

Exploring the Relationship Between Pupil/Teacher Ratio and Reading Scale Scores in the United  
States and a Comparison of the East and West Regions

Julian Narchet and Joshlyn Thomas

University of Central Florida

## Exploring the Relationship Between Pupil/Teacher Ratio and Reading Scale Scores in the United States and Comparison of the East and West Regions

The United States generally acknowledges the year of 1963 as the beginning of the modern middle school movement. Around this timeframe, educational scholars proposed the name of “middle school” and began to distinguish the middle school characteristics that many Americans are familiar with today (Schaefer, Malu & Yoon, 2016).

By the 1980s, with over 4,000 middle schools operating within the country, the movement saw advancements in curriculum related to student engagement, block scheduling, team teaching, and school counseling (Schaefer, et al., 2016).

According to a study conducted by Chamberlain, Daniels, Madden, & Slavin (2007), this reform in middle schools has dominated conversations nationally and focuses on “such structural elements as grade span (6-8 instead of 7-9), interdisciplinary teaming, site-based decision making, flexible scheduling, and reducing ability grouping.”

As of the 2012-2013 school year, The National Center for Educational Statistics, cites approximately 24,280 middle (or secondary schools) in existence in the United States. These middle schools vary by location, provide students the educational fundamentals of reading, writing, mathematics, and science, and offer each respective student a different experience to travel along with them upon completion.

Because of the variability in middle schools amongst different regions in the United States, a broadly perceived need for more focus on instruction and curriculum, especially related to literacy, has been prominent (Chamberlain, et al., 2007). The location of the school, resources available, funding received, and teacher efficiency are all factors that contribute to the success of

middle school students. These factors also impact student assessments within each respective district, state, and on a national level.

Generally, the proper tools provided to students upon entering middle school are noted to assist them on assessment procedures, literacy levels, and overall, to equip them in successful lives post-middle school. These tools can be defined as a nurturing classroom environment, inclusive of a small pupil/teacher ratio, available and up-to-date resources (technology and books), and the middle school's commitment to each student personally succeeding. These general tools will lead into the purpose of the researchers' study.

### **Purpose**

The purpose of this study is to thoroughly examine the construct of middle school, specifically its effect on eighth grade students and their reading test scores as measured by the National Assessment of Educational Progress. The researchers of this study seek to compare regional differences amongst states and readily available resources offered within classrooms, and to what effect these factors have upon eighth grade reading test scores.

### **Research Questions**

1. Is there a significant relationship between pupil/teacher ratio and eighth graders' reading scale scores (as measured by the National Assessment of Educational Progress)?
2. Is there a significant difference in reading scale scores (as measured by the National Assessment of Educational Progress) between states located in the Eastern and Western regions of the United States?

The researchers intend to find if there is a significant relationship between pupil/teacher ratio and eighth graders' reading scale scores (as measured by the National Assessment of Educational

Progress), and if the East and West coast regions have a significant difference in reading scale scores (as measured by the NAEP).

### **Operational Definitions**

The following terms are considered throughout the research as operational definitions: pupil/teacher ratio, reading scale scores, National Assessment of Educational Progress, and Eastern and Western states of the United States.

**Pupil/teacher ratio** is calculated by the total reported students divided by the full-time equivalent (FTE) classroom teachers. The national average for this ratio is 16.07 (NCES, 2006b). The NAEP measures **reading scale scores** through a scale score based on statistical procedures called item response theory (IRT). The national average for eighth grade reading scores is 261 (NAEP - Overview, 2016). The **National Assessment of Educational Progress** is “the largest nationally representative and continuing assessment of what America's students know and can do in various subject areas” (NCES, 2006a). The states that comprised **Eastern and Western regions** of the United States were defined by geographic location, and were selected at random with a sample size of 24.

### **Methods**

#### **Research Design**

This quantitative research study will utilize a non-experimental research design. The data considered in this study has been sourced from State Education Profiles of the National Center for Education Statistics.

#### **Participants**

This study will consider a sample size of data from the population of the United States. A non-random sample of twenty-four states were chosen from two regions of the United States to represent the population.

For the purpose of this study, the United States will be divided into the aforementioned two regions of East and West. The states representing these regions are listed below

East

1. Connecticut
2. Washington, D.C.
3. Florida
4. Georgia
5. Maine
6. Massachusetts
7. New Jersey
8. New York
9. North Carolina
10. Pennsylvania
11. South Carolina
12. Virginia

West

1. Arizona
2. California
3. Colorado
4. Idaho
5. Montana
6. Nevada
7. New Mexico
8. Oregon
9. Texas
10. Utah
11. Washington
12. Wyoming

Instruments

The data considered from the sample group includes pupil/teacher ratio and the reading composite scale scores of eighth grade students from the 2014-2015 school year.

The pupil/teacher ratio data has been sourced from the Common Core of Data (CCD) which “is a program of the U.S. Department of Education's National Center for Education Statistics that annually collects fiscal and non-fiscal data about all public schools, public school districts and state education agencies in the United States” (“Common Core of Data (CCD)”, 2016).

The reading scale scores come from the National Assessment of Educational Progress (NAEP). NAEP “is the largest nationally representative and continuing assessment of what America's students know and can do in various subject areas” (NAEP - Overview, 2016).

The NAEP has created a scale score rooted in statistical procedures, known as the item response theory (IRT). The IRT is used to summarize student's performance in test exercises that utilize similar knowledge and skills; the reading scale score ranges from 0 – 500 (NAEP - Overview, 2016).

The test exercises used to determine a student's reading scale score with IRT are aimed at assessing the student's ability to locate and recall, integrate and interpret, as well as critique and evaluate. These skills are assessed using multiple-choice questions regarding individual texts and constructed-response questions that pertain to approximately two passages that the test-taker read (NAEP Nations Report Card - More About the NAEP Reading Assessment, 2016).

The students chosen to take the NAEP are sampled in order to represent the student population. The sampling is non-random and is proportionate to the size of the school and grade being assessed. The sampling is done in a multi-stage fashion in which public and private schools are selected within designated areas, then select students designated within relevant grades, and then students are allocated to assessment subjects (NAEP Nations Report Card - More About the NAEP Reading Assessment, 2016).

### **Results**

(Below) The frequency distribution of the collected data illustrates that the most common average figure within the set is the amount of 269.00 at a frequency of four. This average is 16.7% of the total amount amongst all 24 sampled states in regards to reading scale scores. Following this amount, the reading scale scores of 263.00 and 268.00 follow directly after with frequencies of three, and percentages of 12.5. Additionally, on a scale of 0-500, no sampled state had an average above 274.00. The national reading scale score average is 261.00, and as illustrated within the frequency distribution, only five states fall below that national amount.

Frequency Distribution & Measures of Central Tendency/Variance of Scale Scores

**Reading\_Scale\_Score**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 248.00	1	4.2	4.2	4.2
253.00	1	4.2	4.2	8.3
259.00	2	8.3	8.3	16.7
260.00	1	4.2	4.2	20.8
261.00	2	8.3	8.3	29.2
262.00	1	4.2	4.2	33.3
263.00	3	12.5	12.5	45.8
267.00	2	8.3	8.3	54.2
268.00	3	12.5	12.5	66.7
269.00	4	16.7	16.7	83.3
270.00	1	4.2	4.2	87.5
271.00	1	4.2	4.2	91.7
273.00	1	4.2	4.2	95.8
274.00	1	4.2	4.2	100.0
Total	24	100.0	100.0	

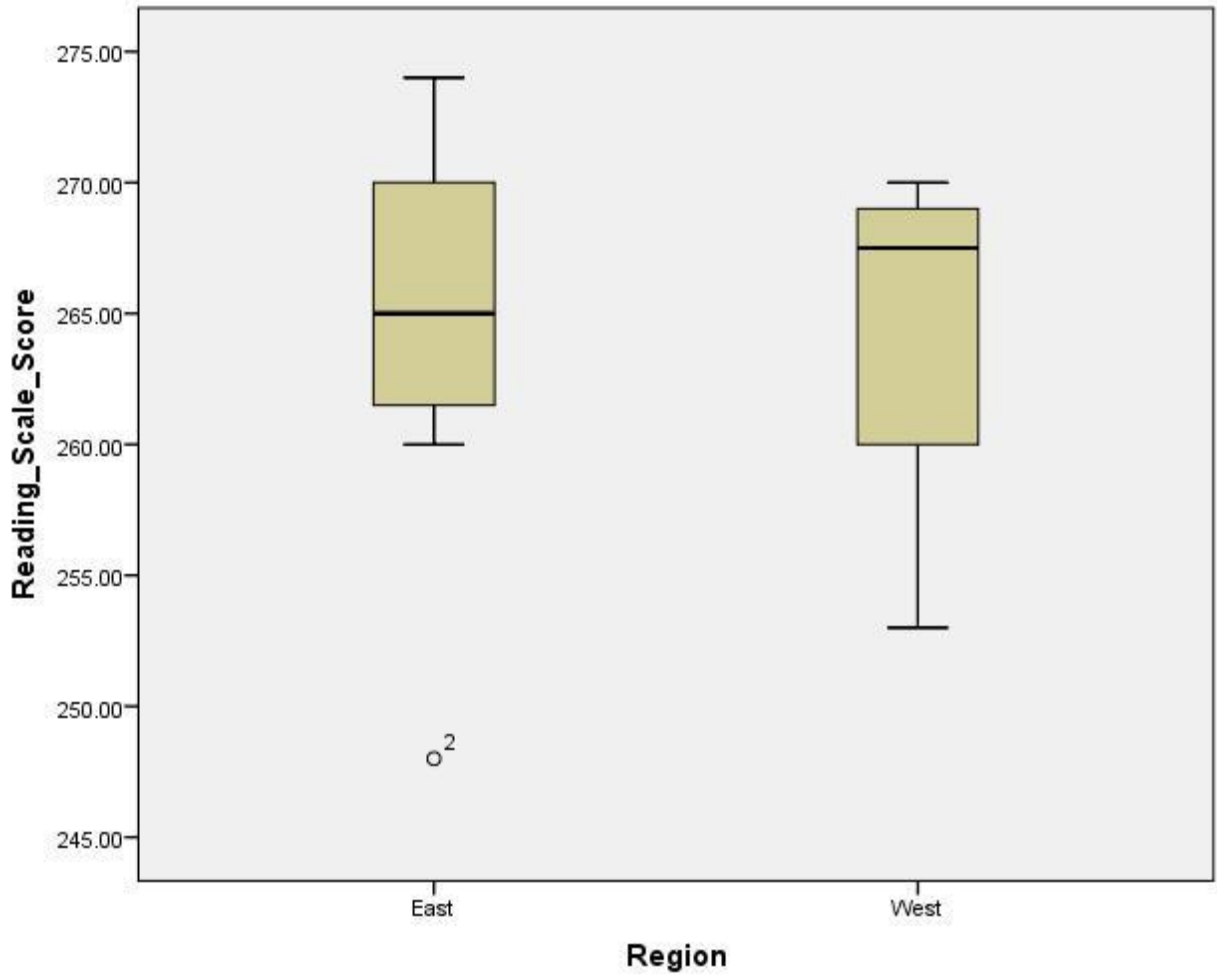
**Statistics**

Reading\_Scale\_Score

N	Valid	24
	Missing	0
Mean		264.7500
Median		267.0000
Std. Deviation		6.20834
Variance		38.543

(Above) The descriptive statistics of the reading scale scores shows that mean and median are significantly close with the amounts of 264.75 and 267.00, respectively. With this information, we can infer that the data points are evenly divided around the mean, and the data is slightly symmetrical and normal. The average scale score for the 24 sampled states is 264.75.

Boxplot



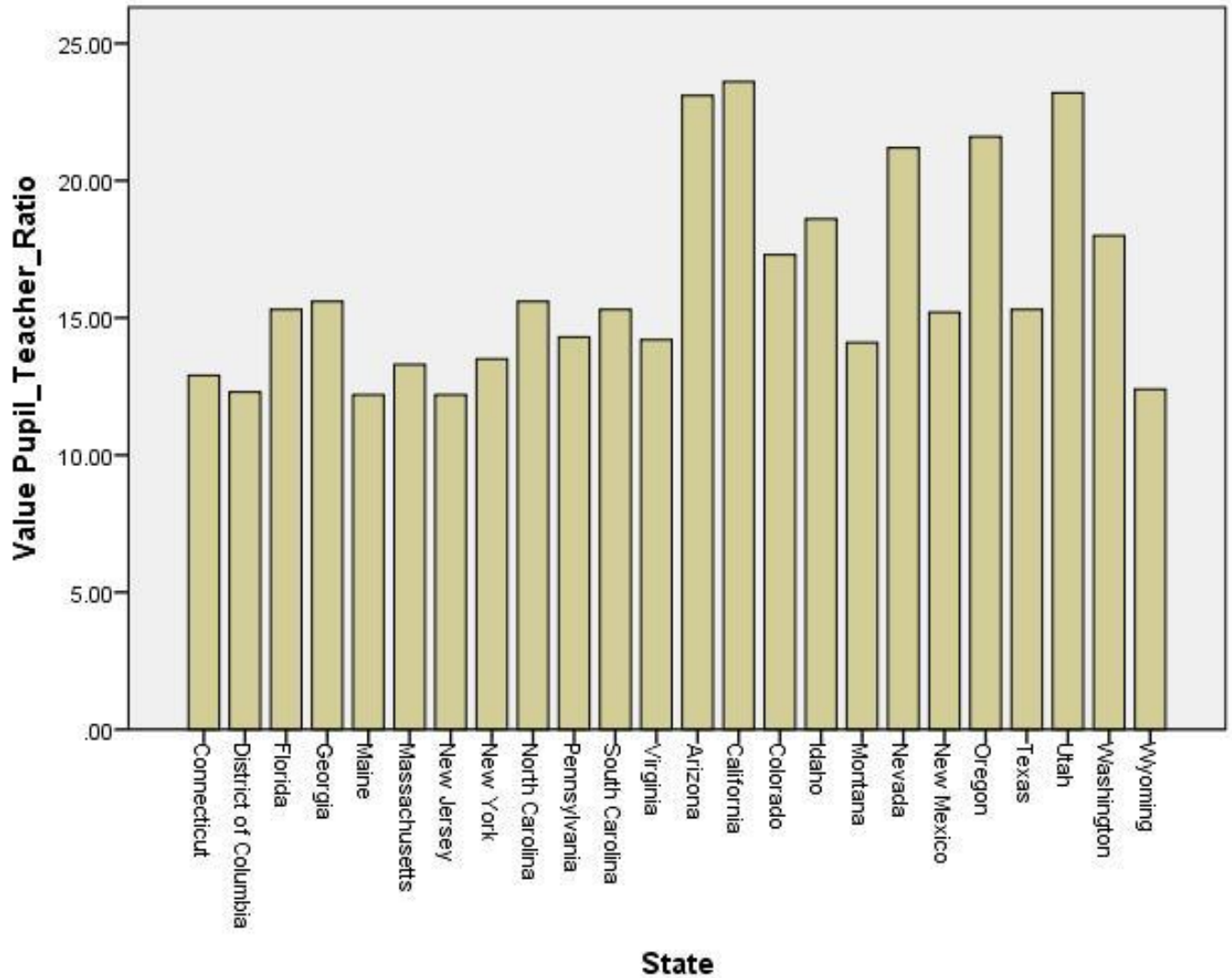
N	Valid	24
	Missing	0
Median		267.0000
Percentiles	25	261.0000
	50	267.0000
	75	269.0000

(Above) The boxplot represents the data spread out and broken into quartiles. Based upon the data collected from the 24 sampled states, Washington, D.C. presents itself as an outlier with a scale score of 248.00. This strikes as a topic of interest because it has the second lowest pupil/teacher ratio. Maine is the only sampled state with an even lower ratio, but it has an

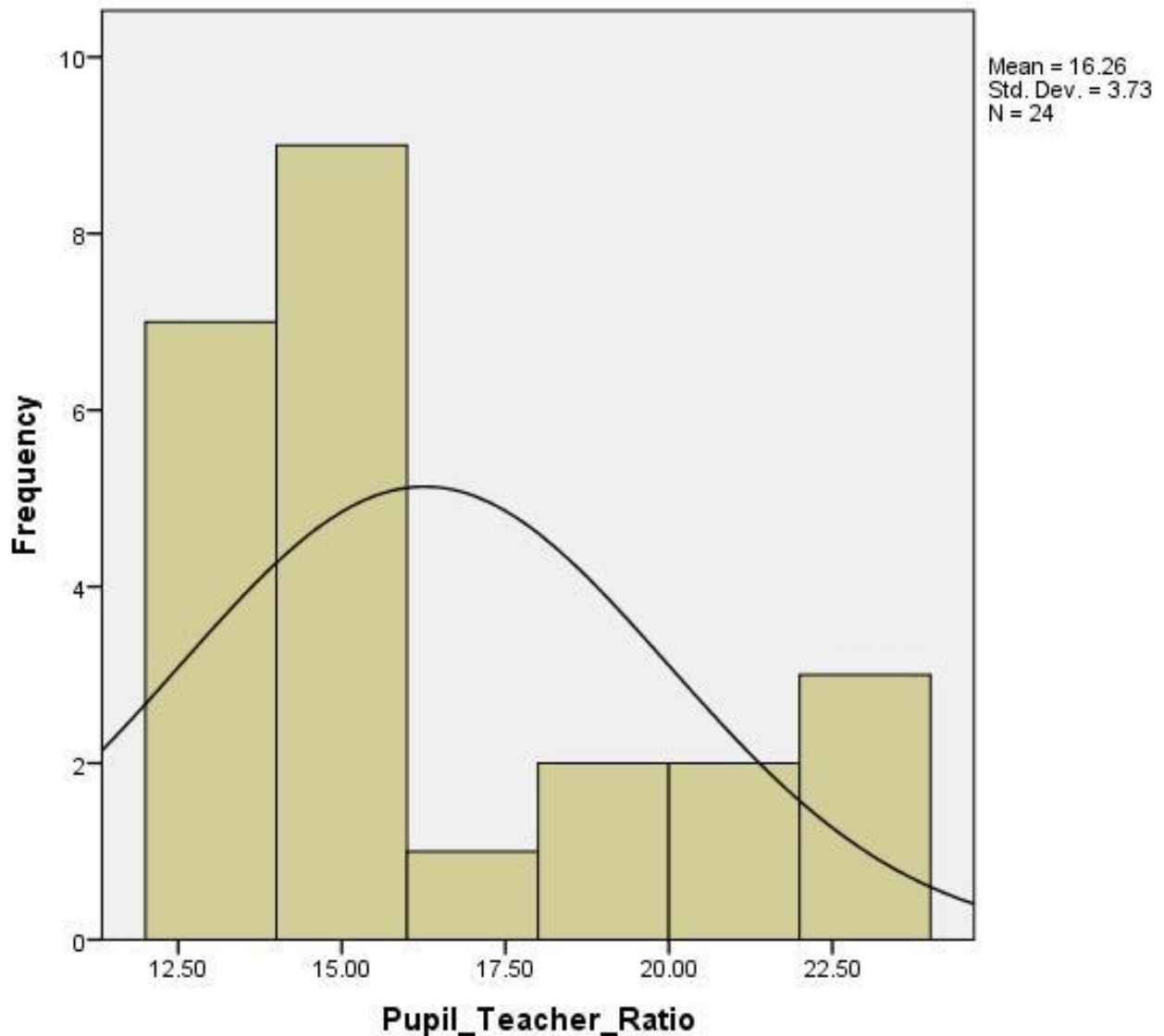


average scale score 268.00. The data of the boxplot is concentrated under the median, making the data skewed left.

### Bar graph



(Above) The pupil/teacher ratio average of the sampled states is 16.26. The bar graph illustrates 16 states which are below the national average of 16.07. The majority of these sixteen states are located on the East coast. California, Utah, and Arizona visually stand out as states with the highest pupil/teacher ratio. In general, the East coast states have lower pupil/teacher ratios, however, there isn't a significant difference in averages between the two regions.

Histogram

(Above) As the data within the histogram illustrates, there are more states with lower pupil/teacher ratios (under 17.5) than those with higher ones, demonstrating a positive skewness. The peakedness is mesokurtic with a standard normal distribution and a slightly lower peak.

Research Question 1:

Research question 1 sought to determine if there was a significant relationship between pupil/teacher ratio and the average reading scale scores for each of our sampled states. A

Pearson Correlation was utilized to answer this research question (Appendix A). The determining coefficients in this correlation are a) pupil/teacher ratio and b) the average reading scale scores recorded from the states in our sample.

The null hypothesis is that the correlation coefficients are equal to 0, and the alternative hypothesis is that correlation coefficients are not equal to 0. The hypothesis is symbolized as follows:

$$H_0: \rho_{XY} = 0$$

$$H_0: \rho_{XY} \neq 0$$

After reviewing the scatterplot of the variables, there appeared to be a weak relationship. The line of best fit drawn through the scatterplot reveals a slightly negative linear relationship between the variables.

The assumption of independence has been violated and there is an increased chance of a Type I or Type II error due to the states chosen for this study being part of a non-random sample.

The two correlation coefficients analyzed did not prove to be statistically significant, with  $p$  values of 0.315. Although the relationship between the variables was not statistically significant, there was a weak negative relationship, per Cohen's standards, between pupil/teacher ratio and average reading scale scores ( $r = -0.103$ ,  $n = 24$ ,  $p > .05$ ). These results suggest that as pupil/teacher ratio rises, average reading scale scores fall. The post hoc power of the correlation was extremely low at 0.31.

#### Research Question 2:

An independent samples  $t$  test was conducted (Appendix B) to determine if there is a significant difference between the reading scale scores of the Eastern and Western regions of the United States. The test was conducted using an alpha of .05. The null hypothesis is that the

population means are equal, and the alternative hypothesis is that the population means are not equal. The hypothesis is symbolized as follows:

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

The assumption of normality was tested and met for the distributional shape of the dependent variable for the East. Review of the Shapiro-Wilk's test for normality ( $SW = .917, p = .262$ ), skewness (-1.045) and kurtosis (1.847) statistics indicate that normality is a reasonable assumption for the East. The boxplot showed that the District of Columbia represented an outlier with an average scale score of 248. The histogram, normal Q-Q plot, and detrended normal Q-Q plot's showed signs of non-normality. The histogram did not seem to show a normal bell shaped curve. The points were adhering relatively closely to the diagonal line of the normal Q-Q plot, and the detrended normal Q-Q plot displayed reading scale scores below zero and higher reading scale scores above zero. Overall, the results suggest evidence of non-normality.

The assumption of normality was tested and met for the distributional shape of the dependent variable for the West. Review of the Shapiro-Wilk's test for normality ( $SW = .852, p = .039$ ), skewness (-0.961) and kurtosis (-0.081) statistics indicate that normality is not a reasonable assumption for the West due to low p value. The boxplot did not suggest any potential outliers for the Western region. The histogram, normal Q-Q plot, and detrended normal Q-Q plot's all showed signs of non-normality. The points were not adhering relatively closely to the diagonal line of the normal Q-Q plot, and the detrended normal Q-Q plot displayed lower reading scale scores below zero and higher reading scale scores above zero. Overall, the results suggest evidence of non-normality.

Levene's test indicated that the assumption of homogeneity of variances was not met ( $F = 0.281, p = 0.601$ ). Because there was not random assignment to groups, the assumption of independence was not met, therefore there is an increased chance of a Type I or Type II error.

The test was not statistically significant,  $t(22) = 0.129, p = .899$ . Eastern states have a slightly higher average reading scale score ( $n = 12, M = 264.92, SD = 7.12$ ) than western states ( $n = 12, M = 264.58, SD = 5.47$ ). The 95% confidence interval for the difference between means was -5.039 and 5.706.

The effect size was calculated by  $d$ , or the difference in means divided by the pooled standard deviation, and found to be 0.0525 indicating that there is less than one standard deviation unit difference in reading scale scores between the East and West regions. This is generally interpreted to be very small effect. Post hoc power was computed to be 0.05. These results fail to provide enough evidence to reject the null hypothesis, showing that there is not a significant difference between the average reading scale scores of the Eastern and Western regions of the United States.

### **Discussion**

The findings of this research suggest that there is a relationship between pupil/teacher ratio and eighth graders' reading test scores, however it is not significant based upon a high  $p$  value of 0.315. Though the relationship is not statistically significant, the collected data does show that there is a weak negative relationship between the two variables, suggesting that as pupil/teacher ratio rises, average reading scale scores fall. This assumption is illustrated by states such as Maine, Connecticut, and Massachusetts that show a below average pupil/teacher ratio, and an above average eighth grade reading score.

Furthermore, the findings of this research suggest that there is not a significant difference between the average reading scale scores of the Eastern and Western regions of the United States. The independent samples  $t$  test concluded that the means for both Eastern and Western regions were extremely close at 264.92 and 264.58, respectively. Through calculated the difference in both means, and then dividing that figure by the averaged standard deviation, the researchers found that there is less than one standard deviation unit difference in eighth grade reading scale scores between the East and West regions, at an amount of 0.0525. This, in turn, illustrates no significant difference between the two regions and the sampled states, but does indicate some practical evidence to alter for future research.

### **Limitations & Future Research**

When conducting this study, the researchers found several limitations, that if corrected, could further benefit future research on these topics. The first limitation was based upon the data provided within the National Center for Education Statistics from the Common Core of Data. The data reports pupil/teacher ratio that is inclusive of elementary and secondary school, and is not specific to only traditional middle school ratios (sixth through eighth grade), and even more specifically, just the ratios of eighth grade students and respective teachers. Further studies could delve deeper into this construct to access a more precise ratio, and to bypass a blanketed number of pupils to teachers in more than the grade level being researched.

The second limitation involves the student population of East versus West coast states. If there are more students, and specifically eighth graders residing on either coast compared to the other, it has the possibility to impact averages and succeeding research. For future research, examining the total amount of participants from each state, and choosing states with regions that are similar in population sizes, may eliminate any discrepancies.

Lastly, the third limitation centers around the eighth grade reading scores for each state within the selected regions. These reading scores do not consider socio-economic status of students, and how resources available, or lack thereof, can impact student testing. A more thorough selection of participants that paint an accurate picture of students from various backgrounds might help to frame the narrative for future research, as it pertains to educational access and attainment within today's society.

### References

- Chamberlain, A., Daniels, C., Madden, N. A., & Slavin, R. E. (2007). A randomized evaluation of the success for all middle school reading program. *Middle Grades Research Journal*, 2(1), 1-21.
- Common Core of Data (CCD). (n.d.). Retrieved October 16, 2016, from <http://nces.ed.gov/ccd/>
- NAEP Nations Report Card - More About the NAEP Reading Assessment. (n.d.). Retrieved October 16, 2016, from <http://nces.ed.gov/nationsreportcard/reading/moreabout.aspx>
- NAEP - Overview. (n.d.). Retrieved October 16, 2016, from <http://nces.ed.gov/nationsreportcard/about/>
- National Center for Education Statistics (2006a). *National Assessment of Educational Progress: The Nation's Report Card*. Retrieved November 20, 2016, from <http://nces.ed.gov/nationsreportcard/>
- National Center for Education Statistics (2006b). *State Education Data Profiles*. Retrieved November 20, 2016, from <http://nces.ed.gov/programs/stateprofiles/>
- Schaefer, M. B., Malu, K. F., & Yoon, B. (2016). An historical overview of the middle school movement, 1963-2015. *Research In Middle Level Education Online*, 39(5), 1-27.



APPENDICES

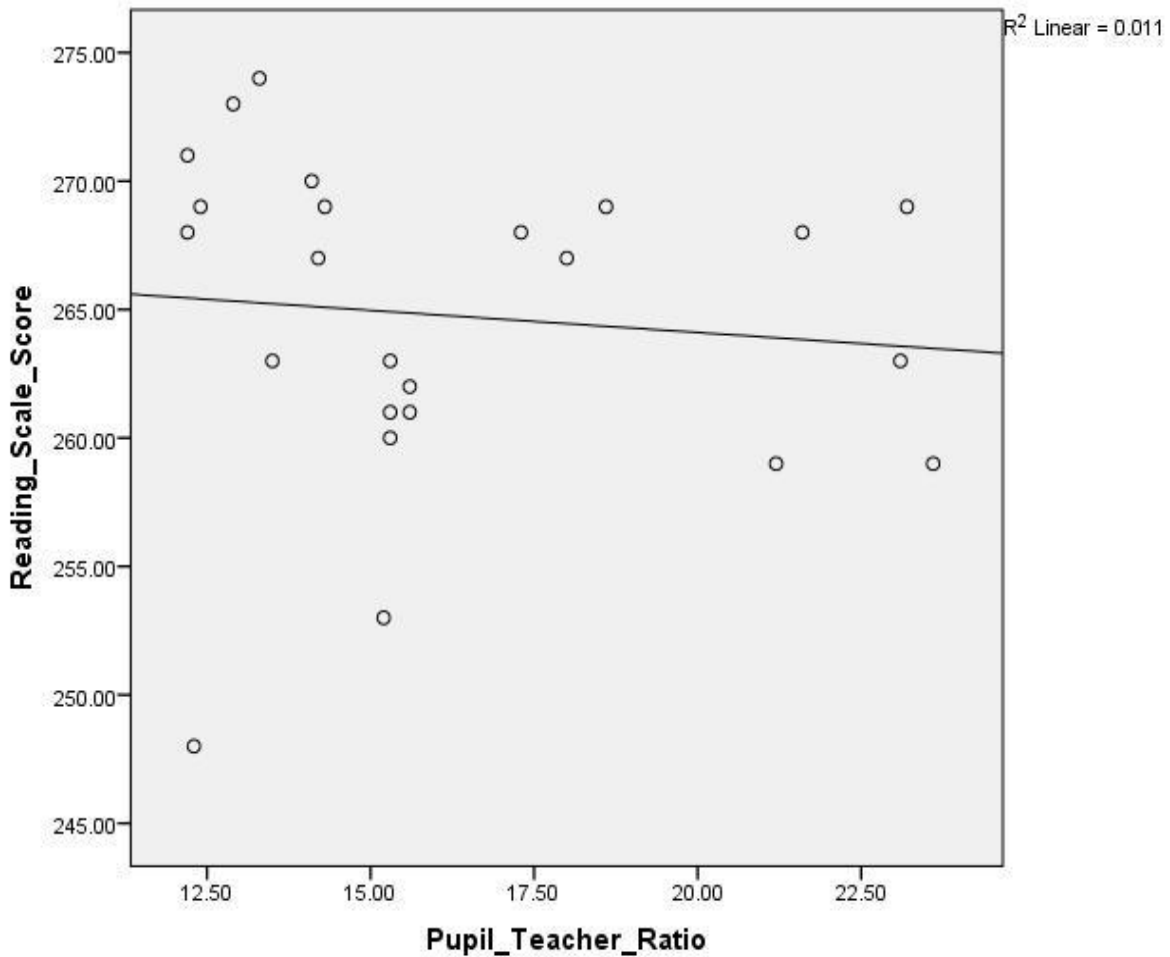
APPENDIX A:

**Correlations**



**Correlations**

		Reading_Scale_ Score	Pupil_Teacher_ Ratio
Reading_Scale_Score	Pearson Correlation	1	-.103
	Sig. (1-tailed)		.315
	N	24	24
Pupil_Teacher_Ratio	Pearson Correlation	-.103	1
	Sig. (1-tailed)	.315	
	N	24	24



APPENDIX B:

**T-Test**

**Group Statistics**

	Region	N	Mean	Std. Deviation	Std. Error Mean
Reading_Scale_Score	East	12	264.9167	7.11539	2.05404
	West	12	264.5833	5.46823	1.57854

**Independent Samples Test**

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Reading_Scale_Score	Equal variances assumed	.281	.601	.129	22	.899	.33333	2.59053	-5.03910	5.70577
	Equal variances not assumed			.129	20.633	.899	.33333	2.59053	-5.05981	5.72648

**Independent Samples Test**

		t-test for Equality of Means			
		Std. Error Difference	95% Confidence Interval of the Difference		
			Lower	Upper	
Reading_Scale_Score	Equal variances assumed	2.59053	-5.03910	5.70577	
	Equal variances not assumed	2.59053	-5.05981	5.72648	