EXAMINING SECONDARY STUDENT ACHIEVEMENT IN LARGE AND SMALL HIGH SCHOOLS IN VIRGINIA

MICHAEL J. BROWN
Radford City Public Schools, Virginia

GLEN I. EARTHMAN
Virginia Polytechnic Institute and State University

ABSTRACT

The purpose of this study was to examine large and small high schools in Virginia to try to understand if the high school student population size influenced the student achievement of eleventh grade students based on identified predictor variables. From the literature review, the main research question, five guiding questions, and a methodology were developed that would best aid in the analysis of the data. Data were collected from the Virginia Department of Education for the 2012-2013 school year that consisted of eleventh grade Virginia Standards of Learning assessments, socioeconomic status, student attendance, minority population, and teacher quality. Hierarchical multiple regression was the statistical method used to analyze the data for the research questions.

The results of the study indicate there is a significant relationship between socioeconomic status, student attendance, minority status and student achievement. However, when student population size was introduced, the result for socio-economic status was not significant. The overall conclusion regarding socioeconomic status and student achievement is that the issue is not rooted in the size of a high school population, but in the school as a whole.

Overall, results of the study indicate that there is a relationship between a high school student population size and student achievement when statistically controlling for selected variables. From the data derived in this study, students in large high schools academically perform better than small high schools in Virginia when selected variables are controlled.

INTRODUCTION

The average number of school buildings per district has fallen in the United States since the 1940's (Lawrence, 2004, p. 41). Meanwhile, student population in the United States has increased steadily since 1985 and is projected to set new records every year from 2012 to 2021 (U.S. Census Bureau, 2010). The increase in student population since the 1940's has led to larger schools throughout the United States (Lawrence, 2004, p. 41). With an increase in student population size, educational administrators encounter numerous challenges in regards to providing optimal learning conditions and student achievement (Stewart, 2009, p. 20).

Educational administrators are tasked with planning for and facilitating appropriate means in which to ensure success for the students they serve (Brown, Finch, & MacGregor, 2012, p. 207). Students' academic achievement as it relates to the size of a school population has been debated and studied by scholars for a number of years. A review of relevant literature notes some of the benefits and detriments of school population size as it relates to student achievement, curriculum, and cost–

efficiency (Fowler & Walberg, 1991; Lee & Smith, 1997; Leithwood & Jantzi, 2009; Lindahl & Cain, Sr., 2012; Program Review & Investigations Committee, 2006; Roby, 2004; Stewart, 2009).

RESEARCH QUESTIONS

The main research question of this study consists of the following: What is the relationship between a high school student population size and student achievement when statistically controlling for selected variables? To examine the relationship that a high school student population may have with student achievement, the following research sub-questions were utilized to guide this study:

1. What is the relationship between high school student population size and student achievement as measured by student performance on the Virginia Standards of Learning (SOL) in English reading and writing, U.S. History, Chemistry, and Algebra II assessments when socioeconomic status, student attendance, minority population, and teacher quality are statistically controlled?
2. What is the relationship between high school student population size and student achievement as measured by student performance on the Virginia SOL English reading and writing, U.S. History, Chemistry, and Algebra II assessments when only socioeconomic status is statistically controlled?
3. What is the relationship between high school student population size and student achievement as measured by student performance on the Virginia SOL English ng and writing, U.S. History, Chemistry, and Algebra II assessments when only student attendance is statistically controlled?
4. What is the relationship between high school student population size and student achievement as measured by student performance on the Virginia SOL English reading and writing, U.S. History, Chemistry, and Algebra II assessments when only minority population is statistically controlled?
5. What is the relationship between high school student population size and student achievement as measured by student performance on the Virginia SOL English reading and writing, U.S. History, Chemistry, and Algebra II assessments when only teacher quality is statistically controlled?

SIGNIFICANCE OF THE STUDY

School divisions in Virginia face circumstances that are beyond their control, such as available land area upon which to build school facilities, year-to-year budget fluctuations, student population growth or decline, and changes to standardized testing requirements. Issues that are a constant concern for school divisions include student achievement and accountability. To assist school divisions and superintendents in the constant planning of improving student achievement, relevant data needs to be examined while extracting strategies that are based on sound research methodologies and results. This study helps school divisions in Virginia by providing data that may potentially assist them with formulating strategies in regards to student achievement and the optimal student population size of buildings based on local conditions.

POPULATION

The population for this study consists of formal organizations designated as public secondary high schools that are governed by rules and regulations specified by the Virginia Department of Education. For the 2012-2013 school year, there were 309 public secondary schools in Virginia. Of the 309 public secondary schools in Virginia, only 102 met the defined criteria of this study, which was size of school – 51 large high schools and 51 small high schools. Of the 102 schools identified for this study, 51 met the criteria for large high schools and 51 met the criteria for small high schools

The population for this study is further broken down into high schools that consist of large and small student populations. The largest high school in the study is T.C. Williams with 2,906 students). The smallest high school in the study is Kecoughtan with 1,868 students. The largest high school in the small high schools group is Washington & Lee with 473 students. The smallest high school in Conference 1 is Highland with 67 students.

DATA COLLECTION

All data for this study were obtained and collected from the Virginia Department of Education, either through the utilization of current online data or by e-mail request through appropriate personnel employed by the Virginia Department of Education. Student information came from the Virginia Department of Education statistics and reports, data for researchers, Fall Membership Report. The Fall Membership Report is submitted by local school divisions to the Virginia Department of Education on September 30 of each school year. The fall membership count is limited to one unduplicated active record per student that attends public schools that enroll students.

The variables for this study were broken down into segments (End-of-Course Reading and Writing, U.S. History, Chemistry, and Algebra II SOL assessments, student attendance, socioeconomic status, minority population, and highly qualified teachers) that included the 2012-2013 student cohort who were enrolled in the eleventh grade at the time data were collected.

Standards of Learning (SOL) assessment was defined as the minimum grade level and subject matter educational objectives, described as knowledge and skills necessary for success in school and for the preparation for life that students were expected to meet in Virginia public schools and specified by the Standards of Quality (Virginia Department of Education, 2005).

Student attendance was defined by the percentage of attendance that equaled the average daily attendance divided by the average daily membership for each of the identified large and small high schools in this study.

Socioeconomic status was defined as the percent of students who were eligible for any of the following services: free or reduced price meals, Temporary Assistance for Needy Families, Medicaid, or identified as either migrant or experiencing homelessness. For this study, socioeconomic status was used as a measure of economic advantage or disadvantage within a family structure. A student that received the above-mentioned services fell below the federal poverty guidelines and was considered economically disadvantaged (U.S. Department of Education, 2012).

Minority population was defined via the Federal Race Code, and designated the racial categories that most clearly reflected the student's recognition of his or her community or with which the student most closely identified (U.S. Department of Education, 2012).

Teacher quality was defined by using the criteria for a highly qualified teacher as set forth by the federal government. The federal government considered a teacher highly qualified who had obtained full state certification as a teacher or passed the state teacher-licensing exam, held a license to teach in the state; and did not have certification or licensure requirements waived on an emergency, temporary, or provisional basis. A highly qualified teacher held a minimum of a bachelor's degree and demonstrated subject matter competency in each of the academic subjects in which he or she taught in a manner determined by the state and in compliance with Section 9101(23) of No Child Left Behind (U.S. Department of Education, 2003).

All the data for this study were collected for each formal organization designated as public secondary high school that met the specified school criteria. The data were managed using a Microsoft Excel spreadsheet. All the data for this study were examined thoroughly after entering it into the Microsoft Excel spreadsheet to make sure that it was correct. Once the data were entered into the Microsoft Excel spreadsheet, it was separated into large and small schools. The data then were imported into the latest version of IBM SPSS, a statistical analysis software program. Once the data were entered into IBM SPSS, it was again examined to identify any outliers or other abnormalities that potentially could have affected the proper statistical analysis of the data set.

DATA PROCESSING

After verifying the data were complete and correct in IBM SPSS, a regression analysis was performed. The reason why a regression analysis was performed on the data set is that it was the statistical analysis that has been identified to find relationships among a set of variables. The specific type of regression analysis to be performed on the data set was sequential multiple regression with an alpha level set at 0.05. While regression analysis allowed examination of relationships, sequential multiple regression allowed a closer examination of the dependent variable and a predictor variable while statistically controlling for the effects of various other independent variables on the dependent variable

The first research question was addressed by performing a sequential multiple regression analysis. The sequential multiple regression analysis used the mean scaled scores for the End-of-Course Reading and Writing, U.S. History, Chemistry, and Algebra II assessments for eleventh grade students as the dependent variable and school population size as the primary predictor variable while statistically controlling for socioeconomic status, student attendance, minority population, and teacher quality.

The second research question was addressed by performing a sequential multiple regression analysis using the mean scaled scores for the End-of-Course Reading and Writing, U.S. History, Chemistry, and Algebra II assessments for eleventh grade students as the dependent variable and school population size as the primary predictor variable while only statistically controlling for socioeconomic status.

The third research question was addressed by performing a sequential multiple regression analysis using the mean scores for the End-of-Course Reading and Writing, U.S. History, Chemistry, and Algebra II assessments for eleventh grade students as the dependent variable and school

population size as the primary predictor variable while only statistically controlling for student attendance.

The fourth research question was addressed by performing a sequential multiple regression analysis using the mean scores for the End-of-Course Reading and Writing, U.S. History, Chemistry, and Algebra II assessments for eleventh grade students as the dependent variable and school population size as the primary predictor variable while only statistically controlling for minority population.

The fifth research question was addressed by performing a sequential multiple regression analysis using the mean scores for the End-of-Course Reading and Writing, U.S. History, Chemistry, and Algebra II assessments for eleventh grade students as the dependent variable and school population size as the primary predictor variable while only statistically controlling for teacher quality.

When the results were determined on the five research questions, further statistical analysis was performed, specifically collinearity analysis and diagnostics using IBM SPSS. The collinearity analysis and diagnostics using IBM SPSS was performed to test for multicollinearity. Multicollinearity had the possibility of existing when two or more predictor variables in a regression model were closely related. Additionally, it occurred when one variable was predicted from one or more of the other predictor variables. The purpose of a regression model was to test each predictor variable and analyze what effect it had on the dependent variable, and establish a ranking order to see which one had the most effect on the dependent variable. Collinearity statistics help sort out the variables predicting power and help reduce the issues found if redundant variables were present.

In order to examine student achievement properly, it was determined that all the Standards of Learning scores identified for the study (English reading and writing, U.S. History, Chemistry, and Algebra II) would be combined into a composite score. The composite score (coded as Total Score) for large high schools ranged from a minimum of 2054.78 to 2334.52 with a standard deviation of 59.74 (Table 4.1). Also presented in Table 4.1 were the scores for each of the Standards of Learning assessments identified for use in the study (coded as Reading, Writing, U.S. History, Chemistry, and Alg II). Although a composite score was used to gauge student achievement, the consideration of each Standard of Learning assessment identified for the study aided in the understanding of the composite score (Table 4.1).

The other variables that were examined as a part of the study included socioeconomic status, student attendance, student minority population, and teacher quality. Socioeconomic status (coded as SES) is the percentage of disadvantaged students, which ranged from 4.74% to 60.60% with a mean of 23.48% and a standard deviation of 13.86% (Table 4.1). Student attendance (coded as Attendance) is the percentage of attendance, which ranged from 88.43% to 97.03% with a mean of 94.44% with a standard deviation of 1.55% (Table 4.1). For the variable of minority population, it was determined to sum the entire minority population subgroups identified for this study with the result being the percent of minority students that would be used for the data analysis portion of the study. Minority population (coded as Minority) ranged from 18.20% to 82.98% with a mean of 51.47% and a standard deviation of 16.66% (Table 4.1). Teacher quality (coded as HQ Teacher) for this study ranged from 94.99% to 100% with a mean of 98.22% and a standard deviation of 1.69% (Table 4.1). The range of students that were in the eleventh grade during the 2012-2013 school year ranged from 378 to 798 with a mean of 549.46 with a standard deviation of 92.93 (Table 4.1).

Descriptive statistics were presented in this section for small high schools in Virginia for each of the variables used in this study. The total number of high schools in the data set that met the criteria for small high schools in Virginia was 51. Upon examination of the data, it was determined that one small high school was missing data for socioeconomic status and two small high schools were missing data for Algebra II scores. Since missing data could skew results when a data analysis was performed, those three small high schools were excluded from the study. The removal of three small high schools from the small high school data set brought the total of small high schools to 48.

The composite score (coded as Total Score) for small high schools ranged from a minimum of 2054.88 to 2326.55 with a mean score of 2158.62 and a standard deviation of 49.21 (Table 4.2). Also presented in Table 4.2 are the scores for each of the Standards of Learning assessments identified for use in the study (coded as Reading, Writing, U.S. History, Chemistry, and Alg II). Although a composite score was used to gauge student achievement, the consideration of each Standard of Learning assessment identified for the study aided in the understanding of the composite score (Table 4.2).

The other variables that were examined as a part of the study included socioeconomic status, student attendance, student minority population, and teacher quality. Socioeconomic status (coded as SES) is the percentage of disadvantaged students, which ranged from 14.29% to 68.33% with a mean of 46.89% and a standard deviation of 12.25% (Table 4.2). Student attendance (coded as Attendance) is the percentage of attendance, which ranged from 87.13% to 97.81% with a mean of 93.68% with a standard deviation of 1.95% (Table 4.2). For the variable of minority population, it was determined to sum the entire minority population subgroups identified for this study with the result being the percent minority that would be used for the data analysis portion of the study. Minority population (coded as Minority) ranged from 0% to 77.65% with a mean of 23.12% and a standard deviation of 23.82% (Table 4.2). Teacher quality (coded as HQ Teacher) for this study ranged from 84% to 100% with a mean of 96.84% and a standard deviation of 4.11% (Table 4.2). The range of students that were in the eleventh grade during the 2012-2013 school year ranged from 23 to 120 with a mean of 77.97 with a standard deviation of 25.18 (Table 4.2).

DATA ANALYSIS

For the data analyses portion of this study, hierarchical multiple regression were used to investigate the five research questions. In this study, two models were employed for each research question. In the first model, the predictor variable(s) were entered to determine if they had a significant effect on the dependent variable, which in this case was student achievement. In the second model for each research question, the predictor variable(s) were entered to gauge what, if any, effect each had on the dependent variable along with student population size. For this study, a level of 0.05 was used to test for significance.

Preliminary analyses were conducted via the correlation table (Table 4.3) to explore for potential relationships between the predictor variables and the dependent variable of student achievement. Two predictor variables were found to be significant. The results of the correlation analysis indicated that the predictor variables of socioeconomic status and student attendance were statistically correlated with the dependent variable of student achievement in this preliminary examination. The correlations between the predictor variables of socioeconomic status and student attendance and the dependent variable of student achievement ranged from moderately positive to

strongly negative $r = .325$, $p < .05$ to $r = -.588$, $p < .05$. Of the two-predictor variables that were found to be significant, socioeconomic status was found to be the best predictor of student achievement.

Preliminary analyses were conducted in each research question to ensure that no violations of the assumptions of normality, linearity, or homoscedasticity existed through the examination of Normal P-Plots and Scatterplots. A preliminary analysis was conducted to check for multicollinearity. This was accomplished through the examination of the correlation table (Table 4.3). Two predictor variables were found to be significantly correlated with each other. They were socioeconomic status and student attendance, which were moderately negative $r = -.422$, $p < .05$. Further analysis was conducted to determine if multicollinearity was a problem in each of the five research questions through the examination of the VIF and Tolerance results for each of the predictor variables.

Data Analyses for Research Question 1

What was the relationship between high school student population size and student achievement as measured by student performance on the Virginia Standards of Learning (SOL) in English reading and writing, U.S. History, Chemistry, and Algebra II assessments when socioeconomic status, student attendance, minority population, and teacher quality were statistically controlled?

Hierarchical multiple regression was performed to investigate the relationship between high school student population size and student achievement, after statistically controlling for socioeconomic status, student attendance, minority population, and teacher quality. Preliminary analyses were conducted to ensure that no violations of the assumptions of normality, linearity, or homoscedasticity existed. This was conducted by the creation of a Normal P-Plot and a Scatterplot. After examination of the Normal P-Plot and Scatterplot, no abnormalities were found. An analysis was also conducted to determine if multicollinearity was present. This was accomplished by examining the collinearity statistics located in the coefficients table (Table 4.4). Specifically, the VIF and Tolerance results were examined. The VIF results ranged from 1.02 to 3.33. VIF results greater than 10 start to indicate relatively high levels of multicollinearity. The results for Tolerance ranged from .981 to .300. A Tolerance result of .10 or lower indicates multicollinearity may be a problem. The results of collinearity statistics examination revealed that collinearity was not a problem for the regression models that were obtained.

In model one of the hierarchical multiple regression, all the predictor variables (socioeconomic status, student attendance, minority population, and teacher quality) were entered. This model was statistically significant $F(4, 93) = 15.52$; $p < .05$ and explained 40% of the variance in student achievement (Table 4.5). After entry of school size in model two, the total variance explained by the model as a whole was 41%. The overall F test for this model was significant ($F(5, 92) = 12.94$; $p < .05$) (Table 4.6). The introduction of school size did not add anything significant to the explanation of the variance in student achievement after controlling for the predictor variables (R^2 Change = .013; $F(1, 92) = 1.97$; $p > .05$) and was not significant (Table 4.5).

Data Analyses for Research Question 2

What was the relationship between high school student population size and student achievement as measured by student performance on the Virginia SOL English reading and writing, U.S. History, Chemistry, and Algebra II assessments when only socioeconomic status was statistically controlled?

Hierarchical multiple regression was performed to investigate the relationship between high school student population size and student achievement, after statistically controlling for socioeconomic status. Preliminary analyses were conducted to ensure that no violations of the assumptions of normality, linearity, or homoscedasticity existed. This was conducted by the creation of a Normal P-Plot and a Scatterplot. After examination of the Normal P-Plot and Scatterplot, no abnormalities were found. An analysis was conducted to determine if multicollinearity was present. This was accomplished by examining the collinearity statistics located in the coefficients table (Table 4.7). Specifically, the VIF and Tolerance results were examined. The VIF result was 1.81. VIF results greater than 10 start to indicate relatively high levels of multicollinearity. The result for Tolerance was .551. A Tolerance result of .10 or lower indicates multicollinearity may be a problem. The results of collinearity statistics examination revealed that collinearity was not a problem for the regression models that were obtained.

In model one of the hierarchical multiple regression, the predictor variable of socioeconomic status was entered. This model was statistically significant F (1, 96) = 50.67; $p < .05$ and explained 34% of the variance in student achievement (Table 4.8). After entry of school size in model two, the total variance explained by the model as a whole remained at 34%. The overall F test for this model was significant (F (2, 95) = 25.24; $p < .05$) (See Table 4.9). The introduction of school size did not add anything significant to the explanation of the variance in student achievement after controlling for socioeconomic status (R^2 Change = .002; F (1, 95) = .229; $p > .05$) and was not significant (Table 4.8).

Data Analyses for Research Question 3

What was the relationship between high school student population size and student achievement as measured by student performance on the Virginia SOL English reading and writing, U.S. History, Chemistry, and Algebra II assessments when only student attendance was statistically controlled?

Hierarchical multiple regression was performed to investigate the relationship between high school student population size and student achievement, after statistically controlling for student attendance. Preliminary analyses were conducted to ensure that no violations of the assumptions of normality, linearity, or homoscedasticity existed. This was conducted by the creation of a Normal P-Plot and a Scatterplot. After examination of the Normal P-Plot and Scatterplot, no abnormalities were found. An analysis was conducted to determine if multicollinearity was present. This was accomplished by examining the collinearity statistics located in the coefficients table (Table 4.10). Specifically, the VIF and Tolerance results were examined. The VIF result was 1.06. VIF results greater than 10 start to indicate relatively high levels of multicollinearity. The result for Tolerance was .942. A Tolerance result of .10 or lower indicates multicollinearity may be a problem. The results of collinearity statistics examination revealed that collinearity was not a problem for the regression models that were obtained.

In model one of the hierarchical multiple regression, the predictor variable of student attendance was entered. This model was statistically significant F (1, 96) = 11.36; $p < .05$ and explained 10% of the variance in student achievement (Table 4.11). After entry of school size in model two, the total variance explained by the model increased to 19%. The overall F test for this model was significant (F (2, 95) = 11.33; $p < .05$) (Table 4.12). The introduction of school size explained an additional 9% of the variance in student achievement after controlling for student attendance (R^2 Change = .087; F (1, 95) = 10.20; $p < .05$) and was significant (See Table 4.11).

Data Analyses for Research Question 4

What was the relationship between high school student population size and student achievement as measured by student performance on the Virginia SOL English reading and writing, U.S. History, Chemistry, and Algebra II assessments when only minority population was statistically controlled?

Hierarchical multiple regression was performed to investigate the relationship between high school student population size and student achievement, after statistically controlling for minority population. Preliminary analyses were conducted to ensure that no violations of the assumptions of normality, linearity, or homoscedasticity existed. This was conducted by the creation of a Normal P-Plot and a Scatterplot. After examination of the Normal P-Plot and Scatterplot, no abnormalities were found. An analysis was conducted to determine if multicollinearity was present. This was accomplished by examining the collinearity statistics located in the coefficients table (Table 4.13). Specifically, the VIF and Tolerance results were examined. The VIF result was 1.48. VIF results greater than 10 start to indicate relatively high levels of multicollinearity. The result for Tolerance was .672. A Tolerance result of .10 or lower indicates multicollinearity may be a problem. The results of collinearity statistics examination revealed that collinearity was not a problem for the regression models that were obtained.

In model one of the hierarchical multiple regression, the predictor variable of minority population was entered. This model was found not to be statistically significant F (1, 96) = 1.51; $p > .05$ and only explained 1% of the variance in student achievement (Table 4.14). After entry of school size in model two, the total variance explained by the model increased to 29%. The overall F test for this model was significant (F (2, 95) = 20.16; $p < .05$) (Table 4.15). The introduction of school size explained an additional 28% of the variance in student achievement after controlling for minority population (R^2 Change = .283; F (1, 95) = 38.23; $p < .05$) and was significant (Table 4.14).

Data Analyses for Research Question 5

What was the relationship between high school student population size and student achievement as measured by student performance on the Virginia SOL English reading and writing, U.S. History, Chemistry, and Algebra II assessments when only teacher quality was statistically controlled?

Hierarchical multiple regression was performed to investigate the relationship between high school student population size and student achievement after statistically controlling for teacher quality. Preliminary analyses were conducted to ensure that no violations of the assumptions of normality, linearity, or homoscedasticity existed. This was conducted by the creation of a Normal P-Plot and a Scatterplot. After examination of the Normal P-Plot and Scatterplot, no abnormalities were found. An analysis was conducted to determine if multicollinearity was present. This was accomplished by examining the collinearity statistics located in the coefficients table (Table 4.16). Specifically, the VIF and Tolerance results were examined. The VIF result was 1.05. VIF results greater than 10 start to indicate relatively high levels of multicollinearity. The result for Tolerance was .952. A Tolerance result of .10 or lower indicates multicollinearity may be a problem. The results of collinearity statistics examination revealed that collinearity was not a problem for the regression models that were obtained.

In model one of the hierarchical multiple regression, the predictor variable of teacher quality was entered. This model was found not to be statistically significant F (1, 96) = 3.84; $p > .05$

and only explained 3% of the variance in student achievement (Table 4.17). After entry of school size in model two, the total variance explained by the model increased to 14%. The overall F test for this model was significant (F (2, 95) = 8.18; p <.05) (Table 4.18). The introduction of school size explained an additional 11% of the variance in student achievement after controlling for teacher quality (R^2 Change = .109; F (1, 95) = 12.08; p < .05) and was significant (Table 4.17).

CONCLUSION

The answer to the main research question of is there a relationship between a high school student population size and student achievement when statistically controlling for selected variables, the answer is yes according to the data. Where high school student population size influenced student achievement the most was when the predictor variables of student attendance, minority population, and teacher quality were taken into account. Out of the four variables that were analyzed in the second model of the hierarchical multiple regression, only socioeconomic status was found not to have a relationship with high school population size and student achievement. However, it should be noted that even though no relationship existed between socioeconomic status, student achievement, and high school population size, it does not mean that it should be excluded from consideration.

DISCUSSION

In research question one, the predictor variables of socioeconomic status, student attendance, minority population, and teacher quality were included in the hierarchical multiple regression. In model one, when all of the variables were accounted for, the results indicated that there is a correlation amongst all the variables and student achievement and were found to be statistically significant F (4, 93) = 15.52; p < .05. In model two, when all of the predictor variables were accounted for, they were found not to have a relationship with high school population size and student achievement (Table 4.5). This finding corresponds with two studies that examined similar multiple variables when school size and student achievement are taken into account (Lamdin 1996; Lindahl & Cain, Sr., 2012).

Although no relationship existed in model two of research question one, when all of the predictor variables were accounted for between student achievement and high school population size, model one should not be discounted entirely. The reason why the results for research question one, model one should not be discounted is the fact that school leaders and administrators have to deal with socioeconomic status, student attendance, minority population, and teacher quality in some form or fashion on a daily basis, no matter the size of the school or school division. The fact that these predictor variables were correlated with student achievement is an indication that they are important and should be considered when school leaders and administrators are looking at ways to make overall improvements to their school divisions. For this reason, the predictor variables were separated out for research questions two through five to investigate the possibility that when student population size is factored in, the predictor variables individually may have a bearing on student achievement.

For research question two, socioeconomic status was examined to determine if a relationship existed between student achievement and school population size. The first model of the hierarchical multiple regression produced a significant result. This indicates that there is a relationship between socioeconomic status and student achievement. However, when student

population size was introduced in model two of the hierarchical multiple regression, the result was not significant. This finding is surprising, since a number of studies included in the literature review found that socioeconomic status was a predictor of student achievement when student population size were considered (Fowler & Walberg, 1991; Leithwood & Jantzi, 2009; Program Review & Investigations; Stewart, 2009; Werblow & Duesberry, 2009). A closer look at the results of research question two would indicate that while no relationship existed between socioeconomic status, student achievement, and high school population size, the data tell a different story. In model one of the hierarchical multiple regression, the result was significant F (1, 96) = 50.67; $p < .05$ (See Table 4.8). This indicates that a relationship does exist between socioeconomic status and student achievement. The overall inference in regards to socioeconomic status and student achievement is that the issue is not rooted in the size of a high school population, but in the overall school population as a whole. Ready (2010) stated that "… the least disputed conclusion to emerge from educational research over the past half-century is that socioeconomically disadvantaged children are less likely to experience school success" (p. 271).

In research question three, model two, a relationship was shown to exist between high school population size and student achievement after controlling for student attendance (Table 4.11). This implies that no matter the size of a high school student population, the more frequently and regularly students attend school, the better they will perform academically. The finding that there is a relationship between student attendance, high school population size, and student achievement are similar with studies performed by other researchers (Lamdin, 1996; Roby, 2004; Werblow & Duesbery 2009). The part of the data analysis that caused alarm was when the Beta levels in model two were examined. It revealed that smaller high schools have a lower rate of student achievement on the Virginia Standards of Learning assessments than large high schools. This finding is in contrast to meta-analysis performed by Leithwood and Jantzi (2009), which found in terms of truancy and attendance, smaller secondary schools fare better in terms of academics than larger secondary schools.

The reason why smaller high schools have a lower rate of student achievement on the Virginia Standards of Learning assessments could be a result of the small high school population size that was examined as part of this study. The size range for small high schools was between 473 students and 67 students. In addition, the percentage of small high schools socioeconomically disadvantaged student population was double that of large high schools (Large – M = 23.48 vs Small – M = 46.89). This could explain why students do not perform as well academically in smaller high schools in Virginia. What this indicates is those students that attend small high schools in Virginia are experiencing economic difficulties, and those difficulties are affecting student achievement. School leaders and administrators need to be aware of this and seek to put measures into place that address this problem.

In regards to research question four, model two, the data indicated that there is a relationship between high school population size and student achievement after controlling for the percentage of minority students. Overall, when the size of a high school student population, whether it *was* large or small, was taken into account, student achievement suffered when a school had a larger percentage of minority students. This finding is similar to the study conducted by Lamdin (1996). Lamdin (1996) concluded that schools with a larger minority enrollment would have lower mathematics scores (p. 160). Additionally, another similar finding was presented in the study commissioned by the Kentucky Legislative Research Commission (Program Review and Investigations Committee, 2006). In that study, when race was factored in and analyzed, the results indicated that minority

students did not do as well as Caucasian students on standardized assessments despite the student population size (Program Review and Investigations Committee, 2006).

In this study, large high schools in Virginia (M= 51.47) had a higher percentage of minority students than small high schools in Virginia (M= 23.12). In theory, it could be reasoned that small high schools in Virginia will have a higher percentage of student achievement when the percentage of minority population is taken into account. This is in contrast to Werblow and Duesbery (2009) who found that schools with both large and small student populations showed some gains in terms of student achievement (p.18-21).

In regards to minority population, school leaders and administrators cannot control the level of minority populations in their schools. What they can and do control is the way in which they strive to reach students despite their minority status. Minority students in school today tend to be economically disadvantaged.

The results for the fifth research question, model two, indicated that there is a relationship between high school student population size and student achievement when controlling for teacher quality (Table 4.17). This finding indicates that high schools in Virginia with student populations between 2,906 and 1,868 tend to have a higher percentage of highly qualified teachers and thus students perform better on the Virginia Standards of Learning assessments. This finding is supported by an observation of the Beta level performed in model two of the hierarchical multiple regression for teacher quality ($ß = -.388$, $p < .05$). This finding is also similar to Lindahl and Cain, Sr. (2012) results. In their study, Lindahl and Cain, Sr. (2012) found that as the size of the high school population increased, so did the scores on both the reading and math assessments examined in their study when teacher quality was taken into account (p.10).

In addition, a further indication of the Beta level reveals that as the size of the high school student population decreases when controlling for teacher quality, student achievement on the Virginia Standards of Learning assessments decreases. Essentially, what this observation indicates is that there is a probability that in smaller high schools in Virginia, teacher quality is not equivalent to larger high schools. Most small schools in Virginia are located in rural areas that for the most part struggle to recruit and retain highly qualified teachers. The reasons for this are varied, but one could speculate that in rural school divisions, the salary scales are not equal to larger schools divisions. From data derived in this study, students in large high schools do perform academically better than students in small high schools in Virginia.

REFERENCES

Brown, P. D., Finch, K. S., & MacGregor, C. (2012). A comparison of learning cultures in different sizes and types. *US-China Education Review A: Education Practice, 2(2)*, 206-222

Fowler, W. J., & Walberg, H. J. (1991). School size, characteristics, and outcomes. *Educational Evaluation and Policy Analysis*, *13*(2), 189-202. Retrieved from http://www.jstor.org.ez-proxy.lib.vt.edu:8080/stable/pdfplus/1164583.pdf?acceptTC=true

Lamdin, D. J. (1996). Evidence of student attendance as an independent variable in education production functions. *The Journal of Educational Research*, *89*(3), 155-162.

Lawrence, B. K. (2004). For authentic accountability think small. *Journal Of Education*, *185*(3), 41-50.

Lee, V. E., & Smith, J. B. (1997). High school size: Which works best and for whom. *Educational Evaluation and Policy Analysis*, *19*(3), 205-227. Retrieved from http://www.jstor.org.ezproxy.lib.vt.edu:8080/stable/pdfplus/1164463.pdf?acceptTC=true&acceptTC=true&jpdConfirm=true

Leithwood, K., & Jantzi, D. (2009). A review of empirical evidence about school size effects: A policy perspective. *Review of Educational Research, 79(1),* 464-490. doi: 10.3102/0034654308326158

Lindahl, R. A., & Cain, Sr., P. M. (2012). A study of school size among Alabama's public high schools. *International Journal of Education Policy & Leadership*, *7*(1), 1-27.

Program Review and Investigations Committee. (2006). *School size and student outcomes in Kentucky's public schools* (Research Report No. 334). Retrieved from Legislative Research Commission website: http://www.lrc.ky.gov/lrcpubs/RR334.pdf

Ready, D. D. (2010). Socioeconomic disadvantage, school attendance, and early cognitive development the differential effects of school exposure. *Sociology of Education*, *83*(4), 271-286.

Roby, D. (2004). Research on school attendance and student achievement: A study of Ohio Schools. *Educational Research Quarterly, 28.1*, 3-14.

Stewart, L. (2009). Achievement differences between large and small schools in Texas. *The Rural Educator, 30(2),* 20-28. Retrieved from http://web.ebscohost.com.ezproxy.lib.vt.edu:8080/ehost/pdfviewer/pdfviewer?sid=33b37d9a-c9fd-4f34-91aa-85bb2ecb3025@sessionmgr104&vid=9&hid=125

U.S. Census Bureau. (2010). [Virginia-2010 census results total population by county]. Retrieved from website: http://www.census.gov/geo/maps-data/maps/pdfs/2010pop/va_totalpop_2010map.pdf

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2012). *Digest of education statistics:2011* (NCES 2012-001). Retrieved from website: http://nces.ed.gov/programs/digest/d11/ch_2.asp

U.S. Department of Education. U.S. Department of Education, Office of Vocational and Adult Education. (2003). *From there to here: The road to reform of American high schools*. Retrieved from U.S. Department of Education, Office of Vocational and Adult Education website: http://www2.ed.gov/about/offices/list/ovae/pi/hsinit

Virginia Department of Education. (2005). *Nine out of 10 Virginia schools now fully accredited: Two percent of schools to 92 percent in 8 years* [Press release]. Retrieved from Division of Policy and Communications website: http://www.pen.k12.va.us/VDOE/NewHome/pressreleases/2005/oct25.html

Werblow, J., & Duesbery, L. (2009). The impact of high school size on math achievement and dropout rate. *The High School Journal, 92(3),* 14-23. Retrieved from http://web.ebscohost.com.ezproxy.lib.vt.edu:8080/ehost/pdfviewer/pdfviewer?sid=33b37d9a-c9fd-4f34-91aa-85bb2ecb3025@sessionmgr104&vid=11&hid=125

APPENDICES

Table 4.1

Descriptive statistics for large high schools in Virginia for male and female students in the eleventh grade along with all variables used in study.

Coded As	N	Minimum	Maximum	Mean	Std. Deviation
Total Score	50	2054.78	2334.52	2201.1050	59.74209
Reading	50	423.16	467.43	446.8704	9.99259
Writing	50	407.96	507.58	469.4276	22.17709
U.S. History	50	416.45	491.87	452.6428	15.97711
Chemistry	50	392.33	471.19	424.5052	18.57789
Alg II	50	359.37	437.48	407.6590	15.77058
SES	50	4.74	60.60	23.4802	13.86239
Attendance	50	88.43	97.03	94.5558	1.55760
Minority	50	18.20	82.98	51.4716	16.66726
HQ Teacher	50	94.99	100.00	98.2288	1.69807
All Students G-11	50	378.0	798.0	549.460	92.9306

a. Lg_1 _Sm_2 =1.0

Table 4.2

Descriptive statistics for small high schools in Virginia for male and female students in the eleventh grade along with all variables used in study

Descriptive Statistics

Coded As	N	Minimum	Maximum	Mean	Std. Deviation
Total Score	48	2054.88	2326.55	2158.6254	49.21666
Reading	48	412.09	469.42	433.7350	10.45427
Writing	48	388.75	491.00	446.0190	18.48310
U.S. History	48	399.56	486.94	438.9438	16.46570
Chemistry	48	405.29	469.96	436.3098	18.04803
Alg II	48	345.35	447.67	403.6179	24.08975
SES	48	14.29	68.33	46.8917	12.25975
Attendance	48	87.13	97.81	93.6889	1.95221
Minority	48	.00	77.65	23.1233	23.82642
HQ Teacher	48	84.00	100.00	96.8452	4.11286
Total_Students_Eleventh_Grade	48	23.0	120.0	77.979	25.1823

a. Lg_1 _Sm_2 = 2.0

Table 4.3

Correlations Between all Variables

		Total_ Score	SES	Attendance	Minority	HQ Teacher
Total_Score	Pearson Correlation	1	-.588**	.325**	-.125	.196
	Sig. (2-tailed)		.000	.001	.222	.053
	N	98	98	98	98	98
SES	Pearson Correlation	-.588**	1	-.442**	-.123	-.187
	Sig. (2-tailed)	.000		.000	.226	.065
	N	98	98	98	98	98
Attendance	Pearson Correlation	.325**	-.442**	1	-.004	-.130
	Sig. (2-tailed)	.001	.000		.968	.203
	N	98	98	98	98	98
Minority	Pearson Correlation	-.125	-.123	-.004	1	.039
	Sig. (2-tailed)	.222	.226	.968		.700
	N	98	98	98	98	98
HQ Teacher	Pearson Correlation	.196	-.187	-.130	.039	1
	Sig. (2-tailed)	.053	.065	.203	.700	
	N	98	98	98	98	98

**. Correlation is significant at the 0.01 level (2-tailed).
Total Score = Student Achievement, SES = Socioeconomic Status, Attendance = Student Attendance, Minority = Minority Population, HQ Teacher= Teacher Quality

Table 4.4

Analysis using the coefficients table to check for collinearity between high school student population size and student achievement, after statistically controlling for socioeconomic status, student attendance, minority population, and teacher quality

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1758.655	356.813		4.929	.000		
	SES	-1.827	.313	-.548	-5.830	.000	.731	1.368
	Attendance	3.164	2.997	.098	1.056	.294	.755	1.324
	Minority	-.462	.191	-.196	-2.419	.017	.981	1.020
	HQ Teacher	2.102	1.550	.114	1.356	.178	.909	1.100
2	(Constant)	1814.875	357.211		5.081	.000		
	SES	-1.419	.426	-.425	-3.328	.001	.391	2.558
	Attendance	3.226	2.982	.099	1.082	.282	.755	1.325
	Minority	-.702	.255	-.298	-2.748	.007	.543	1.841
	HQ Teacher	1.777	1.559	.097	1.140	.257	.889	1.125
	Lg_1_Sm_2	-23.902	17.017	-.205	-1.405	.164	.300	3.338

a. Dependent Variable: Total_Score

Table 4.5

Model Summary of Hierarchical Multiple Regression to investigate what is the relationship between a high school student population size and student achievement, after statistically controlling for all predictor variables.
Model Summary

					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	
1	.633[a]	.400	.375	46.32928	.400	15.527	4	93	.000	
2	.643[b]	.413	.381	46.08882	.013	1.973	1	92	.164	

a. Predictors: (Constant), HQ Teacher, Minority, Attendance, SES
b. Predictors: (Constant), HQ Teacher, Minority, Attendance, SES, Lg_1 _Sm_2
c. Dependent Variable: Total_Score

Table 4.6

ANOVA of Hierarchical Multiple Regression to investigate what is the relationship between a high school student population size and student achievement, after statistically controlling for all predictor variables.

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	133310.709	4	33327.677	15.527	.000[b]
	Residual	199615.385	93	2146.402		
	Total	332926.094	97			
2	Regression	137501.609	5	27500.322	12.946	.000[c]
	Residual	195424.485	92	2124.179		
	Total	332926.094	97			

a. Dependent Variable: Total_Score
b. Predictors: (Constant), HQ Teacher, Minority, Attendance, SES
c. Predictors: (Constant), HQ Teacher, Minority, Attendance, SES, Lg_1 _Sm_2

Table 4.7

Analysis using the coefficients table to check for collinearity between high school student population size and student achievement, after statistically controlling for socioeconomic status
Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	2248.839	10.764		208.921	.000		
	SES	-1.961	.276	-.588	-7.119	.000	1.000	1.000
2	(Constant)	2243.728	15.190		147.714	.000		
	SES	-2.081	.373	-.624	-5.584	.000	.551	1.814
	Lg_1 _Sm_2	6.235	13.021	.053	.479	.633	.551	1.814

a. Dependent Variable: Total_Score

Table 4.8

Model Summary of Hierarchical Multiple Regression to investigate what is the relationship between a high school student population size and student achievement, after statistically controlling for socioeconomic status

Model Summary

					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	
1	.588[a]	.345	.339	47.64256	.345	50.676	1	96	.000	
2	.589[b]	.347	.333	47.83495	.002	.229	1	95	.633	

a. Predictors: (Constant), SES

b. Predictors: (Constant), SES, Lg_1 _Sm_2

Table 4.9

ANOVA of Hierarchical Multiple Regression to investigate what is the relationship between a high school student population size and student achievement, after statistically controlling for socioeconomic status

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	115024.007	1	115024.007	50.676	.000[b]
	Residual	217902.087	96	2269.813		
	Total	332926.094	97			
2	Regression	115548.782	2	57774.391	25.249	.000[c]
	Residual	217377.312	95	2288.182		
	Total	332926.094	97			

a. Dependent Variable: Total_Score

b. Predictors: (Constant), SES

c. Predictors: (Constant), SES, Lg_1 _Sm_2

Table 4.10

Analysis using the coefficients table to check for collinearity between high school student population size and student achievement, after statistically controlling for student attendance

Coefficients

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	1186.734	294.731		4.027	.000		
	Attendance	10.555	3.130	.325	3.372	.001	1.000	1.000
2	(Constant)	1462.965	294.512		4.967	.000		
	Attendance	8.181	3.081	.252	2.655	.009	.942	1.062
	Lg_1 _Sm_2	-35.388	11.076	-.304	-3.195	.002	.942	1.062

a. Dependent Variable: Total_Score

Table 4.11

Model Summary of Hierarchical Multiple Regression to investigate what is the relationship between a high school student population size and student achievement, after statistically controlling for student attendance
Model Summary[c]

					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	
1	.325[a]	.106	.097	55.68468	.106	11.368	1	96	.001	
2	.439[b]	.193	.176	53.19188	.087	10.209	1	95	.002	

a. Predictors: (Constant), Attendance
b. Predictors: (Constant), Attendance, Lg_1 _Sm_2
c. Dependent Variable: Total_Score

Table 4.12

ANOVA of Hierarchical Multiple Regression to investigate what is the relationship between a high school student population size and student achievement, after statistically controlling for student attendance.
ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	35250.918	1	35250.918	11.368	.001[b]
	Residual	297675.175	96	3100.783		
	Total	332926.094	97			
2	Regression	64135.355	2	32067.678	11.334	.000[c]
	Residual	268790.738	95	2829.376		
	Total	332926.094	97			

a. Dependent Variable: Total_Score
b. Predictors: (Constant), Attendance
c. Predictors: (Constant), Attendance, Lg_1 _Sm_2

Table 4.13

Analysis using the coefficients table to check for collinearity between high school student population size and student achievement, after statistically controlling for minority population
Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	2191.330	10.736		204.109	.000		
	Minority	-.293	.239	-.125	-1.230	.222	1.000	1.000
2	(Constant)	2336.872	25.239		92.590	.000		
	Minority	-1.169	.247	-.496	-4.730	.000	.672	1.488
	Lg_1 _Sm_2	-75.611	12.227	-.648	-6.184	.000	.672	1.488

a. Dependent Variable: Total_Score

Table 4.14

Model Summary of Sequential Regression to investigate what is the relationship between a high school student population size and student achievement, after statistically controlling for minority population
Model Summary[c]

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.125[a]	.016	.005	58.43091	.016	1.513	1	96	.222
2	.546[b]	.298	.283	49.59791	.283	38.239	1	95	.000

a. Predictors: (Constant), Minority
b. Predictors: (Constant), Minority, Lg_1 _Sm_2
c. Dependent Variable: Total_Score

Table 4.15

ANOVA of Hierarchical Multiple Regression to investigate what is the relationship between a high school student population size and student achievement, after statistically controlling for minority population.
ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5165.697	1	5165.697	1.513	.222[b]
	Residual	327760.396	96	3414.171		
	Total	332926.094	97			
2	Regression	99230.603	2	49615.301	20.169	.000[c]
	Residual	233695.491	95	2459.953		
	Total	332926.094	97			

a. Dependent Variable: Total_Score
b. Predictors: (Constant), Minority
c. Predictors: (Constant), Minority, Lg_1 _Sm_2

Table 4.16

Analysis using the coefficients table to check for collinearity between high school student population size and student achievement, after statistically controlling for teacher quality
Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	1828.135	179.744		10.171	.000		
	HQ Teacher	3.610	1.842	.196	1.960	.053	1.000	1.000
2	(Constant)	2019.108	178.829		11.291	.000		
	HQ Teacher	2.253	1.787	.122	1.261	.210	.952	1.050
	Lg_1 _Sm_2	-39.362	11.321	-.338	-3.477	.001	.952	1.050

a. Dependent Variable: Total_Score

Table 4.17

Model Summary of Hierarchical Multiple Regression to investigate what is the relationship between a high school student population size and student achievement, after statistically controlling for teacher quality

Model Summary[c]

					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	
1	.196[a]	.038	.028	57.74518	.038	3.843	1	96	.053	
2	.383[b]	.147	.129	54.67406	.109	12.088	1	95	.001	

a. Predictors: (Constant), HQ Teacher

b. Predictors: (Constant), HQ Teacher, Lg_1 _Sm_2

c. Dependent Variable: Total_Score

Table 4.18

ANOVA of Hierarchical Multiple Regression to investigate what is the relationship between a high school student population size and student achievement, after statistically controlling for teacher quality

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	12813.534	1	12813.534	3.843	.053[b]
	Residual	320112.560	96	3334.506		
	Total	332926.094	97			
2	Regression	48947.082	2	24473.541	8.187	.001[c]
	Residual	283979.012	95	2989.253		
	Total	332926.094	97			

a. Dependent Variable: Total_Score

b. Predictors: (Constant), HQ Teacher

c. Predictors: (Constant), HQ Teacher, Lg_1 _Sm_2