THE RELATIONSHIPS BETWEEN EDUCATIONAL INPUTS AND MEASURES OF STUDENT ACHIEVEMENT AS OUTPUTS: A MULTIVARIATE APPROACH

by

OLAJIDE OLUWASEFUNMI AGUNLOYE

(Under the Direction of CATHERINE C. SIELKE)

ABSTRACT

There is an unsettled debate about the nature of the relationships between educational inputs and measures of student achievement. The purpose of this study was to describe and to explain the relationships between educational inputs and measures of student achievement as outputs. Quantitative multivariate data on 10 educational inputs and 2 measures of student achievement, covering 2 academic school years in 71 highs schools across 6 school districts, were collected and analyzed as a matrix of vectors.

Data for across-school-districts study were analyzed using multivariate statistical methods involving discriminant analysis, canonical correlation, and multivariate regression. Data on schools-within-districts study were analyzed using a combination of parametric and non-parametric statistical methods involving Pearson's product moment correlation, Kendall's Tau-b, and Sommer's d.

The findings of the study showed significant variations between school districts and school categories on all of the inputs and outputs. There were significant positive relationships between (1) teacher quality and per pupil expenditure as well as percentage of economically disadvantaged students, (2) percentage of students passing the science section of the Georgia

High School Graduation Test (GHSGT) and the percentage of students who qualified for Georgia's HOPE scholarship. There were significant negative relationships between the percentage of economically disadvantaged students and the percentage of students passing the science section of the GHSGT, as well as the percentage of students who qualified for the HOPE scholarship.

There were significantly positive correlations between principal stability and the percentage of students passing the science section of the GHSGT as well as the percentage of students who qualified for the HOPE scholarship. There were no significant relationships between expenditure per pupil and the measures of student achievement. However, when the percentage of economically disadvantaged students (PEDS) is considered as a monetary input, the relationships became significant.

There were schools performing above expectation given their relatively high PEDS. There were schools performing according to expectation given their PEDS value. There were also schools performing below expectation, given their relatively low PEDS. These schools were categorized as 'positively discordant,' 'concordant,' and 'negatively discordant,' respectively.

Through the qualitative portion of the study, principals and Science Department heads offered important suggestions for improving science education.

INDEX WORDS: Education Productivity, Input Variables, Output Measures, Per Pupil Expenditure, Teacher Quality, Class Size, Teacher Development, Principals' Stability, Production Function Models, X-factors, Concordancy, Positively Discordant Schools, Concordant Schools, Negatively Discordant Schools, Georgia High School Graduation Test, Georgia HOPE Scholarship.

THE RELATIONSHIPS BETWEEN EDUCATIONAL INPUTS AND MEASURES OF STUDENT ACHIEVEMENT AS OUTPUTS: A MULTIVARIATE APPROACH

by

OLAJIDE OLUWASEFUNMI AGUNLOYE

B.S., University of Ibadan, Nigeria 1977

M.S., University of Ibadan, Nigeria 1982

Ed. S., University of West Georgia 2001

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF EDUCATION

ATHENS, GEORGIA

2004

THE RELATIONSHIPS BETWEEN EDUCATIONAL INPUTS AND MEASURES

OF STUDENT ACHIEVEMENT AS OUTPUTS: A MULTIVARIATE APPROACH

by

OLAJIDE OLUWASEFUNMI AGUNLOYE

Major Professor: Catherine C. Sielke

Committee: Stephen Olejnik
Sally J. Zepeda

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2004

DEDICATION

This doctoral dissertation is dedicated to the following people:

All my children: Olanireti, Ayodele, Omowumi, and Omowale, my sources of inspiration.

My mother, Mrs. Docas Olaribigbe Agunloye, for her love for me. Mum, I love you too.

My late father, Mr. Adebayo Agunloye, for setting my feet on the path of knowledge.

ACKNOWLEDGMENTS

The research and the writing of this dissertation were possible through the help and guidance of many people. My special appreciation goes to my major professor, Dr. Catherine C. Sielke, for her time, support, guidance, and useful suggestions toward the completion and quality of this study. Dr. Sielke, thank you — lots.

My special appreciation also goes to Dr. Stephen Olejnik. I am especially thankful to him for taking the time to read, suggest, and guide me through the methodology, statistical analysis, and the interpretations of the data. Dr. Olejnik, I have learned a great deal from you. Thank you.

I am grateful to Dr. Sally Zepeda, not only for her unflinching support, but also for taking the time to read and to make useful suggestions that go into improving the quality and making the completion of this study possible. Dr. Zepeda, thank you.

I am also grateful to Dr. Karen Hunt for her useful contributions to some of the thoughts that constituted the starting ingredients for this study. Dr. Hunt, I appreciate you.

My very special thanks go to my friend Genie L. Armour, my relatives (the Tijani, the Oluwadare, and the Agunloye families), and those who are close and dear to me, for their understanding, moral support, and belief in everything I do. Thank you all. It is good to have friends and family members like you.

I am especially grateful to all the superintendents, directors, central office personnel, principals, science department heads, and teachers that participated in this study. In fact, the study would have been impossible without their cooperation and help. I thank you all.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Background to the Study

The perception that student performance is not commensurate with growing expenditure on education in K-12 public schools has been the subject of policy debates in this country (Coleman, 1968; Coleman et al., 1966; Coleman, Easton, & LaRocque, 1998; Clotfelter & Ladd, 1999; Hanushek, 1981, 1989; Klick, 2000). Since the publication of *A Nation At Risk* (National Commission on Excellence in Education, 1983), and following the launch of Sputnik by the Soviet Union in 1957, there has been concern over the stagnant performance in science across the nation despite increases in education expenditures (Hanushek, 1981, 1989).

International comparisons of science achievement among students ages13 and above ranked the United States in the 13th position below countries like South Korea, Switzerland, Soviet Union, Canada, and France (Goldschmidt & Eyermann, 1999; Zehr, 1998). The report by the U. S. Department of Education, National Center for Education Statistics (*NCES*) (2001a) indicated the United States is in 15th position, behind countries like Australia, Canada, England, Hungary, Singapore, and Slovenia, on international comparison (*NCES*, Table13-2). In another report, the National Assessment of Educational Progress (*NAEP*, Table 13-5) also showed that the trend in science performance has been relatively flat over the past decade (U.S. Department of Education, 2000a).

In his remarks at the Mathematics and Science Summit, February 3, 2003, the Secretary for Education, Dr. Rod Paige, pointed out that 82% of 12th graders are not proficient in science (U.S.

Department of Education, 2003a). Paige further stated, "It is undisputed that American 12th graders lag far behind their European and Asian counterparts in mathematics, science, and technology" (Secretary's Remarks at Mathematics Summit, Mathematics and Science Initiative, Speeches, U.S Department of Education, p.1).

Despite the general flat trend in performance, some schools have been reported to make better use of their allocated resources to get more achievement out of their students (Unnever, Kerckhoff, & Robinson, 2000) in what is now generally termed "allocative efficiency" (Grissmer, 2002; Levin & McEwan, 2001). Forty-seven states now have academic standards in all core subject areas including science. All the 50 states have testing programs in science among other subject areas. Forty-five states, including Georgia, require report cards and a pass in science in the state's criterion-referenced test for the award of high school diploma (U.S. Department of Education, NAEP, 1999).

As part of an education reform to improve student achievement, the State of Georgia passed the *A-Plus Education Reform Act* in 2000. The Act mandates a sweeping reform in education, especially in the areas of education performance and accountability. The Act mandated report cards and a system of reward for performance (Official Code of Georgia Annotated, *OCGA* 20-14-25, *OCGA* 20-14-27), overseen by a legislative body known as the Office of Student Achievement (OSA) formally called Office of Education Accountability.

The 2001-2002 Georgia public education report card (Georgia Department of Education, 2003) showed that K-12 expenditures in the state of Georgia, for the fiscal year 2001-2002, was about $13.6 billion. Close to 59% ($7.5 billion) of this was spent on instruction. By extension, this means that a sizable percentage of this total was spent at the high school (grades 9-12) level. The national per pupil average expenditure on high school students was $7,207 compared to

$6,015 for the state of Georgia, during the academic year 2001-2002 (U.S Department of Education, *NCES*, 2003b, Table 3). Moreover, the U.S. Department of Education reported that the average class size across the state is 24.4, compared with the 23.6 national average (U.S. Department of Education, *NCES*, Digest of Education Statistics, 2002a, Table 69). The state is also ranked 14th in the nation in resource allocation to education, but lower in overall student achievement scores (U.S. Department of Education, *NCES*, 2000b).

Many studies have been done on the relationship between educational resource inputs and student achievement (Greenwald, Hedges & Laine, 1996; Hanushek, 1981; Hanushek, 1989; Hanushek, 1996a; Monk, 1992; Klick, 2000). The general consensus is that resource inputs do matter in determining student achievement depending on the way and manner such resources are allocated and utilized. The mode and form of resource allocation is determined by many school-level variables (Murnane & Levy, 1996; Murnane & Nelson, 1984). Therefore school-level variables play an important role in dictating how educational resources could affect student achievement in schools. Some of these variables may include administration stability, quality of leadership, quality of teachers, and other pedagogical parameters, to mention a few. Therefore, there is a need to look at other factors, in addition to money, that impact student achievement in science.

According to Heinbuch and Samuels (1995), very little concerted effort has been be made to establish the nature of the relationship between multiple educational inputs and student achievement in science. The poor performance in science is often blamed on lack of qualified teachers, poor instruction, inadequate curriculum, inadequate resources, or combinations of all of these, without any empirical evidence to support or refute any of them. Much of the prevailing notions of the state-of-science-education are informed by social intuition rather than tangible empirical information. Little or no research has been reported on the effect of multiple educational inputs on the achievement of students in science in the State of Georgia. More specifically the researcher did not find any reported empirical study on the relationship of school-level resources for science education, to student achievement in science in the state of Georgia.

A study was therefore necessary to provide information on the relationships between educational inputs and student achievement in general and particularly in science. The study can also provide information on other input factors, beyond money, that contribute to the performance of students in science and how to shape policy decisions regarding the allocation of resources for science education in the state.

Purpose of the Study

The purpose of this study was to describe and to explain the nature of the relationships between educational inputs and measures of student achievement as outputs. The input variables investigated were: (1) Expenditure Per Pupil (EXPPP), (2) Teacher Quality (TEACHQLT), (3) Percentage of Economically Disadvantage Students (PEDS), (4) School-level Science Department Expenditure Per Pupil (SCIEXPPP), (5) Science Teacher Quality (SCITCHQT), (6) Science Class Size (SCCLSIZE), (7) Science Lab-based instructional activities per Teacher per Week (LABSPWK), (8) the number of Professional Development Activities, per Science Teacher per Year (PDVSCTCH), (9) Principal Stability (PRINSTBY) and (10) Head of Science Department Quality (HEADSQLT).

The achievement output measures investigated were: (1) Percentage of student passing the Science Section of the Georgia High School Graduation Test at first sitting (GHSGT) and (2) Percentage of graduating students qualified for the Georgia HOPE Scholarship (HOPE).

Additional school-level qualitative parameters, including suggestions for improving science education, were also investigated for possible relationships to student achievement in science beyond the quantitative measures. A graphic model representation of these statements of purpose is presented in Figure 1.

Conceptual Framework

The question of whether or not there is a relationship between education resources, including money, and the performance of students is a recurring one in education and policy circles. This question gained national prominence following the Coleman Report (Coleman et al., 1966). The report indicated little or no relationship between the educational achievement of students and expenditure. The report led to a wave of education reform initiatives in most sates aimed at improving standards in order to improve student achievement without necessarily increasing expenditure. The research in this area employs education production function methodology to investigate relationships between educational input variables, such as money, and educational outcomes measures.

Production Function in Education

Production function uses linear regression, correlation, or multivariate-discrminant analysis to indicate the nature of the relationship between educational input and output (Monk, 1992; Murnane, 1975). The use of production function research in education began around 30 years ago. Large-scale production-function research began with studies reported by Coleman et al., (1966) called Equality of Educational Opportunity Report (EEOR), also referred to as the Coleman Report. The study used a representative sample of students and schools across the U.S. to look at the relationship between monetary expenditure on education and student achievement.

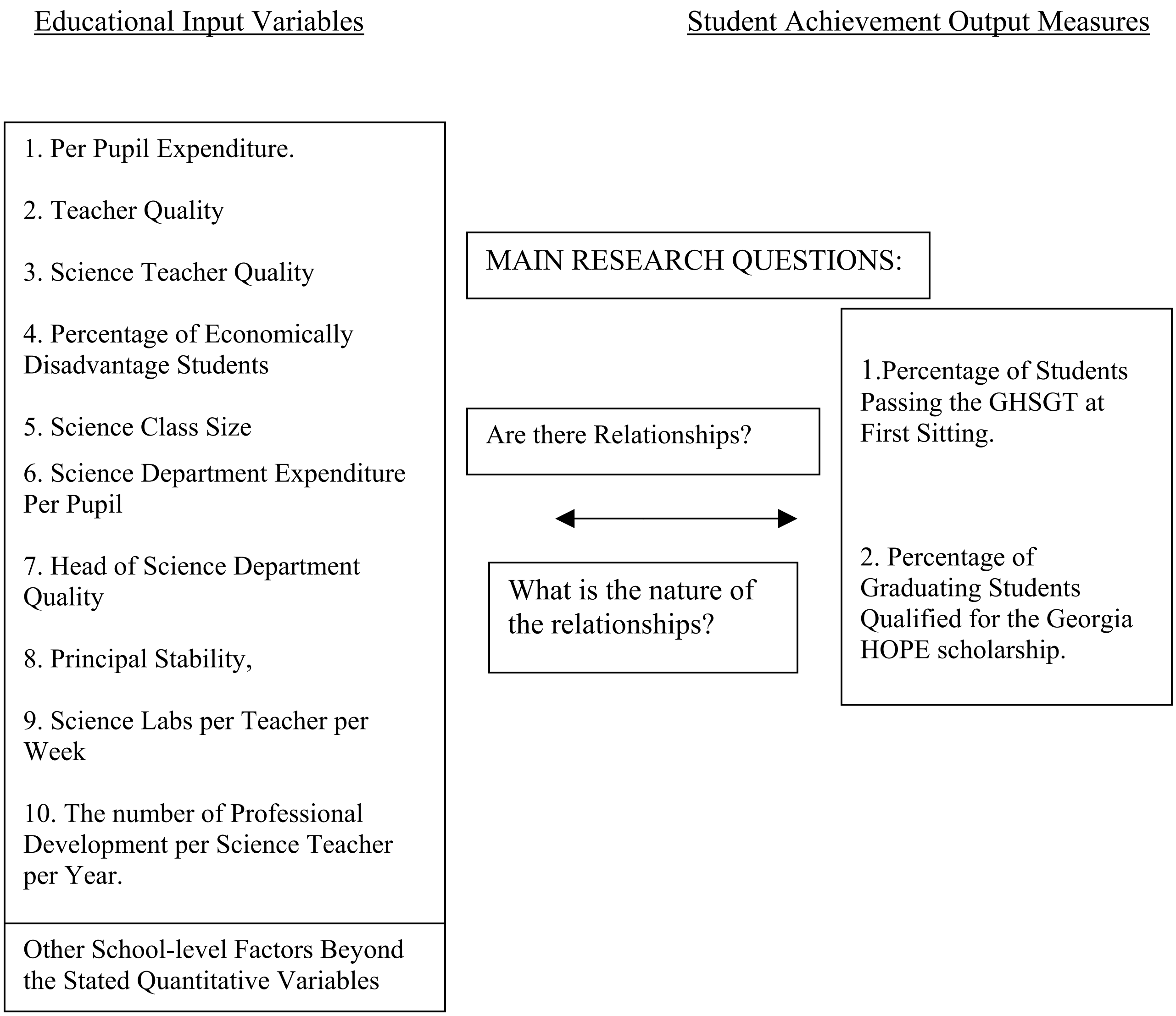


Figure 1: Graphic Representation of the Statements of Purpose

Two other large-scale studies by Mortimore, Simmons, Stoll, Lewis, and Ecob (1989) and Smith and Tomlinson (1989), were carried out in the United Kingdom. The two studies relate various educational inputs, including money, to the performance of students in primary and secondary schools.

Around the same period, Montmarquette and Majseredjian (1989), used a two-way analysis to establish the relationship between educational outcomes in grades 1 and 4 and a combination of observable and unobservable class-level inputs. In a related study (Goldhaber & Brewer, 1997a), explored the relationship between observable-unobservable school-teacher factors and the performance of students in 10^{th} grade mathematics. Over the years, about 400 studies have been reported, and some of these studies have been evaluated and summarized in a meta-analysis study by Hanushek (1997).

Production function model-building in education involves the selection of a set of existing educational input variables along with a corresponding set of output (ex post) educational measures and establishing the relationship between the input variables and the output measures (Monk, 1992). The process of model-building requires the use of correlation and regression analysis to quantify and describe the nature of the relationship (Monk, 1992; Walberg, 1982; Walberg & Fowler, 1989; Walberg & Weinstein, 1982; Wenglinsky, 1997).

Almost all of the studies reported on education production function examined the relationship between two or more educational input variables and one educational output measure, usually student achievement (Hanushek, 1997). The analysis is often done using a univariate analysis approach. The general univariate model is often in the form of:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \ldots\ldots + b_kX_k$$

where Y = single educational output measure, *a* = intercept of the regression line on the output measure, $b_1, b_2, b_3, \ldots, b_k$ = regression coefficients associated with the educational input variables, and $X_1, X_2, X_3, \ldots X_k$ are educational output variables.

In education, there is often more than one output measure for a given set of inputs. Monk (1992) examined the use of production function in education productivity research and concluded that using just one measure of output for a set of educational inputs "abrogates the responsibility of the desired attempt to capture education in its true form." (p. 308). It is therefore imperative that an education production function examines the relationship between sets of input variables and sets of output measures, as composites (Pedhazur, 1997).

In the few studies that have been reported using multiple output measures, the analysis is often done by examining the relationship on each set of output variables separately, using a series of univariate tests (Tatsuoka, 1988). Pedhazur (1997) also noted that some researchers simply calculate the zero-order correlations for all possible pairs of variables using univariate tests. Researchers do this because it is easier and simpler. However, as noted by Pedhazur (1997), the use of these types of multiple univariate tests on a data with multivariate constructs, "… affects the prescribed α level" (p.895). This increases the chance of committing a Type 1 error; where the researcher rejects the null hypothesis when the null is true thus producing a liberal test. Also, analyzing multivariate context data using a series of univarite methods compromises the very essence and richness of a multifaceted phenomenon, such as education production. For these reasons, all the measures of educational output, describe earlier, are used as composite output measures simultaneously. This requires the use of multivariate analysis methods.

In a multivariate context, education production examines the relationship between a set of two or more educational input variables on two or more educational output variables (Tatsuoka, 1973; Tatsuoka, 1988; Pedhazur 1997). The general multivariate model may be expressed as:

$$a_1Y_1 + a_2Y_2 + a_3Y_3 \ldots. + a_kY_k = b_0 + b_1 X_1 + b_2X_2 + b_3X_3 + \ldots\ldots. b_mX_m$$

where, $a_{1,} a_{2,} a_{3,\ldots\ldots,}$ a_k = coefficients associated with the linear composite of educational output measures $Y_1, Y_2, Y_3, \ldots., Y_k$; b_0 = intercept of the composite regression line on the linear composite of output measures; and $b_1, b_2, b_3, \ldots, b_k$ = regression coefficients associated with the linear composite of educational input variables $X_1, X_2, X_3 ,\ldots. X_m$.

Since the purpose of this study was to describe and explain the relationship between many input variables and two measures of student achievement (as outputs), the application of multivariate analysis is appropriate. Based on the multivariate model explained above, the overall production function model for this study may be expressed as:

$$a_1Y_1 + a_1Y_2 = b_0 + b_1 X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5$$

where $X_1, X_2, X_3, X_{4,} X_5$ = the educational input variables, $Y_{1,}$ $Y_2,$ = the two measures of student achievement in science, and $b_{1,}$ $b_{1\ldots.}$= the coefficients or slopes of the composite linear regression model.

Establishing the nature of the relationship between two sets of continuous variables, in a multivariate education production function context, requires a two-step process: (1) Establishing the strength of the relationship between the linear composite of a set of educational input variables and the linear composite of educational output measures using Canonical Analysis (CA), and (2) Establishing the linear regression model (equation), that maximizes the relationship between the educational input variables and the educational output measures, using Multivariate Regression Analysis (MRA).

Canonical Analysis (CA) was developed by Hotelling (1936) to study the relative strength of the relationship between two sets of variables. The conceptual reasoning behind CA is to form two linear combinations of the input variables and the output variables by differentially weighing them to obtain the maximum possible correlation between the two linear combinations (Pedhazur, 1997). The correlation between the two linear combinations is called the Canonical Correlation (C_r).

Multivariate Regression Analysis (MRA) is used to describe the linear model between two or more independent interval variables and two or more dependent interval criterion variables (Stevens, 1972). First, the multivariate model examines the significance of the overall multivariate regression model. Second, the individual interpretation of each independent variable is conducted, in a multiple univariate sense, to see how well the individual predictors compare on each of the dependent variables when considered separately.

The Inputs

Expenditure per pupil was considered as one of the input variables because there is no doubt that the focus of education spending should be instruction at the front-line in the schools. A combination of state funding, district-level budgetary input and school-level inputs such as grants, endowments, corporate partnerships, internal foundations, booster clubs and parent-teacher associations donations are important to consider in looking at school-level expenditure per pupil (Poss, 1993). For the purposes of this study, the focus of the funding source for estimating per pupil expenditure was the state of Georgia's QBE allotment for each school. It is the most reliable source of financial information for schools. To confirm how the state allotment impacts science education at the school level, additional information on school-level direct instructional expenditure on science was also gathered.

Teacher quality, as an input, has also been determined to have some effect on student performance (Figilo, 1999; Hanushek, 1971). Most of the literature on teacher quality (Figilo, 1999; Goldhaber, Brewer, & Anderson, 1999; Goldhaber & Brewer 1997b; Mancebon & Bandres, 1999) only takes into consideration either the years of service of the teachers, or the highest qualification of the teachers separately, as measures of teacher quality. Other research has also extended the definition of teacher quality to include such factors as verbal expression ability and scores on standardized certification tests (Firestone, 1991). A composite of these measures reflects a more representative proxy for teacher quality than just looking at one the measures in isolation.

Education productivity study is not complete without a measure of the percentage of economically disadvantaged students, PEDS (a measure to as socio-economic status of student (SES), as a factor (Alexander, 1997). The State of Georgia educates over 1.4 million students; comprising of 55.5% White, 38.2% Black, 4.0% Hispanic, 2.1% Asian/Islanders, and 0.2% Native Americans. About 23% of the students are from households below the national poverty income limit (Georgia Department of Education, Report Card, 2002), and over 44% of students in public schools in the state are eligible for the federal free-reduced lunch program and therefore classified as economically disadvantaged (Georgia Department of Education, Report Card, 2003). In much of the previous research studies investigated so far, the socioeconomic status of students (SES), has been used as the controlling variable (Figilo 1999; Hanushek, 1986; Hanushek, 1989; Hanushek, 1991; Mancebon & Bandres, 1999; Monk, 1992; Walberg, 1982). In this study, SES is measured by the percentage of students qualified for the free-reduced meal program, and referred to as the percentage of economically disadvantaged students (PEDS). PEDS (a measure of SES), in this study, is used both as an input variable as well as a controlling factor.

The amount of money spent on direct instruction is very important. This is more so in science where hands-lab activities are comprised mostly of specialized purchased equipments and materials. Eighty percent of high school graders reported doing hands-on activities at least once a week (O'Sullivan & Weiss, 1996). Science students, who do hands-on activities at least once a week, are more likely to perform above proficiency level than those who do not (O'Sullivan & Weiss, 1996). Hence a look at the direct school-level expenditure per pupil as measured by the amount of money received by science departments for these activities is important.

The issue of science teacher quality is also entering into the debate on how to improve student achievement, following the publication of the Schools and Staffing Survey, SASS by the U.S. Department of Education, *NCES* in 1994 (Ingersoll, 1999). According to Ingersoll (1999), 20% of secondary grade teachers have no degree in their assigned fields. Part of the Math and Science Initiative (MSI), carved out of the *No Child Left Behind Act* (*NCLBA*), by the Secretary of Education, Rod Paige, focuses on improving science teacher quality through teacher development programs (U.S. Department of Education, 2003a). Although the debate on what constitutes teacher quality in the specific subject areas still remains unresolved (Haertel, 1991; Ingersoll, 1999) a measure of the impact of science teacher quality, as presently constituted, on student achievement, is necessary.

Class size is the number of students regularly in a teacher's room and for whom that teacher is responsible each day of the school year (Finn & Achilles, 1999). Class size has also been viewed as an input variable that affects student performance. Several studies have been published on the effects of class size on student performance (Glass & Smith, 1978; Glass, Cohen, Smith, & Fibly, 1982; Robinson, 1990; Robinson & Wittebols, 1986; Slavin, 1989). The studies opined that class size could affect academic achievement, although significant benefits from class size are only obtained below a critical class size number (usually below 20), especially at lower grades and for economically disadvantaged students. In the State of Georgia, the maximum mandatory class size for science is 28 students. The peculiar nature of science instruction, requiring laboratory and hands-on modules, makes the issue of class size a critical factor to consider in the teaching of science and in student performance.

Science as a subject area that deals with physical and natural phenomena requires that students interact with these phenomena in order to conceptualize them. A majority of high school students (80%) who do hands-on activities, at least once a week are more likely to perform above proficiency level than those who do not (O'Sullivan & Weiss, 1996). An empirical measure of the impact of the number of hands-on activities in a science class per week, on student achievement, is therefore important.

The Carnegie Commission on Science, Technology, and Government (1991) advocated for professional development programs to enhance the content knowledge and teaching skills of science teachers. Part of the science initiative, as contained in the NCLB Act, is the focus on professional development for science teachers. Proper education of students, especially in science requires teachers to gain new knowledge. It has been noted that science educators often lack sufficient content knowledge in instructional strategies to improve instruction without constant professional development (Darling-Hammond & McLaughlin, 1995; Darling-Hammond, Wise, & Klein, 1995). It was for this reason that this study looked at the professional development of science teachers as an input.

Principal's stability is more likely to create a school culture and climate conducive for education to take place (Argyris, 1993). As principals become stable, their organizational maturity improves and they are in better positions to realize their visions for the schools (Argyris, 1993). Studies have shown that there is a link between principal stability and achievement in science (Comber & Kreeves, 1973).

Most of heads of science departments are often teachers as well (U.S. Department of Education, *NCES*, 2002b). The heads of science departments are also often knowledgeable in the requirements of science pedagogy. Therefore, the years of service as a teacher as well as the years of service as heads of departments may impact the way and manner they manage and allocate resources that could impact student achievement.

In the State of Georgia, Article 4A of the *A Plus Education Reform Act* of 2000 mandates the establishment of Local School Advisory Councils (LSAC) with the principals as chairpersons (*OCGA* 20-2-58, 2000a; *OCGA* 20-2-86, 2000b), thus increasing the power of school principals in the governance of schools. Part of he legislative intent of this section of the act is to improve the academic achievement of students while giving schools some level of control over budgetary issues. It is in the pursuance of this legislative intent that the Georgia's Quality-Based Education (QBE) school-site allotment funding practice was initiated.

A part of the NCLB Act strongly emphasizes school performance accountability by mandating an Adequate Yearly Progress (AYP) report for each school and each school district that gets federal funds. The AYP requirement puts the leadership of schools at the forefront of the performance accountability forum, thus making the issue of school leadership an important input in the education accountability equation.

<u>The Outputs</u>

In the State of Georgia, science education at the high school level (grades 9-12) is based on the Quality Core Curriculum (QCC) guidelines from the state (*OCGA* 20-2-140, 1994). The

QCC is a broad set of curriculum outlines or guides on what high school students are supposed to know by the time they finish high school. Based on the QCC, school districts have some degree of flexibility to individualize their curriculum, as long as it is within the QCC guidelines. However, the QCC generally defines what the requirements are for a student to graduate from any public high school within the state.

For science, a student is required to have taken and passed biology, physical science, and a choice between chemistry and environmental science. This essentially means a total of 3 credits made up of one credit each from biology, physical science, and one of either chemistry or environmental science. In addition, a prospective high school graduate is also expected to pass the Georgia High School Graduation Test (GHSGT), including the Science Section of the test before the award of the state's high school diploma (*OCGA* 20-2-281, 1996). The performance of students in the Georgia High School Graduation Test is an important measure of educational output in the state.

The Georgia HOPE scholarship program was established in 1993 to provide money for tuition and books for students, who are residents of Georgia, pursuing post-secondary education programs in colleges and universities in the state ("Georgia HOPE scholarship program", 1997). To qualify for HOPE during a student's first year in college, the student must have an exit Grade Point Average (GPA) of not less than 3.0 on a 1.0-4.0 scale, at graduation from an accredited high school in the state. The award is not means tested, so everyone who meets the GPA criterion gets the scholarship. Therefore, the percentage of students qualifying for the HOPE scholarship is also an important educational achievement measure.

The purpose of this study was to describe and explain the nature of the relationships between educational inputs and measures of student achievement as outputs. Sets of hypotheses were tested to describe and explain the nature of the relationships between educational inputs and measures of student achievement.

Hypotheses To Be Tested

The following hypotheses are to be tested in this study:

Hypothesis #1: There is no significant difference between school districts on the composite of educational input variables and the composite of achievement output measures.

Hypothesis #2: There is no significant difference between school districts on each of the educational input variables and the achievement output measures, considered separately.

Hypothesis #3: There is no significant difference between school districts on the composite of educational input variables and the composite of achievement output measures, after controlling for Percentage of Economically Disadvantaged Students (PEDS).

Hypothesis #4: There is no significant difference between school districts on each of the educational input variables and the achievement output measures, considered separately, after controlling for PEDS.

Hypothesis #5: There is no significant correlation between the composite of educational input variables and the composite of achievement output measures across school districts.

Hypothesis #6: There is no significant correlation between the individual input variables and the achievement output measures, considered univariately, across school districts.

Hypothesis #7: There is no significant linear relationship between the linear composite of input variables and the linear composite of achievement output variables.

Hypothesis #8: There is no significant correlation between additional school-level input variables and achievement output measures.

Hypothesis #9: There is no indication of any other school-level factors**,** beyond the quantitative inputs considered, that differentiate schools on the achievement measures.

Method

This study is a correlational (non-experimental) study, aimed at examining the relationship between the educational inputs on one hand and student achievement as measured by the scores on science section of the Georgia High School Graduation Test (GHSGT) and the percentage of graduating students qualified for the HOPE scholarship on the other. The objective outcome is to use the results for practical applications in guiding agenda and funding policies for science education for the overall purpose of improving student achievement.

Ten input variables comprising of: (1) Expenditure Per Pupil (EXPPP), (2) Teacher Quality (TEACHQLT), (3) Percentage of Economically Disadvantage Students (PEDS), (4) School-level Science Department Expenditure Per Pupil (SCIEXPPP), (5) Science Teacher Quality (SCITCHQT), (6) Science Class Size (SCCLSIZE), (7) Science Lab-based instructional activities per Teacher per Week (LABSPWK), (8) the number of Professional Development Activities, per Science Teacher per Year (PDVSCTCH), (9) Principal Stability (PRINSTBY) and (10) Head of Science Department Quality (HEADSQLT), were investigated for this study.

The two measures of student achievement, considered as inputs, for this study are: (1) The percentage of students who passed the science section of the GHSGT and the percentage of graduating students qualified for the state of Georgia's HOPE scholarship award. Six school districts within the metro-Atlanta area of Georgia were selected for this study. The individual high schools in each of the selected school districts were the sampling units. The six school districts and the schools within the districts were the units of analysis.

The data collection process involved both quantitative and qualitative components. The quantitative components were those in which numeric values are available on record and can be accessed through either archival records or obtained through personal interviews from the appropriate sources. The qualitative data include suggestions for improving science education in the schools. Numeric values attached to the qualitative data are based on the nature of the responses to semi-structured interview questions (Appendix B).

Limitations of the Methodology

Some of the analysis methods used for this study have been used in some previous studies (Goldhaber, Brewer, & Anderson, 1999; Heinbuch & Samules, 1995: Figilo, 1999; Mancedon & Bandes, 1999; Shannon & Davenport, 2001). Based on the information gathered on education production function, there are some limitations to this study:

(1) The study is limited to 10 quantitative input factors due to constraints of resources and time. There are obviously more than these ten input factors that may affect student performance in science at the high school level. Some of these may be school climate, other instructional support services, and parental involvement, to mention just a few. This is why the study also looks into other qualitative factors that may impact student achievement in science. As shown in the literature review, the ten factors selected are considered by the researcher to be important to student performance in general, and in science in particular.

(2) Two measures of student performance are used as the measure of educational outcome. These two measures may not capture all the areas of student educational outcome. Other measures of outcome include attendance, dropout rates, lifetime earnings, service to

community, etc. Again the two selected measures of student performance are considered by the researcher to be closely related to the domain of performance in science.

(3) Only the most recent two years data on the GHSGT and HOPE are used as the performance measures in this study. The two-year data may not completely capture the performance of a school over a longer period of time. However, given the limitations of time and resources, this study limits outcome time scale to the most recent two years. However, this two-year time scale is more reliable than the one-instance single-snapshot performance measure used in previous studies (Mancebon & Barnes, 1999; Figilo, 1999; Hanushek, 1986). This is also why the study includes two measures of student outcome to improve the robustness and generalizability of the results.

(4) The use of the percentage of students eligible under the federal guidelines on free and reduced meal program, as a proxy for the percentage of economically disadvantaged students (PEDS), does not completely capture the total scope of students' contextual environment. However, it is the reasonable approximation of student's SES approved by the Department of Education (Heinbuch & Samuels, 1995; Mancebon & Bandres, 1999; Walberg, 1982)

(5) The data on teacher certification and subjects taught were to find out if teachers have science degrees and are certified to teach science. The questions used for the collection of this did not investigate if the teachers teach in the content areas of their degree majors.

(6) Professional development did not include college or university courses taken by teachers towards the award of higher degrees or certification-add-ons.

(7) The semi-structured interviews were limited to the descriptive responses of principals and heads of science departments. Since the intent was to corroborate the findings of the

quantitative results of the state-district level study and the perspectives of the respondents on how to improve science education in the schools, the qualitative data gathering did involve any triangulation sessions with other school community groups outside of the principals and the heads of science departments.

The Assumptions

For the purposes of this study, some assumptions were made. These assumptions are:

(1) The input and output variables were multivariately normally distributed across the sampling population, and each variable measure in each school is also normally distributed across all the population. Because the sample size is large, violation of this assumption would have little or no effect on the statistical validity of the results of this study.

(2) The variation among the performance measures was the same across all the levels of the factors. The variances of the performance measures are positively proportional to the sample size. The SPSS Version 10.5 computer program used in the analysis of the data allowed the researcher to test, explain, and make adjustment for these assumptions.

(3) The score on the selected variables for each sampling unit were independent of each other. As each school was a distinct unit from each other on the selected variables, this assumption was met in this study.

(4) The multiple regressions of the input factors against specific student performance measures were linear. This was the basis for production function and multivariate regression-correlation analysis. The selection of variables was based on the high possibility of obtaining extreme values for each variable to optimize group separation on the performance measures. (Cox, 1958; Draper & Smith, 1981).

(5) The qualitative data are quantifiable and are quantified for the purposes of ease of analysis and interpretation.

Definition of Terms

In the context of this study, the following terms will be used to indicate the explanations attached to them as outlined below.

Education Productivity: A model measure of student performance relative to educational input variables.

Input Variables: The variable factors and resources into education that directly affect science education at the individual sample schools as explained earlier.

Output Measures: Measurable educational outcomes relative to a set of educational resource inputs.

Per Pupil Expenditure: Per pupil expenditure for instruction at the individual sample schools, based on FTE allocation and school-source funds.

Teacher Quality: A composite index based on teacher qualification (training) and experience.

Class Size: The number of students in a science class and for whom a science teacher is responsible during a complete academic year.

Professional Development Activities: Training activities undertaking by a teacher to improve professional practices in school and classroom settings, excluding college or university courses taken by teachers toward the award of higher degrees or certification add-ons.

Principal Stability: The number of consecutive years the principal of a school has served as the principal in the school.

Achievement Measures: Measurable outcomes of an educational process. In this study it is the combination of the percentage of students passing the science section of the GHSGT at first sitting, and the percentage of graduating students qualified for the Georgia's HOPE scholarship award.

Production Function Model: A mathematical construct that models the relationship between a given set of educational output and a set of educational resource inputs.

Georgia High School Graduation Test (GHSGT): The summative criterion-based graduation test required for high school student in the state of Georgia to be awarded high school diplomas.

Concordancy: Measure of the degree of agreement between the percentage of economically disadvantaged students and the achievement measures in a school.

Positively Concordant Schools: These are schools whose achievements measures are above expectation compare to the relatively high percentage of economically disadvantaged students in the schools.

Concordant Schools: These are schools whose achievements measures are according to expectation compared to the relative percentage of economically disadvantaged students in the schools.

Negatively Concordant Schools: These are schools whose achievements measures are below expectation compared to the relatively low percentage of economically disadvantaged students in the schools.

Organization of the Study

Chapter 1 introduces the context surrounding the issue of education inputs and student achievement. The chapter also introduced the application of production function in finding explanations for the relationship between inputs and student performance measures. The problems surrounding the research and application of production function in education were examined, thus setting the stage for the purpose, the assumptions, and, the limitations of the study.

Chapter 2 reviews the body of available published studies. The chapter summarizes the past and currently available literature with a critique of the various perspectives presented by the different authorities in the field. Some of the questions still to be answered in the area of the relationship between inputs for and student performance, including the methodological problems associated with research in this area are discussed. The chapter concludes by identifying what the current study aims to contribute to the present body of knowledge available in the area of the relationship between education inputs and student performance. Possible implication for policy development and implementation in the funding of science education, were also identified as the outcome foci of the study.

Chapter 3 discusses the research design for the study. The basis for using non-experimental multivariate production function design is explained, and the research questions are revisited. The context of the study including explanations for where, why, and how the research sites were selected, is discussed. The issue of sample size and possible sampling difficulties are examined. A step-by-step plan of how the data for the study is collected and the benefit of using the various statistical techniques used are presented. Finally, the measures taken to ensure reliability, validity and possible generalizability of data analysis and findings are enumerated.

Chapters 4 and 5 present and discuss the results of the study. Conclusions from the findings of the study are also made. Finally some of the implications of the results and findings, for developing a more coherent policy on funding for science education, are suggested. Future areas for research, as extensions of this study, are also presented.

CHAPTER 2

REVIEW OF LITERATURE

The purpose of this study was to describe and to explain the relationships between educational inputs and measures of student achievement as outputs. The relationship between input variables and student performance measures is in the domain of education productivity. This chapter identifies and summarizes the existing body of literature, on education productivity, in general, and education production function on each of the input factors in relation to student performance, in particular.

Education Productivity

Education productivity can be defined as the measure of the educational output(s) relative to a given set of educational input(s) (Monk, 1992). There are various approaches to measuring education productivity, but three are most popularly used. The three most popular are cost-effectiveness analysis, human productivity or human capital, and production function studies.

Cost-effectiveness Analysis (CEA) Studies

Cost-effectiveness studies examine the cost of inputs for a set of educational programs, compare the cost with the benefit or output measures of the respective programs, and determine the relative cost-effectiveness of the programs (Grissmer, 2002; Hanushek, 1999; Harris, 2002; Levin, Glass, & Meister, 1987; Levin & McEwan, 2001). Cost-effectiveness analyses are mostly used as decision-making tools for program evaluation, implementation, or discontinuation (Levin & McEwan, 2002). McEwan (2002) identified one of the major limitations to the application of this approach in education policy decisions as the lack of methodological standards. McEwan

(2002) stated, "…there are few attempts to define methodological standards for a proper CEA in education, or to evaluate whether the extant literature adheres to these standards. As a consequence, researchers have few guideposts when conducting new research, and decision makers cannot determine how existing studies should inform policy" (p.38). The absolute reliability of cost effectiveness study in education policy decisions is still uncertain.

Human Productivity Studies

The focus of human productivity studies is on the totality of inputs into education compared with the life-long earnings and social contributions of the educated individuals to society (Becker, 1964; Cohn & Geske, 1990; Johnes, 1993). It has long been recognized that education has a positive relationship to human productivity and that wages paid are based on the perceived present and future value of the individual's contribution to the value-added position of an organization or society (Johnes, 1993).

Dublin and Lotker (1946) developed a mathematical expression that estimates the net value of an individual by taking into account the cohort-time-series estimate of discounted earnings over the working life of a person in the cohort. Later, Becker (1964) proposed that an investment in education is like an investment in a machine, which can be fitted onto a human body, thereby improving performance and productivity. The future returns after fitting the machine, that is education, are expected to outweigh the cost. Since Becker, the theory of economics, that views humans as capital, has gradually gained grounds in education production and cost considerations, in what is now known as human capital theory (Johnes, 1993). Again there are problems attached to the use of this approach in informing education policy.

Human productivity studies require a longer time, usually over a lifetime, to estimate productivity or return on investment of the cohort under study, and the methodology requires complicated mathematical calculations (Johnes, 1993). The two approaches to education productivity discussed above look at the monetary value of education programs and the productivity of the student at the end of his or her education career through working life. The third approach is the education production function approach.

Education Production Function Studies

Education production function examines the productivity of the resources put into the educational process relative to the educational performance of the students over a given period. Education production function studies use linear regression, correlation, or multivariate-discrminant analyses to indicate the nature of the relationship between input and output (Monk, 1992; Murnane, 1975). The focus of such study is to explain the relationship between selected educational input variables and selected student achievement measures, as outputs, during an educational process. Education production function, therefore, considers resource input against student performance, as output, within a snapshot period of time.

Backgrounds to Education Production Function

The interest and use of production function research in education began around 30 years ago with the large-scale study by Coleman et al. (1966) and reported in the Equality of Educational Opportunity Report (EEOR), commonly called the Coleman Report. The study used a representative sample of students and schools across the U.S. The findings of the study were both revealing and controversial.

First, the findings of Coleman Report (Coleman et. al., 1966) confirmed the replacement of *de jure* segregation with *de facto* segregation in almost all school districts across the country. The report noted that *de facto* segregation was not a “southern problem” only. Second, at every grade level, White majority children performed better than minority children on achievement measures. Third, Asian American children performed as well as, if not better than Whites. Fourth, on the balance, African American students attended schools with fewer facilities than White students. The conclusions from this study were: (1) money matters in determining student achievement, but it was not the only factor, (2) although, the socio-economic background of students is one of the most important determinants of educational achievement, it was not the only determinant. Hence, they concluded that:

> Schools bring little influence to bear on a child's achievement that is independent of his background and general social context; and that this very lack of independent effect means that the inequalities imposed on children by their home, neighborhood, and peer environment are carried along…….Equality of educational opportunity, through the schools, must imply a strong effect of schools that is independent of the child's immediate social environment, and that strong independence is not present in American schools. (Coleman et al., 1966, p. 325)

Therefore, there are other contextual determinants of student achievement that need to be factored into the educational input-output-production mix beyond money.

Two other large-scale studies by Mortimore et al. (1988) and Smith and Tomlinson (1989), were carried out in the United Kingdom. The two studies relate various educational inputs, including money, to the performance of students in primary and secondary schools. They concluded that other factors, in addition to money, contributed to the performance of students in the schools. Around the same period, Montmarquette and Majseredjian (1989), used a two-way analysis to establish the relationship between educational outcome in grades 1 and 4 and a combination of observable and unobservable class-level inputs. They concluded that there was a moderate relationship between unobservable classroom-level inputs and educational outcomes. In a related study, Goldhaber & Brewer (1997a) explored the relationship between observable-unobservable school-teacher factors and the performance of students in 10th grade mathematics.

They came to the same conclusion as Montmarquette and Majseredjian that there was a moderate relationship between observable teacher factors and student achievement.

Wenglinsky, (1997), used a national database from the Educational Testing Service (ETS) to model the relationship between school district expenditures, school resources and student achievement. He applied a multilevel estimation program called Hierarchical Linear Model (HLM) to school-level independent variables and a student-level dependent variable. He concluded,

> Expenditures on instruction and central office administration affect teacher-student ratios, which in turn affects student achievement. On the other hand, capital outlays, school-level administration and teacher education levels were found not to be associated with student achievement. (p. 115)

This statement represents the nature of the debate on whether there is a relationship between educational input and output, and if so what is the nature of the relationship?

The Debate on Input versus Output in Education Production Studies

Hundreds of studies have been reported on education production (Hasci, 2002). Murnane (1975) and Hanushek (1979, 1981) began the debate on whether there is a relationship between educational input and output. Hanushek (1981) argued that despite increased spending on education in the U.S., over several decades, student achievement had stagnated or fallen.

To test the veracity of his position, Hanushek conducted a follow-up comprehensive meta-analysis study on 141 published studies on education productivity (Hanushek1986). He found that only five out of 55 studies on per-pupil expenditure indicated a significant relationship between expenditure and student achievement measures. Three out of the 55 actually indicated a negative relationship. He concluded that there was no strong or systemic relationship between school expenditures and students performance. Hanushek did not stop here. In 1989 he again published another meta-analysis study in which he carried out 187 separate analyses. Again, he concluded that variations in school expenditure did not bear any systemic relationship to student achievement (Hanushek, 1989).

Hedges, Laine, and Greenwald (1994) did a follow-up analysis on Hanushek's studies by examining his methodology and reanalyzing his data. Their conclusion disagreed with Hanushek's. They opined that the data were more consistent with patterns that suggest some positive relationship between the dollars spent on education and education output than patterns of no effect or negative effect. They reported more than twice as many positive estimates than negative estimates of effects of expenditure per pupil and student achievement. They noted that when the studies were combined, the body of literature points to significantly positive effect of education expenditure on students.

Another review of the literature on school resources and student performance, compiled by Alexander (1997), confirmed the earlier conclusion by Hedges et. al. (1994) that school resources were systematically related to student achievement. A later study of Pennsylvania school system by Klick (2000) concluded that dollars, when considered in conjunction with the socio-economic background of students, definitely do make a difference. Klick further concluded, "….the difference is not necessarily in the schools; instead, it is in the home. Poverty, without fail, proved to be a significant determinant of whether or not a student will succeed in school" (p. 85).

Another empirical study reported by Ferguson (1991) indicated that educational resources accounted for between 25% and 33% of the variations among Texas school district in students' scores on statewide-standardized reading examinations. Additional expenditures have also been shown to be associated with standardized test scores (Sander, 1993). Williams (1999), in his review of the influence of education finance analysis on education policy, concluded

> Yes, money matters, but it matters most when it is used thoughtfully in productive ways, backed by systemic evaluations to assess effectiveness. Otherwise, money may make marginal differences but mostly be spent ineffectively, at least in terms of improving student achievement. (p. 236)

Further support for the nature of the relationship between educational input and student achievement was reported in the works of Unnever, Kerckhoff, & Robinson (2000) where they submitted that schools with more expenditure per pupil were associated with higher test scores and higher rates of continuing education after correcting for ability and socio-economic status of students in New York schools. They also found that high wealth districts, as determined by assessed property value, had more resources than low wealth districts. Districts with more resources had more favorable student outcomes, on some outcome measures, than districts with less resources.

These findings were inconsistent with Hanushek's earlier statement that:

> Over the past quarter century, researchers have made the surprising discovery that there is little systemic relationship between resources and student performance. For every study that finds that increases in basic school resources promote higher achievement, another study shows just the opposite. (Hanushek, 1994b, p. 12)

In his commentary in response to Hanushek, Kremer (1995), agrees that, " …the impact of additional resources varies with the circumstances, ... Because different inputs are effective in different circumstances, it may be useful to decentralize spending in education" (p. 251).

Explanations for the Inconsistencies Between Input and Output in Production Function Studies

In general, the results of empirical studies on the effect of educational inputs on student performance are still inconclusive. In some studies, the results indicate that the benefits of improved educational performance exceeds the cost (Greenwald, Hedges, & Laine, 1996; Klick, 2000; Walberg & Fowler, 1989), while in others, the reverse is the conclusion (Hanushek, 1986; Krueger, 1999). Questions still remain that, although it can be argued that resources are associated with student performance, but can this association be explained by other factors outside of the specified model? What is the optimal educational input threshold that maximizes educational output?

Catterall (1997) summarized the difficulty inherent in showing that money matters in education. Catterall noted that the typical economic production function approach is not adequate for assessing all the educational input and output variables. He also indicated that the problem is compounded because most production function studies focus on resource availability rather than on productive capacities such as skills and time related efforts that actually determine education outcomes. Michell (1998) extended Catterall's argument and suggested four conceptions of education as a productive enterprise made up of many functional attributes. He argued that functionally, education can be seen as a product, a service, an investment, and a cultural identity tool.

Michell, (1998) then concluded that the 'product' function of education, measured by the knowledge acquired (in grades and test scores) is more rational, easily manifested and documented, and therefore easy to criticize. The 'service' function measured by learning experience, engagement, and participation, is more difficult to quantify. The 'investment' function as measured by 'improved life changes,' 'social mobility,' and 'improved social responsibility,' is also less obvious and difficult to quantify. The last one, 'cultural identity,' measured by parameters such as 'respect for cultural values,' 'morality,' and other abstract indices, are even less tangible and thus very difficult to quantify.

These varied functional variables may be the reason why Card and Krueger (1996) further contended:

> Although school resources tend to be positively associated with educational attainment, the relationship is not always concrete to the set of data or the empirical specifications in the studies. (1996, p. 27).

They also concluded that:

> A study of the economic outcomes of education requires the incorporation of a theoretical framework that embraces the diverse interactions between the student's family background, school inputs, and educational achievement. (Card & Krueger, 1996, p. 28)

Mancedon and Bandres (1999) also enumerated the list of exogenous variables that contribute to both education input and output. Among the variables they identified were the characteristics of the education production process itself, some of the immeasurable intangible outputs of education, the cumulative nature of human capital, and the personalized experiences of the individuals. These personalized attributes can also be considered as both input and output in the education process.

In summary, Mancebon and Bandes stated:

> These considerations suggest that the education sector has a production process that is hard to disentangle … and difficult to determine output that is itself influenced by numerous elements which lie outside the formal education context (the socio-economic environment of family, innate abilities, accumulated human capital, etc.). As a consequence, the attempt to achieve a single, universal and exportable specification of education production technology is shown to be a truly controversial task. (p. 134)

These diverse variables highlight the heterogeneity and the non-standardized nature of the education and the "tacit and idiosyncratic character of education production" (Murnane & Nelson, 1984, p. 258). Therefore, most of the later studies on education production function have tended to focus on disaggregating interacting inputs and outputs with the hope of making individualized or partial inferences, specific to particular situations under study.

In line with the concept of disaggregation of outputs and inputs, Coleman, Easton, and LaRocque (1998) conducted and reported a study that made a distinction between the productive use of resources and the mere presence of such resources in individual schools. They opined that schools that used available resources effectively showed better and improved student achievement relative to the available resources, than schools that did not use available resources effectively.

Murnane and Levy (1996), studied 16 elementary schools in East Austin, Texas. Each of the schools were given $300,000 a year for 5 years to spend at their discretion, but with the objective of improving student achievement. At the end of the five years, the results indicated that 14 out of the 16 schools used the money given to them in ways that made little or no difference in student achievement. These 14 schools used the money on extra teachers to reduce class sizes without changing what goes on in the classroom, where learning takes place. The remaining two schools used their money in ways that showed a significant increase in student achievement. These two schools spent their money on directly engaging students, teachers, and parents to focus on improving student achievement, leading to improved instructional practices in the classrooms.

Rice (1997) disaggregated the costs and outcomes of programs by identifying and justifying the distribution of cost and outcome across stakeholders including teachers, students, support services, and administrators. She concluded that each of the stakeholder categories had elements of waste in relation to education costs and student outcomes. She contended that these elements of waste diminished the possible impact of educational inputs on student achievement.

Variable Specification in Education Production Function Studies

One of the criticisms that have been levied against the use of production function in education productivity research centers on the lack of adequate and relevant variable specification (Fortune & O'Neil, 1994). The critics say that researchers often fail to control for nested errors among input factors (Griffin & Ganderton, 1996; Hanushek, Rivken, & Taylor, 1996). Nested errors are errors that may arise due to confounding between variables, which may occur when there is possible relationship between the factors considered and some other factors that are not specified as variables in the production function.

Another criticism is the need to establish possible co-variation among the components of the production function considered. To establish a statistical relationship between two variables, there must be a possible covariance association between them. Co-variation is not often significantly present among sets of data characterized by internal homogeneity, for example within one school or within a school district (Saber, 1984). Hanushek (1996b) found that studies of production function that use cross-state variations in school resource typically find significantly positive effects, whereas studies that use within-state or district variation are more likely to find the positive relationship insignificant or even a negative association.

<u>The Need to Track Resources to School-level</u>

Most of the earlier studies on the relationship between money and student achievement have focused on expenditures either at the state or district level (Hanushek, 1997). One of the criticisms levied by Card and Krueger (1996) against the empirical studies reported, in their review of literature on educational inputs and student achievement, was that most of the studies relied more on district or state-level data rather than on school-level or classroom-level information.

Looking at school-level data allows one to follow inputs to where they are used and also increases the chances of exposure to dramatically different education input availability and utilization, which increases variability. Few studies have gone as far as looking at the relationship between educational input and outcome measures at the individual school level (Cooper et al., 1994; Heinbuch & Samuels, 1995; Murnane & Levy, 1996).

Heinbuch and Samuels (1995) used the School-Site Allocation Model (SSAM), developed by Cooper et al. (1994) to examine and describe the influence of school-level input such as teacher experience and per pupil expenditure on SAT scores of students in New York Schools. Their results suggested that student performance on the SAT could be expected to increase with increased spending on instruction in the classroom. In one of their concluding statements, they stated,

> Our findings suggest that more bang for the educational buck can be achieved if policy makers and educators redirect their attention to the efficiency of current fund flows from the central administrative level to the actual classroom settings, in which education is to occur and where productivity can be revealed. (p. 238)

The results of the study of 14 schools in the East Austin, Texas, by Murnane and Levy (1996), discussed earlier, also lay credence to the need to trace resources to the school or even to classroom level.

The Inputs and Outputs

Almost all of the studies reported on education production function consulted so far examined the relationship between two or more educational input variables and one educational output measure, usually student test scores (Figilo, 1999; Hanushek, 1997; Hasci, 2002). However, education, as a process, requires various forms of input and also produces various measures of output. Therefore, the use of a single variable input or a single output measure does not capture the whole scope of education production function.

The purpose of this study was to describe and to explain the relationships between educational inputs and measures of student achievement as outputs. This study used ten input variables and two educational output measures. The inputs were: (1) Expenditure Per Pupil (EXPPP), (2) Teacher Quality (TEACHQLT), (3) Percentage of Economically Disadvantage Students (PEDS), (4) School-level Science Department Expenditure Per Pupil (SCIEXPPP), (5) Science Teacher Quality (SCITCHQT), (6) Science Class Size (SCCLSIZE), (7) Science Lab-based activities per Teacher per Week (LABSPWK), (8) the number of Professional Development Activities, per Science Teacher per Year (PDVSCTCH), (9) Principal Stability (PRINSTBY) and (10) Head of Science Department Quality (HEADSQLT). Additional school-level qualitative parameters, including suggestions for improving science education, were also investigated for possible relationship to school performance categories.

The two measures of student achievement, considered as outputs, for this study were: (1) The percentage of students who passed the science section of the Georgia High School Graduation Test at first sitting (GHSGT) and (2) the percentage of graduating students qualified for the state of Georgia's HOPE scholarship award (HOPE).

The Inputs

Inputs, as considered in this study, are the educational resources, both quantitative and qualitative, that go into the educational process to achieve a desired educational outcome for students. The literature on each of the inputs considered in this study varied from one input to another.

Expenditure Per Pupils (EXPPP)

A systemic effort to determine the relationship between expenditures and student achievement gained public awareness after the publication of Equity of Educational Opportunity (Coleman et al., 1966). As discussed earlier, Coleman and others, found that resources including expenditures had only a small impact on student achievement. Since then, Hanushek (1979, 1981, 1986, 1989, 1991, 1994a, 1994b) has published a series of synthesis analysis of the literature on education production function and concluded that the data he used did not provide enough evidence of a systemic relationship between educational inputs and student achievement.

However in a follow-up reanalysis of Hanushek's data, Greenwald, Hedges, and Laine (1996) concluded that a broad range of school inputs was positively related to student achievement. They wrote,

> After reanalysis of Hanushek's evidence, our position was that the data he (Hanushek) assessed on the relations between school resource inputs and student outcomes, including achievement, were substantially more consistent and positive than he believed. We found that the typical relation between input and outcome in the data he considered was positive and large enough to have important implications for educational policy. (p. 363)

In a more recent study, Klick (2000), used data from 501 school districts in Pennsylvania to look at the relationship between funding and achievement while controlling for economic background of each school's student population. The input variables he considered were total expenditure per pupil, percentage of student body coming from low-income households, size of school district, and occurrence of strikes within the district. The finding of the study showed that funding made a difference. However he noted that poverty, "….proved to be a significant determinant of whether or not a student will succeed in school" (p. 85).

The conclusions from these study indicated that yes, per pupil expenditure (money) could make a difference in student achievement, but the difference is not automatic (Hasci, 2002). This view was better summarized in Hasci's comments, "...we need to know how to use money wisely, and unfortunately school finance evaluations offer limited guidance on that question" (p.194). It is therefore important to examine closely other inputs, in addition to money, in a production function study.

Teacher Quality

Teacher quality, as an input, has been determined, in some studies, to have some effect on student performance (Figilo, 1999; Hanushek, 1971). Most of the literature consulted so far (Boardman, Davis, & Sandy, 1977; Figilo, 1999; Fried, Lovell, & Schmidt, 1993; Goldhaber, Brewer, & Anderson, 1999; Gyimah-Brempong & Gyapong, 1992; Mancebon & Bandres, 1999) only took into consideration either the years of service of the teachers, the highest qualification of the teachers, or teacher salary separately, as measures of teacher quality. Other researchers have also extended the definition of teacher quality to include such factors as verbal expression ability and scores on standardized teacher certification tests (Firestone, 1991).

The National Center for Education Statistics (U.S. Department of Education, *NCES*, 1995), first released, in the 1990s, a major data source on teacher quality known as the Schools and Staffing Survey (SASS). The SASS was prepared based on sampling of 55,000 teachers from over 11,000 schools. The SASS provides data on the nature and structure of staffing in the education sector across all educational establishments in the country. The latest report, released in May 2002 (U.S. Department of education, *NCES*, 2002b), examined the quality of teachers through the parameters of qualification, experience, and out-of-field teaching. The SASS reports indicated that there is little consensus on what constitutes teacher quality. As a compromise, SASS presented a wide range of indicators of teacher quality including, teacher preparation training, qualifications, experience, in-field/out-of-field teaching, and other job conditions.

In the area of science education, the National Science Foundation (NSF) commissioned a study called the National Survey of Science and Mathematics Education (NSSME) and reported in Oakes (1990). An extensive study by Oakes (1990) using the NSSME data showed a high prevalence of out-of-field teaching among science teachers. The conclusion from this study is that out-of-field teaching in science had some negative implications for the quality of science teachers and science education.

The multidisciplinary nature of science requires that a science teacher be adequately knowledgeable in the content of the specific area of science taught. A teacher whose major content area at college is physics may not be able to competently teach biology. It is believed that one of the reasons why science achievement is low is due to the prevalence of out-of-field teaching (Ingersoll, 1999). In his commentary on the problem of teacher quality Ingasoll (1999) wrote, "at the 12th-grade level, 41% of public school students in physical science classes are not taught buy someone with either a major or a minor in chemistry, physics, or earth science" (p. 21).

The present study measures teacher quality as a composite of both experience and highest qualification. The composite measure removes the unnecessary upward bias that results from long service with minimal qualification, or from maximum qualification with minimum length of service and experience.

Percentage of Economically Disadvantaged Students

An education production function study is not complete without a measure of the socio-economic status of the students (SES) as a factor (Alexander, 1997). Almost all the studies in education production function involving the use of the socio-economic background of the students conclusively evidenced the significant impact of SES on student achievement (Coleman, et al., 1966; Greenwald, Hedges, & Laine, 1996; Grismmer, Flanagan, & Williamson, 1997; Hanushek, 1996a, 1996b; Heinbuch & Samuels, 1995; Klick, 2000; Monk, 1992; Wenglinsky, 1997). Klick (2000) concluded that, "Poverty, without fail, proved to be a significant determinant of whether or not a student will succeed in school" (p.85).

Many measures of SES have been identified (Ferguson & Ladd, 1996; Mancebon & Bandres, 1999; Michel, 1991). Some of these measures include number of students eligible under federal Free-Reduced Meal programs, parents occupation, parents education level, crime rate, average property value, proportion of households that own houses, etc. However, Gyimah-Brempong and Gyapong (1992) in their study of education production functions pointed out that using many measures of SES, simultaneously, as independent variables to predict an educational output often create co-linearity problem. They suggested the use of only one measure of SES that serves as the best proxy for all the other SES measures. In their study, they used the median family income as a proxy for SES.

Other production function researchers (Hamilton, 1983; Hanushek, 1997; Summers & Wolf, 1977) also argued that SES should be included as input because it influences the educational outcome by complementing the purchased school inputs. In much of the previous research studies investigated so far, SES has been used as the controlling variable (Figilo, 1999; Hanushek, 1986, 1989, 1991; Mancebon & Bandres, 1999; Monk, 1992; Walberg, 1982). These researchers also pointed out that the inclusion of SES made it difficult to identify the independent effects of school-related resources on education output.

SES may also bear a relationship to the amount of resources available for science education in the schools. In other words the SES is a measure of the relative 'quality' of the student body in each of the sample schools. Therefore, unlike previous studies, SES is considered as an input as well as a confounding factor in this study; instead of being considered only as a intervening variable (Heinbuch & Samuels, 1995; Michell, 1991). For the purpose of this study, SES was measured as the percentage of students eligible for the federal Free-Reduced Meal program in the schools, now referred to as the Percentage of Economically Disadvantaged Students (PEDS).

The State of Georgia educated over 1.4 million students comprising of 55.5% White, 38.2% Black, 4.0% Hispanic, 2.1% Asian/Islanders, and 0.2% Native Americans during the academic year 2001-2002 (Georgia Department of Education, Report Card, 2003). About 23% of the students are from households below the national poverty income limit and over 44% of students in public schools are on free-reduced lunch program (Georgia Department of Education, Report Card 2003). The percentage of economically disadvantaged students (PEDS) is therefore an important factor to consider in the any education production function relating to education in the state of Georgia.

Class Size

Class size is the number of students regularly in a teacher's room for whom that teacher is responsible each day of the school year (Finn & Achilles, 1999). Class size has been viewed as an input variable that affects student performance. The largest experimental study on the effects of class size reported so far was the Student-Teacher Achievement Ratio study (Project STAR), commissioned in 1985 by the Government of the State of Tennessee under the then Governor Lamar Alexander (Word et al., 1990).

The Tennessee STAR experiment involved almost 12,000 students in grades K-8 in 329 classrooms from 46 school districts in the State of Tennessee. In addition to other outcome measures, student performances in a battery of standardized tests were also measured. The findings of the STAR experiment included the following. First, class size produced significant effects on student achievement in lower grades, but the effects diminished with increasing grade-level of the students. Second, the benefit of the effect of class size is more significant among students from low socio-economic backgrounds. Third, the effects of class are most significant for class sizes less than 20, typically between 17 and 20. Several studies on the effect of class size on student achievement have since been carried out, using the data from the STAR experiment (Finn, & Achilles, 1999; Grissmer, 1999; Hanushek, 1999; Hanushek, 1998; Mostella, 1995; Nye, Hedges, & Konstantopoulos, 1999; Robinson, 1990; Rice, 1999). The findings of these studies also pointed to the need for reduction in class sizes below 20 to elicit any significance gains in student achievement.

However, the issue of the critical class size, required to maximize student achievement, still remains unresolved especially for students in higher grades above K-3. Speaking to this issue, Hanushek (1999) stated,

> The results show effects that are limited to very large (and expensive) reductions in kindergarten or possibly first grade class sizes. No support for small reductions in class size (i.e., reductions resulting in class size greater than 13-17 students) for reductions in later grades is found in the STAR results. (p. 143).

Rice (1999) also stated, “Class size has an impact on the use of class time, both instructional and non-instructional. The effect varied by subject area, type of student, and the amount of time teachers spend planning for class” (p. 225).

In the State of Georgia, the mandated maximum class size for science is 28 students. However, the peculiar nature of science instruction, requiring laboratory and hands-on modules, makes the issue of class size a critical factor to consider in the teaching of science and in student performance considerations in science. No reported literature on the effect of class size on student performance in science in the state of Georgia was found.

Lab-based Instruction in Science as an Important Input

Science is a knowledge domain that deals with the understanding of observable physical and natural phenomena. As a result, pedagogy in science requires constant interaction with these observable phenomena. These interactions are only possible through physical, often tactile interactions with pedagogical materials, to enable students to conceptualize scientific principles and methodologies. Hence the frequency of student interaction with physical manipulatives may be an important factor in ensuring understanding and impacting achievement. There is limited reported literature on the relationship between the number of hands-on pedagogical activities in science and student performance. O'Sullivan and Weiss (1996) reported that the majority of high school students who did hands-on activities at least once a week were more likely to perform above proficiency level than those who did not. This study also examined the impact of the frequency of hands-on science laboratory activities as an input in the production mix.

Professional Development

The Council for School Performance (1998) defined professional development in education as, "An organized (on-the-job) learning opportunity for teachers to acquire knowledge and skills to help them become more effective teachers" (p. 3). According to Irving, Dickson Jr., and Keyser (1999), the series of curriculum reforms in the country have stressed moving from traditional instruction that emphasizes memorization to inquiry-oriented methods. Stofflett and Stoddart (1994) reported that science instruction in schools was dominated by lecture, demonstration, textbook reading, and memorization at the expense of seriously engaging students' interest in the conceptual understanding of science. For the foregoing reasons, traditional professional development practices have been shown to be ineffective in promoting student achievement in science and mathematics (Darling-Hammond & McLaughlin, 1995; Edwards, 1997; Stofflett & Stoddart 1994).

Darling-Hammond and McLaughlin (1995) suggested that, “staff (professional) development must focus on the serious task of learning the skills and perspectives assumed by new visions of practice and unlearning the practices and beliefs about students and instruction that have dominated their (educators) professional lives to date” (p.597). In Their book, *A New Vision for Staff Development,* Sparks and Hirsh (1997) stressed the need for major shifts in staff development practices. They recommended shifts, namely:

> From individual development to individual development and organization development. From fragmented efforts to staff development driven by clear, coherent strategic plan for school district, each school, and departments that serve schools. From district-focused to school-focused approaches to staff development. From a focus on adult needs and satisfaction, to focus on student needs and learning outcomes, and changes in on-the job behavior. (p. 12).

There is limited reported study on the effect of staff development and student achievement in the state of Georgia. The study conducted in Georgia by the Council for School Performance (1994) on the effects of staff development on student achievement, reported non-significant impact of staff development on student achievement, as presently conducted in the state’s educational arena. Therefore, there is a need to examine professional development as part of the input into the education production function mix and its impact on student achievement in science.

Principals’ Stability, Leadership Quality, and Resource Allocation

There is very limited reported study on principal stability and student achievement. Argyris (1993) noted that individuals in an organization, including the leaders, need time to ‘learn’, ‘mature’, create a state of independence, have a deeper interest in what they do, and refine visions. As leaders become more stable, their organizational 'maturity', increases (Argyris, 1993). The increased maturity often translates to increased organizational performance. In their review of studies across 14 countries, Comber and Kreeves (1973) indicated that the school environment modifies student habits and, on the aggregate, predicts student scores in science across these countries.

In his study of allocative efficiency, Anderson (1996) used extensions of the Data Enveloping Analysis (DEA), to identify effective schools relative to resource allocation and utilization. He associated stable leadership, stable student population, attendance, and school climate with effective schools. Leithwood and Janzi (1999) examined the relationship of principal leadership to student engagement as a measure of output. They concluded there was only a weak relationship between principal leadership and student achievement. Leithwood and Janzi (1999) also showed that schools in which principals have some control over allocation of resources, were better performing schools that those where principals have less control over resource allocation, especially money.

Levin (1997), in his study on the concept of x-efficiency, first proposed by Leiberenstein (1966), examined the role of non-financial variables (x-factors) like stable leadership, motivation, and other organizational dimensions on the allocative efficiencies of schools. He concluded that there were improved allocative efficiencies in schools that focus on improving these x-efficiencies. Burtless, (1996) in his book, *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, stated that, "…variations in the level of school spending are less important than effective organization of school resources in determining whether spending differences have important consequences for student outcome" (p. 11).

Therefore, stable principals are more likely to create the appropriate climate for learning to occur with some influence on resource allocation and thus influence student achievement.

Heads of Science Departments' Quality

A search of the literature did not reveal any reported study on the influence of the quality of heads of science departments on student achievement. Most heads of science departments are also science teachers (U.S. Department of Education, School and Staffing Survey, SASS, 2002). Therefore their influence on student is dual, both as classroom teachers as well as heads of the science departments. A measure of the heads of department quality should take to consideration both years of experience as teachers as wells as the years of experience as heads of the science departments.

This study is therefore also designed to describe and explain any empirical quantitative relationships between certain school-level leadership variables and student achievement.

<u>The Outputs</u>

Another problem faced by researchers in education production function studies is the determination of which outcomes constitutes the best proxies or estimates of student performance. As Mancebon and Bandes (1999) stated, "The peculiarities of education production, underscore the almost impossibility of arriving at a single and universal theoretical specification of education outputs that are valid for all educational situations" (p. 134). Therefore it is reasonable to consider a measure of output specific to the educational reality of the situation that is objectively acceptable to the audience who are to benefit from the research. With respect to secondary education, the specification of education production function output with reference to the cognitive domain (Bloom, 1976; Carroll, 1963), particularly to performance in specific subject area is appropriate.

Mancebon and Bandres (1999) gave three reasons why the use of performance in a specific subject area is particularly attractive in determining output measures for secondary school students. First, at the secondary school level, students' source of learning of specific subjects, especially science, is most likely to be from school-based pedagogy, with minimal instructional input from outside-of-school settings. Second, in secondary schools, there is more interconnection between subjects; and science in particular calls on different skills that the students have learned from the other disciplines of knowledge (Madaus, Kellaghan, Rakow, & King, 1979). Therefore, a student performance measure in science is a reasonable and appropriate estimate of education output.

Finally, because academic performance reflects directly on the main objectives and the public perception of secondary schools, this measure is often the acceptable yardstick for education productivity at the secondary school level (Fogelman, 1984a, 1984b). The evaluation of schools by students, parents, community, and policy makers is based on test scores (Hanuschek, 1979).

In this study, the test scores of high school students in the science section of the Georgia High School Graduation Test (GHSGT) was used as one of the three measure of output (student performance) in the production function model.

The Georgia High School Graduation Test (GHSGT)

The Georgia High School Graduation Test (GHSGT) is administered by the Georgia Department of Education (Georgia Department of Education, Report Card, 2003). The test includes assessment in the core subject areas of Science, English, Mathematics, and Social Studies. To obtain a pass in each of the core subject area, a student must score between 500-600 on a scale of 400-600. The requirement for the award of a Georgia High School Diploma is a pass in all of the subject areas specified, including science, as well as a pass in the Georgia High School Writing Test (GHSWT). The performance Report Card from the state reports only the percentage of student passing the GHSGT at first sitting. Repeat test-takers are not counted as part of the performance scores. No available literature was found on the relationship between educational input and student performance in the science section of the GHSGT.

The Georgia HOPE Scholarship

The Georgia HOPE Scholarship Program was established by an act of legislation in 1993 (Council for School Performance, 1999). The legislative intent of the program was to provide tuition and book assistance to Georgia residents pursuing post-secondary education in the state's universities and colleges ("Georgia HOPE scholarship program," 1997). To qualify for HOPE during a student's first year in college, the student must have an exit GPA of not less than 3.0 at graduation from high school in the state. The award is not means tested, so everyone who meets the sole GPA criterion gets the scholarship. Therefore, the percentage of graduating students qualified for this scholarship is an important and relevant measure of student achievement.

There is limited available study on the HOPE scholarship. One study commissioned by the Andrew Young School of Policy Studies and Applied Research Center, Domestic Programs Studies (2001), only evaluated the HOPE scholarship in terms of the program's affective impact on student attending public colleges and Universities. Another study commissioned by the Council for School Performance (1999), reported increased expectation for higher education, by students, as the significant impact of HOPE on students. Again, no reported study was found on the relationship between educational inputs and percentage of HOPE recipients from the high schools in the state of Georgia.

None of the literature consulted so far has done any empirical studies on the effect of educational resources on student performance and productivity in science at individual high school level, in the state of Georgia. This research began the study of production function in education by constructing a theoretical framework that combined the relevant contextual factors of educational input (resources) with educational outputs. From the theoretical framework constructed, a set of model(s) was built and used to test a set of hypotheses on whether and which recourses mattered in the performance of student.

Production Function Models

Production function model studies use a range of multivariate analyses including canonical analysis and multivariate regression. Canonical Analysis (CA) was developed by Hotelling (1936) to study the relative strength of the relationship between two sets of variables. The conceptual reasoning behind CA is to form two linear combinations of the input variables and output variables by differentially weighing them to obtain the maximum possible correlation between the linear combinations (Pedhazur, 1997).

The correlations between the linear combinations are collectively called the Canonical Correlation (R_c). Gyimah-Brampong and Gyapong (1992) used canonical regression in place of the separate regression model as a subset of the multiple regression analysis. The use of canonical regression was appropriate for their study because they were interested in more than one dependent variable as measures of educational output. Unlike most of the other production function models cited earlier, the focus of the present study was on more than a single education output measure as the dependent variables. Therefore, canonical and multivariate regression analysis was appropriate for the study.

Multivariate regression is used to describe the linear model between two or more independent interval variables and two or more dependent interval criterion variables (Stevens, 1996). First, the multivariate model examines the significance of the overall multivariate regression model. Second, the individual interpretation of each independent variable is conducted, in a multiple univariate sense, to see how well the individual predictors compare on each of the dependent variables when considered separately. Stevens (1996) stated, that "…. although the multivariate tests take into account the correlations among the dependent variables, the regression equations are those that would be obtained if the dependent variables were regressed separately on the set of predictors" (p 131). Thus, in deriving the prediction equations, the correlations among the dependent variables are ignored, or not taken into account. To establish the nature of the relationship between the composite of the input variables and the composite of the performance measures, the multivariate significance of the model is tested first. Then the univariate significance of the relationship between each of the performance measures is tested separately against the each of the input variables.

Summary

The purpose of this study was to describe and explain the relationships between educational inputs and measures of student achievement as outputs. This chapter identified and summarized the literature on education productivity in general and each of the input factors in relation to student performance in particular. Research showed that there is still a lot of controversy on the issue of the relationship between education input variables and student performance. Most importantly, the review also indicated that despite the multivariate nature of education input and output, most of the previous research is dominated by single input and output (univariate) measures. Little or no previous work had been done to establish the multivariate relationship between sets of quantitative and qualitative input variables and high school student performance. Chapter 3 presents the rationale for the study with a detailed definition of the variables identified, the hypotheses tested, the context, and the data collection process.

CHAPTER 3

METHODOLOGY

The Rationale

The purpose of this study was to describe and to explain the relationships between educational inputs and measures of student achievement as outputs. The study investigated both quantitative and qualitative educational inputs and their relationships to two output measures of student achievement.

The research design selected for this study is a non-experimental design to identify, describe and explain the stated relationships. This design was chosen because the researcher examined existing data on an existing population without the application of any experimental manipulation or interventions. In a non-experimental design, causal inferences may not be made (Cook & Campbell, 1979; Pedhazur & Schmelkin, 1991). Cause-effect inferences may not be made from the results, except in situations where relationships between the input and the output variables are highly significant both statistically and practically (Pedhazur & Schmelkin, 1991). However, the benefit of this type of design is that larger samples, based on availability, can be collected thus improving reliability and generalizability (Kerlinger, 1986). The study is organized, conducted, and analyzed at two levels: across school districts and across schools within districts.

Across-Districts Study

The purpose of the across-districts level study was to describe and explain the relationships between quantitative educational inputs and outputs across school districts. The input variables investigated at this level were: (1) Expenditure Per Pupil (EXPPP), (2) Teacher Quality (TEACHQLT), and (3) Percentage of Economically Disadvantage Students (PEDS). The achievement output variable measures investigated at the district level were: (1) Percentage of student passing the Science Section of the Georgia High School Graduation Test (GHSGT) at first sitting and (2) Percentage of graduating students qualified for the Georgia HOPE Scholarship (HOPE).

The Context of the Across-Districts Study

The school districts used in the across-districts study are part of the public school districts in the north central region of the state of Georgia. The school districts are located in the metropolitan area of the largest city in the state. Six of the largest school districts in the metropolitan area were selected based on their sizes. The six school districts selected are the largest school districts in the state, in terms of student enrolment. These six school districts account for over 30% of the total number of students enrolled in public schools in the state (Georgia Department of Education, Report Card, 2003).

The six school districts had at total enrolled population of about 480, 000 students during the 2001-2002 academic year. There were 77 high schools in the six school districts with a total enrolled students population of about 125,000 comprising 26.3% of the totals enrolled students in the six districts in the 2001-2002 academic year (Adapted from the Georgia Department of Education, Report Card, 2003). The school districts were the units of analysis for the across-districts studies. The data used for the across-districts study were on the individual high schools in the six school districts. The individual schools were the sampling units.

Sampling for the Across-Districts Study

The school selection criteria for the sampling units for the across-districts study were based on the number of high schools in the districts with complete performance and financial report records in the archival database of the state of Georgia's Department of Education, on all the variables considered. The selected high schools also had complete records for the two consecutive school years covering 2000-2001 and 2001-2002 on all of the variables considered. High schools that exist to run special programs for non-traditional students, like open campuses and alternative education schools, were not included in the sampling units. Non-traditional students, as used in this study, refer to students who are severely challenged as to render them incapable of inclusion in regular high schools or students who cannot fit into regular high school program of education due to social or other contingent factors.

Sample Size for Across-District Study

The across-districts study data were analyzed using multivariate analysis methods. The issue of sample size in multivariate analysis deserves special consideration. The issue of desired sample size, for multivariate analysis (MANOVA), has been considered by Tatsuoka (1973), Lauter (1978), and Kres (1983). However, like the issue of desired sample size for univariate Analysis of Variance (ANOVA) addressed by Cohen (1988), Kraemer and Thiemann (1987), and Green (1991), the desired sample size for multivariate studies is still more or less based on the rule of the thumb (Kres, 1983). The sample size rules, for both multivariate and univariate analyses of variances, depend on: (1) the desired power, or effect size, of the statistics of interest (that is, the acceptable proportion of variation in the response variable to be explained by the grouping variable, (2) the probabilities of Type I (β) and Type II (α) errors, and (3) the number of grouping variables (p) involved. Effect sizes are usually denoted by R^2, η^2, or ω^2.

Lauter (1978), in his tables, suggested, that to do a reliable multivariate analysis test at α=0.05, and minimize the probability of a Type I (β) error at 0.30, for a medium effect size of 0.13, the minimum group size should be between 3p to 4p where p is the number of output variables. This study involved two output variables, for the purposes of multivariate analysis. Therefore, the desired minimum group size should be between 6 (3x2) and 8 (4x2). Each of the six school districts represented a group. Therefore, the desired sample size should be between 36 (6x6) and 48 (6x8). The objective minimum sample size needed for across-district data in this study was between 36 and 48 schools. The actual sample size used in this study was 71 high schools. Therefore, there were enough samples for robust multivariate analyses.

Data Collection for Across-Districts Study

The data on the input and output variables were collected on 71 high schools, within the 6 school districts, from the electronic archival Report Card records from the State of Georgia's Department of Education's web sites at: *http://techservices.doe.k12.ga.us/reportcard* and *http://techservices.doe.k12.ga.us/reports/finacial*.

To protect the anonymity of the school districts and for the purposes of ease of analysis, the six school districts were labeled TL, YT, BB, KL, LT, and WN in this study. Data for two consecutive school years, covering academic years 2000-2001 and 2001-2002, were collected. These were the latest two years for which both public education report cards were available at the time the data were collected during the months of July to August 2003.

Explanation and Determination of Inputs and Outputs for Across-Districts Study

All the data on the three input variables and the two measures of student achievement were collected from the archival records of the Georgia Department of Education as explained in the following sections.

Expenditure Per Pupil (EXPPP)

Expenditure Per Pupil (EXPPP) was determined by taking the Quality Based Education (QBE) earnings (state funds) for each sample high school in each district and dividing it by the total number of enrolled students in the school, at the October Full Time Equivalent (FTE) count, for the appropriate academic year. QBE earnings are state funds allocated to schools based on specified criteria. These are often called allotments. QBE state funds were chosen as the proxy for determining per pupil expenditure because all public schools in the state of Georgia get QBE allotments. The QBE allotments are made directly to the individual schools from the state with a stipulation that at least 90% of the allotment must be spent at the school site during the allotment year; otherwise the school district returns the balance to the state. Almost all schools and school districts in the state ensure that their full QBE allotments are spent at the school site, during the academic year of the allotment. Therefore, QBE allotment provides one of the best proxies for estimating per pupil expenditure in the sample schools in the sample districts.

EXPPP for each school in each district was determined by dividing the total state QBE funding allotment for each sample school (TSQBEFA) obtained from the Financial Reports of the Georgia Department of Education, by the number of enrolled students (N) in the school for each of the two academic years:

$$PPEXPNT = TSQBEFA/N$$

For example for a school with 2000 students (N), receiving QBE allotment funding amount of $5,000,000 (TSQBFA), the PPEXPNDT is calculated as $2500 (5000000/2000).

Teacher Quality (TEACHQLT)

Teacher Quality (TEACHQLT) was determined by transforming the state's data on teacher Training (Qualification) and Experience (T&E) for each of the sample schools. An example of the state's T&E data used is shown in Appendix A.

In previous studies (Figilo, 1999; Mancebon & Bandres, 1999), only the highest qualification of teachers or teachers' experience, considered separately, were used as measures of teacher quality. Using only one of these measures, as proxy for teacher quality, produces a downward or an upward bias depending on the relativity of qualification to experience for each teacher considered. The determination of teacher quality, in this study, builds on these previous works by using a combination of teacher qualification and teacher experience to determine teacher quality. The transformation calculation to obtain the values for TEACHQLT was calculated using the formula:

$TEACHQLT=(N_{q1}Q_1+ N_{q2}Q_2+ N_{q3}Q_3+ N_{q4}Q_4+ N_{q5}Q_5) +(N_{e1}E_1+ N_{e2}E_2+ N_{e3}E_3+ N_{e4}E_4+ N_{e5}E_5)/N_{ts}) \times 100$

where: Q_1 to Q_5 = Paraprofessional, Bachelor, Master, Specialist, and Doctoral qualifications respectively. Q values range from $Q_1=1$ to $Q_5=5$;

N_{q1}- N_{q5} = Number of teachers with the corresponding Q (qualifications) respectively;

E_1 to E_5 = Years of experience in each of corresponding Q categories respectively. E values range from E_1 =1 (less than 1year experience), E_2 =2 (1-10 years experience) E_3=3 (11-12 years experience) E_4=4 (21-30 years experience), and E_5=5 (greater than 30 years experience).

N_{e1}-N_{e5} = Numbers of teachers with the corresponding E (experiences) respectively;

N_{ts} = Total number of students enrolled in the school for the corresponding school year.

For example, the transformation calculation for a typical sample high school is shown in Table 1.

Percentage of Economically Disadvantaged Students (PEDS):

The numeric values for the Percentage of Economically Disadvantaged Student (PEDS) were the percentages of students eligible for the federal free and reduced price meal program in the schools as reported in the state's archival records.

Percentage of Students Passing the Science section of the GHSGT:

The numeric values for percentage of students passing the science section Georgia High School Graduation Test (GHSGT) at first sitting (GHSGT) were also the actual numbers retrieved from the state's archival data.

Percentage of Graduating Students Qualified for HOPE (HOPE):

The numeric figures for the percentage of graduating students qualified for the HOPE scholarship for each high school were again the actual numbers from the state archival records as reported by the schools at the end of July of the relevant school year. The qualification for the HOPE scholarship, as it relates to graduating students, is the attainment of a minimum average of 3.0 GPA at graduation.

<u>Hypotheses Tested for the Across- District Study</u>

The following hypotheses were tested for the across-district study data:

<u>Hypothesis #1</u>: There is no significant difference between school districts on the composite of educational input variables and the composite of achievement output measures.

<u>Hypothesis #2</u>: There is no significant difference between school districts on each of the educational input variables and the achievement output measures, considered separately.

Table 1: [a] Transformation Calculation for Teacher Quality

Qualification	Number of teachers (N_q)	x	Qualification Wt. (Q)	=	Value
Bachelor	12 (N_{q1})		2 (Q_2)		24
Master	20 (N_{q2})		3 (Q_3)		60
Specialist	2 (N_{q3})		4 (Q_4)		8
Doctorate	1 (N_{q4})		3 (Q_4)		3
Others	2 (N_{q5})		1 (Q_4)		2

Experience (E)	Number of Teachers (N_e)	x	Experience Wt. (E)	=	Value
< 1year (E_1)	1 (N_{e1})		1 (E_1)		1
1-10 years (E_2)	6 (N_{e2})		2 (E_1)		12
11-20 years (E_3)	12 (N_{e3})		3 (E_1)		36
21-30 year (E_4)	12 (N_{e4})		4 (E_1)		48
< 30years (E_5)	7 (N_{e5})		5 (E_1)		35
Total value ……………………………………………………………					233

If the total number of enrolled students for the school year (N_{ts})=627, then
TEACHQLT index = (233/627) x100 = 37.20

[a] Transformation is based on the Training & Experience (T& E) Scale for Teachers used by the Georgia Department of Education.

Hypothesis #3: There is no significant difference between school districts on the composite of educational input variables and the composite of achievement output measures, after controlling for Percentage of Economically Disadvantaged Students (PEDS).

Hypothesis #4: There is no significant difference between school districts on each of the educational input variables and the achievement output measures, considered separately, after controlling for PEDS.

Hypothesis #5: There is no significant correlation between the composite of educational input variables and the composite of achievement output measures across school districts.

Hypothesis #6: There are no significant correlations between the individual input variables and the achievement output measures, considered univariately, across school districts.

Hypothesis #7: There is no significant linear relationship between the linear composite of input variables and the linear composite of achievement output variables. Each of these hypotheses is represented in a schematic diagram in Appendix D.

Schools-within-Districts Study

The schools-within-districts study was conducted as a follow-up to corroborate or confirm the findings of the across-district study. The overall objective was to describe and explain the relationships between school-level quantitative input variables, on one hand, and the two output measures of student achievement on the other, as further explanation for the findings of the across-districts study.

The Context of the Schools-within-Districts Study

In order to get the required information for the schools-within-districts study, direct personal interviews with the school principals and the heads of science departments in the high schools within the six school districts were conducted. Each of the school districts was approached for official permission to conduct the study.

The departments of Research and Development, in each of the six school districts, were contacted and briefed of the intent to conduct the study in the high schools within their school districts. The necessary applications and documentations, including a prospectus of the study, were completed, packaged, and sent to each of the research departments for the research boards' review for the approval process.

Three out of the six school districts agreed to participate in the study. The reasons given by the three remaining school districts for not participating varied. The reason given by the first school district (coded WN) for not participating was that the study did not fit their current areas of research interest. The second school district (coded BB) declined to participate because their staff, including the principals, were too busy to have time for the study. The third school district (coded LT) refused to participate citing too many research studies going on at the time with no room for an additional research study.

The schools-within-districts study was subdivided into two sections. The purpose of the first section of the study was to describe and explain the relationship between school-level quantitative input variables and the two measures of student achievement across the schools. The purpose of the second section of the study was to describe and explain any indication of other school-level factors (x-factors) that differentiate schools on the achievement output measures beyond the quantitative inputs considered.

The quantitative input variables investigated at this level were: (1) School-level Science Department Expenditure Per Pupil (SCIEXPPP), (2) Science Teacher Quality (SCITCHQT), (3) Science Class Size (SCCLSIZE), (4) Science Lab-based instruction per Teacher per Week

(LABSPWK), (5) the number of Professional Development Activities, per Science Teacher per Year (PDVSCTCH), (6) Principal Stability (PRINSTBY), (7) Head of Science Department Quality (HEADSQLT), and (8) Percentage of Economically Disadvantage Students (PEDS).

The qualitative input variables investigated were the suggestions for improving science education, in the schools.

The achievement output variable measures investigated for the schools-within-districts study were: (1) Percentage of student passing the Science Section of the Georgia High School Graduation Test at first sitting (GHSGT) and (2) Percentage of graduating students qualified for the Georgia HOPE Scholarship (HOPE) as determined for the across-district study data.

Explanation and Determination the Variables in the Schools-within-Districts Study

The data produced from the personal interviews were both quantitative and qualitative. The quantitative data were comprised of (1) School-level Science Department Expenditure Per Pupil (SCIEXPPP), (2) Science Teacher Quality (SCITCHQT), (3) Science Class Size (SCCLSIZE), (4) Science Lab-based instructional activities per Teacher per Week (LABSPWK), (5) the number of Professional Development Activities, per Science Teacher per Year (PDVSCTCH), (6) Principal Stability (PRINSTBY), (7) Head of Science Department Quality (HEADSQLT), and (8) the Percentage of Economically Disadvantaged Students (PEDS), as previously determined.

The qualitative data were generated from the responses to the semi-structured questions contained in the interview instrument used. Descriptive adjectives and phrases associated suggestions for improvement, as contained in the responses, were counted and used as the quantitative basis for analysis. Summary of the descriptive adjectives and phrases are presented in Chapters 4 and 5.

School-level Science Department Expenditure Per Pupil (SCIEXPPP)

The numeric values for SCIEXPPP data were computed by dividing the amount of money actually received by the head of science department to run the department, per academic year, by the total number of students (N) in each of the sample schools. For example if the head of science received $5000 in hand, per academic year to run the department in a school with 1,500 enrolled students, then the SCIEXPPP = 5000/1500 = $3.33 per pupil. The total number of students (N) in each school was used because every Georgia high school student has to take science classes and pass the science section of the Georgia High School Graduation Test to get a high school diploma from their respective schools. Also, a student normally takes one science subject-class per each semester per academic year.

Science Teacher Quality (SCITCHQT)

The numeric values for SCITCHQT data were obtained as in TEACHQLT for district-level data using similar transformation process as detailed in Table 1. The corresponding qualification and experience values were those obtained and recorded for science teachers in the sample schools during the semi-structured interview (see Appendix B). The calculations were similar, as discussed for the across-district study.

Science Class Size (SCCLSIZE)

The numerical values for SCCLSIZE data were the average number of students per science class per teacher for the academic year.

Science Labs per Teacher per Week (LABSPWK)

The numeric values for the number of labs per week were the average number of laboratory-based hands-on activities per week per science teacher, as reported by the heads of science departments during the personal interview sessions.

Professional Development Activities, per Science Teacher per Year (PDVSCTCH)

The numeric values for PDVSCTCH were the number of professional staff development events attended per science teacher per academic year in each of the sample schools.

Principal Stability (PRINSTBY)

The numeric values for PRSTABTY were the number of years a principal has served in the sample school as the principal of the school.

Head of Science Department Quality (HEADSQLT)

The numeric values for HEADSQLT were computed using the formula:

$$HEADSQLT = Y_t \times 2(Y_{hod}),$$

where Y_t = number of year as a teacher before becoming head of science department, and Y_{hod} = number of years as head of Science Department. The number of years as head of Science Department is weighted by an arbitrary factor of 2, relative to the number of years as science teacher, because the responsibilities as head of Science Department carry more weight than the responsibilities as a science teacher (U.S. Department of Education, 2002b). Moreover, almost all the heads of Science Departments have teaching responsibilities, in addition to their responsibilities as heads of the science departments. Thus, it was assumed that the position as head of department is worth twice as much as the position of as teacher. For example for a head of department with who has 10 years as a teacher before becoming head of department (Y_t) and has served 5 years as head of department (Y_{hod}), the HEADSQLT = 10 x 2(5) = 100.

GHSGT & HOPE

The numerical values used for the schools-within-district-level data on GHSGT and HOPE, were as computed for the across-district study for the schools.

Suggestions for Improving Science Education

The numeric values used for the suggestions data were counts of descriptives contained in the responses to the semi-structured questions in the interview questionnaires (see Appendix B). The questions relate to suggestions for improving science education in the sample schools. Further details on the descriptive adjectives are shown and explained in chapters 4 and 5.

<u>Sampling for the Schools-within-Districts Study</u>

The school selection criterion for the sampling units for the schools-within-districts study was based on the number of high school principals, in the three consenting school districts, who agreed to let their schools participate in the school level study.

Sample Size for Schools-within-Districts Study

Individual principals in the high schools from the three consenting school districts were approached with requests for permission to conduct the study in their schools. A total of 35 principals in 35 high schools in the three consenting school districts were approached with the request. Twenty-eight out of the 35 high schools (80%) agreed to participate in the study. This was the sample size used for schools-within-districts study.

<u>Data Collection for the Schools-within-Districts Study</u>

The data for the schools-within-districts study were collected through personal face-to-face interviews with the principals and science department heads in the consenting schools using the interview instruments shown in Appendix B. The data collected were both quantitative and qualitative.

The Interview Methodology

The step-by-step methodology used in the interview process was done using the guide provided by Thomas (1999). Where necessary, modifications and adaptations were made to accommodate the peculiarities of this study. The purpose of the interview was to solicit answers to a set of questions aimed at providing the appropriate information for the school-level qualitative input variables mentioned earlier. The target groups were the school principals and the heads of the Science Departments from the consenting schools.

The contents of the survey instruments were divided into two format typologies (see Appendix B). The first typology was comprised of fixed-response requests for the quantitative data of interest. The second was comprised of open-ended response questions to obtain the qualitative data of interest. The open-ended questions were designed to elicit descriptives or phrases that were used as explanatory variables to support the findings of quantitative data. The identified descriptives were later translated into fixed responses by counting the number of respondents who used the descriptives or phrase in their answers to the semi-structured questions. Care was taken to ensure that the open-ended questions were not leading questions and any attempt to suggest responses was carefully avoided (Thomas, 1999). A cover letter, in a consent format, was also developed to introduce the participants to the purpose and nature of the study (see Appendix C).

Pilot testing of the interview instrument was conducted on three respondents each from the principals' group and heads of science departments' group. The purpose of the pilot testing was to seek item clarity, format comfort, and any other suggestions for improvement (Thomas, 1999). Further improvements were made to the instrument based on the information from the pilot test. To increase response rate and enhance the reliability of the responses, the data collection was done through personal face-to-face interviews with the respondents.

Before the formal interviews were conducted, IRB approval was granted, from the Human Subjects Office of the University of Georgia, to ensure the protection of the rights of the participants. The application process for the IRB approval involved the completion of the application form and the compilation of all required documents, including the approval documents from the school districts that agreed to participate in the study.

Hypotheses Tested by the Schools-within-Districts Study

The following two hypotheses were tested for the schools-within-districts study. The first hypothesis is a statistical hypothesis based on the school-level quantitative data, and the second hypothesis is also a statistical hypothesis based on the ranked quantitative data for PEDS, GHSGT, and HOPES, collected on the schools.

Hypothesis #8: There are no significant correlations between each of the school-level input variables and the achievement output measures.

Hypothesis #9: There is no indication of any other school-level factors**,** beyond the quantitative inputs considered, that differentiate schools on the achievement measures.

Qualitative data on suggestions for improving science education in the schools were also collected.

Data Analysis

Across-District Study Data Analysis

For the purposes of analysis, the hypotheses for the district level were categorized into three categories as follows:

Category 1

Category 1 consists of those hypotheses analyzed using Multivariate Analysis of Variance (MANOVA) techniques. The hypotheses are:

Hypothesis #1: There is no significant difference between school districts on the composite of educational input variables and the composite of achievement output measures.

Hypothesis #2: There is no significant difference between school districts on each of the educational input variables and the achievement output measures, considered separately.

Hypothesis #3: There is no significant difference between school districts on the composite of educational input variables and the composite of achievement output measures, after controlling for Percentage of Economically Disadvantaged Students (PEDS).

Hypothesis #4: There is no significant difference between school districts on each of the educational input variables and the achievement output measures, considered separately, after controlling for PEDS.

The use of multivariate analysis is appropriate for the following reasons: (1) education inputs and outputs operate in simultaneously multivariate manners to make the process of education a complex system. (2) The data collected were vectors comprised of replicates corresponding to each of the two school years on each of the variables involved for each of the sample schools in each the of six school districts.

MANOVA explores the characteristics of the linear composites of all the input and output variables and develops a set of statistics on the basis of which group separations or differences can be identified, characterized, described, and explained (Harris, 1985; Stevens, 1996; Tatsuoka, 1988). From the overall MANOVA, the following is interpreted: (1) the significance of the overall group separation or difference on the bases of the multivariate composites, by examining the Wilks' Λ value and its F-statistic; (2) the significance of each input and output variable at defining the group differences by looking at their individual F-statistics and the corresponding effect sizes; (3) the contribution magnitude of each of the input and output variables to the group differences by looking at the standardized weights and the structure coefficients of each variable. A univarate ANOVA cannot be used to analyze the linear composites of both input and output variables, simultaneously (Tatsuoka, 1988).

The focus of this study was not to check for variation from year to year within the two-year period covered by the study. The purpose was to describe and explain the relationships between the variables using the years as replicates in a multivariate sense**.** Also, the data for each year were based on different cohorts (sets) of students. To for check consistency in the p-values across the two years, additional crosschecking analyses were done. The data for each academic year were analyzed separately. The results of the crosscheck showed that the significant p-values were consistent with the results of the multivariate analysis reported in this study. The inputs and outputs were also analyzed separately, and the significant p-values were also consistent with the results of the multivariate analysis reported in this study. The tests of independence and correlation across the two years also confirmed the independence of the variables and the insignificant correlation between the sets of data across the two years. These results are consistent with Stevens' conclusion that in multivariate tests, the correlations among the variables, across replicates, are usually ignored (Stevens, 1996).

Category 2

Category 2 consists of those hypotheses to be analyzed using Multivariate Correlation Method, Canonical Correlation Analysis (CA). The hypotheses are:

Hypothesis #5: There is no significant correlation between the composite of educational input variables and the composite of achievement output measures across school districts.

Hypothesis #6: There are no significant correlations between the individual input variables and the achievement output measures, considered separately, across school districts.

Canonical correlation analysis (CA) was appropriate to test these hypotheses because of the multivariate nature of education and the multivariate nature of the data collected, as discussed earlier (Hair, Anderson, Tatham, & Black, 1992; Hand & Taylor, 1987; Tatsuoka, 1988). Canonical Analysis (CA), developed by Hotelling (1936), is used to study the relative strength of the relationship between two composite sets of variables. The composite scores are created from a larger set of mathematically combined scores.

The conceptual reasoning behind CA is to form two linear combinations of the input variables and the output variables by differentially weighing them to obtain the maximum possible correlation between the two linear combinations (Pedhazur, 1997). The correlation between the two linear combinations is called the Canonical Correlation (C_r). The square of the canonical correlation (C_r^2) is an estimate of the variance shared between the linear combinations of the input variable and the output measures. After obtaining the maximum C_r, additional C_r's are calculated depending on the minimum of the number of input (p) variables or output measures (q). In this study, the number of possible C_r's (canonical functions) = 2; the number of educational achievement output measures. The number of C_r's selected finally for meaningful interpretation depends on their statistical and practical significance.

From the overall multivariate canonical analysis, the following is interpreted: (1) the stability of the overall canonical model by examining the Wilks' Λ value and its F-statistic; (2) the stability of each of the three canonical correlation values by looking at their F-statistics; (3) the extent of the contributions of the correlation between each of the three outcome measures and composite of the five input variables by examining the individual squared canonical correlations; (4) the contribution of each of the input variables to the performance measure by looking at the standardized weights for each variable; (5) the variable(s) that contribute most to the composite model by examining the squared structure coefficients.

Category 3:

Category 3 consists of those hypotheses analyzed using Multivariate Regression Analysis (MRA) methods. The single hypothesis tested with this method was:

Hypothesis #7: There is no significant linear relationship between the linear composite of input variables and the linear composite of achievement output variables.

Multivariate Regression Analysis (MRA) was used to describe the linear relationship between multiple independent interval variables and multiple dependent interval criterion variables (Stevens, 1972). The linear composite multivariate regression model for this study may be written as:

$$a_1 GHSGT + a_2 HOPE = b_0 + b_1 EXPPP + b_2 TEACHQLT + b_3 PEDS$$

where; *EXPPP, TEACHQLT, & PEDS* = the linear combination of input variables, *GHSGT & HOPE* = linear combination of achievement output measures; b_0 = intercept of the composite regression line on the composite of performance measures; b_1, b_2, b_3, = regression coefficients associated with the input variables; a_1, a_2 = regression coefficients associated with the output variables.

Multivariate regression analysis produced statistical values for both multivariate as well as univarite measures. To determine whether the overall multivariate model is significant on the regression of the composite of the input variables with the performance measures, the Wilks' Λ value and its F-statistic, for the multivariate regression was examined. The contribution of each of the input variables to the composite of performance measures was confirmed by looking at the Eta Squared values (η^2) a measure of effect size. The individual regression model for each performance measure on the composite of the input variables was inferred from the univariate analysis output produced as part of the multivariate analysis. The univariate linear regression model was constructed. The univariate linear regression models tested by the MRA analysis can be written as:

$$GHSGT = a + b_1\, EXPPP + b_2 TEACHQLT + b_3 PEDS$$

$$HOPE = a + b_1\, EXPPP + b_2 TEACHQLT + b_3 PEDS$$

Each of the above multiple linear (univariate) regression models, was checked for: (1) How well each of the overall models predicted the relationship between each performance measure and the input variables by consulting the overall squared correlation (R^2) and its standard error; (2) The reliability of the prediction of the relationship shown by the R^2, by examining its F-value, and the associated p-value; (3) The input variables that significantly contribute to the outcome measure in each of the regression models by consulting the b-weights and their associated p-values or the confidence interval values; (4) How the individual input variables ranked in their effect in the explained variance of each of the performance measures by examining their respective Beta weights. The Beta weight for each model allows the comparisons between the unit changes in the performance measures with respect to each of the input variables; and (5) The contribution of each of the input variables to the overall explanation of the variation in the performance measures by calculating and looking at the squared structure coefficients.

Schools-within-Districts Data Analysis

For the purposes of analysis the schools-within-districts-level data were analyzed in two categories:

Category 1

Category 1 was the analysis of the quantitative school-within-districts data collected. Only one school-within-districts level hypothesis was tested in this category.

Hypothesis #8: There are no significant correlations between each of the quantitative schools-level input variables and the achievement output measures.

Due to the logistic limitations involved in personal interviews and the resultant sample size (N=28) of the school-level data collected, both parametric and non- parametric statistical methods were used to investigate the nature of the relationships between the eight school-level input variables and the two output measures. Pearson's product moment correlation (a parametric statistic) and Kendal's tau-b (a non-parametric statistic) were used to establish the magnitude and the direction of the relationships between the input variables and the output measures. Schools were then grouped into performance categories using Somers'd statistic. Further details and explanations of these parametric and non-parametric approaches are discussed in chapter IV where the results of this study are reported.

Category 2

Category 2 was the analysis of the qualitative school-level data collected. Only one hypothesis was tested in this category.

Hypothesis #9: There is no indication of any other school-level factors**,** beyond the quantitative inputs considered, that differentiate schools on the achievement measures.

The data used to test this hypothesis consisted of the numerical values obtained for PEDS, GHSGT, and HOPE. The data were ranked according to the output measures and paired with their corresponding PEDS values.

Qualitative data on suggestions for improving science education in the schools were also collected. The qualitative data collected consisted of the number of participants who expressed certain descriptive adjectives and phrases in response to the personal interview questions on suggestions for improving science education in the qualitative section of the interview instrument (see Appendix B).

The response descriptives counted for suggestions for improving science education were “improve science labs,” “teacher development,” “cooperative science instruction with external experts,” “public engagement,” “focus on science at lower grades,” “improve reading and writing,” and “improve science curriculum.”

All the data were summarized in a table and each response category was presented in bar graphs as detailed in Chapter 4.

Reliability, Validity, and Generalizability

Reliability, validity, and generalizability are important considerations in any study involving the use of measurement instruments and sampling from a population (Green, Salkind, & Akey, 2000; Shannon & Davenport, 2001).

Reliability

Reliability indicates the consistency of the results from a measurement instrument used in the data collection for a study (Huck, 2000, Shannon & Davenport, 2001). A good instrument must produce results that are consistent over time, consistent in item, and consistent across the sample population (Shannon & Davenport, 2001).

The source of information for the data used in the across-districts study was from the state of Georgia’s Department of Education. The reporting format of the information from the Department of Education is consistent across all the schools and across all school districts in the state. Therefore, the state instrument, from where the data were collected, was consistent across all the sample schools across the districts. This ensured the reliability of the data from this source.

The same interview instrument with the same item content was used on all of the sample respondents throughout the interview process. Each respondent was asked the same set of open-ended questions as contained in the instrument (see Appendix B). Moreover, the interview instruments comprised of sections soliciting both quantitative and qualitative data for a more robust interpretation and inference. All of these steps were taken to ensure the internal consistency and hence the reliability of the results inferred from the use of the interview instrument.

Validity

Validity expresses the accuracy of the results of a measurement instrument. "A measuring instrument is valid to the extent to which it measures what it purports to measure" (Huck, 2000, p.100). The state's database, from which most of the quantitative data were obtained, is an accurate source for the data collected. Also, enough care was taken to ensure that the questions in the interview instrument for the schools-within-districts study were precise enough to elicit reliable information for the questions asked from the human subjects. Initial pilot interviews of a sample of prospective participants was conducted and necessary adjustments were made to the instrument before the full study interviews were carried out.

Moreover, the qualitative-oriented questions in the interview instruments were open ended. The semi-structured nature of the questions precludes any biases. Also the information and responses to the questions were obtained through personal face-to-face interviews, to confirm the accuracy of the data. Where necessary, school records were consulted to provide accurate answers to the questions. All of these precautions were taken to ensure the validity of the instruments.

Generalizability

Generalizabilty is the extent to which the results of the study involving a sample from a given population can be generalized to a larger population similar to the population from which the study sample was taken (Huck, 2000). The results of this study may be generalized for the following reasons. First, it is a non-experimental design that relies on preexisting data for analysis and inferences. Second, the sample size, as determined, is large enough to infer close approximation to the target population thus enhancing the generalizability of this study. Third, the detailed methods used in the analysis of the data is reliable, making the results robust enough to be of more generalized application. However, cause-effect inferences may not be made from the results.

Summary

The purpose of this study was to describe and explain the relationships between educational inputs and measures of student achievement as outputs. This chapter presented the rationale for the study with a detailed definition of the variables identified and measured for the purposes of generating the appropriate findings and conclusions for this study. The hypotheses tested were also presented with full illustration of the hypothesized pathways. The context, regarding the units of study, was described and the details of data collection were also presented. Sample size and the rationale for determining the sample size were illustrated with full explanation of the methods used for data analysis and interpretation. The next two chapters present the results, findings, interpretation, discussions, and conclusions from the study. Chapter 4 presents the results of the quantitative and qualitative data collected. Chapter 5 presents a discussion of the results, the conclusions, recommendations, and suggestions for further consideration and future research.

CHAPTER 4

FINDINGS OF THE STUDY

The purpose of this study was to describe and to explain the relationships between educational inputs and measures of student achievement as outputs. To fulfill this purpose, the study was organized, conducted, and analyzed at two levels. The first level described and explained the relationships between input and output variables across school districts. The second level described and explained the relationships between input and output variables across schools within the districts.

Across-Districts Study Findings

The across-districts-level data were collected on sample high schools in the six school districts from the electronic archival Report Card and Financial Report records on the schools from the State Department of Education's web sites at: *http://techservices.doe.k12.ga.us/reportcard* & *http://techservices.doe.k12.ga.us/reports/finacial*.

For purposes of anonymity, the six school districts involved in this study are labeled TL, YT, BB, KL, LT, and WN. Data were collected from the web sites between the months of June 2003 and August 2003. The data collected included Expenditure Per Pupil (EXPPP) based on the State's QBE (Quality-Based Education) funding for each of the selected high schools in the six school districts; overall school Teacher Quality (TEACHQLT); the percentage of economically disadvantaged students (PEDS) as measured by the percentage of students eligible, under the federal guideline, for the free and reduced lunch program; the percentage of students passing the science section of the Georgia High School Graduation Test at first sitting (GHSGT); and the percentage of graduating students qualified for the HOPE scholarship program (HOPE). Data were collected on a total of 71 high schools in the six school districts.

Descriptive Statistics of Across-district Study Data

The data on the across-district-level variables, consisted of the three input variables (EXPPP, TEACHQLT, and PEDS) and the two achievement output measures (GHSGT and HOPE) were vectors comprising of replicates corresponding to school years 2000-2001 and 2001-2002 on each of the stated variables. The two years were the latest school years on which complete data, on the stated variables, were available.

As explained in Chapter 3, initial test-run analysis confirmed both the independence and the insignificant correlation of the variable values across the two years. Since the purpose of this study was to describe and to explain the relationships between the input and output variables and not how the variables vary over time, this initial test-run analysis was necessary.

The statistical analysis was done as multivariate analysis using Multivariate Analysis of Variance (MANOVA), Multivariate Analysis of Covariance (MANCOVA), Canonical Correlation Analysis (CA) and Multivariate Linear Regression (MLR) programs on SPSS 10.5 version. The multivariate analysis methods used the pooled (combined) statistics of the variables, over the two years, as the basis of the analysis. This provided a more robust interpretation of the outcomes.

The means and the corresponding standard deviations of all the variables across the six school districts are shown in Table 2. Two sets of observation elements, corresponding to the two school-year periods, were collected on each of 71 high schools in the six school districts. A total of 142 sets of observations, totaling 710 (2x71 schools x 5 variables) elements in a matrix, were collected. Each of the six school districts represented a cell in the vector matrix.

Table 2: Descriptive Statistics for the Input and Output Variables.

School District	EXPPP Mean (*SD*)	TEACHQLT Mean (*SD*)	PEDS Mean (*SD*)	GHSGT Mean (*SD*)	HOPE Mean (*SD*)	Sample size Observations (n)	Schools[a] (N)
TL	2702.82 (233.17)	34.47 (5.73)	66.31 (14.88)	57.27 (16.03)	49.71 (13.81)	22	11
YT	2924.79 (323.61)	27.49 (3.09)	39.56 (10.47)	63.64 (7.98)	54.81 (7.33)	14	7
BB	2786.73 (223.82)	32.16 (3.39)	10.87 (10.39)	79.81 (11.86)	68.54 (9.56)	26	13
KL	2799.26 (227.41)	29.57(4.47)	40.06 (17.43)	65.53 (14.49)	61.33 (10.38)	34	17
LT	2402.05 (154.10)	33.23 (3.27)	22.83 (19.50)	73.85 (15.00)	66.95(16.32)	20	10
WN	2988.62 (233.74)	31.58 (3.54)	14.32 (11.74)	76.80 (8.31)	62.20 (9.67)	26	13
Avg.	2773.13 (288.05)	31.48 (4.53)	31.56 (23.77)	69.92 (14.94)	61.16 (12.91)	142	71

[a]Total number of schools sampled = 71

School district WN (*M* = 2988.62, *SD* = 233.74) had the highest EXPPP in dollars. School district YT (*M* = 2924.79, *SD* = 323.61) had the next highest EXPPP. School district LT (*M* = 2402.05, *SD* = 154.10) had the lowest EXPPP. The EXPPP for the remaining three school districts did not differ significantly from the overall mean EXPPP value (*M* = 2773.12, *SD* = 288.05) for the six school districts.

School district TL (*M* = 34.47, *SD* = 5.73) had the highest mean for TEACHQLT; while school district YT (*M* = 27.49, SD = 3.09) had the lowest mean TEACHQLTY. The mean TEACHQLT for the entire six districts was, *M* = 31.48, SD =4.53 The variation in TEACHQLT indices among the school districts was minimal.

There was a relatively high degree of variation between school districts on the PEDS measure. School district TL (*M* = 66.31, *SD* = 14.88) had the highest PEDS score, almost double the mean PEDS value of *M* = 31.56, *SD* = 23.77 for the entire six school districts. School district BB (*M* = 10.87, *SD* = 10.39) and WN (*M* = 14.32, *SD* = 11.74) had the lowest mean PEDS.

On GHSGT, school district BB (*M* = 79.8, *SD* = 11.86) had the highest mean percentage; while district TL (*M* = 57.27, *SD* = 16.03) had the lowest mean score. District TL showed the largest variability (*SD* = 16.03) in GHSGT while District WN showed the smallest variability (*SD* = 8.31). For the entire six districts, the mean GHSGT was *M* = 69.92, *SD* = 14.94. This overall mean score compared favorably with the state's average GHSGT of 70.00 over the two years.

On HOPE measures, school district BB (*M* = 68.54, *SD* = 9.56), had the highest mean percentage; while TL (*M* = 49.71, *SD* = 13.81) had the lowest mean HOPE value. The overall mean HOPE value for the entire sample was *M* = 61.16, *SD* = 12.91. The graphical plots of the estimated marginal means for EXPPP, TEACHQLT, PEDS, GHSGT, and HOPE are shown in

Appendix E. The graphs further confirm the differences in the mean values between the districts on the stated composite of measures.

Box's M Test was also conducted to evaluate whether the variances and covariance among the variables are the same across the schools in the district groupings. The Box's M Test value of 175.799 was significant at $F_{(75,18791)} = 2.113$, $p<0.01$ with a log determinant = 26.686. Ordinarily this would have been a source of mild concern for the violation of the assumptions of homogeneity of covariance matrix and multivariate normality. However, given the large enough number of schools sampled (71), a violation of these assumptions was not considered a problem in the multivariate analysis of the results in this study. In fact, Green, Salkind and Akey (2000) cautioned that the results of Box's M test should be interpreted cautiously in that a non-significant result may be due to small sample size and a lack of statistical power. Therefore, a significant value here may be the result of the large sample size taken in order to get the desired robust statistical power needed for meaningful interpretations of the results as previously described in Chapter 3.

Moreover, a closer examination of the determinants of each group (district) TL, YT, BB, KL, LT, and WN, indicated log determinant values (/*B*/) of 27.30 (N = 11), 24.23 ($N = 7$), 24.74 ($N = 7$), 26.41 ($N = 7$), 23.50 ($N = 7$), and 25.01 ($N = 7$) for each of the groups, respectively. The sample sizes (N) were the number of sample units (high schools) in each of the six school districts. The correlation between the log determinants (/*B*/) and the sample sizes (N) for each of the groups was positive (r = .487) an indication that the multivariate tests were conservative. Therefore, there was no need to be concerned about the violation of the assumptions of homogeneity of covariance matrix and multivariate normality.

Inferential Statistics of Across-district Study Data

In order to fully describe and explain the nature of the relationships between the input variables (EXPPP, TEACHQLT, and PEDS) and the two output variables (GHSGT and HOPE) at the district level, the following null hypotheses were tested:

Hypothesis #1: There is no significant difference between school districts on the composite of educational input variables and the composite of achievement output measures.

Hypothesis #2: There is no significant difference between school districts on each of the educational input variables and the achievement output measures, considered separately.

Hypothesis #3: There is no significant difference between school districts on the composite of educational input variables and the composite of achievement output measures, after controlling for Percentage of Economically Disadvantaged Students (PEDS).

Hypothesis #4: There is no significant difference between school districts on each of the educational input variables and the achievement output measures, considered separately, after controlling for PEDS.

Hypothesis #5: There is no significant correlation between the composite of educational input variables and the composite of achievement output measures across school districts.

Hypothesis #6: There are no significant correlations between the individual input and output variables, considered univariately, across school districts.

Hypothesis #7: There is no significant linear relationship between the linear composite of input variables and the linear composite of achievement output variables.

Test of Hypothesis #1

Hypothesis: There is no significant difference between school districts on the composite of educational input variables and the composite of achievement output measures.

This hypothesis was tested using a multivariate analysis of variance (MANOVA). Table 3 shows the results of the multivariate test of significance between the districts on the composite of variables. The MANOVA test was performed to find out if there are significant differences between the districts on the linear combination of the input and output variables and to determine the extent to which each of the input and output variables, as a linear composite, contributes to the separation of the districts. The follow-up dimensionality test revealed the magnitude and extent of the contributions of the variables to the differences between the districts.

The analysis established that the difference between districts was statistically significant on the composite vectors of the input and output measures. The test indicated a Wilks' $\Lambda = .08695$ at $F_{(25,652)} = 18.295$, $p < .001$. The groupings explained 43.5% of the differences in the linear combinations (composites) of the inputs and outputs ($\eta^2 = .435$). These results show that the districts were different on the linear combination of EXPPP, TEACHQLT, PEDS, GHSGT, and HOPE.

Test of Dimensionality

The test of dimensionality and structure analysis of the districts differences, on the linear combination of the variables, showed that five discriminant functions, with respective Eigenvalues (λ) = 3.528, 1.009, 0.215, 0.039, and 0.001, were produced out of the six-district groupings. Three out of the five possible functions carried enough cumulative percentage (99.164%) with their respective Eigenvalues (λ=3.528, λ=1.009, λ= 0.215) significant enough to define the separation of the districts on the composite of the input and output measures.

The standardized discriminant function coefficients indicated that the first of the three significant functions was a socioeconomic construct, defined by PEDS, with standardized discriminant function coefficient = 1.798. The second construct was financial defined by both

Table 3: Multivariate Test of District Differences on the Composite of Inputs and Outputs

Statistic	Value	Hypothesis *df*	Error *df*	*F*	*p*	η^2
Wilks Λ[a]	.0869	25	652	18.295	0.000	.435

Function	Eigenvalue(λ)	Cumulative %	Canonical Correlation	*p*
1	3.528	73.62	.883	.000
2	1.009	94.67	.709	.000
3	0.215	99.16	.421	.000

Standardized Coefficients

Variables	1	2	3
EXPPP	.605	.841	.231
TEACHQLT	.632	.800	.613
GHSGT	.648	.179	.994
HOPE	.143	.152	.891
PEDS	1.798	.407	.134

[a] Significant at α= .05, p< .001

EXPPP with standardized discriminant function coefficient = .841, and TEACHQL with standardized discriminant function coefficient = .800. The third construct was achievement defined by both GHSGT with standardized discriminant function coefficient = .944, and HOPE with standardized discriminant function coefficient = .891.

Given the relatively low eigenvalue(λ=0.215), and the very low marginal cumulative percentage contribution (4.49%) of the third function to the dimensionality of the discriminant functions, the results indicate that the differences between the districts were mostly defined by the first two functional constructs; socioeconomic (PEDS) and financial (EXPPP and TEACHQLT).

Test of Hypothesis #2

Hypothesis: There is no significant difference between school districts on each of the educational input variables and the achievement output measures, considered separately.

This hypothesis is tested using the univariate results from the multivariate analysis of variance discussed above (see Table 3).

Levene's Test of homogeneity-of-variance was done to test the null hypothesis that the error variances of the dependent variables are equal across the groups (districts). This background test was necessary to test the assumption of homogeneity of variances across the variables. The test indicated non-significance for EXPPP $F_{(5,136)} = 0.907$, $p = .479$; and TEACHQLT $F_{(5,136)} = 1.863$, $p = .105$; but significance for PEDS $F_{(5,136)} = 6.103$, $p < .001$; GHSGT $F_{(5,136)} = 5.611$, $p < .001$; and HOPE $F_{(5,136)} = 7.124$, $p < .001$. Given the large sample size (71 schools) used in this study, the values for the Levene's Test was not a serious concern for the violation of homogeneity. All the variances were reasonably equal across the groups. Moreover, a closer inspection of the grouping variances and their respective sample sizes indicated positive relationships across all the variables EXPPP (*r* = .783), TEACHQLT (*r* = .256), PEDS (*r* =. 304) GHSGT (r = .284), and HOPE (r = .054); confirming no violation of the homogeneity and multivariate normality assumptions.

The univariate test results (Table 4) indicated that there were significant differences between the six districts on all input and output variables. For EXPPP (*M* = 2773.13, *SD* = 288.05), $F_{(5,136)}$ =16.516, $p< .001$, $\eta^2 = .378$; for TEACHQLT (*M* = 31.48, *SD* = 4.53) $F_{(5,136)} = 7.369$, $p<0.001$, $\eta^2 = .214$; for PEDS (M = 31.56 *SD* = 23.77) $F_{(5,136)} = 46.349$, $p<0.001$, $\eta^2 = .630$; for GHSGT (*M* = 69.92, *SD* = 14.94) $F_{(5,136)} = 10.571$, $p< .001$, $\eta^2 = .378$; and for HOPE (*M* 61.16, *SD* = 12.91) $F_{(5, 136)} = 8.439$, $p<0.001$, $\eta^2=0.237$ (see Table 2 also). The result showed that each of the individual composite vectors of predicator and indicator variables was significantly different among the six school districts.

Figure 2 shows a graphic representation of the summary of the multiple (Bonferroni) comparisons between the districts on each of the input and output measures. On EXPPP, WN school district was statistically significantly higher than all the other school districts except for district YT. EXPPP in school district LT was the lowest and statistically significantly different from any of the remaining five school districts. There were no statistically significant differences between TL, BB, KL, and YT on the means of EXPPP.

District TL had a statistically significantly higher TEACHQLT when compared with districts KL and YT. There were no statistically significant differences between districts WN, BB. LT, and TL on the means of TEACHQLT measure. District YT had the lowest TEACHQLT compared with the other school districts, except for district KL where there was no statistically significant difference between the two districts.

Table 4: Univariate Test of District Differences on Inputs and Outputs Individually

Variables	Hypothesis *MS*	Error *MS*	*F*	*p*	η^2
EXPPP [a]	884015.450	53523.6652	16.51635	0.000	0.378
TEACHQLT [a]	123.31089	16.73420	7.36879	0.000	0.214
PEDS [a]	100043.1886	216.68517	46.43922	0.000	0.630
GHSGT [a]	1762.06210	166.68144	10.57144	0.000	0.378
HOPE [a]	1113.08073	131.88859	8.43955	0.000	0.237

[a]Significant at $\alpha = .05$, $p < .001$

Variables	Mean Comparisons of Variables Across School Districts					
	└──┘ Indicate no significant difference in means between districts on the indicated variables at α= .05.					
	Low					**High**
EXPPP	LT 2402.05	TL 2701.82	BB 2786.73	KL 2799.26	YT 2924.79	WN 2988.62
TEACHQLT	YT 27.49	KL 29.57	WN 31.58	BB 32.16	LT 33.23	TL 34.47
PEDS	BB 10.87	WN 14.32	LT 22.83	YT 39.56	KL 40.06	TL 66.31
GHSGT	TL 57.27	YT 63.64	KL 65.53	LT 73.85	WN 76.80	BB 79.81
HOPE	TL 49.71	YT 54.81	KL 61.33	WN 62.20	LT 66.95	BB 68.54

Figure 2: Multiple Mean Comparisons of Inputs and Outputs Across School Districts.

District TL had a significantly higher PEDS when compared with any of the other school districts. There was no statistically significant difference between district KL and district YT on the PEDS measure. Similarly, there was no statistically significant difference between district BB and district WN on the means of their PEDS values.

On GHSGT, district TL mean score was significantly lower when compared with the remaining school districts except for districts YT and KL where there was no significant differences. There were no significant differences on the means of GHSGT between districts LT, WN, and BB. Districts WN and BB had a significantly higher GHSGT than districts TL, YT, and KL.

District TL mean HOPE score was significantly lower when compared with the remaining districts, except for district YT. There were no statistically significant differences between districts KL, WN, TL, and BB on the HOPE measure. District YT was also not significantly higher than district KL and WN on the HOPE measure. Although district BB showed a significantly higher HOPE scores when compared with some of the other districts, the difference was not significant when compared with districts KL, WN, and LT.

Test of Hypothesis #3

Hypothesis: There is no significant difference between school districts on the composite of educational input variables and the composite of achievement output measures, after controlling for Percentage of Economically Disadvantaged Students (PEDS).

This hypothesis was tested using Multivariate Analysis of Covariance (MANCOVA) with PEDS as the covariate. The test involved two steps. First, to find out if PEDS was related to the GHSGT and HOPE on the composites of EXPPP and TEACHQLT across the six school districts. Second, if so, to establish the nature of the relationship between PEDS and the other variables and make the necessary adjustments to their means.

A test of homogeneity of regression slope was carried out. This is an interaction test of equality of slopes for the relationship between PEDS (the covariate) and the other variables to establish the basis for a multivariate analysis of covariance (MANCOVA) and to adjust the means of the other variables across districts as necessary. The results (Table 5) show the Wilks' $\Lambda = .81346$ was not significant at $F_{(20,422)} = 1.35572$, $p = .140$. This was an indication that the slopes were equal and that PEDS was related to GHSGT and HOPE, given the composites of EXPPP and TEACHQLT, across the six school districts. There was, therefore, a basis to adjust the means of the other variables for a more powerful and reliable analysis of the results.

The results of the multivariate test, to confirm group differences on the pooled relationship between PEDS and the other variables, are shown in Table 6. The test was performed to confirm the pooled relationship between PEDS and the other variables and the effect on the separation (differences) between the school districts after controlling for PEDS. From the analysis, the difference between the districts, on the basis of the other variables, after controlling for PEDS, was statistically significant with Wilks $\Lambda = 0.25303$ and at $F_{(4,132)} = 97.42045$, $p < .001$. The result confirmed that PEDS contributed to the district differences on the vector of the other variables. Therefore adjustments were made to the observed means after controlling for PEDS. Table 7 shows the adjusted means and their corresponding observed means, for each of the school districts, on the vector of the remaining composite of measures.

Table 7 indicates that after controlling for PEDS, school district WN had the highest EXPP=$3084.03 and district LT has the lowest EXPP=$2452.23. School district LT has the highest TEACHQLT=33.88 and district YT has the lowest TEACHQLT=27.01 after adjusting

Table 5: Test of Homogeneity of Regression Slope using PEDS as Covariate

Statistic	Value	Hypothesis *df*	Error *df*	*F*	*p*
Wilks Λ [a]	0.81346	20	422	1.35572	.140

[a]Significant at α= .05, p= .140

Table 6: Test of Differences Between Districts on the Composites of the other Variables after Controlling for PEDS

Statistic	Value	Hypothesis *df*	Error *df*	*F*	*p*
Wilks Λ[a]	0.25303	4	132	97.42045	.000

[a]Significant at $\alpha = .05$, $p < .001$

Table 7: Observed and Adjusted Means of Inputs and Outputs after Controlling for PEDS.

	EXPPP		TEACHQLT		GHSGT		HOPE	
District	Obs. Mean	Adj. Mean[a]	Obs. Mean	Adj. Mean[a]	Obs. Mean	Adj. Mean[a]	Obs. Mean	Adj. Mean[a]
TL	2702.82	2521.88	34.47	32.12	57.27	81.63	49.71	69.70
YT	2924.79	2887.65	27.49	27.01	63.64	68.64	54.81	58.91
BB	2786.73	2900.47	32.16	33.63	79.81	64.49	68.54	55.97
KL	2799.27	2757.99	29.57	29.03	65.53	71.08	61.33	65.89
LT	2402.05	2452.23	33.32	33.88	73.85	67.10	66.95	61.40
WN	2773.13	3084.03	31.48	32.82	76.81	63.96	62.20	51.65

[a] Adjusted Means based on the pooled relationship between PEDS and the other variables. Wilks Λ = .25303, $F_{(4,132)}$= 97.42045, α= .05, $p < .001$

for PEDS. On the GHSGT, school district TL has the highest score at 81.63% while school district WN has the lowest at 63.96% when adjustment is made for PEDS. Correspondingly, after adjusting for PEDS, school district TL has the highest HOPE score at 69.70% while district WN has the lowest HOPE score of 51.65%.

After controlling for PEDS, the multivariate test of significance between the districts on the composite of the remaining variables is shown in Table 8. The result indicated that the districts were still significantly different on the composite of the remaining variables with Wilks Λ=0.23512 at $F_{(4,132)} = 12.005$ $p < .001$, $\eta^2 = .747$. The composite of remaining variables explained 74.7% of the difference between the districts after controlling for PEDS.

Dimensionality

The statistical test for dimensionality (Table 8) indicated that four possible functions with Eigenvalues (λ) 1.597, 0.388, 0.144, and 0.031, respectively, describe the dimensionality of the district differences after controlling for PEDS. Three out of the four possible constructs (functions) carried enough cumulative percentage (99.550%) with Eigenvalues (λ) =1.597, λ=0.388, and λ=0.144 respectively, after controlling for PEDS, to significantly define the dimensionality and structure of the separation of the districts on the composite of the remaining input and output measures.

The standardized discriminant function coefficients and the values of the correlation between the dependent and their corresponding canonical variables, after controlling for PEDS, indicated that the first of the three significant constructs was financial, defined by EXPPP with a standardized discriminant function coefficient = .989, explaining 47.61% (structure r = 0.690) of the functional characteristics. The second construct was defined by and TEACHQLT with standardized discriminant function coefficient = .813, explaining 58.82% of the characteristics of

Table 8: Multivariate Test of District Differences on the Composites of Inputs and Outputs after Controlling for PEDS

Statistic	Value	Hypothesis *df*	Error *df*	*F*	*p*	η^2
Wilks Λ[a]	.23512	20	439	12.005	0.000	.747

Function	Eigenvalue(λ)	Cumulative %	Canonical Correlation
1	1.528	73.93	.784
2	0.388	91.90	.529
3	0.144	98.55	.354
4	0.031	100.00	.174

	Standardized Coefficients			Correlation Between Dependent & Canonical Variables		
Variables	1	2	3	1	2	3
EXPPP	.989	.107	.279	.690	.081	.435
TEACHQLT	.551	.813	.218	.196	.767	.331
GHSGT	.225	.185	.928	.248	.352	.895
HOPE	.437	.533	.227	.353	.600	.106

[a]Significant at $\alpha = .05$, $p < .001$

the function (structure r = .767) and GHSGT with standardized discriminant function coefficient = .533, explaining 36.00% of the characteristics of the function (structure r = .600). The third function is defined by GHSGT with standardized discriminant function coefficient = .928, explaining 80.10% of the characteristics of the function (structure r = .895).

Hence, after controlling for PEDS, the district differences were significantly determined by the other variables, with per pupil expenditure being the number one determinant of the differences between districts. Compared with the other variables however, the difference between the districts on TEACHQLT was minimal, given its standardized discriminant function coefficients and its canonical correlation. Therefore, PEDS may be a determinant of TEACHQLT.

Test of Hypothesis #4

Hypothesis #4: There is no significant difference between school districts on each of the educational input variables and the achievement output measures, considered separately, after controlling for PEDS.

After controlling for PEDS, the univariate *F*-test for differences between the districts on the each of the vector of variables, considered separately, indicated a highly significant difference between the districts on each of the measures (see Table 9). The difference between the districts on EXPPP was highly significant with $F_{(4,135)} = 21.596$, $p < .001$, $\eta^2 = .444$; the difference on TEACHQL was highly significant with $F_{(4,135)} = 8.482$, $p < .001$ $\eta^2 = .239$; the difference on GHSGT was highly significant with $F_{(4,13)} = 7.061$, $p < .001$, $\eta^2 = .207$; and on HOPE the difference between the districts was also highly significant with $F_{(4,132)} = 9.611$, $p < .001$, $\eta^2 = .263$.

Table 9: Univariate Test of District Differences on Inputs and Outputs Individually After Controlling for PEDS.

Variables	Hypothesis *MS*	Error *MS*	*F*	*p*	η^2
EXPPP	1031269.49	47752.11	21.596[a]	0.000	.444
TEACHQLT	134.16	15.81	8.482[a]	0.000	.239
GHSGT	396.61	56.17	7.061[a]	0.000	.207
HOPE	553.14	57.55	9.611[a]	0.000	.263

[a]Significant at $\alpha = .05$, $p < .05$

Test of Hypothesis #5

Hypothesis #5: There is no significant correlation between the composite of educational input variables and the composite of achievement output measures across school districts.

The analysis for the multivariate correlation test of the hypothesis was done using the Canonical Correlation syntax in the Macros file of SPSS 10.5 version. Table 10 summarizes the results of the canonical correlation analysis. The overall multivariate canonical correlation was defined by only one function, out of two possible functions, with a very strong overall Canonical Correlation (*Cr*) = .85, explaining 75.42% of the total correlation. The strong correlation was also confirmed by the highly significant result of multivariate test that the remaining canonical correlations were zero, with Wilks Λ=0.274 at Chi-squared,$\chi^2_{(6)}$=178.886, $p < .001$.

The standardized coefficients and canonical loadings indicated that the single function was defined by PEDS with standardized coefficients = .993, and canonical loadings = .998; EXPPP with standardized coefficients = .071, and canonical loadings = .157; TEACHQLT standardized coefficients = .012, and canonical loadings = .170. The PEDS explained the largest proportion of the overall correlation between the composite of input variables (EXPPP + TEACHQLT + PEDS) on one hand and the composite of output variables (GHSGT + HOPE) on the other hand.

Test of Hypothesis # 6

Hypothesis: There are no significant correlations between the individual input and output variables, considered univariately, across school districts.

The results of the univariate analysis for the relationships between the variables considered individually (univariately) are shown in Table 11. The result indicated significant positive correlations between; (1) PPEXPNT and TEACHQLT, Correlation coefficient (*Cr*) = .1450, *p* =

Table 10: Canonical Correlations and Dimensionality of the Relationship Between the Composites of Inputs and Outputs.

Canonical Function	Wilks	χ^2	*df*		*Cr*	η^2
1	0.274[a]	178.886	6	.000	.851	.754
2	0.989	1.472	2	.479	.103	.011

Input Variables	Standardized Coefficient	Canonical Loadings
EXPPP	.071	.159
TEACHQLT	.012	.170
PEDS	.993	.998

[a]Significant at $\alpha = .05$, $p < .01$

Table 11: Correlation Coefficients of the Univariate Relationships Between Individual Inputs and Outputs

Variaes	EXPPP	TEACHQLT	PEDS	GHSGT	HOPE
EXPPP	__	.1450[a]	.0908	- .1151	-.1644[a]
TEACHQLT		__	.1727[a]	-.1329	-.1499[a]
PEDS			__	-.8343[a]	-.7428[a]
GHSGT				__	.7727[a]
HOPE					__

[a] Significant at $p < .05$

.043; (2) TEACHQLT and PEDS, *Cr*= .1727, p = .020; (3) GHSGT and HOPE, *Cr* = .7727, p< .001.

There were significant negative correlations between: (1) GHSGT and PEDS, *Cr* = -0.8343, p< .001; (2) HOPE and PEDS, *Cr*= -0.7428, p< .001; (3) TEACHQLT and HOPE, *C*r = -0.149, p = .037; and (4) EXPPP and HOPE, *Cr* = -0.1644, p = 0.025. Although, there were negative relationships between: (1) TEACHQLT and GHSGT, *Cr* = -0.1329, p =0.057; and (2) EXPPP and GHSGT, *C r*= -0.1151, p< .086 these relationships were weak and non-significant statistically.

Test of Hypothesis #7

Hypothesis: Hypothesis #7: There is no significant linear relationship between the linear composite of input variables and the linear composite of achievement output variables.

Due to the fact that the district-level study involved three input (independent) variables and two achievement output (dependent) measures, a multivariate regression analysis was done using GLM multivariate commands with a special request for multivariate test of the entire model on SPSS 10.5 version.

Table12 shows the multivariate test results. The result indicates that the overall composite of input variables (EXPPP, TEACHQLT, and PEDS) had a highly significant linear relationship with the overall composite of output measures (GHSGT and HOPE) with Wilks Λ= .274 at $F_{(6,274)}$ = 41.647, p<0.001 explaining 47.7% (η^2 =0.477) of the multivariate linear relationship.

However the multivariate test of the individual input variables (EXPPP, TEACHQLT, and PEDS) on the linear composite of output variables (GHSGT and HOPE) indicated that only PEDS made a highly significant contribution to the linear relationship, with Wilks Λ= .286 at

Table 12: Multivariate Test of Linear Regression (Relationship) Between the Linear Composites of Inputs and Outputs.

Input Variables	Wilks Λ	Hypothesis *df*	Error *df*	*F*	*p*	η^2
All Predictors: Overall Model	0.274[a]	6	274	41.647	.000	.477
EXPPP	0.979	2	137	1.440	.241	.021
TEACHQLT	0.998	2	137	0.121	.886	.002
PEDS	0.286[a]	2	137	170.783	.000	.714

[a] Significant at $\alpha = .05$, $p < .001$

$F_{(2,137)}$ = 170.783, p< .001. PEDS explained 71.40% (η^2=0.714) of the multivariate linear relationship on the linear composite of GHSGT and HOPE.

The parameter estimates detailing the coefficients for the univariate linear relationship is summarized in Table 13. The linear relationship between GHSGT and the composite of EXPPP, TEACHQLT, PEDS, had overall R^2 = .698 with *Adjusted* R^2 = .691, explaining 69.1% of the multivariate linear association. The coefficients of the regression lines were β = -.002 for EXPPP; β = -.057 for TEACHQLT; and β = -.524 for PEDS. The coefficients of both EXPPP and TEACHQLT were not significantly different from zero; indicating no significant contribution to the linear model. The linear multivariate regression model for GHSGT may then be written as:

$$GHSGT = 90.698 – .002EXPPP – .056TEACHQLT – .524PEDS.$$

Since the slopes of EXPPP and TEACHQLT are not statistically significant, after running the regression model was run with PEDS as the only input variable, the linear regression model can be more correctly written as:

$$GHSGT = 86.464 – .524PEDS$$

The linear relationship between HOPE and the composite of EXPPP, TEACHQLT, PEDS, has overall R^2 = .561 with *Adjusted* R^2= .552, explaining 55.2% of the multivariate linear association. The slopes of the regression lines were β = -.004 for EXPPP; β = -.027 for TEACHQLT; and β = -.398 for PEDS. The slopes of both EXPPP and TEACHQLT were not significant at α = .05; hence they did not contribute significantly to the linear model. The linear multivariate regression model for HOPE may be written as:

Table 13: Univariate Regression Coefficients of the Linear Relationships Between Individual Inputs and Outputs

GHSGT		HOPE	
R^2	*Adjusted* R^2	R^2	*Adjusted* R^2
.698	.691	.561	.552

Input Variables	Parameter (β_o , β)	Coefficients	*SE*	*t*	*p*	95% CI Lower	95% CI Upper	η^2
GHSGT	Intercept	90.698	7.828	11.586	.000	75.21	106.17	.493
	EXPPP	-0.002	0.002	-0.885	.378	-0.007	0.003	.006
	TEACHQLT	-0.057	0.158	0.359	.720	-0.256	0.369	.001
	PEDS[a]	-0.524	0.030	-17.508	.000	-0.583	-0.465	.690
HOPE	Intercept	86.545	8.152	10.617	.000	70.43	102.663	.450
	EXPPP	-0.004	0.003	-1.690	.093	-0.009	0.001	.020
	TEACHQLT	-0.027	0.165	-0.162	.871	-0.352	0.299	.000
	PEDS[a]	-0.398	0.031	-12.766	.000	-0.459	-0.336	.541

[a] Significant at α= .05, p< .001.

HOPE = 86.545 – .004EXPPP – .027TEACHQLT – .398PEDS. Since the slopes of EXPPP and TEACHQLT were not statistically significant, after running the regression model with PEDS as the only input variable, the linear regression model can be more correctly written as:

$$HOPE = 73.885 - .403PEDS.$$

The regression plots of the residuals of the output variables against the input variables are shown and explained in Appendix F.

Schools-within-Districts Study Findings

Three out of the six school districts used for the across district-level study agreed to participate in the research at the school-level. Twenty-eight out of a total of 35 regular high schools (80%), in the 3 participating school districts, agreed to participate in the research at the school-level. In addition to the GHSGT, HOPE, and PEDS data collected from the state archival records for the corresponding school years, other school-level data on the sample schools from the three school districts were also collected, as explained in Chapter 3.

The data were collected through personal interviews with the principal, the head of the science department, and where necessary the principal also sought the assistance of the bookkeeper, in each of the sample schools. The school-level data collected are contained in the interview questionnaire forms shown in Appendices B and C. The data were collected between the months of July 2003 and November 2003. Twenty-six school principals and 28 heads of Science Departments in the 28 high schools, who agreed to participate, from the three participating school districts, were personally interviewed. For analytical purposes, the school-level data were further divided into two categories.

The first category consisted of the data analyzed for the quantitative description of the nature of the relationships between the measured variables at school level. The variables investigated in the school-level data were: Actual school-level Science Department Expenditure Per Pupil (SCIEXPPP); Science Teacher Quality (SCITCHQT); Science Class Size (SCCLSIZE); Science Labs per Teacher per Week (LABSPWK); the number of Professional Development Activities, per Science Teacher per Year (PDVSCTCH); Principal Stability (PRINSTBY); and Head of Science Department Quality (HEADSQLT). The PEDS, GHSGT, and HOPE values, used for the school-level data, were as computed for the across-district study data for the corresponding schools. The numerical values for SCIEXPPP, SCITCHQT, SCCLSIZE, LABSPWK, PDVSCTCH, PRINSTBY, PEDS, GHSGT, and HOPE were as explained in Chapter 3.

The second category consisted of the data analyzed for both quantitative and qualitative explanations to support the findings of quantitative analysis for the across-district study and the schools-within- districts quantitative study data. The schools-within-district qualitative analysis examined the science education assessments dimensions in the individual schools that could be used to further explain the nature of the relationship between student performance in science and the input variable(s).

Descriptive Statistics of Schools-within-Districts Study Data

A total of 28 schools were analyzed for the 8 school-level input variables explained earlier, except for PRINSTBY for which only 26 school principals provided the relevant data. For reasons of sample size, relative to the number of input variables considered, both parametric and non-parametric methods were used to describe the nature of the relationships between GHSGT/HOPE and the eight school-level input variables.

Pearson’s Product Moment Correlation, Kendall’s tau-b, and Somers’d statistics, from the Crosstabs Programs on SPSS Version 10.1, were used to assess and describe the nature of the relationships. Pearson's product moment correlation and Kendall's tau-b measured the magnitude and direction of the relationship between GHSGT/HOPE, on one hand, and each of the input variables, on the other hand. Somers' d measured the degree of concordance-discordance between GHSGT/HOPE and each of the eight input variables. Kendall's tau-b was used in conjunction with Pearson's product moment correlation to show that similar inferences were possible from the two statistics despite the fact that Kendall's tau-b is a non-parametric measure while Pearson's product moment correlation is a parametric statistic.

The means and the corresponding standard deviations for each of the eight input and the two output variables considered are shown in Table 14. There was not a great deal of variation in SCCLSIZE (M = 25.82, SD = 2.06) across schools. There was a lot of variation in SCIEXPPP (M = 5.18, SD = 4.46); SCITCHQT (M = 43.95, SD = 8.58); LABSPWK (M = 1.46 , SD = 0.69); PDSCTCH (M = 2.61, SD =1.3); PRINSTBY (M = 3.12 , SD =2.63); HEADSQLY(M = 24.96, SD =11.33); PEDS (M = 42.92, SD = 18.77); GHSGT (M = 64.57 , SD =14.30); and HOPE (M = 56.28 , SD = 13.85).

The shape of the distribution of each of the variables is indicated by the skewness statistics. SCITCHQT (skewness=1.800), SCIEXPPP (skewness=2.908), PRINSTABY (skewness=2.414), and LABSPWK (skewness=1.935) showed extreme skewness to the positive side of the distribution. This is an indication that there are some relatively high values for each of these variables. SCIEXPPP in particular has the highest skewness value; an indication that a few schools may be spending a relatively higher amount per pupil in their science departments than the other schools. In one particular school, the science department gets a sizable amount of money from its corporate sponsors and grants, thus making its per pupil expenditure on science about $23.00. This is relatively very high compared to the other schools and disproportionately

Table 14: Descriptive Statistics of the Quantitative School-level Inputs and Output.

Variables**	*M*	*SD*	Skewness	N*
SCIEXPPP	5.18	4.47	2.908	28
SCITCHQT	43.94	8.58	1.800	28
SCCLSIZE	25.82	2.06	-0.540	28
LABSPWK	1.46	0.69	1.935	28
PDVSCTCH	2.61	1.31	0.907	28
PRINSTBY	3.12	2.63	2.414	26
HEADSQLY	24.96	11.33	0.709	28
PEDS	42.92	18.77	0.623	28
GHSGT	64.57	14.30	-0.397	28
HOPE	56.28	13.85	-0.511	28

*N=Number of schools (28).

**The variables are the eight inputs and two outputs at the school-level.

higher that the mean value of $5.18, across all the schools. All values were included in the analysis, even when they appear as outliers.

The mean SCIEXPPP (*M* = 5.18, *SD* = 4.46) was lower than the overall QBE allotment EXPPP (*M* = 2773.13, *SD* = 288.05) across the six school districts, constituting only about 0.19% of the total expenditure per pupil based on state QBE funding. SCITCHQT (*M* = 43.95, *SD* = 8.58) was higher than the mean overall TEACHQLT (*M* = 31.48, *SD* = 4.53) across districts. On the average science teachers were more likely to have qualifications above the bachelor level. The mean PEDS (*M* = 42.92, *SD* = 18.77) in the 28 schools was higher than the mean PEDS (*M* = 31.84, *SD* = 4.53) across the six school districts. The schools involved in the schools-within-districts study had a relatively high percentage of economically disadvantaged students than the average across the six school districts. Also the means of GHSGT (*M* = 64.57, *SD* = 14.30) and HOPE (*M* = 56.28, *SD* = 13.85), were generally lower than the average GHSGT (*M* = 69.92, *SD* = 14.94) and HOPE (*M* = 61.16, *SD* = 12.91) across the districts.

<u>Inferential Statistics of Schools-within-District Study Data</u>

The focus of the quantitative data collected at school-level was to test one hypothesis. Hypothesis #8: There were no significant correlations between each of the eight school-level input variables and each of the two achievement output measures.

<u>Test of Hypothesis #8</u>

The nature of the relationships between GHSGT and the eight input variables are summarized in Table 15. Using the Bonferoni approach to control for Type I error across the eight input variables, a *p*-value of less than .006 (.05/8 = .006) was required for statistical significance. The correlation coefficients computed indicate that only two out of the eight input

Table 15: Correlations Between GHSGT and School-level Inputs.

School-level Variables	Pearson's Correlation (R)			Kendall's tau-b			
	Value	Statistic (T)	p	Value	Statistic (T)	p	η^2
GHSGT * SCIEXPPP	-.091	-0.464	.081	.035	0.233	.102	.990
GHSGT * SCITCHQT	-.094	-0.480	.078	-.027	-0.189	.106	.990
GHSGT * SCCLSIZE	.183	0.950	.118	.133	0.797	.536	.406
GHSGT * LABSPWK	.240	1.258	.027	.214	1.421	.019	.059
GHSGT * PDVSCTCH	.123	0.631	.067	.024	0.164	.109	.151
GHSGT * PRINSTBY	.463 [a]	3.326	.001	.486 [a]	3.345	.001	.625
GHSGT * HEADSQLT	.079	0.405	.086	.041	0.305	.095	.658
GHSGT * PEDS	-.971[a]	-6.594	.000	-.537 [a]	-4.312	.000	1.000

a = significant at p< .006 after controlling for family-wise error for the eight input variables tested.

variables showed statistically significant linear relationships with GHSGT at the tested Bonferoni α-level level. There was a highly statistically significant negative correlation between GHSGT and PEDS (Pearson's *R*= - .971, $t_{(27)} = 6.594$, $p< .001$, Kendall's tau-b = -.537, $t_{(27)} = 4.312$, $p <.001$); GHSGT as a dependent variable explained almost 100% of the relationship ($\eta^2 = 1.00$). The positive correlation between GHSGT and PRINSBTY was also statistically significant (Pearson's *R*= .463, $t_{(27)} = 2.560$, $p = .002$, Kendall's tau-b = 0.486, $t_{(27)} = 3.345$, $p = .001$); GHSGT as a dependent variable explained almost 62.5 % of the relationship ($\eta^2 = 0.625$).

There was also a positive correlation between GHSGT and LABSPWK (Pearson's *R*= .240 $t_{(27)} = 1.258$, $p = .027$, Kendall's tau-b = 0.214, $T_{(27)} = 1.421$, $p = .019$), but the correlation was not statistically significant at the stringent Bonferoni test level of .006. GHSGT as a dependent variable explained only 5.9 % of the relationship ($\eta^2 = 0.059$). However, given the relatively high value of the correlation coefficient (.240), the sample size, and the stringent Bonferoni alpha level for the tests, the positive correlation between GHSGT and LABSPWK is of practical significance. Therefore, there was a moderate indication of a practically significant positive relationship between GHSGT and LABSPWK.

Further details of the regression plots of GHSGT against each of the eight input variables and the intercorrelations between all the school-level variables are shown in Appendices F and G respectively. Closer examination of the partial regression plots of GHSGT, as dependent variables against LABSPWK, indicate two outliers that could have influenced the value of the correlation coefficient. This is not surprising because the distribution of the LABSPWK showed extreme skewness (1.935) to the positive side of the distribution, as discussed before. Because the value for the outlier point for LABSPWK was a genuine value obtained during the interview sessions, the value was included in the regression analysis.

The nature of the relationships between HOPE and the eight input variables is summarized in Table 16. Again, using the Bonferoni approach to control for a Type I error across the eight input variables, a p-value of less than .006 (.05/8 = .006) was also required for statistical significance. As in GHSGT, the correlation coefficients computed indicate that only two out of the eight input variables showed statistically significant linear relationships with HOPE. There was a highly statistically significant positive correlation between: HOPE and PRINSTBY (Pearson's $R = 0.400$, $t_{(28)} = 2.136$, $p = .005$, Kendall's tau-b = 0.327, $t_{(28)} = 2.467$, $p = .002$); and a statistically significant negative correlation between HOPE and PEDS (Pearson's $R = -.742$, $t_{(27)} = 5.638$, $p < .001$, Kendall's tau-b = -0.519, $t_{(27)} = 4.396$, p< .001). HOPE, as a dependent variable explained almost 54.2 % ($\eta^2 = 0.54.2$) and 100% ($\eta^2 = 1.000$) of the relationship to PRINSTBY and PEDS, respectively.

Although the positive relationships between HOPE on one hand and LABSPWK (Pearson's $R = 0.295$, $t_{(28)} = 1.574$, $p = .016$, Kendall's tau-b = 0.216, $t_{(28)} = 1.497$, $p = .017$); PDVSCTH (Pearson's $R = 0.263$, $t_{(28)} = 1.390$, $p = .022$, Kendall's tau-b = 0.168, $t_{(28)} = 1.158$, $p = .031$) on the other hand, were not statistically significant, given the stringent Bonferoni α-level, the positive relationships may be of practical significance. There is therefore a moderate indication of positive relationships between HOPE and LABSPWK/PDVSCTH.

GHSGTS and HOPE showed similar relationships with all the eight variables. This is not surprising, since there was a significantly strong correlation between GHSGT and HOPE, (Pearson's Correlation $R = 0.806$, $p < .001$). The significantly negative relationships between GHSGT/HOPE and PEDS at the school-level echo exactly the same relationships between the three variables as in the across-district-level study. Both the across-district-level and schools-

Table 16: Correlations Between HOPE and the School-level Inputs.

School-level Variables	Pearson's Correlation (*R*)			Kendall's tau-b			
	Value	Statistic (*T*)	*p*	Value	Statistic (*T*)	*p*	η^2
HOPE * SCIEXPPP	-.149	-0.770	.056	-.019	-0.121	.113	.904
HOPE * SCITCHQT	-.227	-1.190	.030	-.027	-0.186	.107	.949
HOPE * SCCLSIZE	.007	0.035	.122	.056	0.319	.094	.651
HOPE * LABSPWK	.295	1.574	.016	.216	1.497	.017	.302
HOPE * PDVSCTCH	.263	1.390	.022	.168	1.158	.031	.296
HOPE * PRINSTBY	.400 [a]	2.136	.005	.327 [a]	2.467	.002	.542
HOPE * HEADSQLT	.058	0.296	.096	.040	0.359	.090	.811
HOPE * PEDS	-.742[a]	-5.638	.000	- .519 [a]	-4.396	.000	1.000

a = significant at p< .006 after controlling for family-wise error for the eight input variables tested.

within-distracts-level quantitative data indicated that PEDS was the most statistically significant predictor of GHSGT and HOPE. The Pearson's intercorrelation coefficients for all the 10 school-level variables are shown in Appendix G, and the regression plots of the residuals are shown in Appendix F.

The Nature of the Relationship Between PEDS and GHSGT/HOPE

The nature of the relationships between PEDS and GHSGT/HOPE was an investigation based on Hypothesis #9:

Hypothesis #9: There is no indication of any other school-level factors**,** beyond the quantitative inputs considered, that differentiate schools on the achievement measures. To test this hypothesis, further analysis was done to investigate the nature of PEDS' influences on GHSGT and HOPE. The additional quantitative analyses were done by ranking the schools in descending order of their values on the GHSGT and HOPE compared to their corresponding PEDS values. A directional measures test, using Sommer's d, was performed to determine the extent of concordance or discordance between the GHSGT/HOPE values and their corresponding PEDS values for each of the 28 schools.

Somer's d measures the proportional difference between the number of concordant (P) and discordant (Q) observation pairs. When the direction of one member of an observation pair is upward while the direction of the other member is downward, the pair is said to be discordant (Q). When both members of a pair of observations move in the same relative direction, they are said to be concordant (P). The difference between P and Q is used to calculate the statistic for Somers' d based on the number of observations.

The results of the Somer's d test indicate that over 53% of the schools were significantly discordant in the values of GHSGT compared to their corresponding PEDS values, (Somer's d =

-0.537, $t_{(27)} = 4.312$, $p < .001$). Similar results were also obtained for HOPE versus PEDS (Somer's d = - 0.519, $t_{(27)} = 4.396$, $p < .001$). Since there is a strong and significantly positive correlation between GHSGT and HOPE, all further analysis, findings, and interpretations were focused on GHSGT as the principal achievement output variable; the interpretations of which could be extended for HOPE.

Based on the concordancy results, the schools were grouped into three categories (see Table 17), labeled as 'positively discordant schools', 'concordant schools', and 'negatively discordant schools'. The first category comprised of schools that performed above expectation on the GHSGT compared to their PEDS values; these were schools with relatively high GHSGT values despite their relatively unfavorable high PEDS (high percentage of economically disadvantaged students) values. These were the schools that were labeled 'positively discordant schools'. The second category comprised of schools that performed according to expectation compared to their PEDS values. These were the schools that were labeled 'concordant schools'.

The third category is comprised of schools that performed below expectation compared to their PEDS values; these were schools with relatively low GHSGT values given their relatively favorable low PEDS (low percentage of economically disadvantaged students) values. These were labeled 'negatively discordant schools'.

A total of 15 out of the 28 sampled schools (53.5%) were discordant. Eight out of the 15 discordant schools were positively discordant, while the remaining 7 schools were negatively discordant (Somer's d = -0.537, $t_{(27)} = -4.312$, $p < .001$). The rest of the sampled schools (13 out of 28) were concordant.

Therefore beyond the measured quantitative factors, especially money and PEDS, there appear to be other factors that differentiate student achievement in the schools. The next level of

Table 17: GHSGT Ranking and Concordancy Grouping Relative to PEDS

Schools' Ranking	GHSGT	PEDS	Positively Discordant[a]	Concordant[c]	Negatively Discordant[b]
1	91.00	17.50		**	
2	82.00	18.05		**	
3	80.50	48.75	**+ve**		
4	79.50	30.50	**+ve**		
5	78.50	16.20			**-ve**
6	77.00	45.45	**+ve**		
7	76.00	46.10	**+ve**		
8	75.50	31.20	**+ve**		
9	72.00	27.15		**	
10	72.00	25.80		**	
11	71.50	28.45		**	
12	70.50	38.50		**	
13	70.00	44.65	**+ve**		
14	69.50	29.75			**-ve**
15	69.50	28.70			**-ve**
16	66.50	33.65			**-ve**
17	61.00	49.85		**	
18	61.00	55.60	**+ve**		
19	60.00	36.63			**-ve**
20	57.00	42.00		**	
21	54.00	58.50	**+ve**		
22	51.00	26.30			**-ve**
23	49.50	52.90		**	
24	48.50	75.05		**	
25	45.00	75.35		**	
26	45.00	74.20		**	
27	40.50	61.20			**-ve**
28	34.00	84.00		**	

a = (+ve) schools performing significantly above expectation compared to their PEDS values.
c = (**) schools performing according to expectation compared to their PEDS values.
b = (-ve) schools performing significantly below expectation compared to their PEDS values.
Somers'd = -0.537, T= -4.312, $\propto$= .05, p< .001.
% Discordancy = 53.57%.

analysis was designed to examine some of the suggestion for improving science education in the schools based on the school concordancy categories.

Findings of the Qualitative Suggestions for Improving Science Education

For purposes of logical interpretation of the qualitative findings, the qualitative data comprised of the principals' and heads of Science Departments' suggestions for improving science education in the schools.

Due to the relatively small sample size (28 schools), inferential statistical analysis was inappropriate to statistically establish the nature of the relationship between the qualitative domains and the school concordancy categories. Therefore, only descriptive inferences were made on the nature of the relationship based on counts of the descriptive responses to the open-ended questions. The counts of descriptive suggestion responses, as elicited in the open ended-questions, were used as the qualitative school-level data and analyzed as such, for each of the concordancy categories identified, as explained in chapter 3.

Principals' and heads of Science Department's suggestions for improving science education relative to school concordancy categories are shown in Table 18 and Figures 3(a) and 3(b). "Improving science laboratories" and "teacher development" were priority suggestions from principals across all the schools. Principals also drummed the need for "public engagement" to enlighten the public about the benefits of science education in the schools and in our daily lives.

Principals in positively discordant and concordant schools also suggested more "focus on science at lower grades before students enter high schools." Only principals in concordant schools suggested "improving reading and writing in order to improve science education in the

Table 18: Suggestions for Improving Science Education Relative to School Concordancy Categories

Improvement Dimensions	**Response Descriptive Adjectives**	**Response Counts in Positively Discordant**	**Response Counts in Concordant Schools**	**Response Counts in Negatively Discordant**	**Total**
Principals' Suggestions for Improving Science Education in the Schools	Improve Science Labs	6	3	4	13
	Teacher Development	3	3	2	8
	Cooperative Science Instruction with External Experts	0	2	5	7
	Public Engagement	2	1	2	5
	Focus on science at Lower Grades	1	1	0	2
	Improve Reading & Writing	0	2	0	2
	Improve Science Curriculum	0	1	0	1
Heads-of-Science Departments' for Improving Science Education in the Schools	Improve Science Labs	4	9	5	18
	Teacher Development	5	9	1	15
	Cooperative Science Instruction with External Experts	0	4	3	7
	Public Engagement	1	3	1	5
	Focus on science at Lower Grades	2	2	1	5
	Improve Reading & Writing	2	0	1	3
	Improve Science Curriculum	1	2	1	4

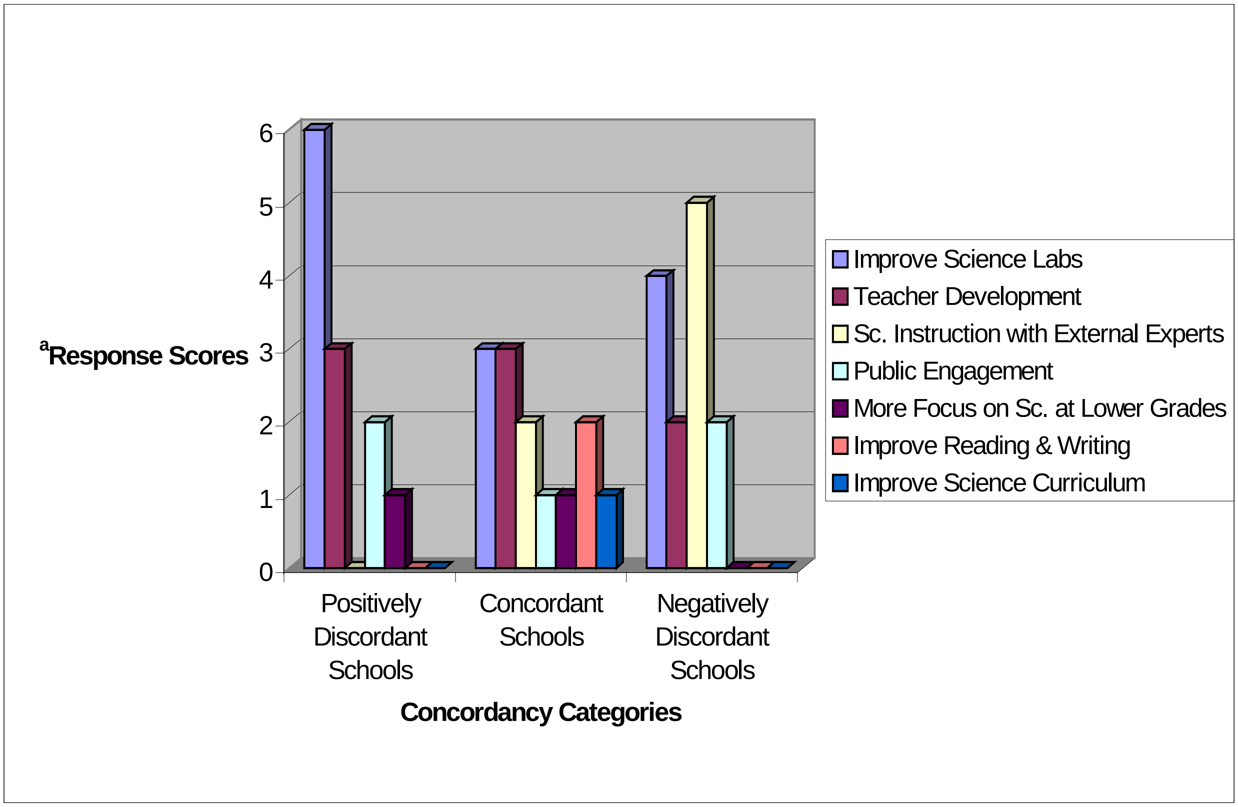


Figure 3a: Principals' Suggestions for Improving Science Education

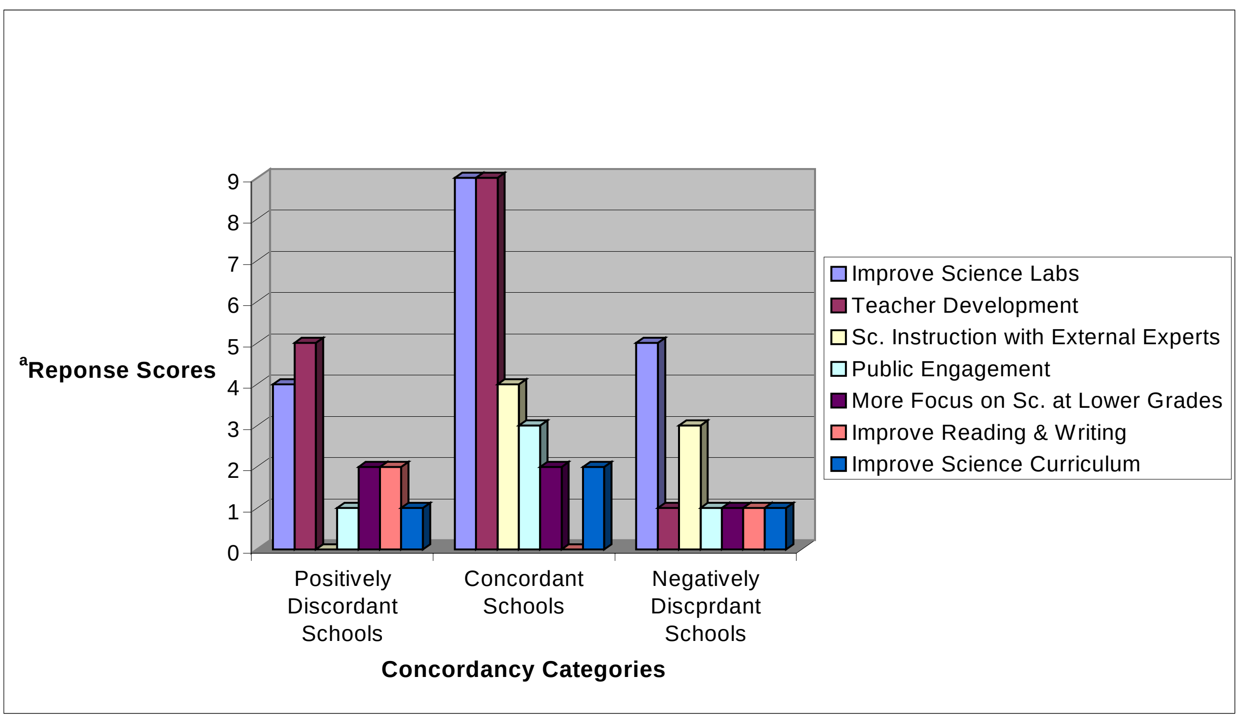


Figure 3b: Heads of Science Departments' Suggestions for Improving Science Education

schools". The principals in concordant and negatively discordant schools also suggested "instructional partnerships" with external experts in the fields of science from both private and public sectors to work with science teachers in cooperative teaching, "to bring science alive in the classrooms and let the kids see the relevance of science in their daily lives." The suggestion that external experts be involved in science instruction was a priority suggestion from principals in negatively discordant schools.

Like the principals, the "need to improve science laboratories and teacher development" were priority suggestions from heads of Science Department. Also, like the principals, the heads of science department in concordant and negatively discordant schools emphasized "the need for involving external experts" especially from universities hospitals, science related industries and agencies in "cooperative instructional delivery in science classrooms to bring science alive and real." The suggestion for "public engagement" to enlighten the public about the benefits of science and science education also resonated from heads of science departments across all the schools.

Post-test Epilogue of Hypotheses Tested

Nine hypotheses were proposed for testing in this study. Given this number of hypotheses, short statements about the findings of the test of each hypothesis is therefore necessary at this point.

Hypothesis #1

Hypothesis: There is no significant difference between school districts on the composite of educational input variables and the composite of achievement output measures.

Post-test Finding: There were statistically significant differences between school districts on the composite of the educational input variables and the composite of achievement output measures considered in this study. This null hypothesis was therefore rejected.

Hypothesis #2

Hypothesis: There is no significant difference between school districts on each of the educational input variables and the achievements output measures, considered separately.

Post-test Finding: There were statistically significant differences between school districts on each of the individual educational input variables and the individual achievement output measures when considered separately in this study. This null hypothesis was therefore rejected.

Hypothesis #3

Hypothesis: There is no significant difference between school districts on the composite of educational input variables and the composite of achievement output measures, after controlling for percentage of economically disadvantaged student.

Post-test Finding: There were statistically significant differences between school districts on the composite of educational input variables and the composite of achievement output measures, after controlling for the percentage of economically disadvantage students. This null hypothesis was therefore rejected.

Hypothesis #4

Hypothesis: There is no significant difference between school districts on each of the educational input variables and the achievement output measures, considered separately, after controlling for the percentage of economically disadvantaged students.

Post-test Finding: There were statistically significant differences between school districts on each of the educational input variables and the achievement output measures, considered separately, after controlling for the percentage of economically disadvantaged students. This null hypothesis was therefore rejected.

Hypothesis #5

Hypothesis: There is no significant correlation between the composite of educational input variables and the composite of achievement output measures across school districts.

Post-test Finding: There is statistically significant correlation between the composite of educational input variables and the composite of output measures across school districts. This null hypothesis was therefore rejected.

Hypothesis #6

Hypothesis: There are no significant correlations between the individual input variables and the output measures, considered separately, across school districts.

Post-test Findings: There was a statistically significant positive correlation between expenditure per pupil and teacher quality. There was a statistically significant positive correlation between teacher quality and the percentage of economically disadvantaged students. There was also a statistically significant positive correlation between the percentage of students passing the science section of the Georgia High School Graduation Test at first sitting and the percentage of graduating students qualified for the HOPE scholarship.

There was a statistically significant negative correlation between the percentage of students passing the science section of the Georgia High School Graduation Test at first sitting and the percentage of economically disadvantaged students. There was a statistically significant negative correlation between the percentage of graduating students qualified for the HOPE scholarship and percentage of economically disadvantaged students. There was a statistically significant negative correlation between per pupil expenditure and the percentage of graduating students qualified for the HOPE scholarship. There was also a statistically significant negative correlation between teacher quality and the percentage of graduating students qualified for the HOPE scholarship.

The correlation between per pupil expenditure and the percentage of students passing the science section of the Georgia High School Graduation Test at first sitting was not statistically significant. The correlation between teacher quality and the percentage of students passing the science section of the Georgia High School Graduation Test at first sitting was not statistically significant. The correlation between expenditure per pupil and percentage of economically disadvantaged students was not statistically significant. Therefore this null hypothesis was therefore rejected in part and accepted in part.

Hypothesis #7

Hypothesis: There is no significant linear relationship between the linear composite of input variables and the linear composite of achievement output measures.

Post-test Findings: There was a statistically significant linear relationship between the linear composite of input variables and the linear composite of achievement output measures. However the linear model was mostly defined by the percentage of economically disadvantaged students. This null hypothesis was therefore rejected.

Hypothesis #8

Hypothesis: There is no significant correlation between school-level input variables and the achievement output measures.

Pos-test Findings: There was a statistically significant negative correlation between the percentage of students passing the science section of the Georgia High School Graduation Test at first sitting and the percentage of economically disadvantaged students. There was also a statistically significant negative correlation between the percentage of graduating students qualified for the HOPE scholarship and the percentage of economically disadvantaged students.

There was a statistically significant positive correlation between the percentage of students passing the science section of the Georgia High School Graduation Test at first sitting and principals' stability. There was also a significantly positive correlation between the percentage of graduating students qualified for the HOPE scholarship and the principals' stability.

However there were no statistically significant correlations between the percentage of students passing the science section of the Georgia High School Graduation Test at first sitting the percentage of graduating students qualified for the HOPE scholarship as output measures, on one hand, and the remaining school-level input variables. Therefore this null hypothesis was rejected in part and accepted in part.

Hypothesis #9

Hypothesis: There is no indication of any other school-level factors, beyond the quantitative inputs considered, that differentiate schools on the achievement measures.

Post-test Findings: There was indication of other school-level inputs (x-factors) beyond the quantitative inputs that differentiate schools on the achievement output measures. Principals and heads of science departments also made useful suggestions based on the achievement concordancy category of the schools. Therefore this null hypothesis was rejected.

The conclusions of this study are summarized in the post-test inferential schematic diagrams as shown in Appendix H.

Summary

This chapter reported the results of the quantitative and qualitative data collected during the course of this study. Both descriptive and inferential statistical techniques were used to elicit the patterns and the nature of the relationships between the investigated educational inputs and measures of student achievement in science as outputs. A post-test epilogue of the nine hypotheses tested was also presented. The next chapter (Chapter 5) presents a discussion of the results, the conclusions, recommendations, and suggestions for further consideration, based on the findings of this study.

CHAPTER 5

DICUSSION AND CONCLUSIONS

Discussion

The purpose of this study was to describe and to explain the relationships between educational inputs and measures of student achievement as outputs. Seventy-one high schools in six metropolitan school districts in the north central area the state of Georgia were selected for the study.

Educational quantitative input variables including overall per pupil expenditure, science department expenditure per pupil, overall school-level teacher quality, science teacher quality, number of professional development trainings per science teacher, science class size, principals' stability, head of science department quality, number of science-lab-based instructions per week, and percentage of economically disadvantaged students, were determined and measured. Qualitative data on suggestions for improving science education were also collected and analyzed.

The quantitative student achievement output measures used were the percentage of students passing the science section of the Georgia High School Graduation Test at first sitting and the percentage of graduating high school students qualified for the State of Georgia's HOPE scholarship award.

In this study, the socio-economic status of students, measured as the percentage of economically disadvantaged students, was treated both as an integral part of the education input-output mix and as a controlling factor. When treated as an integral part of the input-output mix, as done in this study, the positioning of socioeconomic status of students, in the input-output mix, gets the deserved attention with wider implications for performance and achievement assessment in education. Klick (2000) argued for the inclusion of socioeconomic status of students as an educational input, stating that the background economic status of a student is a significant determinant of educational success.

Multivariate methods were used in the collection and analysis of the data for this study in order to capture and take into consideration the multivariate nature of educational inputs and achievement outputs. Educational inputs and outputs do not usually interact in a simple bivariate or univariate manner. Educational inputs and outputs come together in a mix to make the educational system what it is, a multivariate system. Monk (1992) concluded that using just one measure of output for a set of educational inputs “abrogates the responsibility” of the desired attempt to capture education in its true multivariate form. Pedhazur (1997) also noted that analyzing a multivariate context data by using a series of univariate methods, “compromise the very essence” and richness of a multifaceted phenomenon like education.

<u>Differences Across School Districts</u>

The findings of this study showed that the school districts were different on the composite of the input and output variables. The school districts were also different on each of the input and output variables considered separately. The socioeconomic background of students was the number one education production variable that distinguishes districts. Other education production variables, like expenditure per pupil and student achievement, are important only after the socioeconomic background of students was first taken into consideration.

The differences in per pupil expenditure between districts may be due to the differential weights attached to various categories of students in each of the schools in the districts, based on the nature and types of educational programs offered. For example students in special programs like the gifted, vocational studies, and special education, get more funding, per pupil, above the base Full Time Equivalent (FTE) QBE funding for the regular high school students.

Therefore the differences in per pupil expenditure between schools, and hence between school districts, may be due to differences in the types of programs and the relative numbers of students in these programs in the different schools and school districts. The findings of this study may therefore be a ‘litmus pointer’ to the type of programs and the profile of the students in the schools and school districts. Schools with relatively high overall per pupil expenditure, as determined in this study, may have relatively more students in these types of special programs that attract more money under the QBE funding due to differential weighting.

From the results of the means comparisons of all the variables across the six districts, the highest per-pupil-spending school district was not the best performing districts on the two achievement measures considered. Also the lowest per-pupil-spending school district was not the least performing school district on the two achievement measures. This is an indication that there may be other factors, in addition to money, that determine student achievement.

Although teacher quality as measured in this study does not vary much across the school districts, there was still statistically significant difference between school districts on this measure. The lack of great variation in teacher quality across the six school districts may be due to the following reasons. Teacher quality, as measured in this study, may not be adequate enough to elicit high enough variability among the six school districts. This study used a mathematical combination index, constructed from the parameters considered as measures of teacher quality by the Georgia Department of Education, based on highest academic qualification and years of experience of teachers. The measure of teacher quality, as used in this study, did not include consideration for science teachers who actually teach the subjects in their specialized areas such as physics, chemistry or biology, and those who do not. Therefore, the measure of teacher quality, used this study, is either not a good enough proxy measure to elicit significantly high variation between schools and school districts, or may be simply due to lack of enough variation in teacher quality across the districts.

In previous literature, the proxies for teacher quality in education production studies have been years of experience, highest qualification or salary, considered separately (Figilo, 1999; Firestone 1991; Goldhaber & Brewer, 1997b, Goldhaber, Brewer, & Anderson, 1999). None of the proxies seemed to be appropriate measures of teacher quality, in educational terms. Other on-the-job parameters, such as in-field teaching experience and classroom practices may be needed to elicit high enough variability in teacher quality among school districts. This is an area for further study

The findings in this study showed that school districts are still greatly separated on socio-economic basis, possibly akin to the *de facto* segregation of school districts on racial lines as identified by Coleman et al. (1966). This socio-economic *de facto* segregation may also influence some other factors that impact student achievement. Significant variations in socioeconomic status measures are usually characteristic of heterogeneous urban-suburban communities similar to the ones in which the school districts used for this study are located (Aaronson, 1999). The significant disparities in the socio-economic status of students across the school districts call for differentiated educational policy agenda and policy formulation to better reflect the important position of this variable in the education production mix.

There was a statistically significant difference between school districts on the percentage of students passing the science section of the GHSGT at first sitting, as a measure of educational output. Again the school districts with relatively high teacher quality and percentage of economically disadvantaged students also had significantly lower percentage of students passing the GHSGT at first sitting. The six school district studied are the largest six school districts in the state of Georgia, accounting for 33.31% of the average student population of enrolled students in the state during the two academic years covered by the study (Georgia Department of Education, 2002, 2003). The 22.53% differential in performance scores between the highest and the lowest scoring school districts is of practical significance. The 22.53% also translates to the percentage of potential graduates in the lowest achievement school district who could not graduate because they did not pass the science section of GHSGT at first sitting. These are additional students who may need remedial initiatives and hence additional money to fund such initiatives, again an additional burden on the financial resources of the target schools and school districts. A closer study and examination of possible cause(s) of the differences between school districts, on this performance measure, is therefore not only necessary but also imperative.

The statistically significant difference between school districts on the percentage of graduating students qualified for the HOPE scholarship, as a measure educational output, is also noteworthy. Again the school district with the lowest percentage of students passing the science section of the GHSGT at first sitting also has a significantly lower percentage of HOPE scholars than the remaining five school districts. The difference of 18.83% between the district with the highest and the lowest HOPE percentage is of practical significance. The Georgia HOPE scholarship program was established by a legislative act in 1993 to fund postsecondary education for Georgia students who make and keep an average of at least 3.00 GPA from high school through their four-year college education (Council for School Performance, 1999; “Holding on to HOPE", 2003). Therefore, the great disparities in the percentage of potential HOPE scholars across school districts should be an issue of education agenda and policy concern.

Recent development about the HOPE program indicates depleting reserves in the fund ("Holding on to HOPE," 2003). This is currently raising fears that the available HOPE fund, and the future stream of revenue into the fund, may not be enough to sustain all future awards, based on the present criteria. A panel, set up by the governor's office, is currently reviewing the criteria for the award of the scholarship with a view to finding ways to streamline the number of potential qualifiers and recipients of the award. One of the propositions on the table, to streamline the award, is linking the award to scores on the Scholastic Aptitude Teats (SAT), as additional criterion for qualification for the award ("Perdue, SAT, and HOPE," 2003). The implication of this is discussed later when we examine the nature of the relationships between the HOPE and the other variables in the educational production mix, as elicited in this study.

The school districts were also statistically significantly different on the combination of input and output variables, after controlling for the percentage of economically disadvantaged students. After controlling for the percentage of economically disadvantaged students, the most significant input variable that differentiated districts is expenditure after which came achievement. Researchers who had previously worked on education productivity have also expressed this view that money is important in education, after the socioeconomic status of students is taken out of the production equation (Alexander, 1997; Figilo, 1999; Hanushek, 1986, 1989, 1991; Mancebon & Bandres,1999; Monk, 1992; Walberg, 1982). Hence, we can say that money matters, but it only matters in the absence and at the exclusion of the socio-economic circumstances of students. This again makes a case for the socio-economic background of students to be considered as an input when allocating resources for education.

Also after controlling for the percentage of economically disadvantaged students, the findings indicated significant differences between the districts on the each of the remaining five variables. The difference between the highest QBE-allotment school district jumped from an unadjusted value of $586.57 to an adjusted value of $631.00. This again confirms the earlier conclusion that differences in expenditure per pupil may also be due to the differences in the type of educational programs offered and the profile of the students, as reflected by the weighted funding nature of the QBE funding formula. It is also an indication that the socio-economic background of students factors into the type of programs offered by schools and school districts, and hence the funding profile.

After adjusting for the effect of the percentage of economically disadvantaged students, three school districts had higher teacher quality indices than the school with the highest teacher quality index before the adjustment. There is therefore an indication of a relationship between percentage of economically disadvantaged students and teacher quality as determined in this study. The school districts with high percentage of economically disadvantaged student tended to have teachers with either more years of experience or higher academic qualifications or both. The implication of this finding is discussed further when the relationships between the variables are discussed.

After controlling for the effects of the percentage of economically disadvantaged students, and adjusting the percentages of students passing the science section of the Georgia High School Graduation Test at first sitting, and the percentage of graduating students qualified for the HOPE scholarship, the lowest performing school district, before the adjustment, became the highest performing school district on these two achievement output measures after the adjustment. In fact the, the two highest performing school districts before the adjustment, became the two worst performing school districts on the HOPE measure, after the adjustment. This finding has relevance in school performance reporting to the public.

At the state level, the State of Georgia's Office of Student Achievement (OSA), formally Office of Education Accountability (OEA), an arm of the state's Governor's Office, was established July 1, 2000, by Georgia legislature under the *A Plus Education Reform Act* (OCGA, 20-14-25). The OSA provides annual report cards on student achievements in public schools and school districts across the state of Georgia, on standardized test. The current practice of reporting educational achievement by the OSA does not statistically control for the effect of the percentage of economically disadvantaged students before reporting the performances of schools and schools districts. Although, the current reporting practice makes some attempt at disaggregating achievement measures on the basis of race, gender, and PEDS, disaggregating performance on the basis of these factors is not the same as statistically controlling for the factors. Disaggregation is simply a classification process, while statistical control for factors is an equalization process. A review of the reporting system by the OEA may be necessary to provide more objective performance assessment information about schools and school districts to the public.

While not advocating for mediocrity under the guise of being economically disadvantaged, the current way of reporting education performance, without controlling for the socio-economic background of students, disregards the possible value-added gains that some of the schools with high percentage of economically disadvantaged students make in improving student achievement while unfairly rewarding the gains of schools with low percentage of economically disadvantaged students. Therefore, moving students from the supposedly 'low-performing' schools to supposedly 'high-perforating' schools, based on the current reporting practices, as called for by the NCLB Act, could be a recipe for misplaced judgment which may most likely produce the exact opposite of what the NCLB Act intended: no-child-left-behind in the journey towards equality of educational opportunities.

The “high performing’ schools, with low percentages of economically disadvantaged students, may lack the necessary ingredients to make much value-adding gains in educating students from economically disadvantaged backgrounds. Without paying more attention to the circumstances of the economically disadvantaged students, simply moving them from the so called ‘low performing schools’ to the so called ‘high performing schools’ is not a panacea for significantly improving the educational achievement of this group of students. On the contrary, such move may end up being a counterproductive measure.

The Relationships Between Inputs and outputs Across School Districts

The significant positive relationship between the overall composite of the input variables and the composite of the two output measures was defined by socio-economic, financial, and achievement considerations respectively, in descending order of magnitude and effect, as indicated the findings of this study. This result again confirms that money matters in education, but only after the socio-economic situation of students are taken to into consideration.

When each of the variables was considered separately, there was a significant positive correlation between per pupil expenditure and teacher quality. This positive relationship is not surprising. As stated before, the current parameters used for defining teacher quality is based on highest qualification (training) and experience, both of which are used in determining teachers’ salaries and compensations in the state of Georgia. High quality teachers, based on the current parameters, mean higher personnel expenditure and more money per pupil on education. The differences in teacher quality may also account for the differences in expenditure per pupil across school districts.

The significant positive correlation between teacher quality and the percentage of economically disadvantaged students is a contrary finding to the popular anecdotal notion that schools with high percentage of economically disadvantaged students have high teacher turnover, and therefore, could not have high quality teachers based on the current criteria of highest qualification and experience. This study shows that schools with high percentage of economically disadvantaged students tended to have more teachers with higher qualifications and higher experience. This may be due to the fact that experienced teachers with higher qualification may be more comfortable in these types of schools than less experienced teachers with lower qualifications or teachers who are new to the teaching profession.

This finding also makes a possible case for more funding to schools with higher percentages of economically disadvantaged students in order for them to continue to keep and attract higher academically qualified and experienced teachers. This again links to the issue of expenditure per pupil. Since a regular student who is economically disadvantaged is more likely to be taught by a more qualified and experienced teacher, districts with relatively high percentages of economically disadvantaged students are likely to spend more per teacher unit. The political, policy, and financial implications of this are worth examining more closely. There should be a way to compensate schools with high percentage of economically disadvantaged students for this higher expenditure per teacher unit.

The statistically significant positive correlation between the percentage of students passing the science section of the Georgia high school graduation test and the percentage of graduating students qualified for the HOPE scholarship has important practical significance. It makes a case for more attention to science education. The higher the percentage of students who take and pass science at first sitting in the Georgia High School Graduation Test, the higher the possibility of having a higher percentage of students qualifying for the HOPE scholarship. There is therefore a high possibility that students with higher ability in science increase their chances of qualifying for the HOPE scholarship significantly. It also makes a case that the school districts that are not doing so well in science may be shortchanging their students on their HOPE scholarship possibilities.

The strong negative correlation between percentage of economically disadvantaged students, on one hand, and two measures of achievement outcome on the other, again reinforces the need to pay more attention to the socioeconomic status of students as an integral part of educational input, not just as a controlling factor. The result also indicated that increased teacher quality and expenditure per pupil does not necessary translate to increased percentage of students qualified for the HOPE scholarship.

The findings of this study showed a statistically non-significant positive correlation between expenditure per pupil and measures of student achievement, without controlling for the percentage of economically disadvantaged students. This finding is congruent with some of the previous findings on money and educational achievement (Coleman et. al., 1966; Hanushek, 1997) which contended that increasing overall expenditure in the presence of the socio-economic background of students, does not necessarily improve student achievement.

However, as the findings of this study also indicated, on the aggregate the schools that get more money tended to perform better on the two measures of student achievement considered. This is the unresolved dilemma that money and education has posed for education funding and the resulting debate for years. While some education funding researchers (Coleman, 1966;

Hanushek, 1997; Unnever et al., 2000) and the constituencies that get more money for education, believe that money does not matter, other education researchers (Greenwald, Hedges, & Laine, 1996; Klick 2000; Mancebon & Bandres, 1999; Monk, 1992; Tatsuka, 1988; Walberg, 1982; Walberg & Fowler, 1989; Walberg & Weistein, 1982; Weglinsky, 1997) and the constituencies that get less money for education, believe that money matters.

Hanushek (1986) began this debate. In his study on economics of schooling (Hanushek, 1986, Hanushek, 1989) indicated that there appears to be no strong or systematic relationship between school expenditures and student performance. A reanalysis of Hanushek's data by Hedges, Laine, and Greenwald (1994) concluded that there were more instances of positive relationships between educational resources and student achievement, than indicated earlier by Hanushek. Also a follow-up comprehensive independent study by Greenwald, Hedges, and Laine (1996), concluded that there was a systematic positive relationship between money and student outcomes.

The findings of this study still tended to reside in this ambivalent and inconclusive relationship between money and student achievement. This study indicated that there is no significant positive relationship between the student achievement measures and funding allotment in the presence of high percentage of economically disadvantages students. However, in the absence of economically disadvantaged students, money takes priority as a factor because high achieving schools, on the aggregate, tend to spend more money per student compared to low achieving schools when the socioeconomic status of students is considered in the input-output mix. This is exactly the finding of Coleman et al.(1966); Coleman (1968); Monk (1992); Greenwald, Hedges, & Laine (1996); Klick (2000); Hasci (2002). Also since the low achieving schools become the best achieving schools after controlling for PEDS, the socioeconomic status of students is therefore an important determinant of student achievement.

Realistically, the percentage of economically disadvantaged students is also mostly a money issue. Economically disadvantaged students, come to school with possible financial deficiencies. The schools they go to, in turn, inherit these financial deficiencies, albeit indirectly. If this realistic premise is acceptable, then the socioeconomic status of students should be part of the money input into education. Schools with high percentage of economically disadvantaged students are therefore inherently financially deficient.

As the findings of this study also indicated, there is a strong and significant relationship between PEDS and student achievement. Acceptance of the above-stated premise, that the issue socio-economic status of students is a money issue, is an indication that money matters in determining the educational achievement of students. Therefore, when the socioeconomic background of students is considered as a monetary input, there is definitely a significant relationship between money and student achievement. This conclusion is in line with the conclusions of Monk (1992), Greenwald, Hedges, Laine, and Ricahrd (1996), Klick (2000), and Hasci (2002) that money matters in determining the educational achievement of students.

The Production Function Models

The findings of this study also indicated significant positive linear relationship between the overall composite of the input variables and the composite of the two output measures. However, the linear models were mostly defined by the percentage of economically disadvantaged students. Therefore, the significant regression model equations developed for each of the two output variables were:

$$GHSGT = 86.464 - 0.524PEDS$$

$$HOPE = 73.885 - 0.403PEDS$$

These equations indicated that across school districts, the schools could probably have 86.464% of their students passing the science section of the GHSGT at first sitting and 73.885% of their graduating students qualifying for the HOPE scholarship, if there are no economically disadvantaged students in the schools. From the equations, a unit increase in the percentage of economically disadvantaged students may likely produce 0.524% drop in percentage of students passing the science section of the GHSGT and a 0.403% drop in the percentage of graduating students qualifying for the HOPE scholarship. This finding has practical significance.

For example, the 55.92% (66.31-10.39) difference in the percentage of economically disadvantaged students between the school district with the highest percentage of economically disadvantaged students and the school districts with lowest may potentially produce a 29.30% (0.524 x 55.92) difference in the percentage of students passing the science section of the GHSGT, and a 22.54% (0.403 x 55.92) difference in the percentage of students qualifying for the HOPE scholarship between the two school districts. The practical implications of this possible difference in the measures of student achievement may most likely translate to a strain on the financial resources of schools with high percentage of economically disadvantaged students and the future of the students these types of schools serve.

The strong linear relationship between the percentage of economically disadvantaged students (PEDS) and the two outcome measures also calls for caution in the current debate on whether to include additional qualifying criterion for the award of the HOPE scholarship. Including additional standardized exam-based criterion for the award of the scholarship may further disenfranchise economically disadvantaged students on the HOPE scholarship award. As this study has indicated, when the socio-economic background of the students is removed form the performance equation, student achievement measures data changes with the pre-control high performing schools becoming the low performing post-control schools, and the pre-control low performing schools becoming the high performing post-control schools. A more objective approach may be to limit the HOPE scholarship award to a set percentage of the top performing graduating students from each high school. Using this criterion will still factor in the element of achievement, given the limitations imposed by the socioeconomic status of the students. In this way, the degree of disenfranchisement, due to the ambient effect of socioeconomic status on student performance, can be minimized in qualifying for the HOPE scholarship.

The socio-economic status (SES) of students, as measured by the percentage of economically disadvantaged students, in this study, is a social issue that educational establishments may not in the best position to do much about. As a social policy issue, only the policy makers and the larger society can influence SES. Therefore, the immediate practical applications of these regression model equations are for policy makers and educators to use when reporting the educational achievements of our public high schools. Additionally, the equations can be used as a guide to approximately predict high schools' performance in the science section of the Georgia High School Graduation Test and the percentage of graduating students who could potentially qualify for the HOPE scholarship, based on the on the percentage of economically disadvantaged students in the schools. The regression model equations can also be useful in emphasizing the importance of SES as one of the most important educational inputs. As Klick (2000) had rightly noted, "Poverty ... proved to be a significant determinant of whether or not a student will succeed in school" (p. 85).

Relationships Between Inputs and Outputs at the Schools-within-districts Level

The nature of the relationships exhibited, at the school level, by both the percentage of students passing the science section of the GHSGT and the percentage of graduating students qualified for the HOPE scholarship, on one hand, with percentage of economically disadvantaged students on the other, were similar to the relationships exhibited at the district level. The non-parametric tests used for the school-level analysis produced similar results as the parametric multivariate results used for the district level analysis. The similarity in the outcomes of the two types of analyses are a cross-validation of the appropriateness of the multivariate, parametric, and non-parametric tests used for the analyses in this study. Only two of the input variables significantly correlate with the output measures, statistically. There was a positive correlation between principals' stability, as measured by the number of years a principal has spent in a school, and student achievement. Given the limitations of this study, the positive relationship between principal stability and the two achievement output measures is a novel finding.

The positive link between principal stability and the two student achievement measures may be related to the immaturity-maturity continuum concept posited by Argyris (1993). According to Argyris, individuals in an organization, including the leaders, need time to learn, mature, create a state of independence, have a deeper interest in what they do, and refine their visions. As principals become more stable, their organizational maturity increases and they are in better positions to align their visions. Stable principals are more likely to have time to create conducive climates and cultures for learning to take place. In their review of studies across 14 countries in the 1960s, Comber and Kreeves (1973), indicated that school environment modify students habits and, on the aggregate, predicts student scores in science across these countries.

There were also significantly negative relationships between the two measures of student achievement and the percentage of economically disadvantaged students across schools within districts, similar to what was observed across schools districts. The implications of this finding are similar to those discussed for the findings across school districts.

The lack of a statistically significant relationship between science department expenditure per pupil and student achievement in science, at the school level, may be due the relatively small proportion of direct instructional money, per pupil, allocated to the science department, compared to the QBE funding allotment to the school as a whole. The mean science department expenditure per pupil of $5.18 constitutes only 0.19% of the average QBE allotment funding across schools. Given the material and equipment requirements of science instruction, this is a relatively small amount of instructional money to make any impact on student achievement in science. The special weighting for science labs, above the other regular education programs, under the QBE funding, was removed. Therefore, science education, in competition with other subject disciplines, is apparently in for a long period of inadequate funding and hence continued mediocrity in science achievement in our high schools.

The non-significant relationship between teacher quality and the percentage of student passing the science section of the GHSGT at first sitting also has practical implications. A very specialized academic discipline area like science requires teachers with specialization in the specific subject areas. One of the main concerns in science education has been that science educators often lack sufficient content knowledge (Monk, 1992). Monk (1992) reported a positive relationship between student achievement and the number of courses the teachers had taken in the subjects they taught.

In recent years there has been concern about the increasing number of science teachers who teach out of the subject-field of their specialization (U.S. Department of Education, NCES, SASS, 2002b). Although all the science teachers in the schools studied had degrees in science, the measure of teacher quality, as determined by the state of Georgia's Department of Education, and transformed for use in this study, did not include in-field or out-of-field of specialization teaching assignment as a factor. Therefore the measure of teacher quality, as a composite of experience and highest academic qualification, may not be an effective measure that impacts student achievement in science.

The statistically non-significant relationship between science class size and student achievement may be due to the small variability in science class size across schools. The mandated maximum class size for science in the state of Georgia is 28 students. As shown in this study, the average science class size was 24 students. Research has shown that the benefits of smaller class sizes only become significant at class levels between 17 and 20 and for students in the lower grades (Grissmer, 1999; Hanushek, 1999; Mostella, 1995; Word et al., 1990). However, because of the specialized natured of science instruction, requiring hands-on activity, a lower class size may have an impact on improving the quality of science instruction. A reduction in class size requires additional teachers and hence, additional expenditure. The current budget crunch, for the fiscal year 2004, in the state of Georgia, is putting pressure on education spending and thus making any reduction in class size almost impossible in all subject areas including science.

Although the positive relationship between the number of lab-based activities per week and student achievement is statistically non-significant at the very stringent alpha-level (0.006), the value of the correlation coefficient is high enough to be of practical significance. Previous research by O'Sullivan and Weiss (1966) had reported that the majority of high school students, who did hands-on lab-based activities, at least once a week, were more likely to perform above proficiency level than those who did not. The finding of this study tends to agree with this earlier finding. A stronger relationship may be established between student achievement in science and the number of hands-on labs per week with improvements in the quality of the laboratory facilities and professional development for science teachers. As Irving, Dickson, and Keyser (1999) noted, the quality of instruction in science needs significant improvement in order to make significant gains in student achievement. This ties in with the issue of professional development for science teachers.

Similarly, the relationships between the numbers of professional development activities per science teacher per academic year and the measures of student achievement were not statistically significant, however the value of he correlation coefficients were high enough at the stringent test alpha-level (0.006) to be of practical significance. This finding agrees with findings reported by the Council for School Performance (1994) indicating a positive relationship between staff development and student achievement in Georgia public schools.

The statistically non- significant relationship between the numbers of professional development activities per science teacher per academic year and the measures of student achievement may also be tied to the quality of the professional development activities for science teachers. According to Stofflett and Stoddart (1994) professional development that emphasizes constructivist practices that merges content knowledge with appropriate and relevant practical hands-on activities is the key to make a difference in improving the achievement of student in science. The U.S. Department of Education, *NCES* (2002b) reported that only 40% of full-time teachers indicated participating in professional development that focus on in-depth study in their main teaching fields. There is therefore a need to ensure that the type and quality of professional development for science teachers match their pedagogical interests and focus on solving the problems they encounter in their classrooms.

Indications of other School-level Factors Beyond the Quantitative Variables and Achievement

The findings of this study showed possible indications of other school-level factor (x-factors) beyond money and economic factors that affect student achievement. There were schools that had high percentages of economically disadvantaged students and yet performing above expectation; these were the 'positively discordant schools.' There were also schools that had low percentage of economically disadvantaged students and yet performing below expectation. These were the 'negatively discordant schools.' Therefore, socioeconomic status, as defined by the percentage of economically disadvantaged students, and other quantitative inputs, may not be the only variables responsible for variations in the student achievement measures considered in this study across all schools.

In his study of allocative efficiency, Anderson (1996) used extensions of the Data Enveloping Analysis (DEA) to identify effective school relative to resource allocation and utilization. He associated attendance, stable leadership, stable student population, and school climate with effective schools. A further study on the types and nature of the x-factors that may impact student achievements at the individual school level, beyond money, is suggested.

Suggestions for Improving Science Education In The Schools

The findings of this study indicated that improving science laboratory facilities and teacher development appear to be the consensus priority suggestions on how to improve science education. The findings also showed that, although there was a positive relationship between the numbers of hands-on labs per week per science class and the percentage of students passing the science section of the GHSGT, the relationship was not statistically significant. Hands-on science labs are observable school-level instructional variables. The issues regarding the effects of observable instructional variables on student achievement have been pointed out in the works of Goldherber, Brewer, and Anderson (1999). Their studies concluded that although the incremental contributions of observable variables to student achievement were noticeable, they were not significant.

The reason for the non-significant impact of hands-on labs on student achievement may be due to the types and quality of the hands-on labs conducted in the schools. The problem may be tied to the issues of lack of science laboratory facilities, science teacher quality, science teacher development, or combination of all of these factors. The non-significant relationship between the number of hands-on lab instructions and achievement in science may be what principals and heads of science department have in mind in suggesting the need for improvements in laboratories and teacher development as a priority in improving science achievement. There is therefore a need for more funding for science to improve facilities, materials, and professional practices of science teachers in order to significantly improve student achievement in science.

The Mathematics and Science Initiative (MSI), carved out of The No Child Left Behind Education Act, NCBL, by the Secretary of Education, Rod Paige (U.S. Department of Education, 2002), focuses on three major areas of science education: improvement in instruction, public engagement, and improvement in science teacher development. More focus on improving the quality of science teachers through science teacher development, in areas relevant to the applications and use of new experiments and practical models of science instruction, may go a long way in establishing a more significantly positive relationship between hands-on lab-based instructional activities and student achievement in science.

The issues of: (1) public engagement; (2) more focus on science education at lower grades; (3) improving reading and writing; and (4) improving science curriculum, as ways of improving the achievements of students in science, also cut across the schools. As stated before, part of the focus of the science initiative, under the NCLB Act, is public engagement involving the education and enlightenment of the pubic, especially parents, about the relevance of science to our daily lives and the importance of science education to our children. Improving reading and writing also tie with curriculum and instructional issues, which the science initiative also calls for.

The findings of this study indicate that curriculum is a priority suggestion from principals and heads of science department for improving science education in the state of Georgia. John Dewey (1938) (as cited in Tanner, 1999) was of the view that curriculum should be an experiential and interest-based construct for students, with teachers as facilitators of learning. This view has also been expressed by Tanner and Tanner (1990) who view curriculum "as the planned and guided intended learning experiences and outcomes of education" (p.7). These assertions are more true to the area of science, where the consideration for engagement of students' interest and educational outcomes are of paramount importance. Therefore attention to the curriculum issue in science, especially a curriculum that focuses on the tactile and daily experiences of students, is highly suggested.

One other suggestion for improving science education, that distinguishes the concordant and negatively discordant schools from the positively discordant schools, is the call for the involvement of external science experts in the cooperative delivery of science instructions in the schools. This is a completely new and radical revelation. Although this suggestion may be tied to the public engagement issue, it should be seen more as an issue bordering on a 'cry for instructional help', or a cry for a change in the curriculum from the concordant and negatively discordant schools, as one of the alternative ways of stimulating student interests and improving student achievement in science in these schools.

The suggestion for "need to involve external experts in science instruction" is considered a curriculum issue. Therefore, science curriculum, as presently constituted, may be in need of some improvement. The state of Georgia is currently developing a new science curriculum (Georgia Department of Education, 2003). Policy makers may be linking the relatively low performance of students in science as a signal to realign the curriculum towards what they consider the essential knowledge to attain the desired outcome in science. The role of the different levels of government as sources of curriculum for schools has been well documented (Glatthorn, 1998). In addition to changing the curriculum, the focus of the state government and educators should also be on those factors that contribute to improved achievement in science at the school level, including socio-economic, financial, and other school level factors, as identified by the findings of this study.

The need for the involvement of external science experts also ties with need for teacher development and may also relate to the lack of statistically significant positive correlation between the number of hands-on lab-based instructions and student achievement in science. All of these may border on the quality of the hands-on lab-based instructions. Science teachers may not be versed enough in the best practices for effective lab-based instructions to significantly impact student achievement in science. As Irving, Dickson, and Keyser (1999) noted:

> To break the didactic teaching-learning cycle, a professional development collaboration between government education professionals; university-based schools, colleges, and departments of education (SCDE); and public school systems that emphasizes the importance of merging content knowledge with teaching practices. (p. 412)

A few governmental agencies, like the National Institute of Health (NIH), National Aeronautical and Space Agency (NASA), and the Institute for Theoretical Physics (ITP), all have educational outreach programs designed to bring science knowledge, in their respective agencies, into the classroom. However, these programs only deal with workshops for teachers, educational videos for science instruction, and publications for science teachers and students. While these programs are laudable, part of the findings of this study has shown that actively bringing the practicality of science into the classrooms, by these and any other external body, may likely go a long way in helping to improve the quality of hands-on lab-based instruction, and possibly improve student achievement in science.

Conclusions

The purpose of this study was to describe and to explain the relationships between educational inputs and measures of student achievement as outputs. The following conclusion may be made form the findings of this study:

First, educational inputs and outputs are multivariate and simultaneous events, and should be studied and analyzed as such. This study has shown that there are relationships between the composite of educational inputs and outputs that are not necessarily revealed when inputs and outputs are treated univariately and separately. A large number of reported research studies on education productivity have so far been conducted and reported by researchers in the fields of econometrics and statistics (Dolan & Schmidt, 1987; Hanusheck, 1996a, 1996b; Monk, 1992; Unnever, Kerckhoff, & Robinson, 2000; Walberg, 1982; Wenglinsky, 1997). Educators and education researchers should be more involved in developing multivariate models and methods for analyzing the multifaceted nature of education as a process, in line with what was done in this study, for more meaningful, robust, and relevant conclusions that may be more useful in shaping education agenda and policies.

Second, this study also indicated differences between schools and school districts in expenditure per pupil, which may be linked to differences in the types of educational programs and numbers of students in these programs across the schools and school districts. However, as indicated in this study, there is also a need for the state to weigh funding for science education, especially for economically disadvantaged students, above the current regular basic education funding provided under the QBE funding formula, just like it is presently done for certain special programs. The additional funding could be used for involving external experts in science instruction in schools with high percentage of economically disadvantages students, as suggested by the principals and heads of science department in these schools. Also the separate funding for science that used to come directly from the state should also be re-introduced to differentially fund science education in the state's public schools.

Third, this study has also indicated that expenditure is important, but it only becomes important after the socioeconomic backgrounds of students are taken out of the equation. This again makes a case for why there should be a way of factoring in the financial deficiencies resulting from the socio-economic class of students as an integral part of the financing side of the funding equation as an educational input. The possible additional funding should be allocated directly to the science departments to be used in implementing adequate science curriculum involving differentiated instructional practices and external experts in science instruction in schools with high percentage of economically disadvantaged students. This may ensure a more equal playing field for educational opportunities in science education. Further research is therefore suggested to collect and make a case for the inclusion of the socioeconomic background of students as an input consideration in money allocation for science, rather than the present view of the socio-economic background of students as an external controlling factor in education policy decisions and funding practices for science.

Fourth, this study has also contributed to the debate regarding the relevance of the current measures of teacher quality based on salary, experience, or highest academic qualification, as it relates to student achievement. The lack of meaningful relationship between the current measures of teacher quality and student achievement, despite the positive relationship between the present definitions of teacher quality and per pupil expenditure, calls for further research into a more relevant measure of teacher quality that bears more direct relationship to student achievement, particularly in the area science education.

Fifth, although there appeared to be no statistically significant relationship between teacher quality and the two measures of student achieving, however, the schools with higher teacher quality performed better than school with lower teacher quality after controlling for the socio-economic background of students. Since the schools with high percentages of economically disadvantaged students tended to have higher teacher quality as presently defined, and hence spend more expenditure per teacher unit, part or partial compensatory measures could be considered for these schools to offset the relatively higher cost per teacher unit.

Sixth, as revealed by this study, the current amount of money that goes directly to science departments for science education and instruction, at the school-level, is a pittance when compared with the overall school-level educational expenditures. If the state of Georgia is serious about science education, and given the special nature of science, the funding for science instruction should be weighted above the other regular programs, like some of the other special educational programs. The money for direct instruction for science should be allocated directly to the science departments in the schools, where science educators can have control to ensure that it is spent in a way and manner that are most likely to impact student achievement.

Finally, beyond fiscal considerations, more focus should also be placed on other school-level factors, including resource allocation practices, school leadership, leadership stability, and classroom practices, which may impact student achievement. As this study indicated, other non-financial factors may have some impact on school performance and student achievement. More research is therefore needed to further investigate other school-level variables dealing with teachers, students, and the classrooms and their impacts on student achievement in science; for more informed agenda and policy considerations for science education in public schools in the state of Georgia.

REFERENCES

Aaronson, D. (1999). The effects of school finance reform on population heterogeneity. *National Tax Journal, 52*(1), 5-29. Retrieved October 27, from *http://proquest.umi.com*.

Alexander, K. L. (1997). Public schools and public good. *Social Forces, 76*(1), 1-30.

Alexander, K. L., & Griffin, L. (1976). School district effects on academic achievement: A reconsideration. *American Sociological Review, 41*(1), 144-152.

Anderson, D. M. (1996). Stretching the tax dollar: Increasing efficiency in urban and rural schools. In P. Lawrence & J. Wattenbarger (Eds.), *Where does the money go? Resource allocation in elementary and secondary schools* (156-177). Thousand Oaks, CA: Corwin Press.

Andrew Young School of Policy Studies. (2001). Evaluating the HOPE scholarship program: Impact on student attending public colleges and universities. Author. Retrieved June 25, 2003 from *http://www.arc.gsu.edu/csp/scp_georgiahope.htm*

Argyris, C. (1993). The individual and the organization: *Some problems of mutual adjustment.* New York: Irvington.

Becker, G. S. (1964). *Human capital*. New York: Columbia University Press.

Bloom, B. (1976). *Human characteristics and school learning*. New York, NY: McGraw-Hill.

Boardman, A., Davis. O., & Sandy, P. (1977). A simultaneous equations model of the education process. *Journal of Public Economics, 7*, 23-49.

Burtless, G. (1996). *Does money matter? The effects of school resources on student achievement and adult success*. Washington, DC: Brookings Institution Press.

Card, D., & Krueger, A. B. (1996). School Resources and student outcomes: An overview of the literature and new evidence from North and South Carolina. *Journal of Economic Perspectives, 10*(4), 31-50.

Carroll, J. (1963). A model of school learning. *Teachers College Record, 68*(8), 723-733.

Carnegie Commission on Science, Technology and Government. (1991). *In the national interest: The federal government in the reform of K-12 math and science education*. New York: Author.

Catherall, J. S. (1997). Reflection on economic analysis and education policy. *Education Evaluation and Policy Analysis, 19*(4), 297-299.

Clotfelter, C. T., & Ladd, H. F. (1999). Recognizing and rewarding success in public schools. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp23-63). Washington D.C: Brookings Institute.

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

Cohn, E., & Geske, T. G. (1990). *The economics of education*. Elmsford, NY: Pergamon Press.

Coleman, J. S. (1968). The concept of equity of educational opportunity. *Harvard Educationl Review, 38*(7), 7-22.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, M. M., Weinfeld, F. D., & York, R. L. (1966). *Equity of educational opportunity*. Washington, D.C: U.S. Government Printing Office.

Coleman, P., Easton, S., & LaRocque, L. (1998). Reconciling equity and excellence: Resource allocation in school districts. In W. T. Hartman & W. L. Boyd (Eds.), *Resource allocation and productivity in education*. (pp 101-128). Westport, CT: Greenwood.

Comber, L. C., & Keeves, J. P. (1973). *Science education in nineteen countries: An empirical study*. Stockholm: Wiley.

Cook, T. D., & Campbell, D. T. (1979). *Quazi-experimentation: Design & analysis issues for field settings*. Chicago, Il: Rand McNally.

Cooper, B. S., Sarrel, R., Darvas, P., Alferno, E., Samuels, J., & Heinbuch, S. (1994). Making money matter in education: A micro-financial model for determining school-level allocations, efficiency, and productivity. *Journal of Education Finance, 20*(1), 66-87.

Council for School Performance. (1994). Effect of staff development on student achievement in Georgia. Author. Retrieved May 14, 2003, from http:*//www.arc.gsu.edu/cps/cps_geogiahope.htm*

Council for School Performance. (1998). Staff development and student achievement: Making the connection in Georgia schools. Author. Retrieved May 14, 2003, from *http://www.arc.gsu.edu/cps/cps_geogiahope.htm*

Council for School Performance. (1999). Hope longitudinal study: Year 2 results. Authur. Retrieved May 14, 2003, from *http://www.arc.gsu.edu/cps/cps_geogiahope.htm*

Cox, R. D. (1958). *Planning of experiments*. New York: Wiley

Darling-Hammond, L., & McLaughlin. M. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan, 76*(5), 389-392.

Darling-Hammond, L., Wise, A. E., & Klein, S. P. (1995). *A license to teach: Building a profession for 21st-century schools*. Boulder, CO: Westview Press.

Dolan, R. C., & Schmidt, W. (1987). Assessing the impact of expenditure on achievement: Some methodological and policy considerations. *Economics of Education Review, 6*(3), 285-299.

Draper, N., & Smith, H. (1981). *Applied regression analysis*. New York: Harper & Brothers.

Dublin, L. I., & Lotka, A. J. (1946). *The money value of man*. New York: Ronald Press Company.

Edwards, N. T. (1997). Integrating content and pedagogy: A cultural journey. *Action in Teacher Education, 19*(2), 44-54

Ferguson, D. F. (1991). Paying for public education: New evidence on how money matters. *Harvard Journal on Legislation, 28*(2), 465-498.

Ferguson, D. F., & Ladd, H. F. (1996,). *Holding schools accountable*. Washington DC: The Brookings Institute.

Figilo, D. N. (1999). Functional form and estimated effects of school resources. *Economics of Education Review, 18*, 241-252.

Finn, J. D., & Achilles, C. M. (1999). Answers and questions about class size: A statewide experiment. *American Educational Research Journal, 27*, 557-557.

Firestone, W. A. (1991). Merit pay and job enlargement as reforms: Incentives, implementation, and teacher response. *Educational Evaluation and Policy Analysis, 13*(3), 269-288.

Fogelman, K. (1984a).Assessment of examination performance in different types of schools. *Journal of the Royal Statistical Society, 147*, 569-581.

Forgleman, K. (1984b). Assessment of examination performance in different types of schools. *Journal of Royal Statistical Society, 146*, 569-581.

Fortune, J. C., & O'Neil, J. S. (1994). Production function and study of educational funding equity: A methodological critique. *Journal of Education Finance, 20*, 21-46.

Fried, H. O., Lovell, C. A. K., & Schmidt, S. S. (1993). *The management of productive efficiency: Techniques and applications*. New York: Oxford University Press.

Georgia Department of Education (2002). *2000-2001 Georgia public education report card, State of Georgia*. Atlanta, GA: Office of Educational Accountability, GA Department of Education. Author.

Georgia Department of Education (2003). 2001-2002 *Georgia public education report card, State of Georgia*. Atlanta, GA: Office of Educational Accountability, GA Department of Education. Author.

Georgia HOPE scholarship program. (1997, February). Georgia HOPE scholarship program: Good intentions, strong funding, bad design. C*opy Postsecondary Education Opportunity Editor, 12-14*.

Glass, G. V., Cohen, L.S., Smith, M. L., & Filby, N. N. (1982). *School class size research and policy*. Beverly Hills, CA: Sage.

Glass, G. V., & Smith, M. L. (1978). *Meta-analysis of research on the relationship of class size and achievement*. San Francisco: Far West Laboratory for Educational Research and Development.

Glatthorn, A. A., (1998). *The principal as curriculum leader*. Thousand Oaks, CA: Corwin Press, Inc.

Goldhaber, D. D., & Brewer, D. J. (1997a). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on education productivity. *Journal of Human Resources, 32*, 505-523.

Goldhaber, D. D., & Brewer, D. J. (1997b). Evaluation the effect of teacher degree level on education performance. In W. Fowler (Ed.), *Developments in School Finance, 1996* (NCES 97-535) (pp. 197-210). U.S. Department of Education, NCES. Washington, DC: U.S. Government Printing Office.

Goldhaber, D. D., Brewer, D. J., & Anderson, D. J. (1999). A three-way error component analysis of education productivity. *Education Economics, 7*(3), 199-208. Retrieved on October 26, 2001 from *http://www.gallelio.org/pqweb*

Goldschmidt, P., & Eyermann, T. S. (1999). International education performance of the United States: Is there a problem that money can fix? *Comparative Education, 35*(1), 27-43. Retrieved on August 21, 2001 from www. *http://www.gallelio.org/pqweb.*

Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research, 26*, 499-510.

Green, S. B., Salkind, N. J., & Akey, T. M. (2000). *Using SPSS for windows: Analyzing and understanding data*. Upper Saddle River, New Jersey: Prentice Hall.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66*(3), 361-396.

Griffin, P., & Ganderton, P. T. (1996). Evidence of omitted variable bias in earnings equations. *Economics of Education Review, 15* (2), 139-148.

Grissmer, D. (1999). Class size effects: Assessing the Evidence, its policy implications, and future research agenda. *Educational Evaluation and Policy Analysis, 21*(2), 321-248.

Grissmer, D. (2002). Cost-effectiveness and cost benefit analysis: The effect of targeting interventions. In H. M. Levin & P. J. McEwan (Eds.), *Cost-effectiveness and educational policy: American Education Finance Association 2002 Yearbook* (pp. 97-110). Larchmont, NY: Eye on Education.

Grissmer, D., Flanagan, A., Kawata, J., & Williamson, S. (2000). *Improving student's achievement: What NAEP state test scores tell us*. Santa Monica, CA: RAND.

Gyimah-Brempong, K., & Gyapong, A. O. (1992). Characteristics of education production functions: An application of canonical regression analysis. *Economics of Education Review, 10* (1), 7-17.

Haertel, E. (1991). New forms of teacher assessment. In G. Grant (Ed.), *Review of Research in Education* (pp. 3-29). Washington, DC: American Educational Research Association.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1992). *Multivariate data analysis with readings*. New York: Macmillan.

Hamilton, S. F. (1983). The social side of schooling: Ecological studies of classrooms and schools. *The Elementary School Journal, 83*(4), 313-334.

Hand, D. J., &Taylor, C. C. (1987). *Multivariate analysis of variance and repeated measures: A practical approach for behavioral scientists*. New York: Chapman and Hill.

Hanushek, E. A. (1971). Teacher characteristics and gains in student achievement: Estimation using micro-data. *American Economic Review, 61*(2), 280-288.

Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources, 14*(1), 351-88.

Hanushek, E. A. (1981). Throwing money at schools. *Journal of Policy Analysis and Management, 1*(1), 91-41.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature, 24*, 1141-78.

Hanushek, E. A. (1989). The impact of differential expenditures on school performance. *Educational Resercher, 18*(4), 45-65.

Hanushek, E. A. (1991). When school finance reform may not be a good policy. *Harvard Journal of Legistlation, 28*, 423-456.

Hanushek, E. A. (1994a). Money might matter somewhere: A response to Hedges, Lane and Greenwald. *Educational Researcher, 23*(4), 5-8.

Hanuskeck, E. A. (1994b). Making America's schools work: This time money is not the answer. *The Brookings Review, 12*(4), 10-14. Retrieved March 12, 2003, from *http://proquest.umi.com/pqdweb.*

Hanushek, E. A. (1996a). School resource and student performance. In, G. Burless (Ed.). *Does money mater: Effect of school Resource on students achievement and adult success* (pp 43-73). Washington, DC: Brooking Institute.

Hanushek, E. A. (1996b). Measuring investment in education. *Journal of Economic Perspectives, 10*(4), 9-30.

Hanushek, E. A. (1997). Assessing the effects of school resources and student performance: An update. *Educational Evaluation and Policy Analysis, 19*, 141-164.

Hanushek, E. A. (1998). *The evidence of class size*. Rochester, NY: University of Rochester, W. Allen Wallis Institute of Political Economy.

Hanushek, E. A. (1999). Some findings from an independent investigation of the Tennessee STAR Experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis, 12*(2), 143.

Hanushek, E. A., Rivken, S. G., & Taylor, L. L. (1996). Aggregation and the estimated effects of school resources. *The Review of Economics and Statistics, 78* (4), 611-627.

Harris, D. (2002). Identifying optimal class sizes and teacher salaries. In H.M. Levin, & P. J. McEwan (Eds.), *Cost effectiveness and education policy: American Education Finance Association 2002 Yearbook* (pp 177-187). Larchmont, NY: Eye on Education.

Harris, R. J. (1985). *A primer of multivariate statistics*. Orlando, FL: Academic Press.

Hasci, T. A. (2002). *Children as Pawns: The politics of education reform*. Cambridge, Massachusetts: Harvard University Press.

Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the affects of differential school inputs on student outcomes. *Education Researcher, 23*(3), 5-14.

Heinbuch, S. E., & Samuels, J. A. (1995). Getting more bang for the education buck. *Public Productivity & Management Review, 18*(3), 233-342. Retrieved August 22, 2001 from *http://proquest.umi.com.*

Holding on to HOPE. (2003, November 9). *The Atlanta Journal & Constitution*, p. E1.

Hotelling, H. (1936). Relations between two sets of variables. *Biometrica, 28*, 321-337.

Huck, S. W. (2000). *Reading statistics and research*. Reading, Massachusetts: Addison Wesley Longman, Inc.

Ingersoll, R. (1999). The problem of underqualified teachers in American secondary schools. *Educational Researcher, 28*(3), 20-31.

Irving, M. M., Dickson Jr., L. A. & Keyser, J. (1999). Retraining public secondary teachers by updating their content knowledge and pedagogical skill. *Journal of Negro Education, 68*(3), 409-419. Retrieved August 24, 2003 from *http://proquest.umi.com*

Johnes, G. (1993). *The economics of education*. New York: St. Martin's Press.

Kerlinger, F. N. (1986). *Foundations of behavioral research*. New York: Holt, Rinehart and Winston.

Klick, J. (2000). Do dollars make a difference? The relationship between expenditures and test scores in Pennsylvania's schools. *American Economist, 44*(1), 81-87. Retrieved August 22, 2001 from *http://proqueat.umi.com.*

Kraemer, N. C., & Thiemann, S. (1987). *How many subjects? Statistical power analysis in research*. New bury Park, CA: Sage.

Kremer, M. R. (1995). Research on schooling: What we know and what we don't – A comment on Hanushek. *The World Bank Research Observer, 10*(2), 247-254.

Kres, H. (1983). Statistical tables for multivariate analysis: *A handbook with reference to applications*. New York: Springer-Verlag.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics, 114*(2), 497-532.

Lauter, J. (1978). Sample size requirements for the T^2 test of MANOVA (tables for one-way classification). *Biometric Journal, 20*, 389-406.

Leibenstein, H. (1966). Allocative efficiency and x-efficiency. *The American Economic Review, 56*, 392-425.

Leithwood, K., & Jantzi, D. (1999). The relative effects of principal and teacher sources of leadership on student engagement with school. *Educational Administration Quarterly, 35*, 679.

Levin, H. M. (1997). Raising school productivity: An x-efficiency approach. *Economics of Education Review, 16*(3), 303-311.

Levin, H. M., Glass, G. V., & Meister, R. G. (1987). Cost-effectiveness of computer-assisted instruction. *Education Review, 11*(1), 50-72.

Levin, H. M., & McEwan, P. J. (2001). *Cost-effeteness analysis: Methods and Applications*. Thousand Oaks, CA: Sage Publications, Inc.

Levin, H. M., & McEwan, P. J. (2002). *Cost-effeteness and education policy*. In H. M. Levin & P. J. McEwan (Eds.), *Cost-effectiveness and education policy* (pp1-16). Larchmont, NY: Eye on Education.

McEwan, P. J. (2002). Are cost effectiveness methods used correctly? In Henry M. Levin, & Patrick J. McEwan (Eds.), *Cost-effectiveness and educational policy: American Education Finance Association 2002 yearbook* (pp 37-53). Larchmont, NY: Eye on Education.

Madaus, G. ft., Kellaghan, T., Rakow, E. A., & King, D. J. (1997). The sensitivity of measures of school effectiveness. *Harvard Education Review, 49*, 207-230.

Mancebon, M., & Bandres, E. (1999). Efficiency evaluation in secondary schools: The key role of model specification and of ex post analysis of results. *Education Economics, 7*(2), 131-152. Retrieved October 27, 2001 at *www.pqweb.org*

Michel, R. C. (1991). Economic growth and income equity since 1989. J*ournal of Policy Analysis and Management, 10* (2), 181-203.

Michell, D. E. (1998). A policy framework for school resource allocation. In W. T. Hartman & W. L. Boyd (Eds.). *Resource allocation and productivity in education*. Westport, CT: Greenwood.

Monk, D. H. (1992). Education resources for improving schools: A case study of teacher allocation in Boston Public Schools. *Education Evaluation and Policy Analysis, 14* (4), 307-332.

Montmarquette, C., & Majseredjian, S. (1989). Does school matter for educational achievement? A two-way nested error component analysis. *Journal of Applied Economics, 4,* 181-193.

Mortimore, P., Simmons, P., Stoll, L, Lewis, D., & Ecob, R. (1988). *School Matters: The junior years*. Somerset: Open Books.

Mostella, F. (1995). The Tennessee study of class size in the early school grades. *The Future of Children, 5,* 113-127.

Murnane, R. (1975). *Impact of school resources on the learning of inner city children*. Cambridge, MA: Ballinger.

Murnane, R. J., & Levy, E. (1996). Why money matters sometimes. New York: Free Press.

Murnane, R. J., & Nelson, R. R. (1984). Production and innovation when techniques are tacit. *Journal of Economic Behavior and organization, 5*, 353-373.

National Commission on Excellence and Equity in Education. (1983). *A nation at-risk: The imperative of educational reform*. Washington, D.C: United States Department of Education.

Nye, B., Hedges, L. V., & Konstantopoulos, S. (1999). The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment. *Educational Evaluation and Policy Analysis, 21*(2), 127-142.

Oakes, J. (1990). Lost talent: *The under preparation of women, minorities, and disabled persons in science*. Santa Monica: CA: The RAND Corporation.

Official Code of Georgia Annotated. (1994). *OCGA, 20-2-140*.

Official Code of Georgia Annotated. (1996). *OCGA, 20-2-281*.

Official Code of Georgia Annotated. (2000a). *OCGA, 20-2-58*.

Official Code of Georgia Annotated. (200b). *OCGA, 20-2-86*.

O'Sullivan, C., & Weiss, A. R. (1996). *Student work and teacher practices in science*. Education Statistics Quarterly, U.S. Department of Education, NCES, Washington, DC: U.S. Governemnt Printing Press.

Pedhazur, E. J. (1997). *Multivariate regression in behavioral research: Explanation and prediction*. 3rd Ed. Toronto, Ontario: Wadsworth.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Perdue wants to add SAT score to HOPE requirement. (2003, November 27, *The Atlanta Journal-Constitution*, A1).

Poss, G. (1993). The quest for equal opportunity: An examination of intra-district fiscal diperisty among selected elementary schools in one Kansas school district. Unpublished doctoral dissertation, Kansas State University.

Rice, J. K. (1999). Cost analysis in education: Paradox and possibility. *Education Evaluation and Policy Analysis, 19*(4), 309-317.

Rice, J. K. (1997). The impact of class size on instructional strategies and the use of time in high school mathematics and science courses. *Education Evaluation and Policy Analysis, 21*(2), 215-229.

Robinson, G. E. (1990). Synthesis of research on effects of class size. *Educational Leadership, 47*(7), 80-90.

Robinson, G. E., & Wittebols, J. H. (1986). *Class size research: A related cluster analysis for decision making*. Arlington, VA: Educational Research Service.

Saber, G. A. F. (1984). *Multivariate observations*. New York: John Wiley.

Sander, W. (1993). Expenditures and student achievement in Illinois. *Journal of Public Economics, 52*, 403-416.

Shannon D. W., & Davenport, M .A. (2001)**.** *Using SPSS to solve statistical problems: A self-instruction guide*. Upper Saddle River, New Jersey: Merrill Prentice Hall.

Slavin, R. E. (1989). Achievement effects of substantial reductions in class size. In R. E. Slavin (ed.), *School and classroom organization* (pp.247-257). Hillside, NJ: Erlbaum.

Smith, D. and Tomlinson, S. (1989). *The school effect: A study of multi-racial comprehensives*. London: Policy Studies Institute.

Sparks, D., & Hirsh, S. (1997). *A New Vision for Staff Development*. Oxford, Ohio: National Staff Development Council.

Stevens, J. P. (1972). Global measures of association in multivariate analysis of variance. *Multivariate Behavioral Research, 6*, 515-524.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Stofflett, R. T., & Stoddart, T. (1994). The ability to understand and use conceptual pedagogy as a function of prior content learning experience. *Journal of Research in Science Teaching, 30*(1), 31-51.

Summers, A. A., & Wolf, B. L. (1977). Do schools make a difference? *American Economic Review, 67*, 639-652.

Tanner, L. N. (1999)**.** *Dewey's laboratory school: Lessons for today*. New York, NY: Teachers College Press.

Tanner, D., & Tanner, L. N. (1990). *History of the school curriculum*. New York, NY: Macmillan.

Tatsuoka, M. M. (1973). Multivariate analysis in education research. In F. N. Kerlinger (ed.), *Review of research in education, 1* (pp. 273-310). Itasca, IL: Paecock.

Tatsuoka, M. M. (1988). *Multivariate analysis: Techniques for educational and psychological research*. New York: Macmillan.

Thomas, S. J. (1999). *Designing surveys that work: A step-by-step guide*. Thousand Oaks: Corwin Press, Inc.

Unnever, J. D., Kerckhoff, A. C., & Robinson, J. J. (2000). District variations in education resources and student outcomes. *Economics of Education Review, 19* (3), 245-259.

U. S. Department of Education, NCES. (1995). *Schools and staffing survey, 1993-1994* (NCES 1995-021). Washington, DC: US Government Printing Office. Author.

U. S. Department of Education, NCES. (1999). *The condition of education 1999*. (NCES 1999-022). Washington, DC: US Government Printing Office. Author.

U.S. Department of Education, NCES. (2000a). *Trends in academic progress: Three decades of student performance* (NCES, 200-469). Washington, DC: U.S. Government Printing Office. Author.

U.S. Department of Education, NCES. (2000b). *The nation's report card 2000: Science 2000* (NCES 2002-451). Washington, DC: U.S. Government Printing Office. Author.

U. S. Department of Education, NCES. (2001a). *The condition of education 2001*, NCES 2001-072, Washington DC: U.S. Government Printing Office. Author.

U. S. Department of Education, NCES. (2001b). *Digest of education statistic 2000*. (NCES 2001-
034). Washington, DC: US Government Printing Office. Author.

U. S. Department of Education, NCES. (2002a). *Digest of education statistics 2001* (NCES 2001-130). Washington, DC: U.S. Government Printing Office. Author.

U. S. Department of Education, NCES. (2002b). *Schools and staffing survey*, 1999-2000 (NCES 2002-313). Washington, DC: US Government Printing Office. Author.

U. S. Department of Education (2003a). Secretary's remarks at mathematics summit (February 6, 2003). Retrieved March 23, 2003 from *http://www.ed.gov/Speeches/02-2003.*

U. S. Department of education, NCES. (2003b). *Digest of education statistics 2002* (NCES 2003-032). Washington, DC: U.S. Government Printing Office. Author.

U. S. Department of education, NCES. (2003c). *Overview of public elementary and secondary schools and districts: School year 2001-2002* (NCES 2003-346). Washington, DC: U.S. Government Printing Office. Author.

Walberg, H. J. (1982). Education productivity: The theory, evidence, and prospects. *Australian Journal of Education, 26*(2), 115-122.

Walberg, H. J., & Fowler, W. J. (1989). Expenditure and size efficiencies of public school districts (Report No. RC-015-786). East Lansing, MI: National Center for Research on Teacher Learning. (ERIC Document Reproduction Service No. ED 274471).

Walberg, H. J., & Weinstein, T. (1982). The production of achievement and attitude in high school social studies. *Journal of Educational Research, 75*(5), 285-292.

Wenglinsky, H. (1997). School district expenditures, school resources, and student achievement: Modeling the production function**.** In W. J. Fowler (ed.), *Developments in School Finance, 1997-1998* (pp. 99-120). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Williams, L. B. (1999). Paradoxes of educational policy and productivity. *Educational Policy, 13* (2), 227-250.

Word , E., Johnson, J., Bain, U., Fulton, D. B., Boyd-Zaharias, J., Lintz, M. N., et. al. (1990). *Student/Teacher Achievement Ratio (STAR): Tennessee's K-3 class size study*. Nashville: Tennessee Department of Education.

Zehr, M. A.(1998). Weak scores, strong economy: How can this be? *Education Week, 17*(29), 1-20.

APPENDICES

APPENDIX A

CERTIFIED PERSONNEL DATA FORMAT: GEORGIA DEPARTMENT OF EDUCATION

CERTIFIED PERSONNEL DATA FORMAT FOR PK-12 TEACHERS**

QUALIFICATION	**NUMBER OF TEACHERS**
4 Yr Bachelor	
5 Yr Masters	
6 yr Specialist	
7 yr Doctoral	
Others*	
EXPERIENCE (Years)	**NUMBER OF TEACHERS**
< 1	
1-10	
11-20	
21-30	
>30	

* Includes One-and Two-Year Vocational Certificates
** As presented by the Georgia Department of Education. Retrieved March 8, 2003 from http://www.doe.k12.ga.us.

APPENDIX B

INTERVIEW INSTRUMENTS

Appendix B1: Interview Instrument for School Principals.

Dear Principal, *ID CODE:*

The tables and comments below solicit responses to some general information about your school and your role as the principal. Please complete by filling in the required information.

General Information

INFORMATION	**NUMBER**
Total School Enrollment	
Average number of graduates per academic year	
Number of years as a teacher prior to becoming a principal	
Total number of years as principal	
Number of years as principal in present school	

Science Teacher Information

QUALIFICATION	**NUMBER OF TEACHERS**
Bachelor	
Master	
Specialist	
Doctoral	
Others (Example: Paraprofessionals)	
EXPERIENCE (Years)	**NUMBER OF TEACHERS**
< 1	
1-10	
11-20	
21-30	
>30	

Science Instructional Fund: Information

INFORMATION	**AMOUNT ($)**
QBE Allocation from District to your School per year	
QBE Allocation for Science instruction in your school per year.	
School-source Fund allocation for science instruction per year	
Fund from others sources for science instruction per year	

Other comments

1. In making decisions about fund allocation to the various departments in the school, which group(s) of people do you depend on to help in the decision-making process?

…………………………………………………………………………………………..

…………………………………………………………………………………………

Continued next page.

2. How are these group(s) selected?

   ………………………………………………………………………………………

   As the principal, what do you think is the role of committees in decision-making on matters relating to the school?

   …………………………………………………………………………………………

3. As the principal, how often do you use committees for decision-making?

   ………………………………………………………………………………………...

4. As the principal, briefly describe how you see your role in the school.

   …………………………………………………………………………………………

5. As the principal, how would you describe your leadership style?

   ……………………………………………………………………………………….

6. As the principal, how do you manage and control information?

   ………………………………………………………………………………………..

7. As the principal, describe how you handle conflicts whenever they arise.

   ……………………………………………………………………………………….

   ………………………………………………………………………………………..

8. As the principal, how would you describe the degree of autonomy you have regarding control over the affairs of your school?

   ………………………………………………………………………………………..

   ………………………………………………………………………………………..

9. As the principal, describe the position of science education in your school.

   ………………………………………………………………………………………..

10. What are your suggestions for improving science education in your school?

   ………………………………………………………………………………………………

Appendix B2: Interview Instrument for Heads of Science Departments

Dear Head of Science Department, *ID CODE:*

The tables and comments below solicit responses to some general information about your department and your role as the head. Please complete by filling in the required information.

General Information about the Science Department

INFORMATION	**NUMBER**
Total number of years as Head of Science Department	
Number of years as a classroom teacher prior to being head of Department	
Average Class Size in the Science Department	
Number of teachers in the Department **certified in the subject area they teach.**	
Average number of professional development training attended/year	
Average # of hands-on/lab-centered instruction per science class per week	

Science Teacher Information

QUALIFICATION	**NUMBER OF TEACHERS**
How many hold Bachelor degrees	
How many hold Masters degrees	
How many hold Specialist degrees	
How many hold Doctoral degrees	
Others (Example: Paraprofessionals)	
EXPERIENCE (Years)	**NUMBER OF TEACHERS**
< 1	
1-10	
11-20	
21-30	
>30	

Science Instructional Fund: Information

INFORMATION	**AMOUNT ($)**
QBE Fund Allocation for Science instruction from school.	
School-source Fund allocation for science instruction	
Funds from others sources for science instruction	

Other comments

1. In making decisions about fund allocation to science classes in the department, which group(s) of people do you depend on to help in the decision-making process?

…………………………………………………………………………………………..

Continued on next page.

2. As the head of the department, how would you describe the level of influence you have on how much money comes to your department for science instruction?

……………………………………………………………………………………………

…………………………………………………………………………………………….

3. As the head of department, briefly describe how you see your role in the department.

………………………………………………………………………………………………

……………………………………………………………………………………………….

4. As the head of department, describe the position of science education in your school.

…………………………………………………………………………………………….

……………………………………………………………………………………………..

5. Suggest major ways to improve science education in your school.

……………………………………………………………………………………………..

……………………………………………………………………………………………...

APPENDIX C

HUMAN SUBJECTS CONSENT FORM

Human Subjects Consent Form

I agree to participate in the research titled ***"The Relationship Between Educational Inputs and Measures of Student Achievement in Science As Outputs"***, which is being conducted by *Olajide Agunloye, Department of Educational Administration & Policy, University of Georgia, Athens, Georgia 30602, Telephone: (770) 322-5366,* under the direction of *Dr. Catherine C. Sielke, Georgia 30602 Department of Educational Administration & Policy, University of Georgia, Athens Georgia. Telephone: (706) 542-9767.* I understand that this participation is entirely voluntary; I can withdraw my consent at any time without penalty and have the results of the participation, to the extent that it can be identified as mine, returned to me, removed from the research records, or destroyed.

The following points have been explained to me:

1) The reason for the research project is to describe and explain the relationship between some input variables and three measures of high school student performance in science. The benefits that I may expect from it are: (i) Being a part of a study that can be used in developing and implementing reform policies at the school. (ii) Being a part of a study that can be used in developing and implementing policies in the school district. (iii) Improve student performance in general and student performance in science in particular.

2) The procedures are as follows:
The sampling units are the individual high schools in the district. Two participants are selected from each school (the principal and the head of science department). Each participant is requested to respond to Survey/Questionnaire soliciting information about fund allocation and the administrative processes involved (copy attached).

3) No discomfort or stresses are foreseen.

4) No risks are foreseen

5) The results of this participation will be confidential. The identity of participants, schools, or school district will not be specifically named in the final report. Auto-tapes are not used on the human research participants.

6) The researcher will answer any further questions about the research, now or during the course of the project, and can be reached by phone at: (770) 322-5366 or (678) 676-6302.

______________________________ ______________________________

Signature of Researcher Date Signature of Participant Date

(770) 322-5366, E-Mail: olagun@aol.com

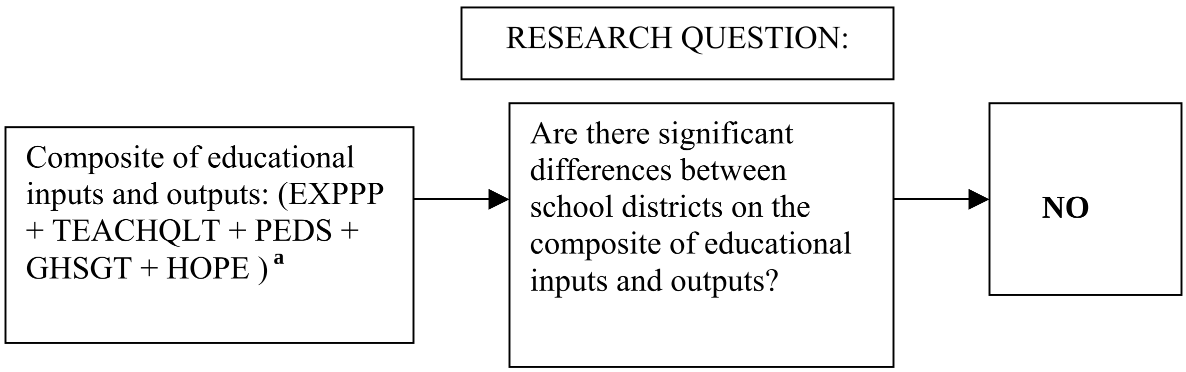
Please sign both copies of this form. Keep one and return the other to the investigator.

Additional questions or problems regarding your rights as a research participant should be addressed to Chris A. Joseph, Ph.D. Human Subjects Office, UGA 606A Boyd Graduate Studies Research Center, Athens Georgia 30602-7411, Telephone: 706-542 3199, E-Mail: IBR@uga

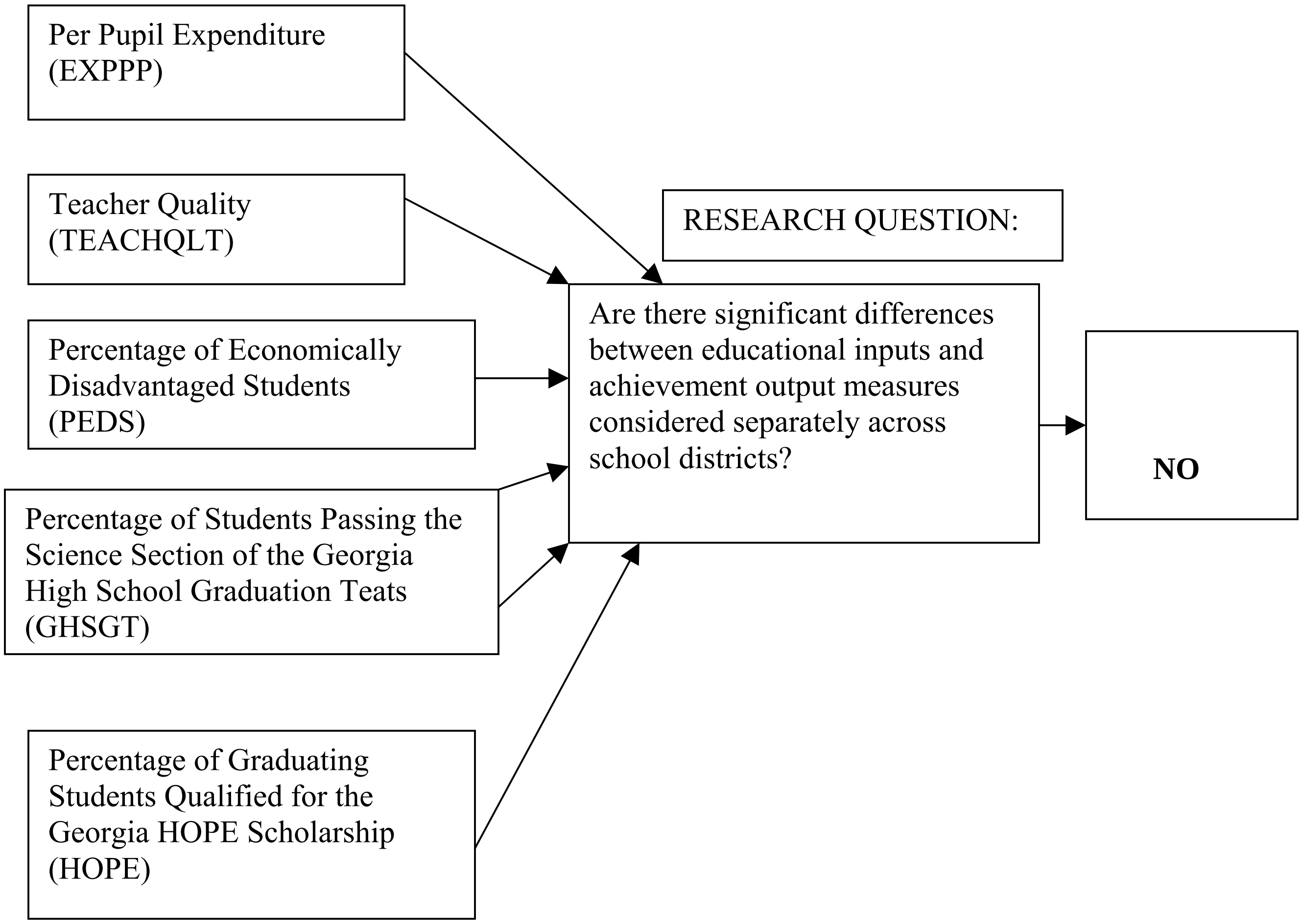
APPENDIX D

PRETEST SCHEMATIC DIAGRAMS OF HYPOTHESES TESTED

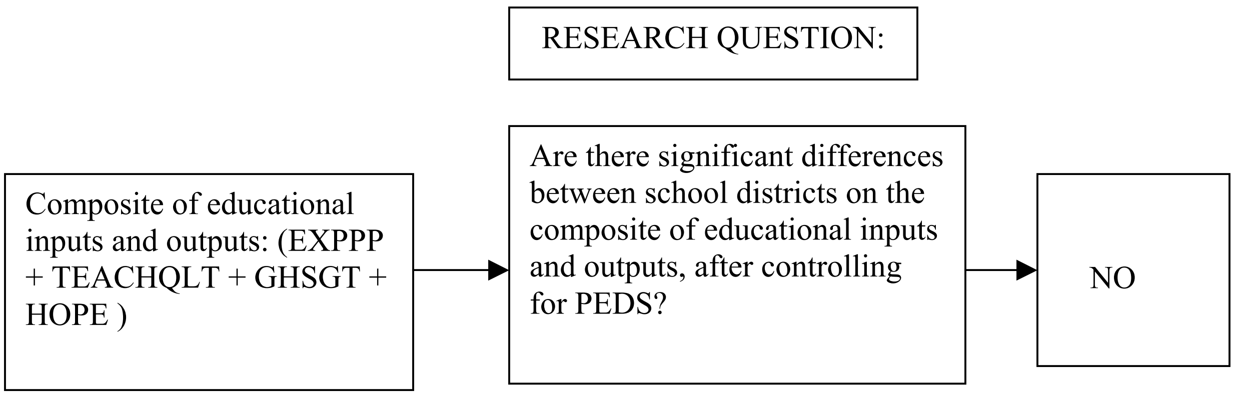
Appendix D1: Pretest Schematic Diagram for Hypothesis #1.



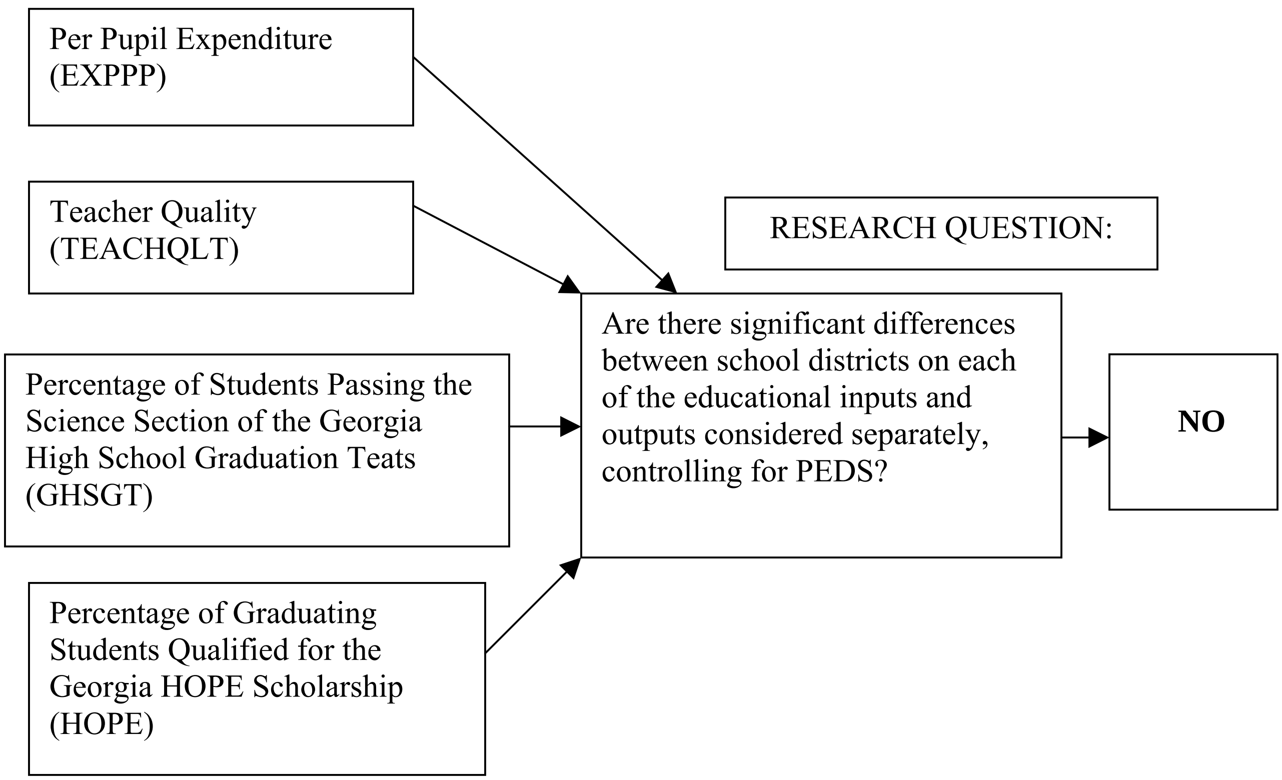
Appendix D2: Pretest Schematic Diagram for Hypothesis #2.



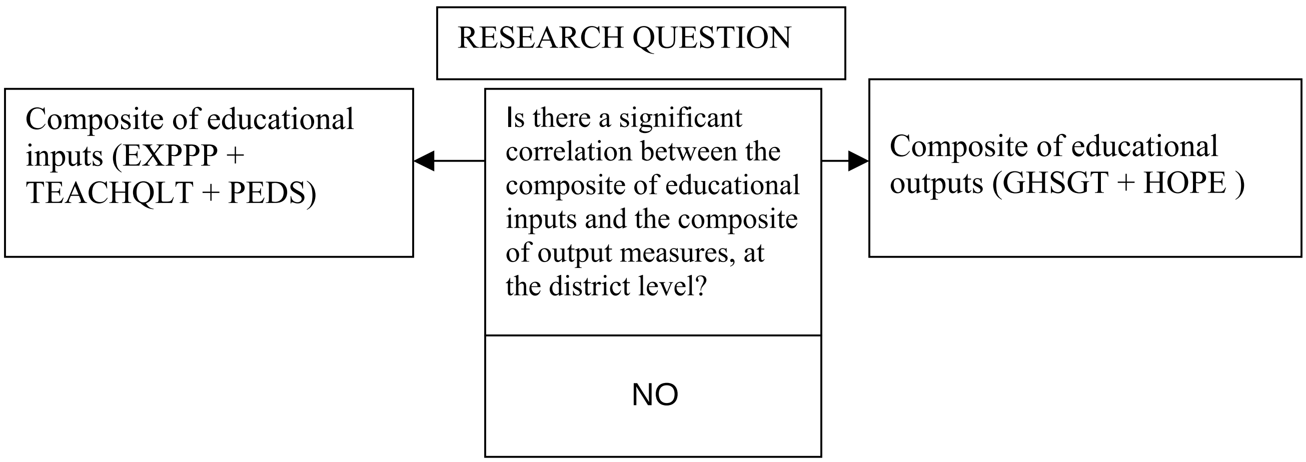
Appendix D3: Pretest Schematic Diagram for Hypothesis #3



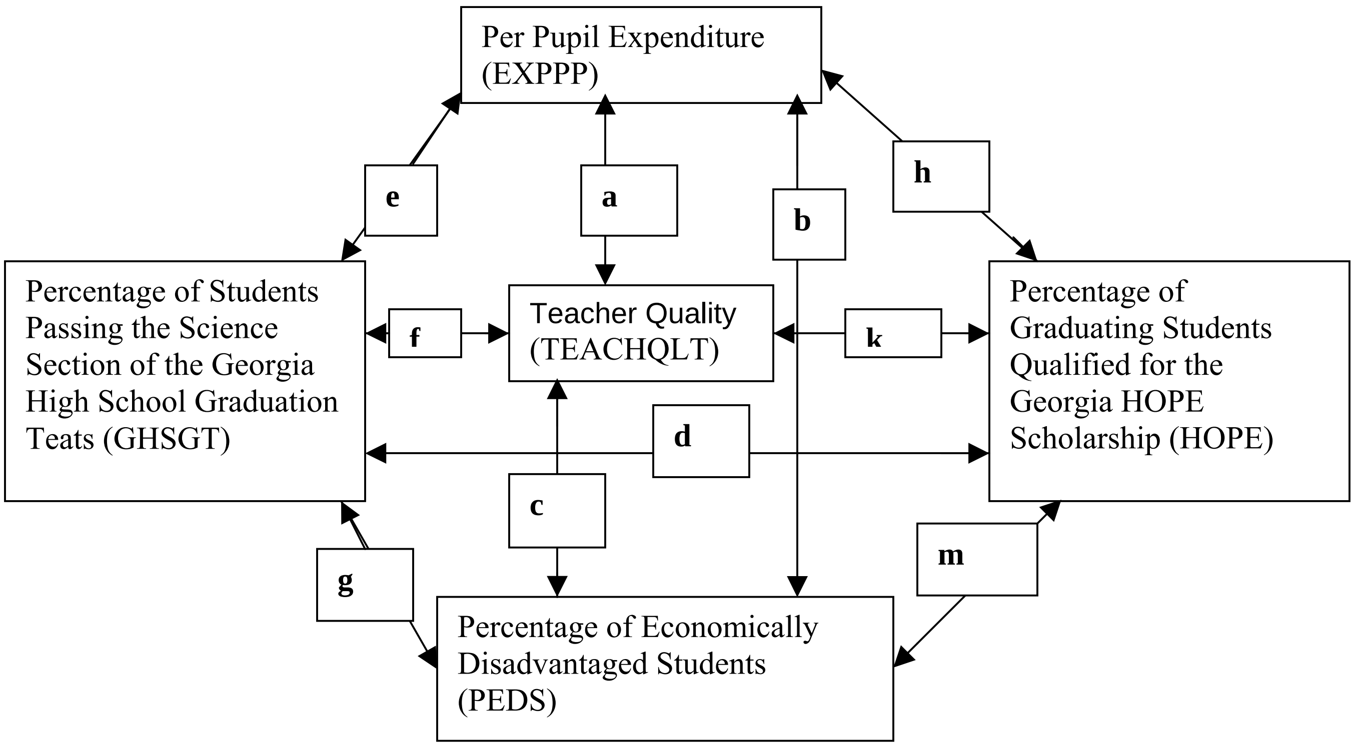
Appendix D4: Pretest Schematic Diagram for Hypothesis #4.



Appendix D5: Pretest Schematic Diagram for Hypothesis #.5

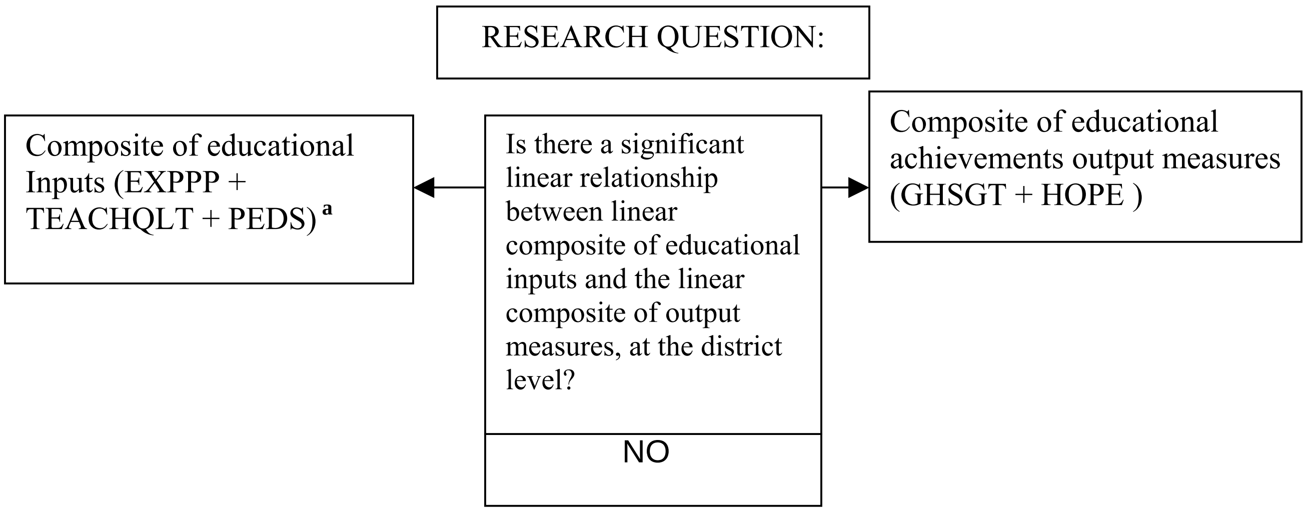


Appendix D6: *Pretest Schematic Diagram for Hypothesis #6.

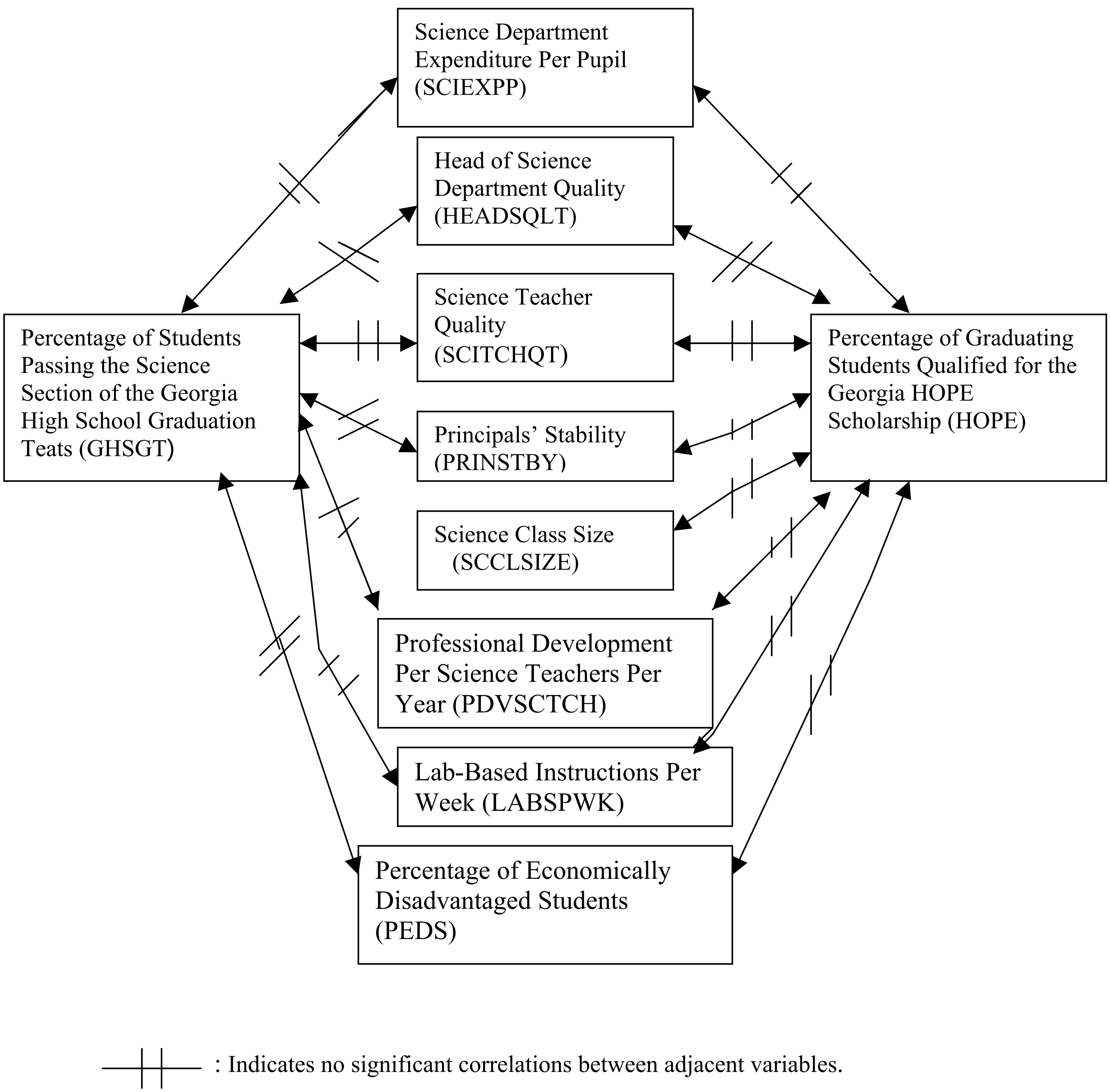


*a, b, c, d, e, f, g, h, k, and m indicate paths of no significant correlations between adjacent variables.

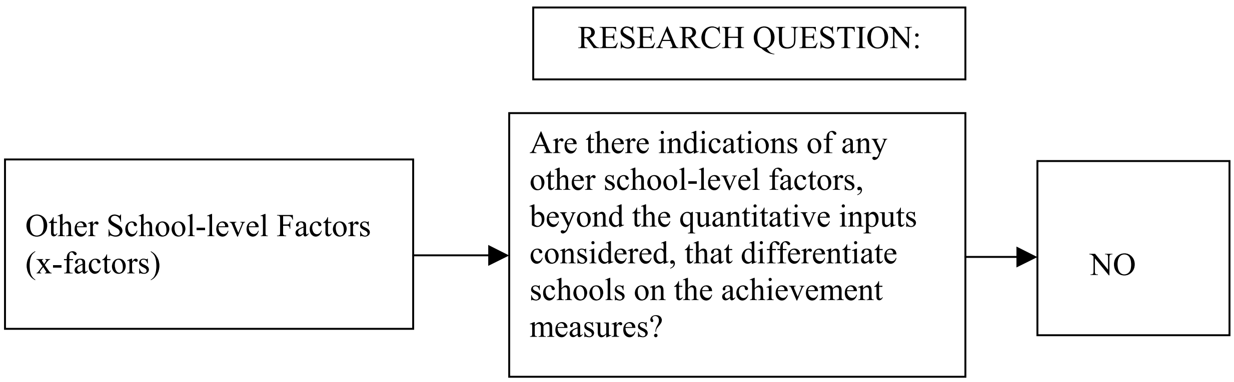
Appendix D7: Pretest Schematic Diagram for Hypothesis #7.



Appendix D8: Pretest Schematic Diagram for Hypothesis #8.



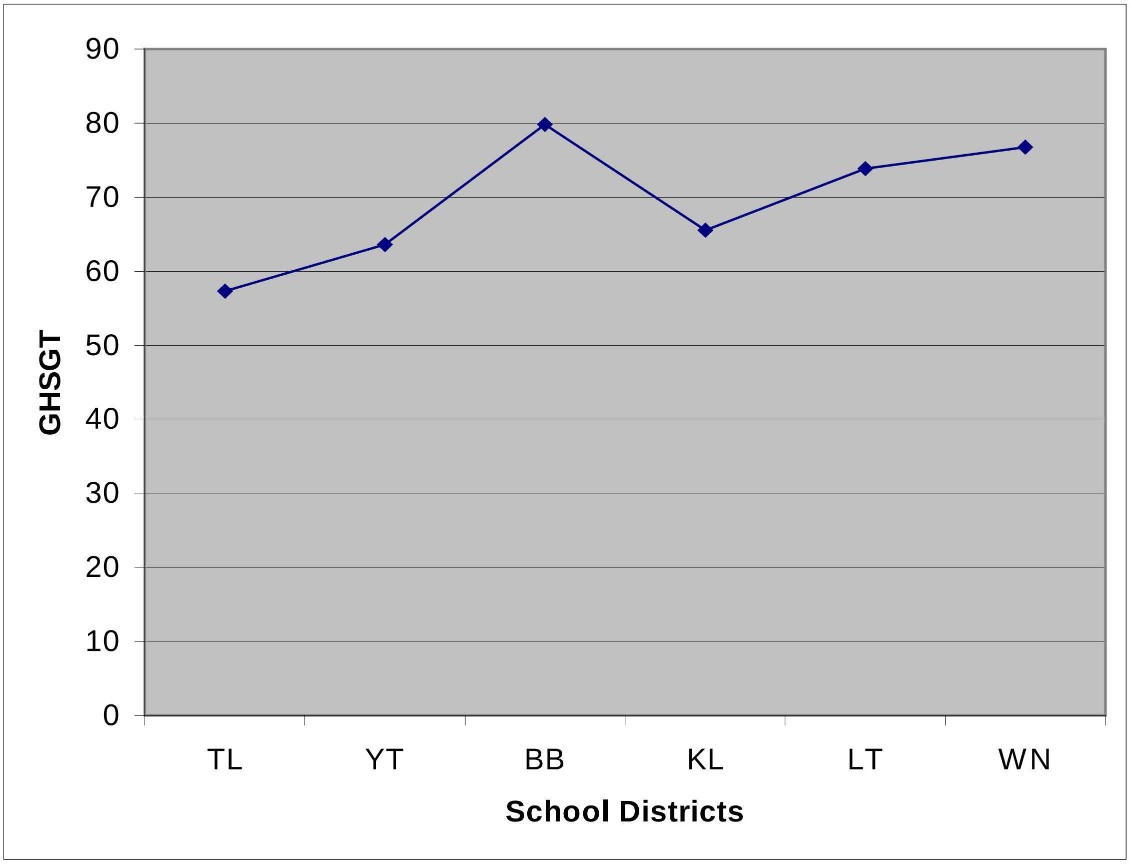
Appendix D9: Pretest Schematic Diagram for Hypothesis #9.



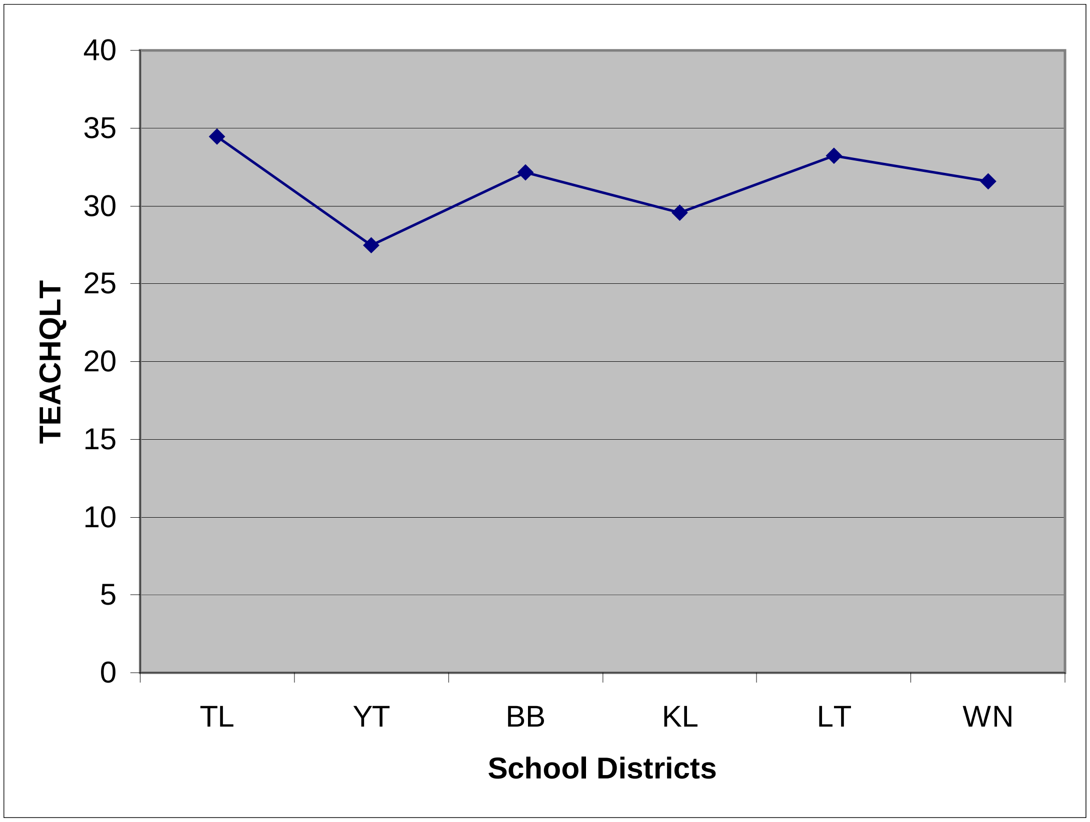
APPENDIX E

PLOTS OF THE VARIATIONS IN THE MEANS OF OUTPUT AND INPUT VARIABLES ACROSS SCHOOL DISTRICTS

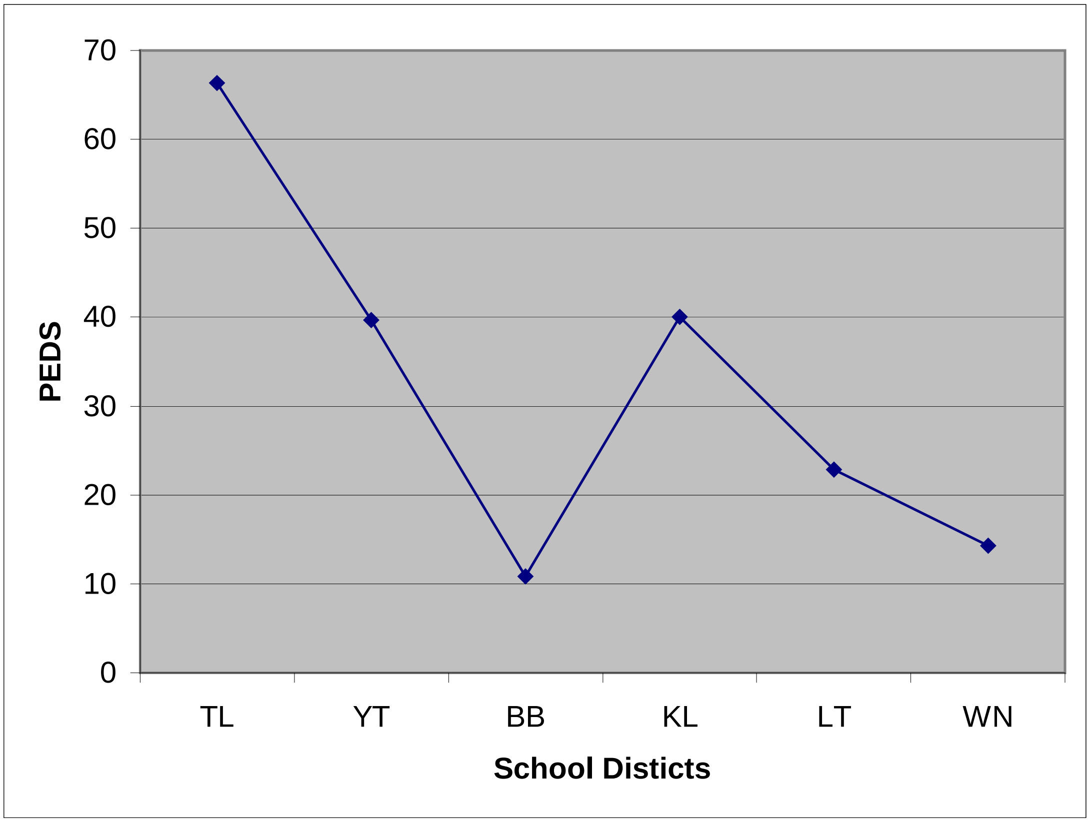
Appendix E1: Variations in EXPPP Across School Districts



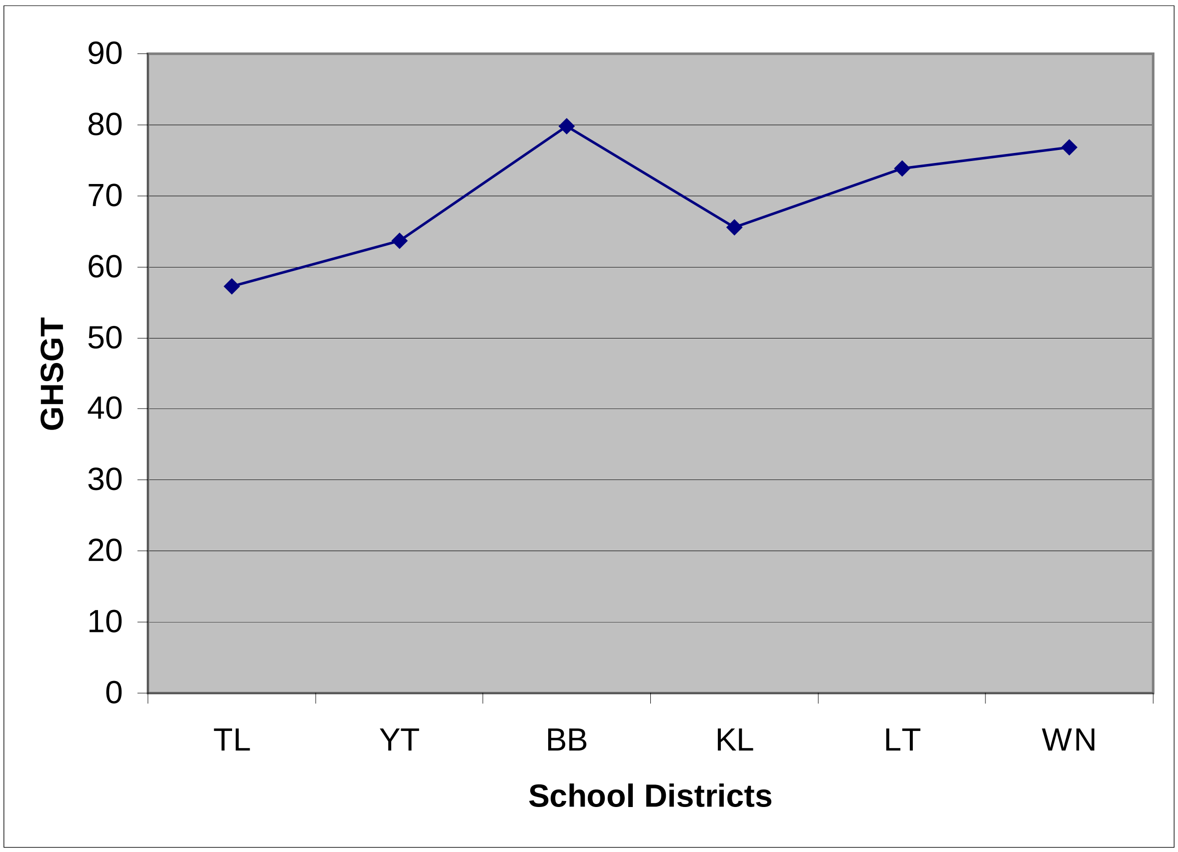
Appendix E2: Variations in TEACHQLT Across School Districts



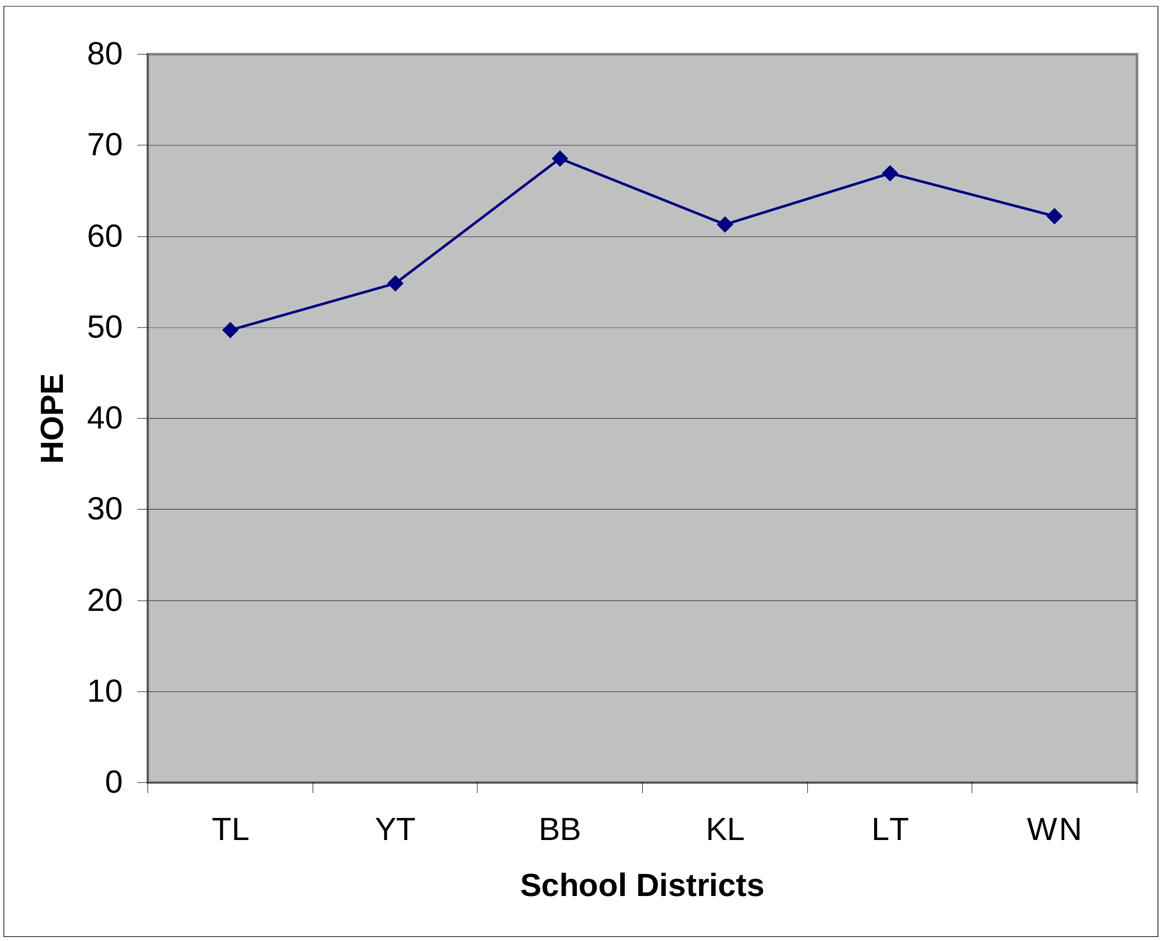
Appendix E3: Variations in PEDS Across School Districts



Appendix E4: Variations in GHSGT Across School Districts



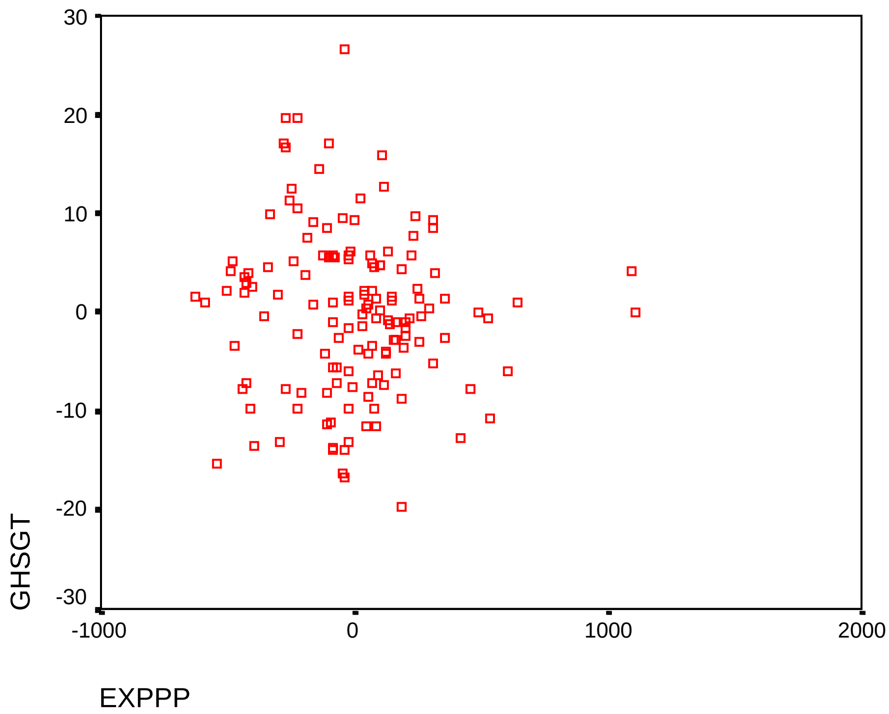
Appendix E5: Variations in HOPE Across School Districts



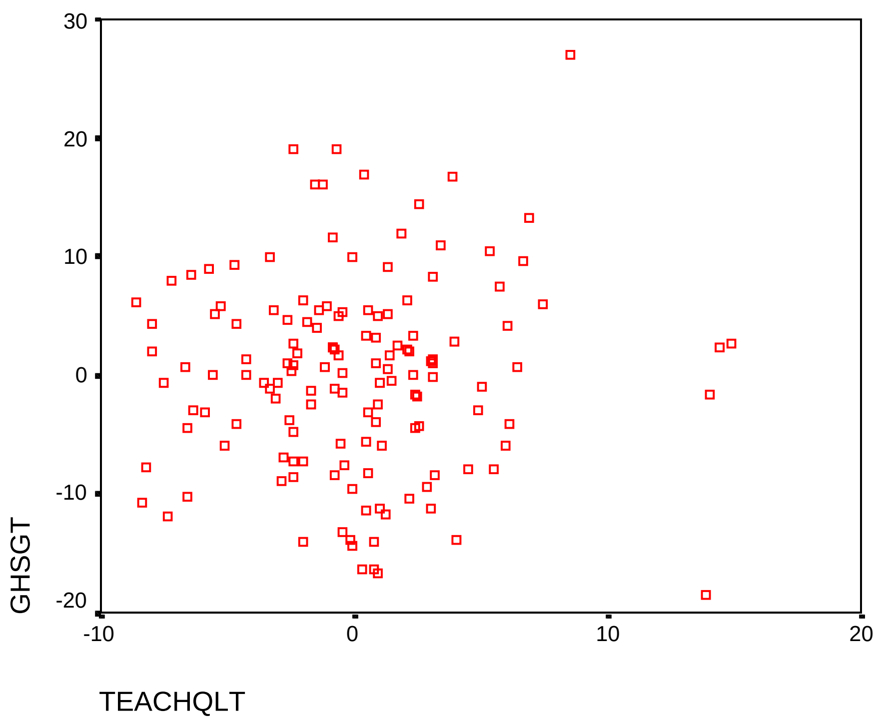
APPENDIX F

RESIDUAL PLOTS OF THE REGRESSION OF OUTPUTS AGAINST INPUTS

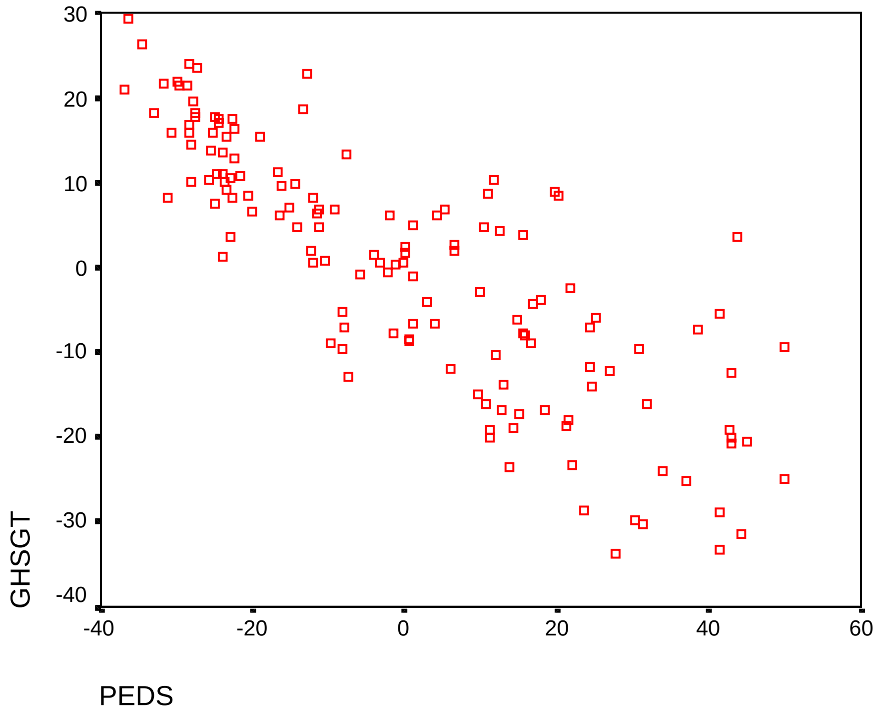
Appendix F1: Residual Regression Plot of GHSGT against EXPPP.



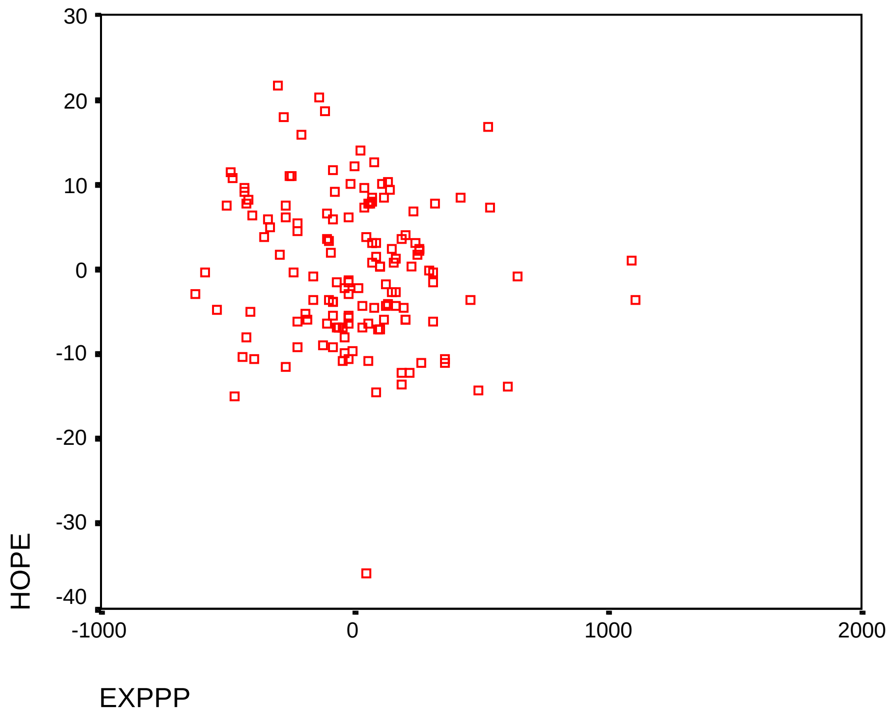
Appendix F2: Residual Regression Plot of GHSGT against TEACHQLT.



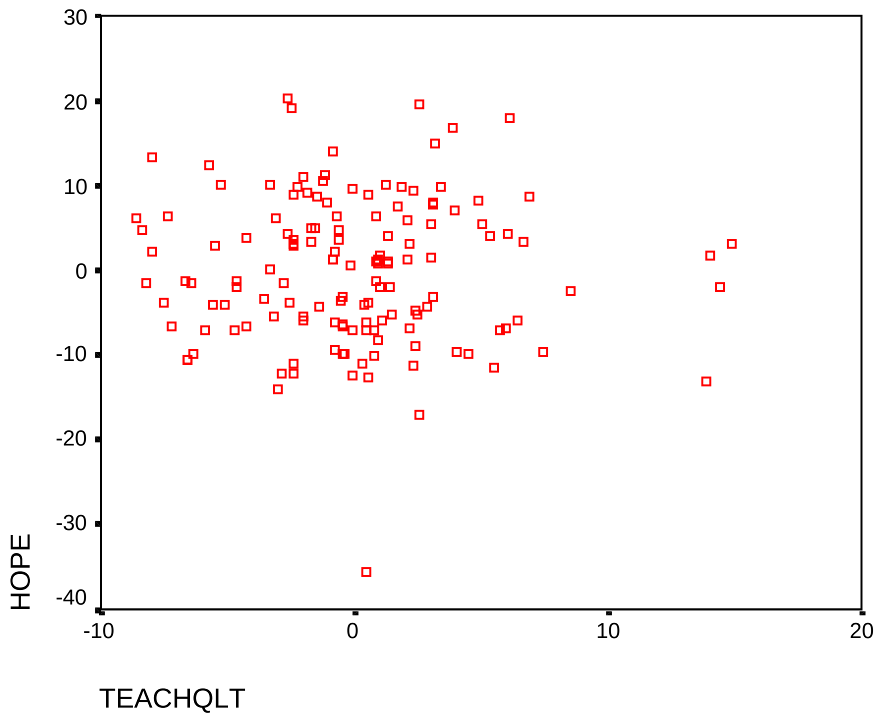
Appendix F3: Residual Regression Plot of GHSGT against PEDS



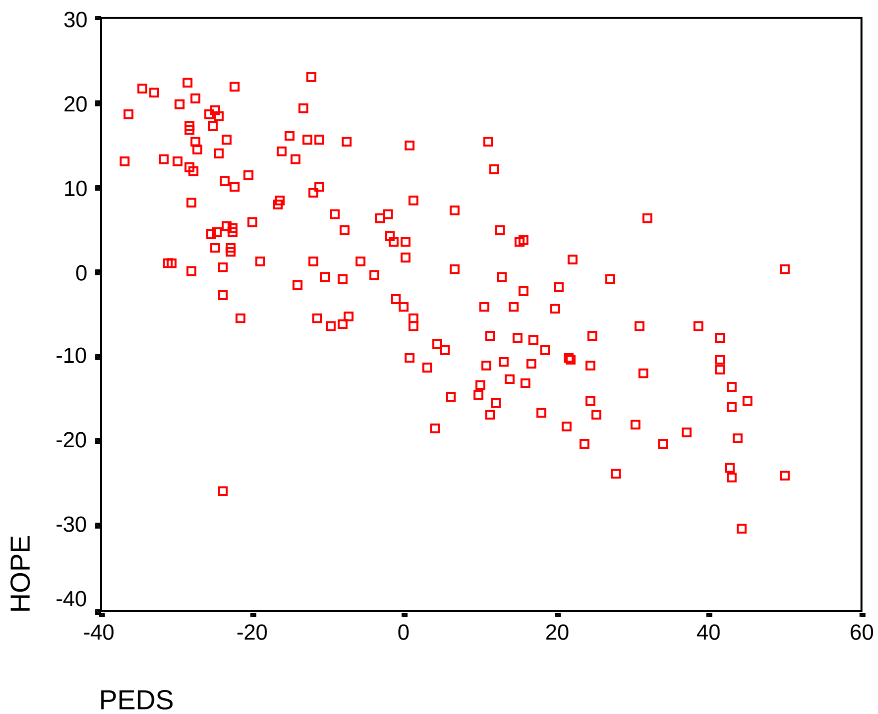
Appendix F4: Residual Regression Plot of HOPE against EXPPP.



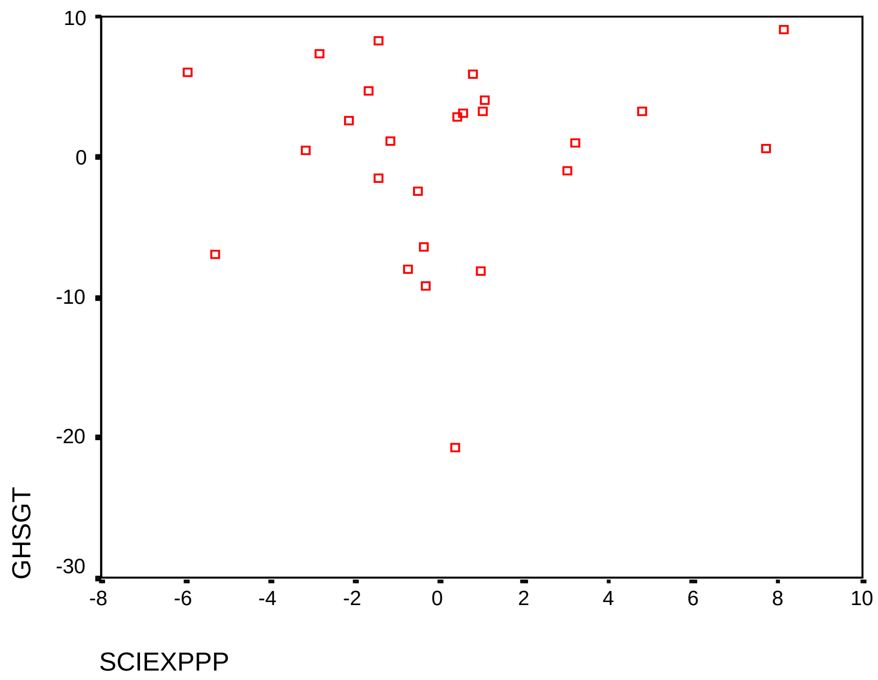
Appendix F5: Residual Regression Plot of HOPE against TEACHQLT.



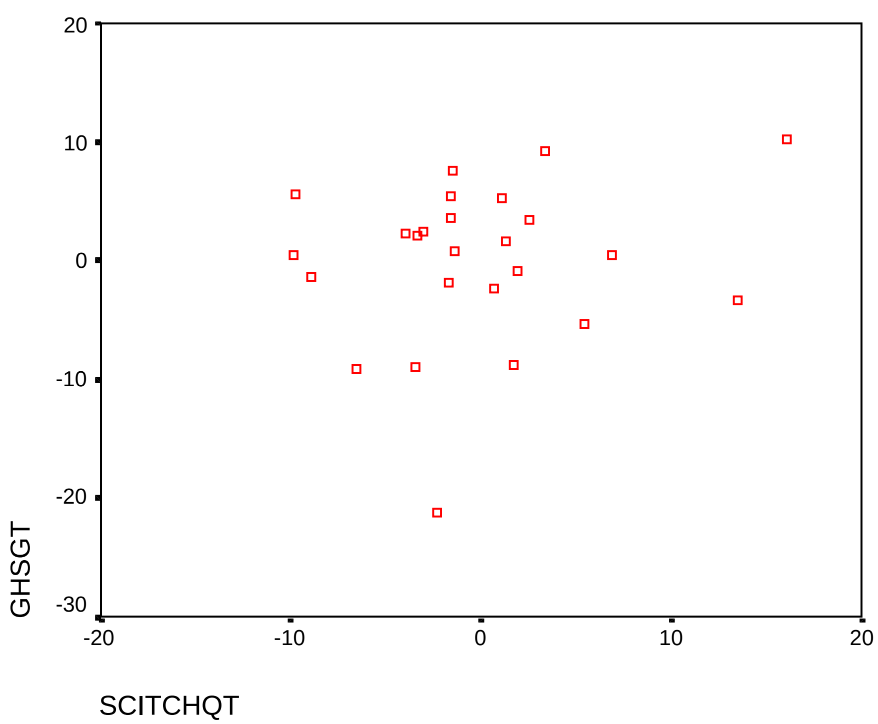
Appendix F6: Residual Regression Plot of HOPE against PEDS.



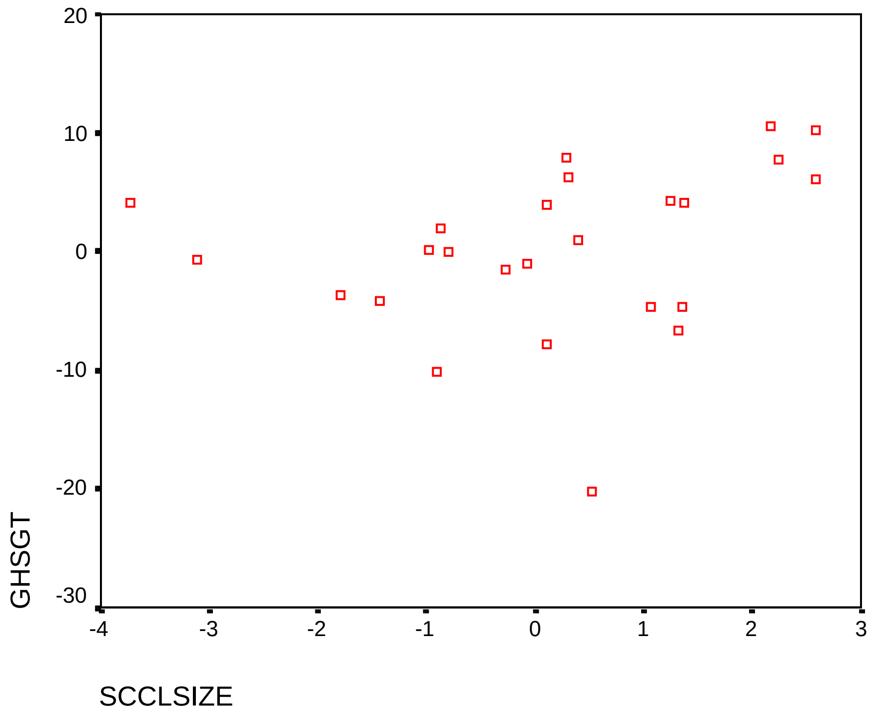
Appendix F7: Residual Regression Plot of GHSGT Against SCIEXPPP



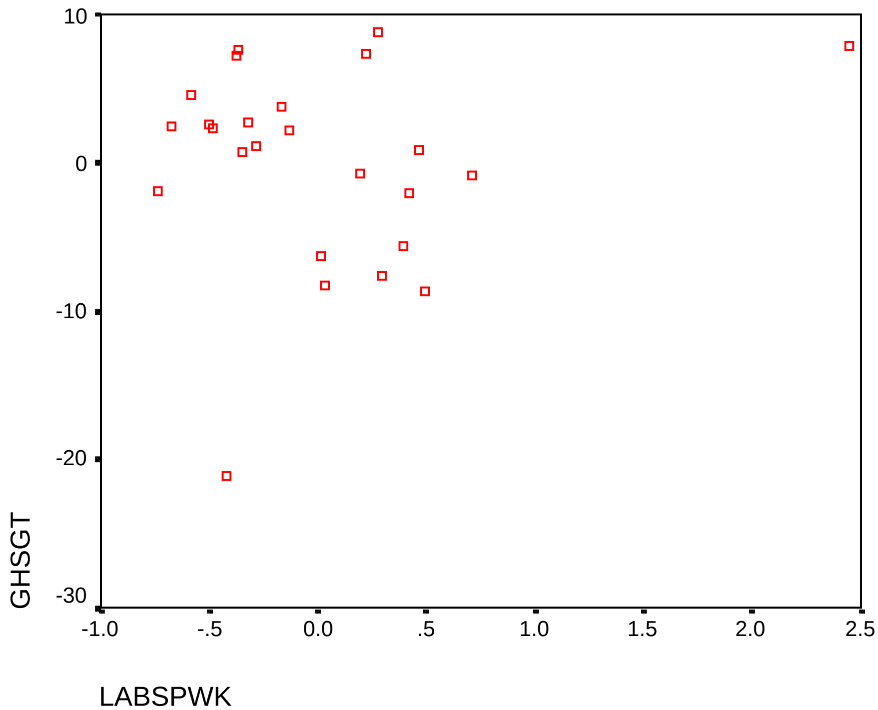
Appendix F8: Residual Regression Plot of GHSGT Against SCITCHQT



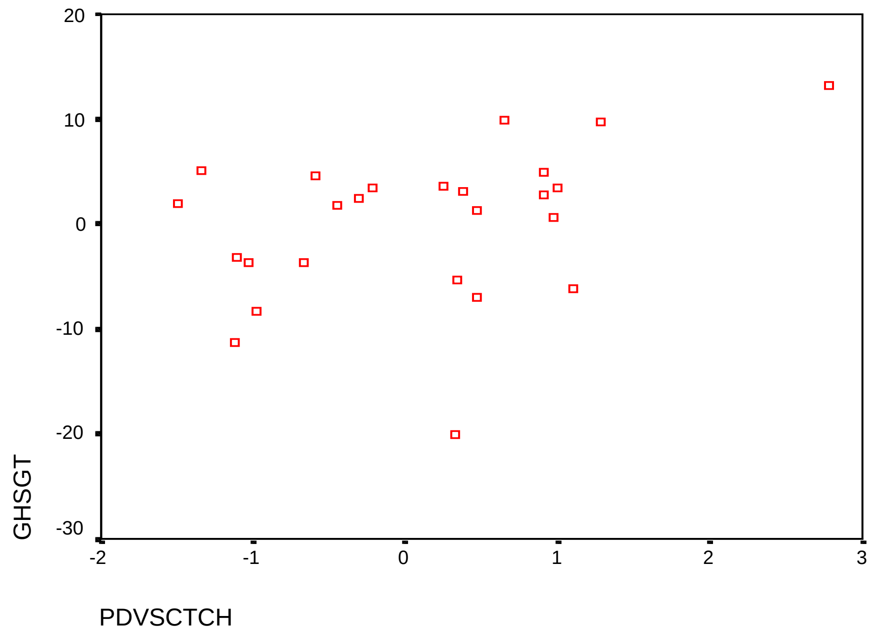
Appendix F9: Residual Regression Plot of GHSGT Against SCCLSIZE



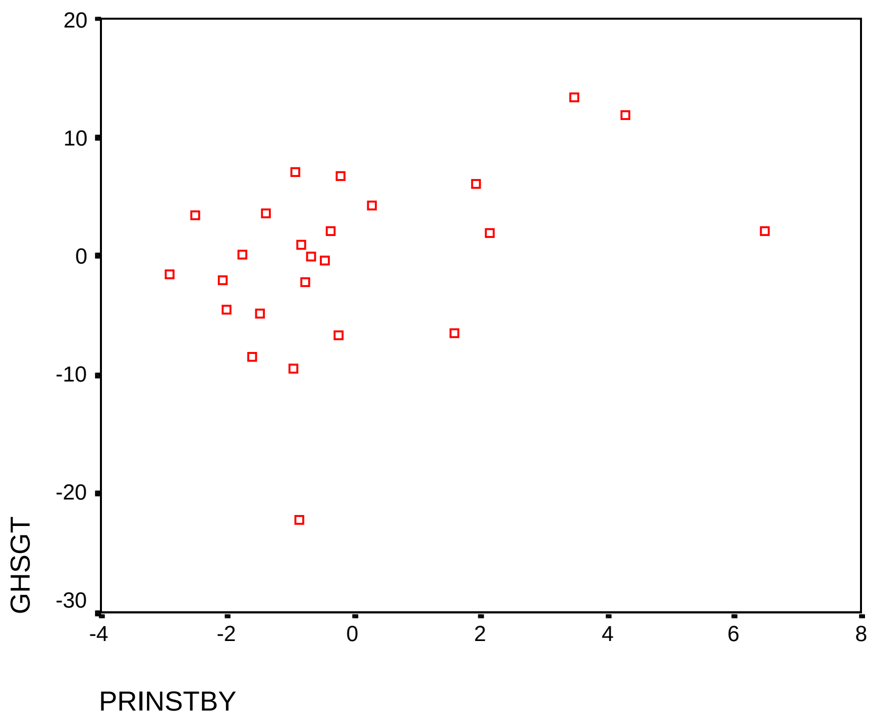
Appendix F10: Residual Regression Plot of GHSGT Against LABSPWK



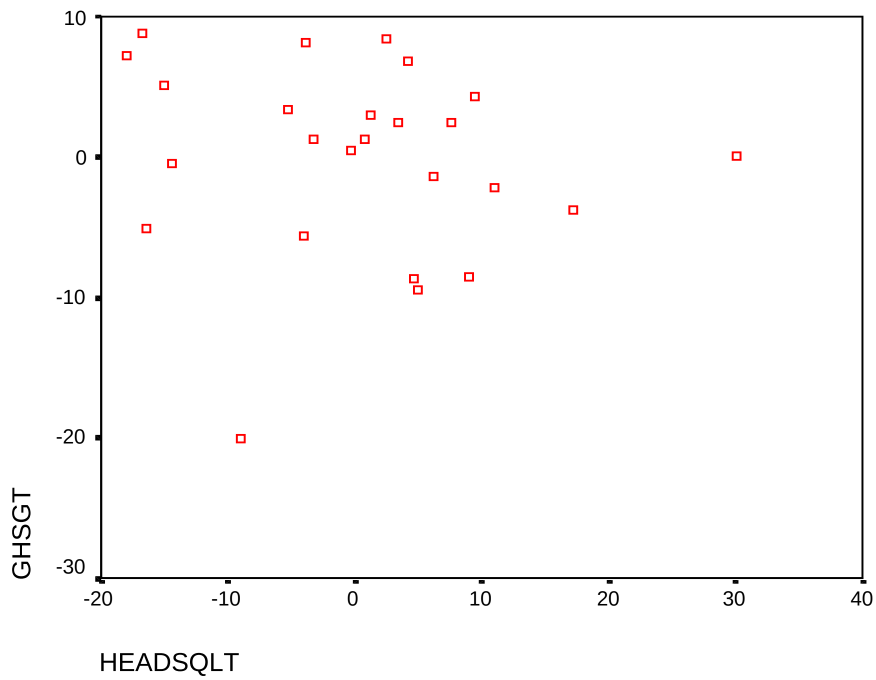
Appendix F11: Residual Regression plot of GHSGT against PDVSCTH



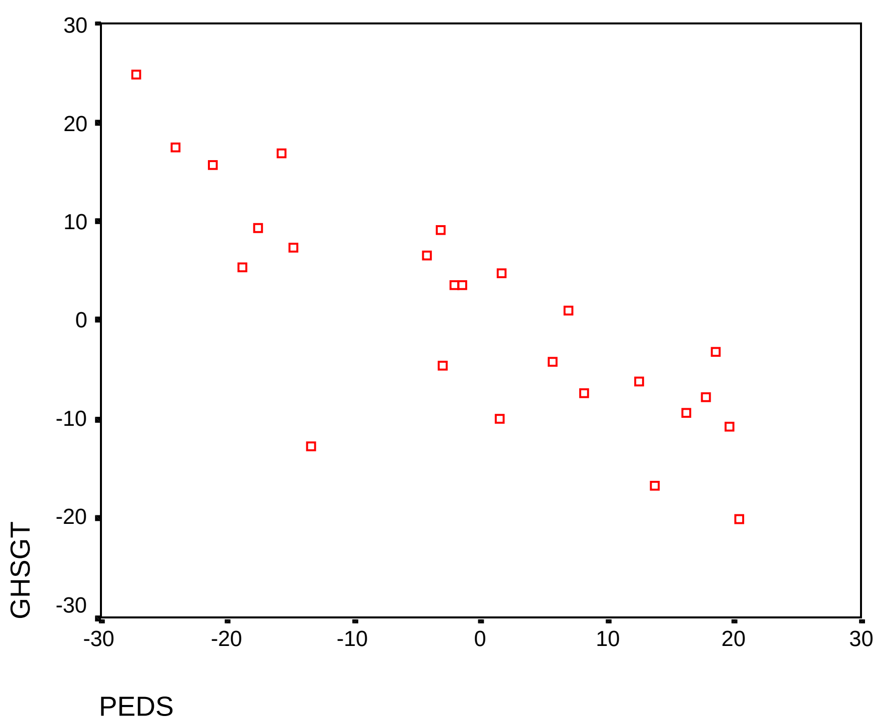
Appendix F12: Residual Regression Plot of GHSGT against PRINSTBY



Appendix F13: Residual Regression Plot of GHSGT against HEADSQLT.



Appendix F14: Residual Regression Plot of GHSGT against PEDS at the School Level.



APPENDIX G

INTERCORRELATION COEFFICIENTS BETWEEN QUANTITATIVE SCHOOL-LEVEL VARIABLES

Intercorrelation Coefficients Between Quantitative School-Level Variables

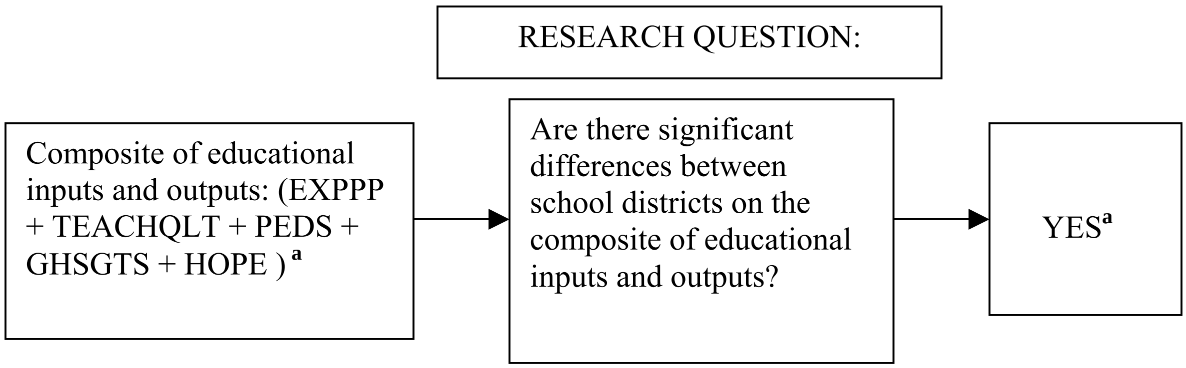
Variables	1	2	3	4	5	6	7	8	9	10
1. SCIEXPPP	__	.656*	-.173	-.178	.114	.106	.015	.309	-.094	-.155
2. SCITCHQT		__	-.174	-.174	.304	.123	-.004	.391	.096	-.231
3. SCCLSIZE			__	.124	-.392	-.269	.058	-.173	.219	.024
4. LABSPWK				__	.149	.118	.273	-.221	.253	.326
5. PDVSCTCH					__	.106	.015	.309	-.094	-.155
6. PRINSTBY						__	.078	-.192	.463*	.400
7. HEADSQLT							__	-.110	.075	.058
8. PEDS								__	-.789*	-.736*
9. GHSGT									__	.806*
10. HOPE										__

* Significant at $p < 0.005$ after controlling for family-wise error for the 10 school-level variables tested.

APPENDIX H

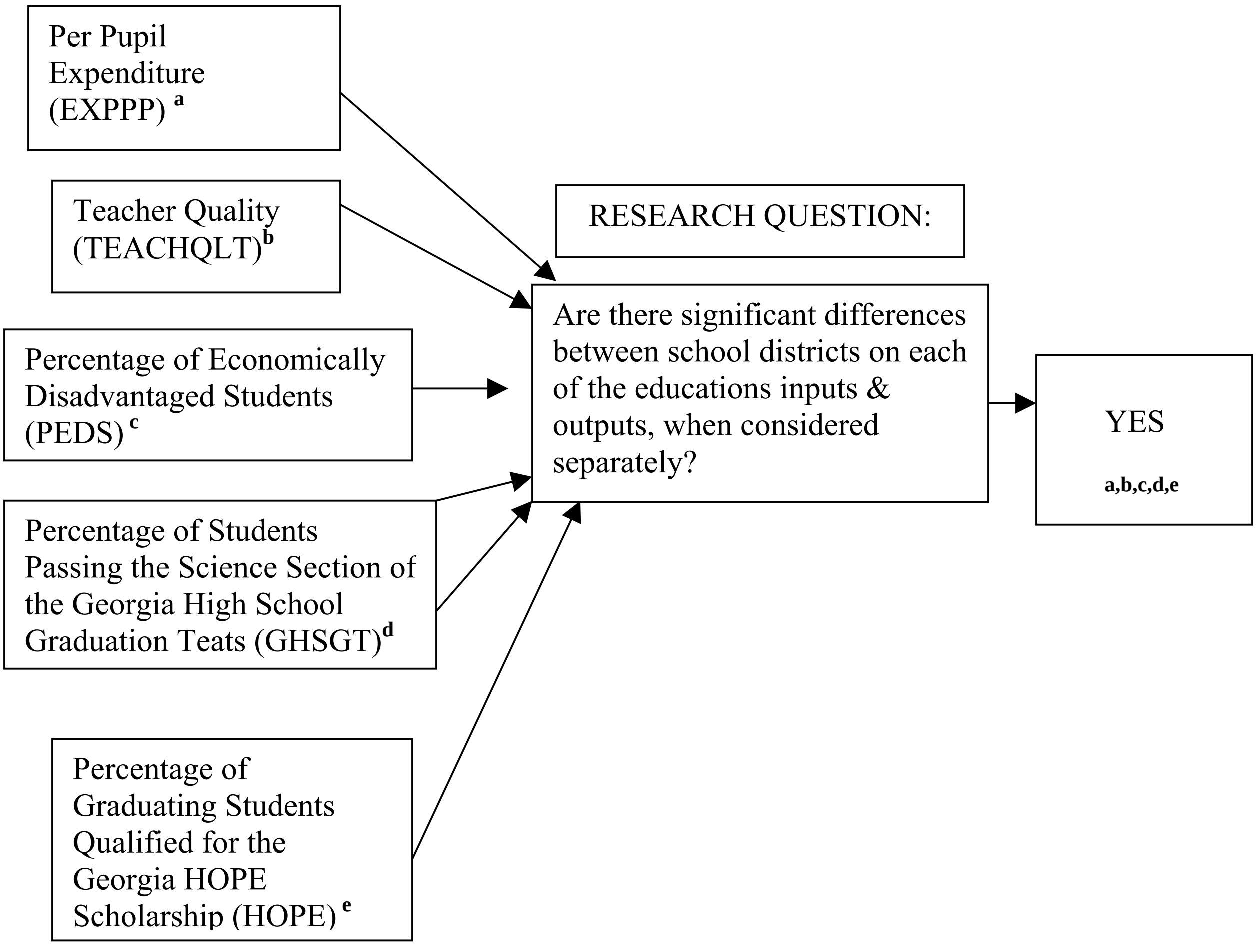
POST-TEST SCHEMATIC DIAGRAMS OF HYPOTHESES TESTED

Appendix H1: Post-test Schematic Diagram for Hypothesis #1.



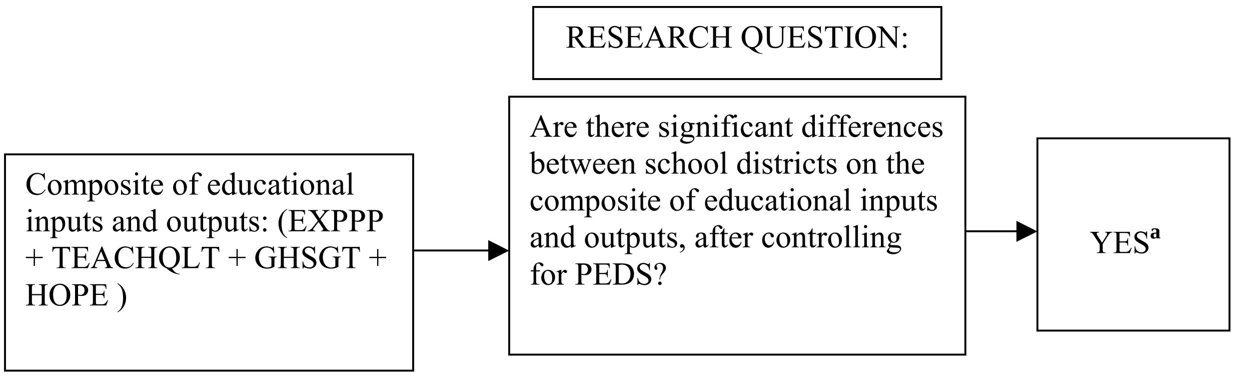
[a] Wilks $\Lambda = 0.08695$ $F_{(25, 652)}$, $\alpha = .05$, $p < .001$

Appendix H2: Post-test Schematic Diagram for Hypothesis #2.



[a]$F_{(5, 136)}$, $\alpha = .05$, $p < .001$
[b]$F_{(5, 136)}$, $\alpha = .05$, $p < .001$
[c]$F_{(5, 136)}$, $\alpha = .05$, $p < .001$
[d]$F_{(5, 136)}$, $\alpha = .05$, $p < .001$
[e]$F_{(5, 136)}$, $\alpha = .05$, $p < .001$

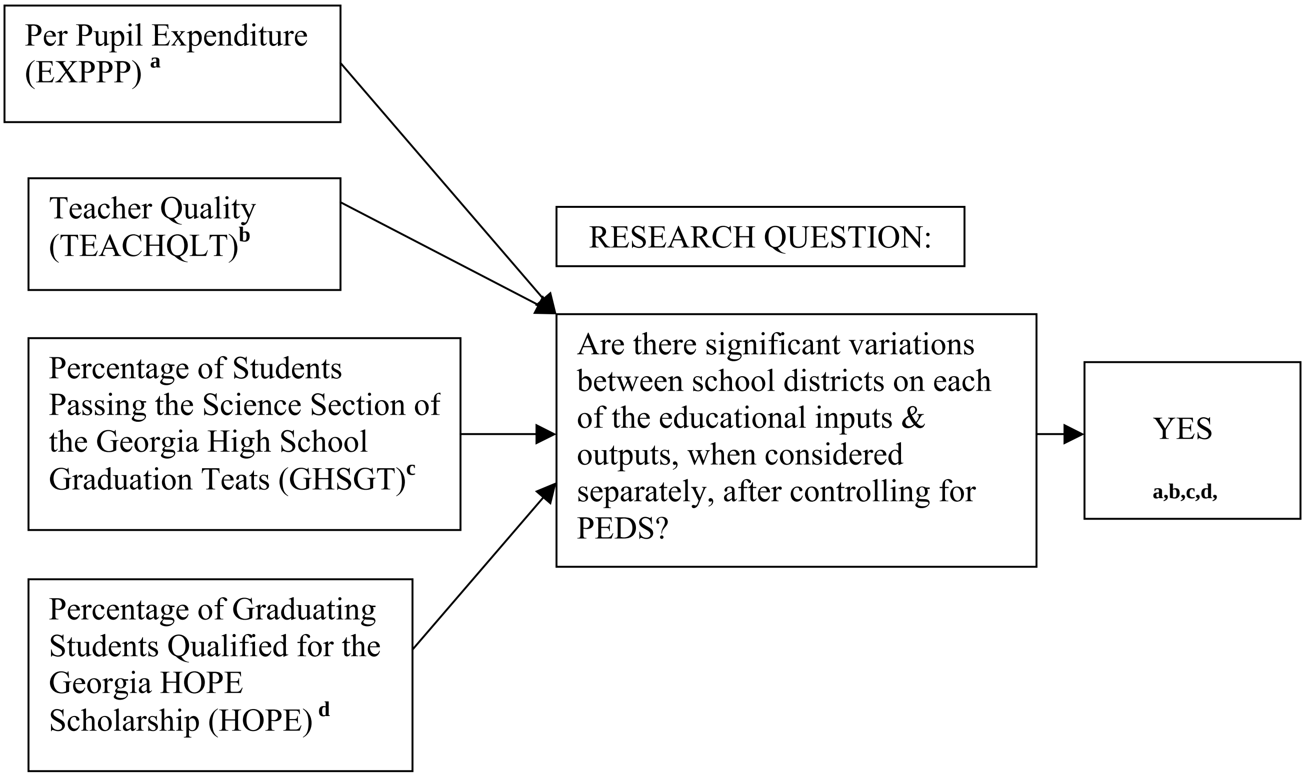
Appendix H3: Post-test Schematic Diagram for Hypothesis #3



[a] Wilks $\Lambda = 0.25303$ $F_{(4, 132)}$, $\alpha = .05$, $p < .001$

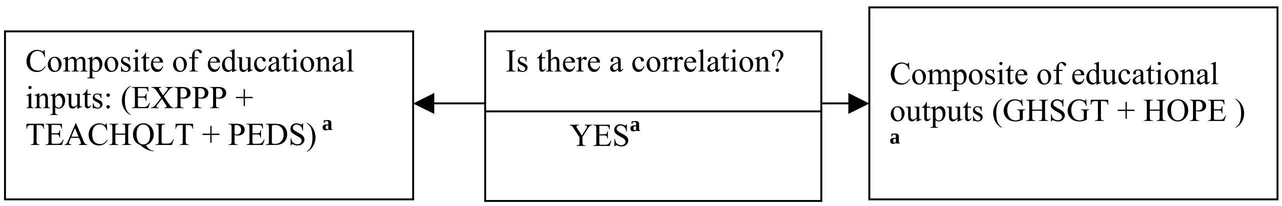
.

Appendix H4: Post-test Schematic Diagram for Hypothesis #4



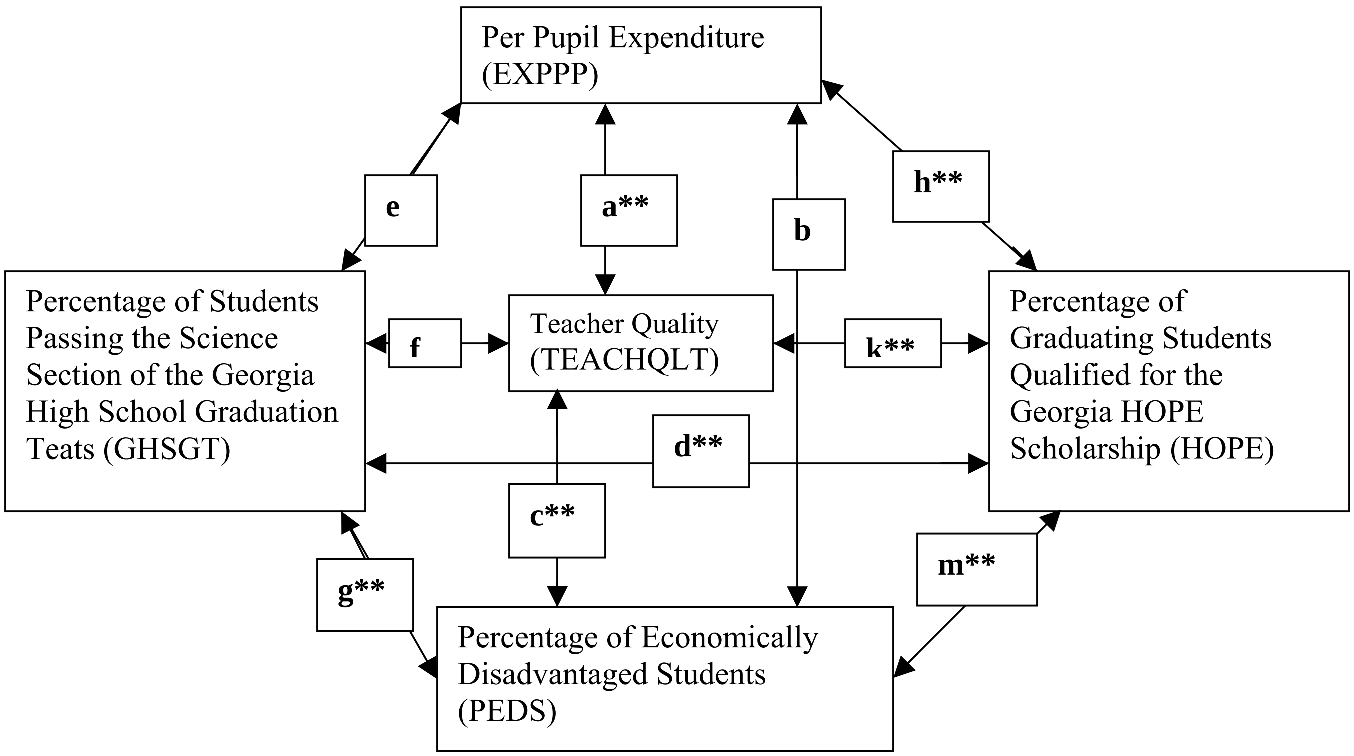
[a] $F_{(4, 135)}$, $\alpha = .05$, $p < .001$
[b] $F_{(5, 135)}$, $\alpha = .05$, $p < .001$
[c] $F_{(5, 135)}$, $\alpha = .05$, $p < .001$
[d] $F_{(5, 135)}$, $\alpha = .05$, $p < .001$

Appendix H5: Post-test Schematic Diagram for Hypothesis #5



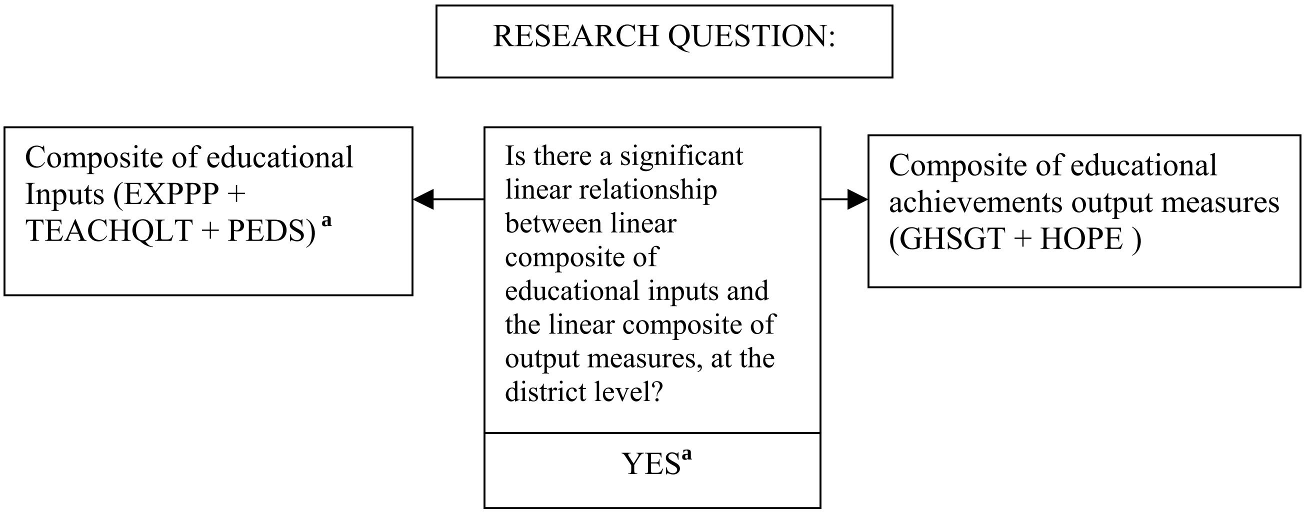
[a]Canonical Correlation = .850, Wilks $\Lambda = 0.274$, $\chi^2_{(6)} = 178.886$, $\alpha = .05$, $p < .001$

Appendix H6: Post-test Schematic Diagram for Hypothesis #6



a** Pearson's correlation = .1450, p= .043 (significant)
b Pearson's correlation = .0908, p= .141
c** Pearson's correlation = .1727, p= .020 (significant)
d** Pearson's correlation = .7727, p< .001 (significant)
e Pearson's correlation = -0.1151, p= .086
f Pearson's correlation = -0.1329, p= .057
g** Pearson's correlation = -0.8343, p< .001 (significant)
h** Pearson's correlation = -0.1644, p= .025 (significant)
k** Pearson's correlation = -0.1499, p= .037 (significant)
m** Pearson's correlation = -0.7428, p< .001 (significant)
↔ Indicate relationships

Appendix H7: Post-test Schematic Diagram for Hypothesis #7



GHSGT = 90.689 – 0.002EXPPP – 0.056TEACHQLT– 0.524PEDS

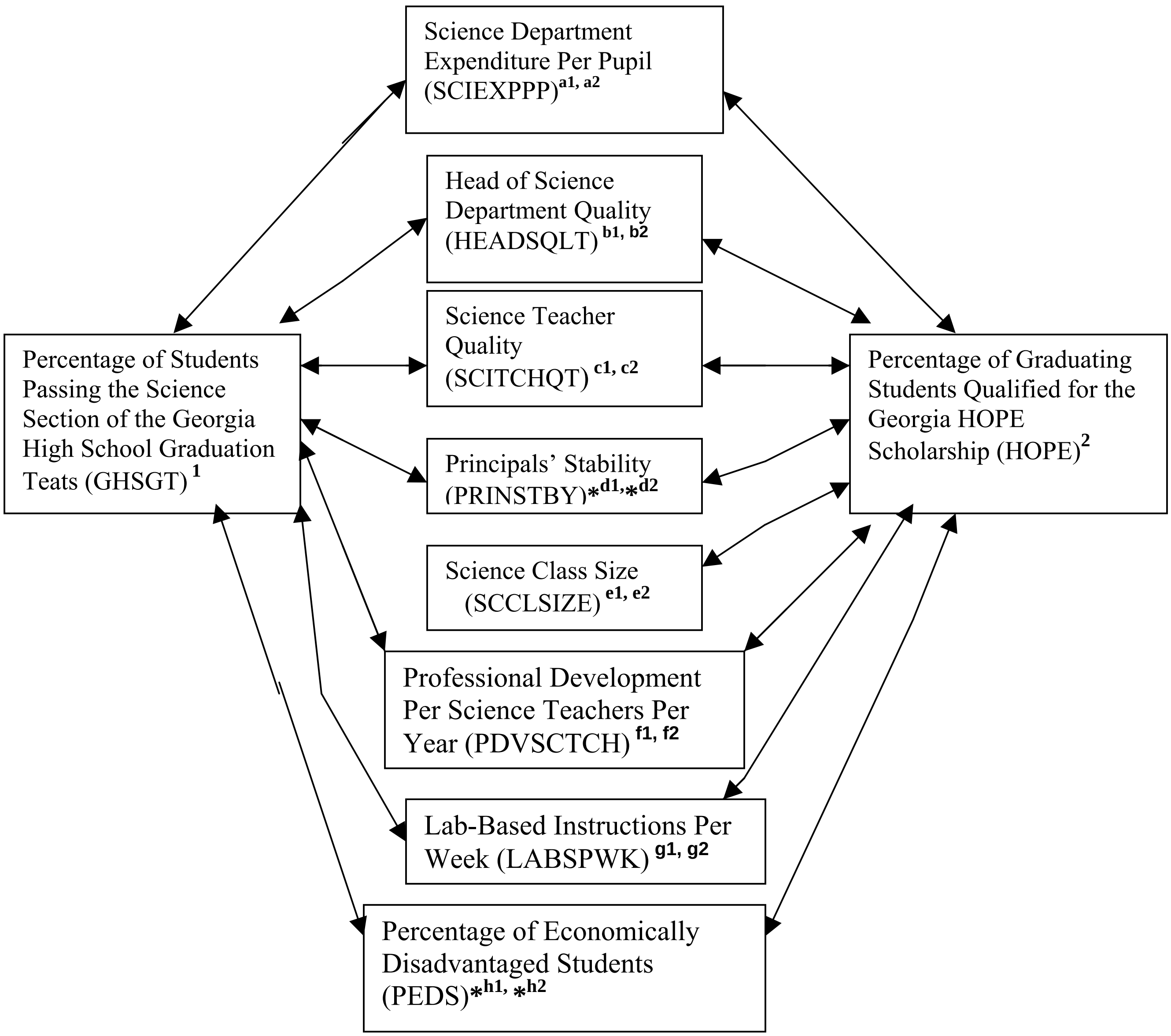
[a]More significantly as: GHSGT = 86.466 - 0.524PEDS

HOPE = 86.545 – 0.004EXPPP – 0.027TEACHQLT– 0.398PEDS

[a]More significantly as: HOPE = 73.885 – 0.403PEDS

[a]Significant: Wilks Λ=0.274 at $F_{(6,274)}$=41.647, α= .05, p< .001

Appendix H8: Post-test Schematic Diagram for Hypothesis #8



Pearson's correlations: $\mathbf{a_1}$ = -0.091, p= .081; $\mathbf{a_2}$ = -0.149, p= .056
Pearson's correlations: $\mathbf{b_1}$ = 0.079, p= .086; $\mathbf{b_2}$ = 0.058, p= .096
Pearson's correlation: $\mathbf{c_1}$ = -0.094, p= .078; $\mathbf{c_2}$ = -0.227, p= .030
Pearson's correlation: *$\mathbf{d_1}$ = 0.464, p= .001; *$\mathbf{d_2}$ = 0.400, p= .005 (significant)
Pearson's correlation: $\mathbf{e_1}$ = 0.183, p= .118; $\mathbf{e_2}$ = 0.007, p= .122
Pearson's correlation: $\mathbf{f_1}$ = 0.123 p= .068; $\mathbf{f_2}$ = 0.263, p= .022
Pearson's correlation: $\mathbf{g_1}$ = 0.240, p= .027; $\mathbf{g_2}$ = 0.295, p= .016
Pearson's correlation: *$\mathbf{h_1}$ = -0.971 p< .001; *$\mathbf{h_2}$ = 0.742, p< .001 (significant)
↔ Indicate relationships

Appendix H9: Post-test Schematic Diagram for Hypothesis #9