent to which the results of this study can be generalized to other states is unknown. Fifth, to analyze trends over several years, researchers are encouraged to conduct longitudinal studies, beginning in Kindergarten and going through Grade 12. A study of this magnitude will allow researchers to connect gender differences with student achievement in multiple grade levels. Last, researchers are also encouraged to conduct mixed and qualitative research studies to provide meaningful data that policymakers and educators can use in making informed decisions regarding educating students based on their gender.

Conclusion

The purpose of this research study was to determine the degree to which differences were present between Grade 4 boys and girls in their reading performance on the Texas state-mandated assessment. Inferential statistical analyses of three years of Texas statewide data revealed the presence of statistically significant differences between boys and girls in their reading performance. In all analyses, girls had better reading scores than boys and higher percentages who met the passing standard. As such, results from this multiyear, statewide analysis are congruent with the extant literature.

References

Ayers, L. (1909). *Laggards in our schools*. New York, NY: Russell Sage Foundation.

Below, J. L., Skinner, C. H., Fearrington, J. Y., & Sorrell, C. A. (2010). Gender differences in early literacy: Analysis of kindergarten through fifth-grade dynamic indicators of basic early literacy skills probes. *School Psychology Review, 39*(2), 240-257.

Catsambis, S. S., Mulkey, L. M., Buttaro, A., Steelman, L. C., & Koch, P. R. (2012). Examining gender differences in ability group placement at the onset of schooling: The role of skills, behaviors, and teacher evaluations. *Journal of Educational Research, 105*(1), 8-20. doi:10.1080/00220671.2010.514779

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Condron, D. J. (2007). Stratification and educational sorting: Explaining ascriptive inequalities in early childhood reading group placement. *Social Problems, 54*, 139–160.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage.

Johnson, R. B., & Christensen, L. B. (2012). *Educational research: Quantitative, qualitative, and mixed approaches* (4th ed.) Los Angeles, CA: Sage.

Klecker, B. M. (2006). The gender gap in NAEP fourth-, eighth-, and twelfth- grade reading scores across years. *Reading Improvement, 43*, 50-56.

McGown, J. A. (2016). *Differences in reading performance of Texas elementary school students as a function of economic status, gender, and ethnicity/race: A multiyear statewide study.* Doctoral Dissertation, Sam Houston State University, Huntsville, TX.

Moore, G., Combs, J. P., & Slate, J. R. (2012). Advanced Placement exams participation and performance: A national study of gender differences. *E-International Journal of Educational Research, 3*(3), 18-32.

Mullis, I., Martin, M., Foy, P., & Hooper, M. (2017). *Progress in International Reading Literacy Study*. Retrieved from http://timssandpirls.bc.edu/pirls2016/international-results/wp-content/uploads/structure/CompletePDF/P16-PIRLS-International-Results-in-Reading.pdf

National Center for Education Statistics. (2015). *Comparative indicators of education in the United States and other G-20 countries*. Retrieved from https://nces.ed.gov/pubs2016/2016100/ind_15.asp

Northwestern University. (2008). Boys' and girls' brains are different: Gender differences in language appear biological. *Science Daily*. Retrieved from www.sciencedaily.com/releases/2008/03/080303120346.htm

Sadker, D., &Zittleman, K. (2005). *Closing the gender gaps*. Retrieved from https://www.naesp.org/sites/default/files/resources/2/Principal/2005/M-Ap18.pdf

Sanchez, C. (2018). *The gap between the science on kids and reading, and how it is taught*. Retrieved from https://www.npr.org/sections/ed/2018/02/12/582465905/the-gap-between-the-science-on-kids-and-reading-and-how-it-is-taught

Schleeter, G. D. (2017). *Differences in the reading achievement of Texas Grade 3 English Language Learners as a function of their economic status, ethnicity/race, and gender: A multiyear statewide study*. Doctoral Dissertation, Sam Houston State University, Huntsville, TX.

Stinnett, M. (2011). Research in reading. *Illinois Reading Council Journal, 39*(3), 72-77.

Tach, L. M., & Farkas, G. (2006). Learning-related behaviors, cognitive skills, and ability grouping when schooling begins. *Social Science Research, 35*, 1048-1079.

Texas Education Agency. (2016). *Student Assessment Division Frequently Asked Questions*. Retrieved from https://tea.texas.gov/student.assessment/staar/

Texas Education Agency. (2017). *The New STAAR Report Card Presentation*. Retrieved from http://www.stamford.esc14.net/page/open/1039/0/Introducing%20the%20New%20STAAR%20Student%20Report-Guide%20for%20Parents%201.pdf

Wright, L. A., & Slate, J. R. (2015). Differences in critical thinking skills for Texas middle school students as a function of economic disadvantage. *Journal of Education Research, 9*(4).