An Analysis of the Virginia Performance Pay Incentives Initiative

by Melissa T. Goodwin

B.S. in Secondary Education English, May 2001, Central Connecticut State University
M.A. in Ed. in Curriculum and Instruction, July 2009, Virginia Polytechnic Institute and State University
Ed.S. in Educational Leadership and Administration, January 2012,
The George Washington University

A Dissertation submitted to

The Faculty of
The Graduate School of Education and Human Development
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Education

May 20, 2018

Dissertation directed by

Marguerita K. DeSander
Associate Professor of Educational Administration, Leadership and Research, Western Kentucky University and Former Assistant Professor of Educational Leadership and Administration

The Graduate School of Education and Human Development of The George Washington University certifies that Melissa T. Goodwin has passed the Final Examination for the degree of Doctor of Education as of February 26, 2018. This is the final and approved form of the dissertation.

An Analysis of the Virginia Performance Pay Incentives Initiative

Melissa T. Goodwin

Dissertation Research Committee:

Marguerita K. DeSander, Associate Professor of Educational Administration, Leadership and Research, Western Kentucky University and Former Assistant Professor of Educational Leadership and Administration

Sharon A. Dannels, Associate Professor of Educational Leadership, Committee Member

Leigh K. Kennedy, Senior Policy and Program Analyst, Strategy, National Education Association, Committee Member

© Copyright 2018 by Melissa T. Goodwin
All rights reserved

Dedication

I dedicate this dissertation to my parents, Jay and Vitaline Balog. Thank you for your unwavering support throughout all of my adventures. You both encouraged me to believe in myself and instilled in me the confidence to move forward and accomplish my goals, no matter how lofty. I will be forever grateful to you for jumping in to help when needed and for having faith in me. I love you!

Acknowledgements

Writing a dissertation is not a singular experience. I am humbled and honored that so many took the time to help me along the way. I am grateful to my chairperson, Dr. Marguerita DeSander, who saw enough value in this research to continue working with me after moving to Western Kentucky University. You were my Yoda throughout this process, and I thank you. Dr. Sharon Dannels was instrumental in fine-tuning the research design and methodology through her eagle eyes and standards of excellence. Thank you for the "SAD" comments! Thank you to Dr. Leigh Kennedy who brought a keen eye about pay for performance and education policy. My readers, Dr. Linda Lemasters and Dr. Jennifer Clayton, were with me throughout the doctoral journey. It was an honor that you shared your expertise during the final steps of the process.

Thank you to the family and friends who cheered me on. To my cohort members, Ana, Jason, and Drew: we began this process as colleagues, but we ended it as friends. I can't wait to work together again someday. To Leslie who encouraged me and laughed with me—it is a blessing to have you as a friend and colleague. I believe that together we will move mountains in our quest for educational equity. To my brothers and sister, Gregory, Mathieu, and Kathleen—thank you for being excited with me. I couldn't imagine sharing a childhood with anyone else. I am so happy we are friends—I love you guys! Finally, a huge thank you to my husband, Thomas, and our brilliant and beautiful daughter, Amelia, for their patience throughout this process. I love you both more than you will ever know. I look forward to our movie nights, camp outings, beach days, and hanging out. Noble Savages unite!

Abstract of Dissertation

An Analysis of the Virginia Performance Pay Incentives Initiative

An equal educational opportunity for all children is the intent of the United States public education system. In 1965, the federal government sponsored Title I funds to public schools to help children from economically disadvantaged backgrounds achieve academically (Cohen & Moffitt, 2009). In 2001, the No Child Left Behind Act (NCLB), an accountability effort connecting student performance, teacher outcomes, and student standardized test scores subsumed Title I. Federal authorization through Race to the Top (2010) deepened the student academic growth and teacher evaluation association. The latest reauthorization of NCLB, Every Student Succeeds Act (ESSA), provides funds to help local districts develop differential compensation and performance pay systems aimed at attracting and retaining teachers and school leaders to low income schools and districts (Every Student Succeeds Act of 2015). The evolving education policy of monetary incentives for positive student achievement outcomes is the focus of this dissertation.

Typical teacher compensation plans are a step/ lane system, based on years of experience with supplements for work toward advanced education and degrees (Figlio & Kenny, 2007; Loeb, Miller & Strunk, 2009; Stronge, Gareis, & Little, 2006). The system shifts with monetary incentives for teachers. Studies show that attempts to incentivize teacher compensation with the expectation of positively impacting student achievement were met with mixed results (Belfield & Heywood, 2008; Fryer, 2013; Kolbe & Strunk, 2012; Martin, 2010; Rice et al., 2012).

In the 2010-2011 school year, the Commonwealth of Virginia piloted a teacher incentive program, the Virginia Performance Pay Incentives Initiative (VPPI). Conclusions in the final evaluation omitted information about the incentive's impact on student pass rates for the Virginia Standards of Learning (SOL) test, the state standardized assessment. The data evaluation process was a quantitative analysis of gain scores in grades 5, 8, and high school, as measured by the SOL. This ex-post facto study used parameters to examine if a difference existed between SOL scores for the population invited to participate in the VPPI, the years before and of the initiative. The results of the study are inconclusive; however, the lessons learned from this analysis are applicable to future evaluation of teacher pay for performance policies.

Table of Contents

Dedication...iv

Acknowledgements...v

Abstract of Dissertation...vi

List of Figures...xi

List of Tables...xii

Chapter 1: Introduction...1

Statement of Potential Significance...4

The Virginia Performance Pay Incentives Initiative...7

Purpose and Research Questions...8

Statement of Potential Significance...10

Monetary Incentives and Teacher Retention...11

Conceptual Framework...13

Summary of Methodology...20

Limitations and Delimitations...23

Summary...28

Definitions of Key Terms...29

Chapter 2: Review of Literature...32

Consulted Databases and Search Terms...33

Description and Critique of Scholarly Literature...36

Federal and State Resources...36

Title I and No Child Left Behind……37

Federal Involvement in Performance Pay Initiatives……..40

Knowledge and Skills Based Pay……46

Career and Ladder Pay...…….47

Performance Pay…….48

Pay for Teaching Hard to Staff Subjects and Schools…….51

Teacher Unions and Alternative Compensation……..52

Studies Investigating Economic Incentive Policies…….63

The Motivation Theory Connection…….71

Virginia Performance Pay Incentives……..72

Motivational Challenges of Implementing Incentive Policies……..85

Conclusion…….90

Chapter 3: Methods……92

Research Questions……94

Research Design…….95

Participants…….99

Data Collection and Handling……..100

Instrument……102

Reliability and Validity……102

Data Analysis……...105

Human Participants and Ethics……106

Summary……..106

Chapter 4: Results……108

Description of the Population……..111

Data Analysis……...112

Analyses of the Research Questions……...…….117

Research Questions……..117

Summary……..124

Chapter 5: Interpretations, Conclusions, Recommendations……...127

Purpose…….130

Summary of Results…….132

Grade 5 Scores…….132

Grade 8 Scores…….132

High School End-of-Course Scores…….132

Interpretation of Findings……133

Using Gain Scores for Teacher Evaluation and Predictors of Student Success…….136

Conclusions and Implications for Future Research……..138

Implications for Teaching and Leadership……..140

Motivation and Teaching…….140

Shifting Education Policy in Virginia……..142

Final Thoughts…….143

References……146

Appendix A……..161

Appendix B……..176

Appendix C……..191

List of Figures

Figure 1 Conceptual Framework……...……..19

List of Tables

Table 1 *VPPI Grants Awarded in 2012*…….27

Table 2 *Percentage of Public School Districts That Used Pay Incentives for Various Reasons, By Selected Public School District Characteristics* ……..47

Table 3 *Percentage of Subject Area Hard to Staff Positions*.……52

Table 4 *TIPA Awards*.....78

Table 5 *Virginia Performance Standards* ……82

Table 6 *Qualifying Criteria of Hard to Staff Elementary Schools in the VPPI*…..…….114

Table 7 *Qualifying Criteria of Hard to Staff Middle Schools in the VPPI*……..115

Table 8 *VPPI Qualifying Criteria of Hard to Staff High Schools in the VPPI*……..116

Table 9 *Grade 5 VPPI SOL Mean, Standard Deviation, and Median Gain Scores*……118

Table 10 *Grade 5 VPPI Effect Sizes*……119

Table 11 *Grade 8 VPPI SOL Mean, Standard Deviation, and Median Gain Scores*…..120

Table 12 *Grade 8 VPPI Effect Sizes*……121

Table 13 *End of Course High School VPPI SOL Gain Scores*…..……..122

Table 14 *End of Course High School Effect Sizes*……...123

Chapter 1: Introduction

Overview

Educating children from poor families has long been a dilemma for educators. In 1965, the federal government enacted a proposed solution through Title I, a program designed to help poor children achieve academic success by providing extra money to public schools (Cohen & Moffitt, 2009). In 2004, the United States Department of Education (ED) amended the name of Title I to *Title I-Improving the Academic Achievement of the Disadvantaged*. Through this amendment, the ED (2004) issued a new statement of purpose that included the following provisions: students meeting at least proficiency on state assessments; aligning the written, tested, and taught curricula; closing the achievement gap between nonminority and minority students; focusing on literacy.

Student achievement was central to this policy, and the focus was the academic achievement of students identified as disadvantaged, from diverse racial or ethnic backgrounds, or needing assistance with developing reading skills (ED, 2004). Research suggests that closing the achievement gap between minority and nonminority students has been a challenge because "high-needs schools are most difficult to staff, and very few incentives currently exist to attract expert teachers, or teachers in general to teach in these settings" (Amrein-Beardsley, 2012, p. 2). This reality left state and school district leaders searching for an answer to the question of how to increase student achievement and entice teachers to their localities so districts could meet the mandates of the law. One possible solution was to find an alternative way to compensate teachers. This solution most often linked student achievement in the form of students' standardized test scores

(Callier, 2010; Gratz, 2011). The focus of this dissertation was to quantitatively evaluate a specific pay for performance initiative, the Virginia Performance Pay Incentives (VPPI), by analyzing the Standards of Learning (SOL) pass rates of all schools invited to participate in the program. In 2012, Virginia instituted a new teacher evaluation model connecting 40% of a teacher's evaluation to student academic growth (RMC Research Corporation, 2012). The implementation of this initiative was through a performance pay pilot initiative in 13 school districts in the Commonwealth of Virginia, with each of the Superintendent's regions represented (RMC Research Corporation, 2012). The *Virginia Performance Pay Incentives (VPPI) In Hard-to-Staff Schools Competitive Grant Application Packet* (2011d) indicated there were 169 (HTS) hard-to-staff schools invited to participate. The final evaluation conducted by the RMC Research Corporation (2012) indicated a total of 8 HTS schools participated. School districts chose to participate by applying for the VPPI grant; to increase the number of schools participating, the Virginia Department of Education (VDOE) included 16 School Improvement Grant (SIG) schools, which received money through the federal government (RMC Research Corporation, 2012). The total number of schools participating in the VPPI was 25. It was not clear in the final evaluation why the entire population invited did not participate in the VPPI or why the final evaluation omitted information about student test scores. Test scores were a key component in determining if teachers received the incentive; therefore, the focus of this dissertation was an analysis of gain scores of the entire population invited to participate in the VPPI to add information presented in the final evaluation of the program.

The traditional teacher compensation system followed by some school districts has been a step system that has remained unchanged for many reasons, which include, but are not limited to: teachers' unwillingness to change the system; teacher union involvement; and the district's inability to sustain an alternative compensation system because of the number of people involved (Levačić, 2009; Marsh & McCaffrey, 2012; Smarick, 2011; Streams, Butler, Cowen, Fowles, & Toma, 2011). This fixed system of teacher compensation has left states and localities searching for a way to recruit qualified and passionate teachers to work in hard-to-staff K-12 schools (Marsh & McCaffrey, 2012; Miles, 2011; O'Donnell, Lambert, & McCarthy, 2008; Petty, Fitchett, & O'Connor, 2012; Sims, 2011). The stagnation of teacher compensation has been exacerbated by the lane system—where all teaching jobs are paid based on time in the teaching profession with supplements for advanced degrees—and step system (Stronge et al., 2006). For example, according to the VDOE (2016a), the top critical teaching shortage areas in Virginia were: (a) special education; (b) elementary education, grades PreK-6; (c) middle education, grades 6-8; (d) mathematics, grades 6-12; (e) school counselor, grades PreK-12; (f) foreign languages, grades PreK-12; (g) English (secondary); (h) health and physical education, grades PreK-12; and (i) history and social science (secondary).

Many versions of monetary incentive programs for teachers have existed, including those linked to student performance, professionalism, years of teaching experience, and subject area expertise (Odden, Pincus, & Goetz, 2010; Stronge et al., 2006). Different forms of teacher compensation drew national attention from politicians and researchers hoping to meet the demands of new accountability measures and keep

quality teachers in the teaching profession using various program designs (Gratz, 2011; Levačić, 2009**;** Rice et al., 2012; Stronge et al., 2006). For this study, an evaluation of using monetary incentives to improve student learning was conducted. Additionally, motivation theory was used to gain insight into what persuades an individual to take certain actions. These factors were evaluated to answer the question: *Can a one-year teacher incentive program positively impact student achievement?* This was a pilot program to analyze a new teacher evaluation system that included a student academic achievement component, and it is unknown if the program was to continue beyond the one year.

Statement of the Problem

The main challenge the VPPI attempted to address was the recruitment and retention of teachers in hard-to-staff schools (HTS) and pilot a pay for performance incentive. This program was intended to positively impact student achievement—a challenge that is not unique to Virginia. Presenting teachers with incentives for increasing student achievement was not new, and the theory behind performance pay systems was that additional compensation inspired teachers to continue effective teaching practices and retained them in the teaching profession (Marsh & McCaffrey, 2012).

The use of financial supplements and incentives to increase student achievement in HTS schools is well-rooted in education policy up to the federal level. In 1965, localities received extra federal funds due to the Title I initiative, designed to support the education of children from poor families (Cohen & Moffitt, 2009). According to Cohen and Moffitt (2009), although Title I was controversial; it was reauthorized every five years. Title I changed significantly in 2001 when it became part of the No Child Left

Behind Act of 2001 (NCLB), the system that provided funding to all school districts based on demonstrated growth in annual student test results (Cohen & Moffitt, 2009). The latest reauthorization of NCLB occurred in 2015 and called the Every Student Succeeds Act (ESSA). The Final Consensus Language of the ESSA (2015b) granted states more authority and required them to administer "high-quality, yearly student academic assessments that include, at a minimum, academic assessments in mathematics, reading/ language arts, and science" (§200.2).

The financial focus, particularly on teacher compensation, has been an education policy highlight in recent years. In 2006, the Bush Administration introduced the Teacher Incentive Fund (TIF), a grant that helped fund differential compensation for teachers and principals in HTS schools. As indicated by Smarick (2011), Congress appropriated $400 million for TIF in 2010 and $950 million in 2011, when the fund also underwent a name change to the Teacher and Leader Innovation Fund (TLIF). The TLIF differed from TIF by including ways to improve teacher effectiveness through teacher educator preparation programs, professional opportunities for teachers, teacher evaluation systems, removal of ineffective teachers, and improvement of professional development to "support school turnaround efforts" (Smarick, 2011, p. 60). An additional change came under the Obama Administration's Race to the Top (RttT) initiative, merging teacher compensation and student academic growth models with a new teacher evaluation system (Smarick, 2011). Race to the Top was a $4.39 billion program funded through the 2009 American Recovery and Reinvestment Act (ARRA). In 2010, the ED indicated that this state-centered program "provide[d] incentives to States to implement large-scale, system-changing reforms that improve student achievement, close achievement gaps, and

increase graduation and college enrollment rates" (para. 1). The federal efforts affected funding at the state and district levels by providing localities with more money for salary incentives, adding to the models of teacher compensation. Specific TIF initiatives are addressed later in this review.

In addition to federal allocations to states, different types of programs exist at the state and local levels. Different salary structures impact these programs. Loeb et al., (2009) indicated that 27 states had a statewide minimum salary schedule, six states had financial incentives for career ladder salary arrangements, and three states had student achievement performance incentives as part of a statewide plan. Studies focusing on the use of federal and state resources in the creation of incentive programs, and an overview of these programs, are in this review. Additionally, an analysis of different models of monetary incentives for teachers is included (Fryer, 2013; Marsh & McCaffrey, 2012; Rice et al., 2012; RMC Research Corporation, 2012). Some models link monetary incentives to student achievement on standardized tests. Examples reviewed within this document include New York City Public Schools, Prince George's County, Maryland, the states of Minnesota and Colorado, and the Commonwealth of Virginia (Fryer, 2013; Minnesota Department of Education, 2016; Paufler & Amrein-Beardsley, 2014; Rice et al., 2012; RMC Research Corporation, 2012). Other reviewed studies focused on monetary incentives using characteristics of economic incentive policies as the unit of analysis (Greenlee & Brown, 2009; Petty et al., 2012) to demonstrate the diversity in monetary incentives programs. Overall, incentive program outcomes vary, and results are mixed. For programs that failed, overarching reasons were that student achievement gains are not made (Fryer, 2013, Paufler & Amrein-Beardsley, 2014, Rice et al., 2012),

incentive amounts are not high enough (Fryer, 2013), or teachers do not know how to increase student achievement (Fryer, 2013, Paufler & Amrein-Beardsley, 2014, Rice et al., 2012).

The Virginia Performance Pay Incentives Initiative

In 2010, the Commonwealth of Virginia piloted a performance pay incentives program called the Virginia Performance Pay Incentives Initiative (VPPI). This was a grant program focused on HTS schools in each of the Superintendent's regions in Virginia. The pilot was the implementation of the Commonwealth's updated teacher evaluation model, which incorporated a standard that tied students' academic growth to teachers' evaluations. The program attempted to positively impact student achievement in HTS and SIG schools by compensating teachers for exemplary ratings on their teacher evaluations. Students' demonstrated academic progress measured through increased test scores determined the rating.

The VDOE distributed lump sums to school districts to give to teachers. Grades 5, 8, and EOC exams were important analysis points because they were the culminating tests that marked the end of students' education at the elementary, middle, and high school levels. The importance of analyzing the SOL pass rates for this study came from an emphasis on the addition of student academic growth as the new portion of the teacher evaluation model. The final evaluation by the RMC Research Corporation (2012) did not indicate if there was a difference between the test scores at schools participating in the VPPI and at schools that were eligible for but did not participate in, the VPPI. To analyze the VPPI, an analysis of the gain scores from all invited schools determined whether a difference existed in student performance between participating and

nonparticipating schools during the 2010-2011 and 2011-2012 school years (i.e., the year before the VPPI and the year of the VPPI).

A gap in data related to student academic achievement emerged in a review of the final evaluation by the RMC Research Corporation (2012) due to missing student test score information. The analysis indicated that the purpose of the study was to "provide VDOE with insights into the functioning of the pilot" (RMC Research Corporation, 2012, p. 1). Additionally, the review specified that the "analyses were qualitative and descriptive, as appropriate to the purpose of the evaluation—not to make definitive statements of whether the pilot worked, but to gain understandings that will assist VDOE in moving forward" (RMC Research Corporation, 2012, p. 1). The evaluation was lacking information about pass rates on the Virginia SOL tests in grades 5 and 8, and on the end-of-course (EOC) high school tests, and in the incentive amounts teachers received.

Purpose and Research Questions

The purpose of this study was to investigate if a one-year teacher incentive program improved student achievement through an analysis of gain scores on standardized tests. The VPPI was the focus of the study, with a specific analysis on the Virginia SOL test scores for all schools invited to participate. The results of this study are not generalizable to other models because the VPPI is unique to Virginia; however, interested parties may be able to utilize their own contexts and experiences to draw conclusions (Paufler & Amrein-Beardsley, 2014).

This study examined the SOL test scores for all schools invited to participate in the VPPI. An investigation into this program was significant both because this was a

pilot of the Commonwealth's new teacher evaluation system and because this was also the first time Virginia had instituted a pay for performance model for teachers in each of the Superintendent's regions. In the evaluation report, the RMC Research Corporation (2012) stated, "The ultimate goals or future outcomes of the VPPI initiative pilot are twofold: academic achievement that soars for all students and exemplary teachers who dedicate their talents to teach in challenging, hard-to-staff schools and schools striving to improve" (p. 16). The main findings of the RMC Research Corporation (2012) final evaluation indicated the following:

1. The program was put into place and finished at each school.
2. The teachers and principals were trained and aided during the program's implementation.
3. The principals were more positive about and had a better opinion of the VPPI than teachers. (RMC Corporation, 2012, p. 30)

The final evaluation did not indicate whether the program achieved the goal of increasing student achievement for students because of the unavailability of Virginia SOL test scores. This missing data raised questions about the success of the program in terms of student achievement as measured by a standardized test and the success of paying teachers for improved student achievement. To find out if a difference existed between the SOL test scores at the participating and nonparticipating schools, gain scores were calculated using SOL data from the year before and the year of the VPPI. At each level, the content areas of reading, writing, science, math, and history were analyzed.

Research Questions:

1. What is the difference in grade 5 Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the incentive?
2. What is the difference in grade 8 Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the incentive?
3. What is the difference in end-of-course high school Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the incentive?

Statement of Potential Significance

The continued analysis of monetary incentives for teachers is important to study because it is a recycled education policy. Additionally, this was the first time Virginia offered an incentive within all of the Superintendent's regions, proposing the program throughout the state. This study examined whether a difference existed between SOL pass rates in grades 5, 8, EOC exams in schools invited to participate in the VPPI. An analysis of the year before the pilot and the year of the pilot took place (i.e., school years

2010-2011 and 2011-2012). The study had a limited timeframe because the VPPI lasted for only one year. To determine if the incentive made a difference in student test score outcomes, data for the year before the incentive and the year of the incentive were gathered, and gain scores were calculated and analyzed for grades 5, 8, and EOC. These scores encompassed culminating years of education at the elementary, middle, and high school grades. For this study, a one point median gain point was used.

A monetary incentive for teachers is an education policy initiative that continues to resurface. The VPPI adds to the list of these types of programs, driving the necessity for this evaluation. Additionally, this study is important because of the ways the VPPI attempted to connect incentives with the teacher evaluation model. The connection between monetary incentives and teacher retention appears below.

Monetary Incentives and Teacher Retention

It was important to gain an understanding of the history of the federal education initiatives aimed at teaching of children in HTS schools to analyze the context of the school, the unit of analysis for this study. To meet the challenges of working in an HTS environment, school finance has been moving to a system that compensates teachers for performance, and there have been many versions of these programs (Clotfelter, Glennie, Ladd, & Vigdor, 2008; Fryer, 2013; Marsh & McCaffrey, 2012; Rice et al., 2012; RMC Research Corporation, 2012). The context of student achievement and investment of federal and state resources devoted to the creation of monetary incentive programs for teachers prompted researchers to examine economic incentive policies that support these initiatives (Greenlee & Brown, 2009; Houck & Kurtz, 2010; Petty et al., 2012). Additionally, research about problems plaguing HTS have been vast, making the goal of

increased student achievement a lofty one (Barela, 2008, Belfield & Heywood, 2008; Edley & Mariano-Florentino, 2013). Therefore, the retention of quality teachers in these schools has been difficult because students face challenges beyond their control, and beyond the control of school personnel (Martin, 2010; Petty et al., 2012). One proposed solution to increase student academic achievement and retain quality teachers in HTS schools has been through economic incentive policies that measure success through a standardized test. According to Gratz (2011), these policies intended to demonstrate student academic success; however, the policies had additional conditions, making them difficult to utilize because multiple goals existed. This literature review addressed these factors.

The traditional salary structure for teachers has been a single-salary schedule consisting of a step and lane salary structure that compensates teachers for years of experience and the attainment of advanced credits or degrees. The single-salary structure had been deemed the fairest way to pay teachers (Stronge et al., 2006). A challenge has existed, however, to successfully recruit and retain teachers for HTS schools. To meet these challenges, one solution has been to pay teachers monetary incentives, with many variations of the model (Belfield & Heywood, 2008; Kolbe & Strunk, 2012; Liu, 2008; Martin, 2010; Odden et al., 2010; RMC Research Corporation, 2012; Salmon, 2010; Stronge et al., 2006). Alternative compensation plans, explored further in Chapter 2, are not new and were in place around the United States in various forms, with different levels of success. Alternative compensation models were important to explore because of the impact on the school budgeting process. Salary is an important factor in the school budgeting process. The different salary structures in place, which utilized money as one

way to recruit and retain a quality teaching force, continue to be prevalent, and led to the necessity for further research. The National Center for Education Statistics (NCES) indicated that teacher salaries constituted 80% of a school district's budget (NCES, 2015).

Conceptual Framework

Legislation has created and influenced education policy related to pay for performance policy initiatives. The results have been different types of programs and varied awards to entice teachers based on the belief that monetary funds will increase teacher motivation, improve teaching practices, and increase student achievement. Some examples of these awards are: knowledge and skills based pay; career ladder pay; performance pay; pay for teaching in HTS subjects and schools (Belfield & Heywood, 2008; Caillier, 2010; Stronge et al., 2006). When analyzing the literature related to pay for performance initiatives, it became apparent that an assumption exists that teachers are motivated by additional compensation. Therefore, teacher motivation deserves consideration. An understanding of motivation theory can help contextualize pay for performance initiatives. The theory of motivation will be explored through the theories of Vroom (1964) and Alderfer (1969).

Vroom (1964) analyzed motivation through expectancy theory and concluded that the following factors impacted employee motivation: personality, skills, abilities, knowledge, and experience. Expectancy theory is based on three beliefs:

1. *Valence*: The value an employee holds toward extrinsic rewards such as money, benefits, and time off, versus intrinsic rewards such as satisfaction.
2. *Expectancy*: An employee's confidence in personal capabilities.

3. *Instrumentality*: An employee's belief that management will fulfill promises about rewards (Vroom, 1964).

Vroom (1964) hypothesized that calculation of an employee's beliefs occurs in the following manner: Motivation = Valence x Expectancy x Instrumentality. This calculation can make predictions about an employee's work ethic, occupation, job satisfaction, and longevity in a job. People are drawn to experiences that will "maximize their pleasure and minimize their pain" (Vroom, 1995, p. 11) and "people are assumed to behave in ways that maximize certain types of outcomes (rewards, satisfiers, positive reinforcements, and so on) and minimize other outcomes (punishments, dissatisfiers, negative reinforcements, and so on)" (Vroom, 1964, p.12). An employee's valence is dependent upon what types of extrinsic rewards the employee finds rewarding, and if the desire for the extrinsic rewards outweighs the intrinsic satisfaction an employee derives from an activity. Expectancy is diverse and encompasses many facets, such as an employee's personal confidence in successfully completing a task, knowledge to successfully complete the task, the personal disposition to accomplish the task, and the necessary resources to see something to completion. Additionally, instrumentality comes from feeling that the reward is deserved, the outcomes are fair, and that there is trust the employee can successfully complete the task. Vroom (1964) theorized that people are motivated to work for the following reasons:

1. To meet personal economic needs for the purchase of goods for survival and goods purchased for social status.

2. The association with production of goods and services, including serving a moral purpose.

3. Social interaction.

4. Social status and the perception of identity outside of work.

As theorized by Vroom (1964), expectancy theory through valence (i.e., the importance of the outcome by the individual), expectancy (i.e., increased effort will increase outcomes), and instrumentality (i.e., positive performance will result in a positive outcome), can lead to not only a monetary outcome for an individual, but also motivation for the "development of skills, the attainment of respect by others, and the opportunity to contribute something useful to society" (p. 52).

In the examination of pay for performance plans outlined in the literature review, overall, teachers did not value monetary incentives (valence) because teachers do not go into teaching for monetary awards. The literature revealed that for many initiatives, teachers lacked an awareness of program structures or program expectations (instrumentality) to earn the incentive. Finally, the literature indicated that overall, teachers lacked an awareness of how to make changes to their current teaching practices (expectancy) to positively impact student achievement. Therefore, a conclusion is that pay for performance programs will not motivate teachers. Vroom's expectancy theory of motivation suggests that if the monetary portion of the incentive does not motivate teachers, or if they do not associate the monetary incentive with a higher social status, they may not be motivated to participate in a pay for performance plan. Valence, instrumentality, and expectancy are key factors in teacher motivation.

In addition to Vroom's motivation theory, Clayton Alderfer's work (1969) focused on the motivation of individuals. Alderfer modified Maslow's theory of hierarchy of needs by nesting the five needs into three needs, referred to as existence,

relatedness, and growth, or ERG (Ko, Rhee, Walker, & Lee, 2013). Existence includes the physical, material, and safety needs. Relatedness encompasses social needs—the needs of individuals to create and sustain positive interpersonal relationships and to feel security as a member of a group. Growth is the need of an individual to be creative, productive, and create meaning in life through tasks (Ko et al., 2013). Along with the ERG principles, *frustration regression is a principal part of* Alderfer's (1969) theory. When a person is not satisfied with growth opportunities, the individual may regress to relatedness needs for interpersonal connections and satisfaction due to the frustration associated with personal or professional development (Alderfer, 1969). The literature suggests that in attracting expert teachers to participate in pay for performance programs, the physical school environment is lacking in one or all of the existence needs (physical, material, or safety) because of the nature of the types of schools that offer incentive programs. Additionally, the relatedness factor is not met because in many instances the pay for performance initiatives are based on an individual content area, leaving out other educators who work with students, such as counselors or special education teachers. Finally, literature suggests that teachers do not have appropriate support systems or misunderstand the incentive program, thereby stopping the opportunity for the growth, and not meeting this need.

This study uses both Vroom (1964) and Alderfer (1969) to conceptualize human motivation. In this study, pay for performance initiatives begin as education policy initiatives. For teachers to be motivated to want to do better, factors from both theories have to come together to see increased motivation. For example, teachers have to value the initiative, believe they can make a change in student achievement, and expect to earn

the incentive. Additionally, they have to feel that they support is provided throughout the process, still be a part of the group with whom they work, and believe their actions will be enough to change student achievement. If any of these pieces is lacking, the intensity of the motivation declines and may go away altogether. Finally, although other motivation theories exist, such as Herzberg's theory of motivators and hygiene factors, for the purposes of this study, Vroom and Alderfer were appropriate. Vroom's theory is an analysis of the process of motivation (Francis, 2017), appropriate for an evaluation of a pay for performance initiative since teaching is a process of steps, analysis, reflection, and more steps. Additionally, Vroom's theory of valence, instrumentality, and expectancy fit well within the lens of analysis of pay for performance and the VPPI, the program analyzed. Vroom addresses the intangible intrinsic factors that motivate a teacher to enter the education profession (valence) as well as their awareness or lack thereof regarding program organization (instrumentality). Additionally, Vroom addresses the systemic issues that arise regarding the changing of teaching practices to create change and earn the incentive (expectancy). Alderfer's theory was an appropriate lens for this study because of the ways the theory of motivation fit with Vroom and the concept of frustration regression. This portion of the theory explores the reasons why a teacher may not feel motivated.

These factors related to motivation are important to consider with regard to pay for performance initiatives. The creation of pay for performance policies is out of a perception that monetary incentives to teachers can mitigate challenges faced in HTS schools. Teachers traditionally earn low salaries compared to other professions with comparable education requirements, and teachers in areas of poverty earn less than their

counterparts in more affluent schools or school districts (Ingersoll & Merrill, 2011). In theory, increasing pay will motivate teachers to be better at their craft. Analyzing motivation through Vroom's theory (1964) of why people work and Alderfer's ERG concept (1969) helps build the framework. The application of what motivates an individual helps to analyze the factors within the program, and the ultimate outcomes.

Figure 1 illustrates the conceptualization of this study. Through the literature review, expectations, values, and outcomes emerged as important factors when considering the connection between motivation theory and pay for performance programs. The federal government has influenced education policy and the types of rewards offered. Pay for performance programs vary, and with those programs come implementation challenges. The inception of the VPPI, for example, was unique because it was the first program of its kind used across the Commonwealth of Virginia. The VPPI lasted one year before it was changed. This study analyzed SOL data from the year before and the year of the program to find out if student achievement improved in one year.

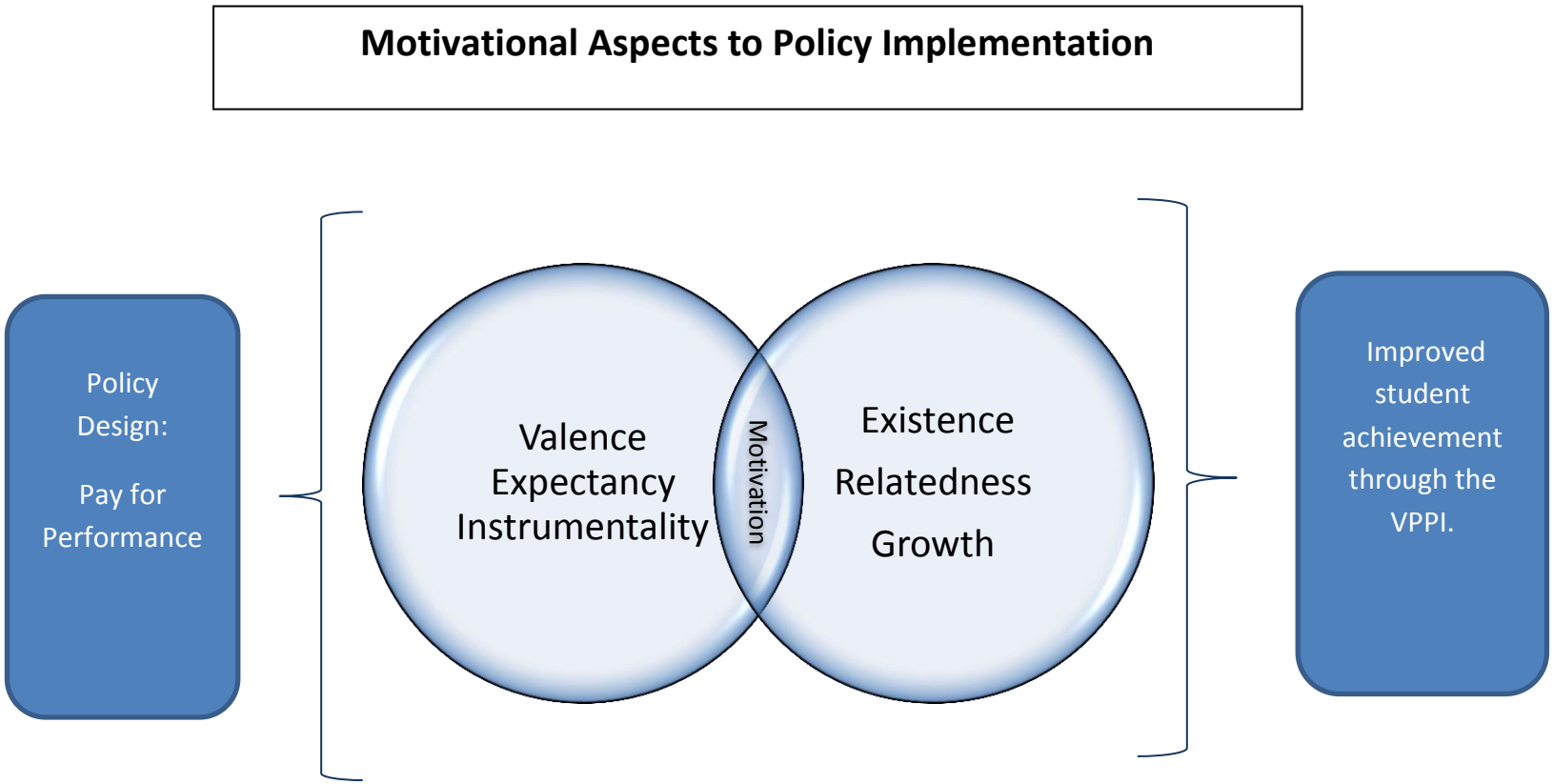


Figure 1. Conceptual framework.

Summary of Methodology

This ex post facto, quantitative study utilized school data obtained from the School Report Card, a publicly available database found on the VDOE website (VDOE, 2011e). The purpose of the study was to analyze SOL test scores for all schools invited to participate in the VPPI to determine whether the initiative positively impacted student achievement. The goals of the initiative were to pilot the implementation of a new teacher evaluation model that focused on increasing student academic growth, and offered teachers an incentive for increased student achievement as measured by the SOL test. The final evaluation by the RMC Research Corporation (2012) addressed the implementation of the new teacher evaluation model but left out information about the impact on student achievement. To fill this gap, the focus of this study was to analyze student achievement to see if a difference existed between the pass rates on the SOL for tests administered in grades 5 and 8 in the subject areas of math, reading, science, and history, and for all EOC tests. The grade level and subject area pass rates were compared across two years. The study included the year before and the year of the pilot (i.e., school years 2010-2011 and 2011-2012). It was unknown what content areas the participating teachers taught. Analysis of gain scores for all tests administered in the tested areas provided information about the impact of the VPPI on the school—the unit of analysis in this study.

On April 15, 2011, an attachment to Superintendent's Memo No. 113-11 went to all Virginia school districts with schools identified as HTS and invited those schools to apply for this grant. Although 169 HTS schools were initially invited to participate in the pilot program, only 9 participated in the initiative. The *Virginia Performance Pay*

Incentives (VPPI) in Hard-to-Staff Schools Competitive Grant Application Packet (2011) indicated that HTS schools were chosen due to the challenges faced in recruiting and retaining quality teachers based on school location (i.e., rural and urban) as well as the low socioeconomic status of the student population. Identification as HTS required schools meet a minimum of four out of the following eight criteria:

- Accredited with warning;
- Average daily attendance rate two percentage points below the statewide average;
- Percentage of special education students exceeds 150% of the statewide average;
- Percentage of limited English proficient (LEP) students exceeds 150% of the statewide average;
- Percentage of teachers with provisional licenses exceeds 150% of the statewide average;
- Percentage of special education teachers with provisional special education licenses exceeds 150% of the statewide average;
- Percentage of inexperienced teachers (0% of teaching experience) hired to total teachers exceeds 150% of the statewide average; and
- School has one or more inexperienced teachers (0 years of teaching experience) in critical shortage area. (VDOE, 2011d, p. 15)

To increase the number of participating schools, the pilot included 16 School Improvement Grant (SIG) schools, which brought the total to 25 schools. SIG schools were schools participating in a federally funded grant project offered through the ED.

States won SIG grants that also served as subgrants to local education agencies for use in raising student achievement (ED, 2016a, para. 1). All elementary, middle, and high schools eligible to participate in the VPPI were part of the study. Data originated from the VDOE's School Report Card page. An analysis of the data for the school years included within the study focused on gain scores calculated from the pass rates. All qualified schools (i.e., HTS and SIG) for the VPPI populated the study; therefore, the data were treated as coming from a population and parameters were reported. A difference of one point in gain scores was considered meaningful due to the accreditation standards set forth by the AYP ratings (VDOE, 2011a). For the school years in this study (i.e., 2010-2011 and 2011-2012), more than 86% of students in reading and 85% of students in math needed to demonstrate proficiency, and 70% of students needed to demonstrate proficiency in science and history. The calculation of frequency polygons, histograms, and box plots took place to summarize the data.

As indicated above, a one point gain score was important to this study. A one point difference can be the deciding factor between a child passing or not passing an SOL test. To be Pass Proficient on an SOL test, a child must get a certain number of correct answers to earn a pass/proficient achievement level or pass/advanced achievement level (VDOE, 2011b). Not passing an SOL test has implications on accreditation. School accreditation ratings are dependent upon student achievement on SOL tests and other factors (VDOE, 2017b). Therefore, the one point gain score was an important factor considered in this study.

Limitations and Delimitations

This study had several limitations. The number of schools invited to participate in the VPPI (N = 178) was a limitation. The initiative included either HTS schools or schools that received SIG money, and included elementary, middle, and high school levels. Information about why schools did or did not choose to apply for the program was not available. Additionally, it was unclear how many teachers at each site participated in the VPPI; based on a presentation to the House Appropriations Committee in 2013, 125 teachers from the HTS schools and 100 teachers from the SIG schools earned an exemplary rating and received the bonus (Hogge, 2013). A total of 782 teachers participated and were eligible for the bonus (Hogge, 2013). Furthermore, the funding components for HTS and SIG schools differed. Awards for teachers in HTS schools came from appropriations in the Virginia general fund, and funding for SIG awards came from federal monies set forth in ESEA guidelines for persistently low-achieving schools. Additionally, funding for the SIG schools was part of a state allocation made available through the 2009 American Recovery and Investment Act.

Pass rates from the subject areas of science, history, English, and math for elementary and middle schools and all EOC tests were represented in this study because achievement data for these content areas are available. This study was an analysis that explored whether a difference in gain scores existed between all schools invited to participate in the initiative the year before program implementation and the year of program implementation. Most importantly, this was a comparison of the schools that participated in the initiative compared to the schools that did not participate. The study cannot extend beyond those years because the VPPI was discontinued, and a new

program implemented. This program, the Strategic Compensation Grants (SCG), greatly varied from the VPPI, so data could not be compared. This study included only HTS and SIG schools in Virginia; therefore, the generalizability of the study was limited. The SCG grant application gave freedom to each school district to design the incentive program based on the district's own goals and objectives (VDOE, 2013). The unit of study was at the school level because teacher-level test score data were not available. It was also not known what grades teachers at participating schools taught, the composition of the classes in terms of students' academic ability (i.e., special education, average, honors, gifted), or how students enrolled in the classes (random assignment, teacher request, or parent request). The limits of the SGC prohibited the use of third year data, further limiting the scope of the study. Not having access to teacher level data, information about the content areas that were part of the third year program, and the expectation for outcomes of the SCG did not provide sufficient data to draw conclusions about the program.

Delimitations included the focus on an education policy specific to the Commonwealth of Virginia and the way the economic incentive was part of the new teacher evaluation system. Research indicated that linking teacher monetary incentives and teacher evaluation systems based on student academic growth became common practices (Caillier, 2010; Fryer, 2013; Gratz, 2011; Marsh & McCaffrey, 2012). The practice of connecting teacher evaluation to student achievement has become more prevalent since 2009 when funds available through federal RttT and TIF dictated that teacher evaluation models be connected to student achievement (Ballou & Springer, 2015). This study was necessary because student achievement data were not part of the

final evaluation by the RMC Research Corporation. In the evaluation by the RMC Research Corporation (2012), principals utilized newly revised performance standards, including the new Standard 7, which tied 40% of a teacher's evaluation to student academic progress. This pilot was for teachers who participated in the new evaluation system, making them the only teachers eligible for the incentive. School participation in the pilot was voluntary.

The criteria for teachers' participation were possession of a Virginia teaching license and licensure to teach that content area. A four-level rating scale—exemplary, proficient, developing/needs improvement, or unacceptable—was used for the teacher evaluation. Teachers who earned an exemplary rating received the incentive, and districts had the freedom to differentiate the incentive based on the teacher's status in the school; this included, for example, "teacher of record, teachers in support positions, or teachers in the SOL tested subjects" (RMC Research Corporation, 2012, p. 13). These factors delimited the results because this was a plan unique to the Commonwealth of Virginia.

A second delimitation was the analyzed grades. Grades 5 and 8 and the EOC tests were the best options because each marked the transition year in a child's education from one school (e.g., elementary) to the next (e.g., middle). The selection of the Pass Proficient performance level descriptor (PLD) was chosen because it indicated "what a student should know and be able to do" to pass the grade 5, grade 8, and EOC SOL tests (VDOE, 2012b, p. 9). The PLDs were scale scores that "are statements of what students should know and be able to do at each performance level given the content standards being addressed" (VDOE, 2012b, p. 9).

Factors not addressed in this study, but which had the potential to influence the outcomes, included: (a) teachers who did not participate in the study; (b) the size of the incentive to each school; (c) teachers' levels of education; (d) relationships between administrators and teachers; (e) years of teaching experience by teachers; and (f) principals' perceptions of teachers and their ability to meet the standards and earn the incentives. The unit of analysis was the school level; therefore, individual teachers' pass rates were absent. At schools where the VPPI took place, not all teachers participated, either because it was not mandatory or they chose not to. This meant that delineation did not exist between teachers who could receive the bonus and teachers ineligible to receive the bonus. Although not all teachers chose to participate, research suggested that financial incentives could motivate low-performing teachers to either improve their teaching performance or leave the teaching profession (Marsh & McCaffrey, 2012).

A phone conversation with the Assistant Superintendent for Teacher Education and Licensure for the VDOE revealed that the names of teachers awarded monetary incentives were unknown to the VDOE because school districts distributed the incentives to those teachers (personal communication, April 2014). Further research found in a presentation by Hogge (2013) to the Virginia House Appropriations Committee displayed the numbers of teachers from each district that received funds and the amounts distributed to those school districts (see Table 1). These teachers received exemplary ratings on their final evaluations to obtain the additional compensation. The final evaluation by the RMC Research Corporation (2012) indicated teachers who received an exemplary rating on their final evaluation received the incentive pay. Out of the 782 teachers who participated, 225 (29%) qualified for the bonus.

Table 1
VPPI Grants Awarded in 2012

School District	# of Eligible Schools	# of Teachers Receiving Bonuses	Grant Amount	HTS or SIG
Accomack	2	11	$41,984	HTS
Caroline	2	44	$230,909	HTS
Colonial Beach	1	8	$24,000	SIG
Dinwiddie	1	6	$32,295	HTS
Fluvanna	3	22	$66,000	SIG
Franklin City	1	4	$12,000	SIG
Greensville	1	13	$69,973	HTS
Hopewell	1	3	$9,000	SIG
Northampton County	2	0	N/A	HTS
Patrick	2	22	$119,492	SIG
Petersburg	2	5	$15,000	SIG
Richmond City	3	23	$69,000	SIG
Roanoke City	4	29 HTS 35 SIG	$224,492	1 HTS 3 SIG

Note. Adapted from "Strategic Compensation Grants (SCG) Initiative Competitive Grants," by Virginia Department of Education, 2013. http://hac.virginia.gov/committee/files/2013/11-20-13/II_a--Strategic_Compensation_Grants.pdf.

As previously stated, the context of this study was an analysis of increased student achievement through a monetary incentive offered through the VPPI. The VPPI served two purposes—to pilot a new standard in the teacher evaluation system based on student academic achievement and to provide the monetary incentive award for meeting that

goal. The final evaluation provided an incomplete representation of the program because no reporting of the student achievement information transpired. This dissertation focused on filling that gap by analyzing student achievement data in the form of gain scores from the SOL test by comparing the participating schools with those invited that did not participate.

Summary

A perceived problem in education is the lack of student achievement in HTS schools. One proposed solution has been to offer a monetary incentive to qualified teachers to increase student achievement. In the 2011-2012 school year, to meet this need, the Commonwealth of Virginia piloted a program for identified HTS and SIG schools: the VPPI. The VPPI offered incentives to these teachers and piloted a teacher evaluation model that linked student academic achievement with the teacher incentives. When the RMC Research Corporation (2012) conducted an evaluation of the program, the focus was on the implementation of the evaluation model to give the VDOE data to demonstrate how the pilot functioned. However, the evaluation neglected to include student achievement data despite the statement that the "ultimate goals or future outcomes of the VPPI initiative pilot are twofold: academic achievement that soars for all students and exemplary teachers who dedicate their talents to teach in challenging, hard-to-staff schools and schools striving to improve" (RMC Research Corporation, 2012, p. 16). This provided the context within which this dissertation examined how pay for performance programs work to solve the problem of lack of student achievement in hard to staff schools. An ex post facto analysis utilized gain scores to evaluate pass rates on SOL tests administered to students in grades 5 and 8, and EOC to determine if a

difference existed in the population of invited schools for school years 2010-2011 and 2011-2012 (i.e., the year before and the year of the VPPI). These scores evaluated whether differences existed in SOL test scores for the entire population of invited schools in the VPPI, which included the participating and nonparticipating schools. This study built upon the existing body of research about pay for performance programs, and thus added to the literature (Butin, 2010). These studies are explored in Chapter 2.

Definitions of Key Terms

- *Adequate Yearly Progress*—Demonstrated by schools that meet the annual objectives of student pass rates on standardized tests.
- *Differential compensation*— For the purposes of this study, differential compensation refers to compensation a teacher receives that supplements the regular salary and is used for the purpose of attracting or keeping the teacher in the specific school building. In this study, receipt of the compensation is contingent on student academic performance. (See also *pay for performance*.)
- *District*—Refers to the centralized entity responsible for the ways in which individual schools operate. In Virginia, these are called *divisions*, but for the purposes of this study all such entities are referred to as districts.
- *End-of-Course Test (EOC)*: The test a student takes at the culmination of a high school course deemed necessary for a verified credit for graduation. In Virginia, students have to earn a minimum of six verified credits to graduate from high school. The credits are: two in English; one in math;

one in a laboratory science, one in History, and one selected by the student.

- *Expectancy*: An employee's confidence in personal capabilities.
- *Hard-to-Staff Schools (HTS)*: Schools with difficulty retaining and recruiting teachers due to limited monetary resources and that serve a higher population of students from economically disadvantaged backgrounds.
- *Instrumentality*: An employee's belief that management will fulfill promises about rewards.
- *Monetary Incentives*: Incentives provided to teachers as part of a compensation program.
- *Motivation Theory:* The analysis of the factors that influence an individual to do something.
- *Pass Advance*—A score of 500 or above on the Virginia Standards of Learning test.
- *Pass Proficient*—A score of 400-499 on the Virginia Standards of Learning test.
- *Pass Rates*—The percentage of students who pass the Virginia Standards of Learning (SOL) test.
- *Pay for performance:* For the purposes of this study, pay for performance refers to compensation a teacher receives that supplements the regular salary and used for the sole purpose of attracting or keeping the teacher in the specific school building. In this study, receipt of the compensation is

contingent on student academic performance. (See also *differential compensation*.)

- *School Improvement Grant (SIG) Schools:* Schools that receive federal funds designated under the School Improvement Grant because of persistently low performance on the SOL tests as determined by the state accreditation standards.
- *Standards of Learning Test (SOL):* Standardized test administered to students in the Commonwealth of Virginia that measure student proficiency in course content.
- *Student Achievement:* The goal of increasing students' academic knowledge; demonstrated by showing growth, meeting, or exceeding benchmarks on standardized test scores.
- *Valence*: The value an employee holds toward extrinsic rewards such as money,

  benefits, and time off, versus intrinsic rewards such as satisfaction.
- *Virginia Department of Education (VDOE):* The Virginia state government agency responsible for programs and services related to education.
- *Virginia Performance Pay Incentives Pilot (VPPI):* The 2011-2012 VDOE program that piloted a new teacher evaluation system by compensating teachers for exemplary student academic achievement. This was only available to certain participants.

Chapter 2: Review of Literature

This literature review is an analysis of pay for performance programs. An evaluation of the different types of rewards for teachers gives insight into the ways in which educators are compensated for participation in such programs. Additionally, motivation theory helps create a framework regarding educator success in such programs. Motivation theory is used to conceptualize the review because incentive programs are driven by education policy. A selected overview of specific pay for performance programs sets the context for such initiatives, along with an evaluation of the literature regarding the challenges of implementing such programs. Specific attention is given to a review of the Virginia Performance Pay Incentives (VPPI) program, the focus of the study.

In 2012, Virginia instituted a new teacher evaluation model that tied 40% of a teacher's evaluation to students' academic growth. To analyze this new model of teacher evaluation, a pilot program was developed which also included a performance pay component. This dual-purpose pilot program occurred in 13 school districts in the Commonwealth of Virginia (RMC Research Corporation, 2012). The *Virginia Performance Pay Incentives (VPPI) in Hard-to-Staff Schools Competitive Grant Application Packet* (2011) indicated a total of 169 hard to staff schools (HTS) were invited to participate. The final evaluation conducted by the RMC Research Corporation (2012) indicated a total of 25 schools participated. Schools chose to participate by applying for the VPPI grant; to increase the number of schools participating, the Virginia Department of Education (VDOE) included School Improvement Grant (SIG) schools,

which received money through the federal government (RMC Research Corporation, 2012).

Incentive pay has been important to analyze because it is an education policy initiative used to increase student achievement that has been proposed as a solution to keep teachers in HTS schools. Incentive pay has been offered as the answer because step system-based teacher pay scales, which most school districts use to pay their teachers, have been stagnant (Stronge et al., 2006). The model of teacher incentive programs has had myriad versions, which are addressed in this review (Belfield & Heywood, 2008; Goldhaber, Gross, & Player, 2010; Gratz, 2011; Rice et al., 2012; RMC Research Corporation, 2011). This idea of extra pay has drawn national attention from politicians and researchers because it cycles around as a solution to increased student achievement and teacher retention.

The literature presented in this chapter is an analysis how pay for performance programs attempt to solve the problem of student achievement in HTS schools. Motivation theories of Vroom (1964) and Alderfer (1969) demonstrate that individuals are motivated based on their personal expectations, what they value, and the anticipated outcomes. In this study, the theory of motivation was associated with these factors because the chosen policy solution to increase student achievement in HTS schools was a monetary incentive for teachers.

Consulted Databases and Search Terms

To investigate the current trends around using teacher incentives as a way to raise student achievement, a review of the literature focuses on teacher incentive plans that had a goal of raising student achievement in HTS schools. This review also encompassed an

overview of education policy in the Commonwealth of Virginia to establish the context and implications of pay for performance programs in the state, as well as an analysis of the VPPI. There is a specific focus on the research from 2008 to the present because of the implications of the 2008 economic crisis on school funding. Information prior to 2008 was included to establish the background knowledge of teacher compensation and pay for performance plans. It was important to review alternative forms of compensation through the lens of current economic realities because of how starkly different school budgets have been since the economic downturn in 2008. The review was conducted using the following keywords to drive the search:

- School finance
- Student achievement
- No Child Left Behind
- School changes
- English language learners
- Standards-based reform
- Local property tax
- Teacher compensation
- Teacher education
- Salary
- Teaching
- Teacher retention
- Teacher pay
- Teacher incentives

- Disadvantaged schools
- School staffing
- *A Nation at Risk*
- Education policy
- Teacher incentive pay
- Education
- Teacher Incentive Fund
- Pay for performance
- Cost adjustments
- Education policy
- School finance
- Title I
- Differential compensation
- Career ladder pay
- Knowledge and skills based pay
- Incentive programs
- ESSA
- Value added models

The following search engines and academic libraries were consulted during the search for current literature:

- EBSCO Host
- JSTOR
- Gelman Library System

- Academic Search Complete
- ERIC
- Google Scholar
- Teacher Reference Center
- Policy File
- Education Source

The initial search resulted in 1,173 articles. To obtain the most recent information about the topic, the search was further narrowed by including only peer-reviewed articles published between 2008 and 2013. As research has evolved, more recent articles were added.

Description and Critique of Scholarly Literature

Federal and State Resources

Economic incentives have been used to help recruit and retain teachers in HTS schools (Amrein-Beardsley, 2012; Kolbe & Strunk, 2012, Petty et al., 2012). This recruitment need is due to an increase in teacher retirement and pressures that stem from a population increase in the number of school-aged children, such as program offerings, class size, special services, and a lack of new teachers entering the teaching profession (Ingersoll & Merrill 2010; Podgursky, 2006). According to Clotfelter et al., (2008), a shortage of qualified teachers, especially in certain subject areas, existed in schools that serve students considered low-income or low performing. In addition, Springer, Swain, and Rodriguez (2014) found that "public schools differ dramatically with respect to perceived working conditions, or attractiveness as places to teach, with schools that have higher concentrations of low-income, non-white, and low-performing students generally

perceived as less desirable places to work" (p. 3). Considering the incentives, policy developed, and questions about how to measure effectiveness arose; as a result, student achievement has been tied to the incentives to justify the additional compensation for teachers (Caillier, 2010; Gratz, 2011).

Title I and No Child Left Behind

According to Cohen and Moffitt (2009), throughout the 1950s and the early 1960s federal support for public education was almost nonexistent, so the responsibility for education was left to each state. Additionally, federal intervention was not welcomed due to divergent civil rights and religious interests throughout the country. The belief was that states should fund public schools. Legislators were divided about federal assistance to public schools because:

> Southern legislators wanted to keep state control of schools; civil rights advocates were opposed to support of segregated schools; Catholics were in favor of federal support of parochial schools; the First Amendment was used by Protestants and supporters of civil liberties to oppose federal aid to public schools. (Cohen & Moffitt, 2009, p. 2)

Despite the divergent opinions, the federal government had a problem it wanted to solve: how to educate children from poor families. The solution to this problem came in 1965 when the Johnson Administration passed the Elementary and Secondary Education Act (ESEA), which included Title I as a focus of the legislation. Title I gave the federal government some authority in public education by providing almost one billion dollars to the Title I initiative. Through this act, the federal government became entwined with public education. It was limited, however, by recognizing states' authority and only mandating that the received monies benefited children from poor families (Cohen & Moffitt, 2009).

Although Title I passed, it faced many challenges. One notable obstacle dating from the policy's inception was curricular; each locality differed, which impacted the effectiveness of what teachers implemented in their classrooms. In addition, "the government remained accountable for something of which it had little influence" (Cohen & Moffitt, 2009, p. 7), except for every five years when the federal government conducted an evaluation of Title I. For Title I, the aims of the policy were to solve the problem of educating children from poor families, and a disconnect emerged when school districts did not know how to utilize the human capital, social capital, and fiscal capital available to them. A proposed solution came when Coleman released a report funded through the United States Office of Education, mandated by Title IV of the 1964 Civil Rights Act. This report, called the *Equality of Educational Opportunity Survey* (EEOS), "measured inequality in education and student achievement outcomes" (Cohen & Moffitt, 2009, p. 68). Coleman discovered that what mattered most in a child's educational success—parental educational and social backgrounds—were not influenced by schools. Nevertheless, the Coleman Report did have limitations; Cohen and Moffitt (2009) indicated that "80 percent of the variation in students' achievement lay within, not among, schools" (p. 69). This meant that Coleman could not compare resources and achievement within schools, because parental education and social background mattered more in a student's achievement than what happened in school (Cohen & Moffitt, 2009). Even with these findings, Title I funds continued to be funneled through public schools.

As time passed, Title I continued to face challenges. Throughout the 1970s and 1980s, comparisons between the performance of students in the United States and other nations occurred. The implication for Title I was that it was part of a larger systemic

issue in a public education system considered to be failing all students. With the issue of the report *A Nation at Risk: The Imperative for Educational Reform* (1983), the prevailing view was that changes in the public education system could help close the achievement gap and alleviate the economic challenges facing the nation. Standards, tests, and accountability were the solutions from which all students could benefit. As a result, Title I gained mass appeal and the support needed to continue (Cohen & Moffitt, 2009).

In 2001, Title I was subsumed by the No Child Left Behind Act (NCLB) as part of the reauthorization of the ESEA. The 5-year period in which Title I should have been reauthorized was to begin in 1999, a presidential reelection year, but nothing changed until a new president officially took office. When George W. Bush became President, he and his administration fulfilled their promise to overhaul the public education system through a bipartisan effort and "aggressive federal intervention to succeed" (Cohen & Moffitt, 2009, p. 147). The intent of NCLB was to align academic standards and create a system of educator accountability through measurements obtained on annual standardized assessments. These results provided data to guide educators in better diagnosing areas of student need. According to Caillier (2010), NCLB held states and localities accountable for student performance based on federal standards and "literally shifted the responsibility for schools from local governments to the states and the federal government" (p. 58). As part of this system, Title I provided more money to school districts. However, a study from 2004-2005 showed only a roughly 5% rise in available funds due to NCLB, with a distribution of Title I funds almost even across school districts regardless of the poverty levels in those districts (Cohen & Moffitt, 2009).

According to Cohen and Moffitt, this was due to the change in education goals from 1965 to 2002, after the first year of implementation under NCLB. Although the aim of Title I had been to educate the poorest students, it still faced challenges because of the uneven distribution of Title I funds. This led educational leaders and policy developers to look toward creative ways to help schools deal with the education of children from low socioeconomic populations. This challenge influenced the creation of pay for performance programs.

Federal Involvement in Performance Pay Initiatives

Economic incentive policies evolved under President Obama's Race to the Top (RttT) program, which provided grants to fund state educational reform initiatives and Teacher Incentive Fund (TIF) grants, which were directed toward differential compensation initiatives for teachers and principals (Gratz, 2011). These federal grant programs required recipients to address student achievement. As school districts began to add their goals, alignment between local, state, and federal measures became difficult. Typical additions to the goals for students included: (a) social and emotional development, (b) good citizenship, (c) appreciation for humanities and arts, and (d) lifelong learners. These qualities were difficult to measure and created challenges when implementing an incentive program. Additionally, Belfield and Heywood (2008) stated that, sometimes, performance pay inaccurately measured the multidimensional nature of the work. Furthermore, the intention of the observation may not be what was measured, such as the quality of relationships that develop between teachers and students, between students in a class, a teacher's classroom management skills, or a teacher's relationships with administrators (Belfield & Heywood, 2008). They further explained that in jobs like

education that required the efforts of many professionals, it was unclear who deserved the credit among "a series of teachers, counselors, specialists, coaches, and a principal" (Belfield & Heywood, 2008, p. 244).

Performance pay raised concerns of economists and measurement experts; the matching of students with teachers was rarely a random event because teachers could request certain students based on academic achievement levels. Challenges existed both within and across schools because students were not randomly assigned to their classes, but instead had been classified based on academic ability, special education status, or parent requests for specific teachers. According to Rice et al., (2012) the focus on using student achievement data, including math and reading outcomes, to measure teacher effectiveness on evaluations caused concerns because of the risk of flaws; this resulted in the sometimes deliberate assignment of students to certain teachers. Caillier (2010) indicated that assignment differences between elementary and secondary teachers could also have an impact on student achievement. He posited that elementary teachers usually taught multiple subjects, were adept in one subject, but the evaluation only pertained to reading and math scores. This made the link between teacher and student achievement more difficult to measure for elementary school educators (Callier, 2010). In middle and high school, teachers typically specialized in one content area, making the link between teacher and student achievement easier to measure. Finally, Caillier (2010) also explained that teachers had many jobs in addition to teaching, which may include being "disciplinarians, counselors, and so forth, and they have to multitask when managing the classroom. Therefore, teachers perform many roles" (p. 59). Additionally, the collaborative nature of elementary schools creates difficulty in connecting student

learning to one individual, and the model leaves out performance measurement for teachers in nontested subjects. Cohen and Moffitt (2009) suggested that the result was rewarding teachers for enlisting students who are academically more able than their peers. This outcome resulted in a mixed message about the effects of economic incentives on teacher performance (Yuan et al., 2013).

The federal government's role had not been limited to the ways funding occurred in schools—it also involved how to spend that money. In 2006, the Bush Administration introduced the TIF. According to the ED (2006), the TIF money was intended to help with:

> Improving student achievement by increasing teacher and principal effectiveness; reforming teacher and principal compensation systems so that teachers and principals are rewarded for increases in student achievement; increasing the number of effective teachers teaching poor, minority, and disadvantaged students in hard-to-staff subjects; and creating sustainable performance-based compensation systems. (para 3)

Smarick (2011) indicated that Congress appropriated $400 million for TIF in 2010 and $950 million in 2011 when TIF changed into the Teacher and Leader Innovation Fund (TLIF). Unlike TIF, TLIF included ways to improve teacher effectiveness through teacher educator preparation programs, professional opportunities for teachers, evaluation systems, removal of ineffective teachers, and improvement of professional development, and also "support[ed] school turnaround efforts" (Smarick, 2011, p. 60). Student growth models were combined with compensation to evaluate teachers under the Obama Administration's RttT initiative. Differential pay for teachers and principals in HTS schools, as originally defined under the original version of TIF in 2006, was no longer a highlight of the new proposal. The trend for differential pay

moved from attracting teachers to HTS schools to compensating teachers for positive student achievement outcomes in HTS schools.

Teacher Compensation and School Funding in the United States

The economic climate has continued to change, especially considering the 2008 recession. Research has suggested that alternative forms of compensation are becoming more common. As indicated by Belfield and Heywood (2008), these alternative forms of compensation within the general workforce have "suggest[ed] that teachers should be a part of that process" (p. 243). Historically, teacher pay scales are on a step scale, and steps and lanes have been most familiar in the United States' teacher salary schedules. A step or lane system is based on a teacher's years of experience with supplements given for additional credits or degrees earned beyond a bachelor's degree (Stronge et al., 2006).

Steps and lanes were not always the case, though; the concept of differential compensation was familiar in the early 1900s, when teachers' pay focused on the grade levels taught, and men received more pay than women (Shepard, 2009; Stronge et al., 2006). As education policy has evolved, merit pay programs have changed to include teacher evaluation based on student academic achievement (Gratz, 2011).

To evaluate the climate of changing education policy connecting teacher evaluation and student achievement, researchers are analyzing teacher perceptions of these programs. Petty et al. (2012) conducted an online survey of high school teachers (N = 189) in economically disadvantaged schools in one southeastern state. This research explored how the teachers understood their work environment and what practices needed implementation to attract and retain teachers in those schools. In addition, researchers analyzed how high school teachers viewed their work environments and looked to gain

an understanding of how the teachers viewed incentives for current practitioners as well as for new and incoming teachers. Findings revealed that 59% of the teachers (n = 78) indicated that money or a financial incentive was the most important factor in attracting teachers to the school and was also the most important factor in retaining teachers at a school. This study had several limitations, however; researchers acknowledged that the limitations of self-reported data from teachers and a response rate of 35% limited generalizability (Petty et al., 2012). An additional limitation is that the findings are from one state. The outcome of teachers believing that money was a factor indicates that their perceptions of the work environment are influenced by monetary incentives.

Greenlee and Brown (2009) used a convenience sample of 97 teachers in an educational leadership program at the University of South Florida. Questions were based on information from a review of the literature regarding teacher recruitment and retention in HTS schools in conjunction with Florida Leadership Standards (Greenlee & Brown, 2009). Researchers used a "categorical, Likert-type scale, ranking, and open-ended survey items" (Greenlee & Brown, 2009, p. 100). They found that augmenting a teacher's salary in the form of bonuses, tuition reimbursement, performance pay, and enhanced retirement benefits would attract teachers to work in HTS schools. Limitations of the study included all-White respondents, which restricted the scope of the survey. Additionally, respondents invited other teachers to participate; however, information about those invited to participate remained unverified by researchers. Additionally, researchers did not indicate their response rates. Although this study was limited in many ways, it is important to note that a monetary supplement was noted as a factor in retaining teachers in the schools.

Martin (2010) utilized information from 3,500 school districts that participated in the 1999-2000 Schools and Staffing Survey (SASS) from the National Center for Education Statistics (NCES) to investigate why some schools offered incentives and others stayed on a traditional pay scale. This study described the distribution of incentive pay. Martin conducted the research through the theoretical lens of compensating wage differential theory, which indicated that districts and schools with less attractive work conditions paid more for quality teachers. A logit regression model was used for the analysis related to the following dependent variables: any incentive, professional development, shortage, and excellence. Independent variables included: (a) student characteristics of free or reduced lunch; (b) race and ethnicity; (c) crime rate; (d) location (urban or suburban); (e) private school proportion; (f) public school religious proportion; (g) public school nonreligious proportion; (h) charter school proportion; (i) number of districts in the county; (j) union; (k) school choice; and (l) magnet program. Results suggested that districts with "higher percentages of Black or Hispanic students, are racially or socioeconomically segregated, have more students or higher crime rates" and are more likely to offer incentives (Martin, 2010, p.1149). In addition, Martin indicated this is consistent with a hypothesis that HTS schools are more likely to offer incentives because "the marginal effect of the segregation of students across schools by socioeconomic status is .373, suggesting that if a district is completely segregated and Black, then it is 37.3% more likely to offer an incentive than a district that is completely integrated" (p. 1050). Martin (2010) deduced that the decision to offer incentives to teachers was not based on the teachers' expertise, but rather the school district's desire to influence teachers to stay in a school.

The four most commonly used teacher compensation models have been: knowledge-and-skills-based, career ladder pay, performance pay, and pay for HTS subjects (Stronge et al., 2006). These concepts are explored in the following sections.

Knowledge and Skills Based Pay

In the knowledge and skills based pay system, teachers receive additional rewards based on their professional knowledge and skills. This model typically requires teachers to demonstrate growth in a certain area identified by the school district. These pay systems gained attention in the late 1990s and early 2000 as school districts adopted the private sector practice of encouraging employees to increase their knowledge and skill base as well as promoting a positive organizational culture (Odden et al., 2010).

In the knowledge and skills based pay system, evaluations of teachers are frequently in the form of classroom observations or portfolio reviews (Stronge et al., 2006). Compensation is either a one-time bonus or a pay scale movement in addition to their yearly increase. One of the most common types of knowledge and skills based pay is tied to National Board Teacher Certification through the National Board for Professional Teaching Standards (NBPTS) as shown in Table 2. Publicly accessible data from the SASS from the NCES (2012) shows percentages of school districts across the United States that used pay incentives. This information originates from questionnaires and was sent to public and private schools, Bureau of Indian Education/tribal schools, and a sample of public charter schools.

Table 2

Percentage of Public School Districts That Used Pay Incentives for Various Reasons, by Selected Public School District Characteristics, N= 16,990 (2011-12)

	To reward teachers who have attained NBPTS certification (%)	To reward excellence in teaching (%)	To recruit or retain teachers in a less desirable location (%)	To recruit or retain teachers to teach in fields of shortage (%)
All public school districts in the United States	24.5	11.3	5.6	13.5

Note. Adapted from the *Schools and Staffing Survey* (SASS), by National Center for Education Statistics, 2012. Retrieved from https://nces.ed.gov/surveys/sass/

Career Ladder Pay

Career ladder pay served as a prevalent form of alternative compensation offered to teachers after the release of the report, *A Nation at Risk: The Imperative for Educational Reform* (Gardner, 1983). In 1983, Wisconsin's governor formed a task force to review teaching and teacher education, and one of the task force's recommendations was the creation and implementation of a career ladder program (Burke & Lind, 1987). In 1984, Rosenholtz and Smylie reported that the implementation of career ladder programs was a public response to the need to improve the quality of teacher candidates and teacher quality to increase student achievement.

Career ladder pay gives educators terms for specific types of professional levels of growth based on level of experience and education. The titles are typically "novice, apprentice, teacher leader, expert teacher, distinguished teacher, or master teacher" (Stronge et al., 2006, p. 77). Teachers increase their levels of experience by earning advanced degrees, completing graduate courses, acquiring special certifications, and

earning additional compensation for duties (e.g., mentor teacher or master teacher; Gratz, 2011). One example of a career ladder program is Minnesota's statewide Q Comp program, explored later in this review.

Performance Pay

Performance pay is different from other forms of economic incentives because it directly ties the incentive to student achievement. These models have been based on student achievement at the teacher level, team teacher level, school level, or department level (Stronge, et al., 2006). Gratz (2011) indicated that the theory supporting performance pay was grounded in the belief that the incentive would encourage teachers, which in turn would improve students' motivation for academic success as demonstrated through higher test scores. Questions about the reliability and validity of the assessments, the measurement of student achievement based on a short period of time, and the intangible factors associated with education makes educators suspicious about this form of incentive (Gratz, 2011; Stronge et al., 2006).

Performance pay is a practice adopted from the business sector and implemented in federal agencies under the Clinton and Bush administrations. President Bush took the implementation further with the 2001 reauthorization of NCLB and the implementation of value added models (VAM) to measure student achievement (Callier, 2010). The VAM approach measures "student gains or change over a specified period rather than criterion-based performance to address concerns about the influence of individual teachers as opposed to collective influence over time" (Stronge et al., 2006, pp. 93-94).

Value added models have generated much debate in the education research and policy communities because RttT incentives included teacher quality measures with a

focus on evaluation of teachers based on student achievement results. The ESSA gave authority back to states regarding the evaluation of teachers and did not dictate that a link between student test scores and teacher evaluations be made. Additionally, the research community has spent much time investigating whether VAM metrics accurately measure a teacher's contribution to student learning. In November 2015, the American Educational Research Association (AERA) made a statement cautioning researchers and policy experts on the use of VAMs in teacher evaluation and teacher preparation. The organization acknowledged that VAMs may be "superior to some other models of measuring teacher impacts on learning outcomes," (para. 2) but it cautioned that "it does not mean that they are ready for use in educator or program evaluation" (para. 5). The AERA (2015) indicated that "there are potentially serious negative consequences in the context of evaluation that can result from the use of VAM based on incomplete or flawed data, as well as from misinterpretation or misuse of the VAM results" (para. 5).

NCLB dictated that state tests only measure grade level standards, which meant that the skills and knowledge from earlier grades had not been measured. However, this indicated that valuable information could have been missing about learning across the grades (Darling-Hammond, 2015). In addition, Paufler and Amrein-Beardsley (2014) caution that VAMs are not random—bias becomes an issue in determining student achievement because larger numbers of certain student populations are in some classrooms. Differences in student backgrounds impact VAMs because of the uneven distribution of poverty, homelessness, familial support, enrichment programs, and healthy and regular meals, as a few examples (Darling-Hammond, 2015; Konstantopoulos, 2014; Martin, 2010; Paufler & Amrein-Beardsley, 2014). Moreover, diversity in student

backgrounds cannot be controlled, creating a ranking order of teachers in schools. For example, some HTS schools that worked to recruit and retain qualified teachers have had to rank teachers. As a result, some effective teachers have been ranked as ineffective; conversely, schools filled with ineffective teachers have had to rank some as effective (Darling-Hammond, 2015).

Goldhaber et al. (2010) also analyzed the use of VAMs within the career trajectories of midcareer female teachers to determine if public schools could retain their best teachers. Researchers analyzed three types of moves in the North Carolina education system: within-district moves, cross-district moves, and exits from the system. The selection of the midcareer period indicated that teachers had a moderate probability to exit the teaching profession. Through an examination of elementary teachers in North Carolina (N = 9,027) from 1996-2002, teacher mobility and student achievement were analyzed to directly link the two. Researchers utilized competing risk models and based all analyses on a discrete time hazard model using one-year increments, "which defined an individual's odds of leaving the position as a function of the baseline hazard [0(t)] and a series of covariates (Xin) that included measures of the teacher's quality, teacher's background characteristics, labor market conditions, and school characteristics" (Goldhaber et al., 2010, p. 65). Teachers with advanced certifications (e.g., NBPTS certification), an advanced academic degree, strong SAT scores, or who teach in schools with fewer low-income students had a greater likelihood of staying at that school. Teachers had a higher probability of leaving schools, however, which included a higher low-income student population, higher African American student population, and lower test scores in math (Goldhaber et al., 2010). In analyzing the link between student

achievement and teacher effectiveness, 44% of the teachers considered to be the lowest performers changed schools at least one time, either within or out of district (Goldhaber et al., 2010). In addition, when weaker teachers left the system, it was a benefit to the overall quality of the system. Therefore, the more effective the teacher, the less likely the teacher left one school for another school. Although this study did directly analyze a differential compensation system, it is important to include in the review because VAMs are used as part of these pay systems. More research is needed about the impact of differential compensation on retaining good teachers in highly disadvantaged schools. A limitation of this study included the lack of follow-up with teachers who left the system, to know whether they continued teaching in a neighboring district or moved to a teaching job in a neighboring state.

Using VAM as an incentive pay model is controversial because of the lack of connectedness between the ways teachers are evaluated (administrator evaluation or student survey data); differences between VAM models; teacher and administrator ignorance about the way the VAM works; and the limited scope of teachers evaluated under VAM models (Amrein-Beardsley et al., 2016).

Pay for Teaching Hard To Staff Subjects and Schools

The push for pay for teaching HTS subjects and in HTS schools stems from concern for students from poor, urban, rural, or minority school districts who do not have the same caliber of experienced or expert teachers as white students from affluent school districts (Opfer, 2011; Podgursky, 2006). Teacher recruitment and retention in the areas of special education, math, and science, and for teachers of English learners (ELs), heightens this concern (Berry, 2004; Podgursky, 2006; Strunk & Zeehandelaar, 2011). A

2015 study by the NCES found that vacancies in HTS positions in high school declined between the 1999-2000 and 2011-2012 school years. Principals self-reported that vacancies for HTS positions were "very difficult" to fill or that a vacancy in a certain subject "could not be filled" (NCES, 2015, p. 20), as displayed in Table 3.

Table 3

Percentage of Subject Area Hard-to-Staff Vacancies

Subject Area	**1999-2000 (%)** **N = 8,430**	**2003-04 (%)** **N = 7990**	**2007-08 (%)** **N = 7570**	**2011-12 (%)** **N= 7480**
Mathematics	23.5	15.7	19.1	9.1
Special Education	20.0	12.9	13.1	7.6
Physical Sciences	15.0	9.8	11.9	6.7
Biology or Life Sciences	13.3	7.5	13.4	5.3
English as a Second Language (ESL)	6.8	4.3	5.8	2.6

Note. Adapted from *National Center for Education Statistics - Percent of all public high schools that had at least one difficult-to-staff teaching position, by subject area: School years 1999-2000, 2003-2004, 2007-2008, and 2011-2012.*

Further analysis about research in this area is provided later in this literature review.

Teacher Unions and Alternative Compensation

The 2015 annual report from the Bureau of Labor Statistics stated that the percentage of wage and salary workers, who were members of unions, and the highest rate of union membership at 35.3%, belonged to an umbrella group made up of teachers, librarians, and protective service occupations. Furthermore, 49% of teachers, or 2.5

million, were members of the teachers' union, down from 53% a decade prior (Bureau of Labor Statistics, 2015). Additionally, research indicated that teachers' unions impact the implementation of alternative compensation programs (Belfield & Heywood, 2008; Goldhaber, 2009; Gratz, 2011; Kolbe & Strunk, 2012; Rice et al., 2012). Goldhaber (2009) suggested that the two largest teacher unions, the National Education Association (NEA) and the American Federation of Teachers (AFT), disapproved of altering teachers' salary schedules, but have had divergent views around specific details related to teacher pay. When economic incentive policies have been implemented with union support, the scope, scale, implementation, and life of each policy varied (Goldhaber, 2009). A summary of the state policies that specifically involved teacher unions follows.

In 2007, the Florida Legislature attempted to implement two pay for performance programs: E-Comp and STAR. Research by Brodsky, DeCesare, and Kramer-Wine (2010) found strong teacher opposition because the main determinant of whether teachers received incentives was students' standardized test scores. The teacher unions filed lawsuits and the policies failed because of little to no support from teacher constituency groups (e.g., teachers and teacher unions; Brodsky et al., 2010).

During the 2007-2008 school year, the New York City Department of Education and the United Federation of Teachers (UFT) unveiled a pay for performance program called the Schoolwide Performance Bonus Program (SPBP). This was school-based program and offered financial rewards to HTS schools at each educational level: elementary, middle, and high. Schools that met student achievement goals received $3,000 for each full-time union-represented staff member in the school. A four-person committee at each school decided the distribution of the bonus, based on (a) the New

York City Public Schools' progress reports that included information about standardized tests and compared scores between schools, as well as (b) school environment, as measured by school-level progress reports of student achievement data and teachers' attitudes and beliefs toward the program (Fryer, 2013; Marsh & McCaffrey, 2012). The design of the program was left to the schools, but it had to meet three criteria: (1) seniority could not impact distribution of incentives; (2) the principal, principal designee, and two UFT members needed to be on the compensation committee; (3) a committee needed to come to a unanimous decision (Fryer, 2013). Fourth and eighth grade state tests were used to select eligible middle schools and high schools, and elementary schools were chosen based on student populations for students identified as English Language Learners (ELL), students identified as special education, or the percentage of school poverty rates. The New York Department of Education identified 438 eligible schools. For reasons unknown to one researcher, the UFT did not allow participation from 34 of the eligible schools, and 75 of the schools were ineligible because they were designated special education schools (Fryer, 2013). This left a group of 396 schools that were eligible for the program.

Two separate research teams evaluated the program: Marsh and McCaffrey (2012) and Fryer (2013). Marsh and McCaffrey (2012) utilized case study and survey methods, and wanted to find out why the program did not reach the goals. Fryer (2013) used regression analysis and wanted to find out the outcomes of a pay for performance program based on student achievement.

In the Fryer (2013) study, the researcher started with the treatment group of 396 eligible schools and randomly selected 212 schools from this group. To participate in the

program, 55% of the full time staff had to agree to participation. Out of the 212 selected, 179 schools had enough votes for participation. To increase the number of participating schools, Fryer (2013) chose 21 additional schools off a wait list, 19 of which had enough votes to participate. The final group of participating schools was 233 participating in the pilot and 163 control schools for the first year. The researcher does not indicate how the control group was chosen. In year two of the program, 198 schools were chosen to come back, but 195 were eligible because the three schools closed due to low performance. It was not clear how the 198 were chosen. Out of the 195 eligible schools, 191 chose to participate a second year. In year three of the program, all 191 schools were chosen to return, and 189 chose to participate. It was not clear why all 191 were chosen to return or why 189 chose to participate.

Incentives for teachers at all three levels were earned based on three points: learning environment, student performance, and student progress compared to a "peer school" and all other schools in the city (Fryer, 2013, p. 382). Additionally, the comparison score to the peer school was weighted three times more in the final evaluation over the other schools in the city, which Fryer (2013) conceded made it difficult to determine how school faculty could find ways to increase scores to earn the incentive.

In the first year of the program, 104 schools received the full bonus because they met the full criteria. Eighteen schools met 75% of the criteria, and received half the bonus. In year two, 154 schools received the full bonus, and seven schools received half the bonus. Year three bonus information was not available to the researcher, and an overall figure of $75 million was given by the New York Department of Education

(Fryer, 2013). This is a weakness of the study because of lack of specific information about bonus information.

To analyze the impact of the incentive, Fryer (2013) used the following student data: race, gender, free and reduced price lunch status, behavior, attendance, course grades, and state math and English scores for students in grades 3-8. Regents exam scores and graduation rates were used for high school students. School level controls were constructed by assigning students to the first school attended in September and "taking mean value of student demographic variables and student test scores in each school" (Fryer, 2013, p. 387). Teacher data were constructed by taking the following information from the teachers' human resources records: race, gender, and teacher value added data, which were available for the 2006-2007 through the 2008-2009 school years (Fryer, 2013). The teacher value added data were "measured in standard deviation units" by the grade level for each year to have a "mean of zero and a standard deviation of one" for cumulative teacher data (Fryer, 2013, 387). Fryer (2013) constructed a counterfactual to restrict analysis to schools selected to participate in the program and "estimate a causal impact of being offered a chance to participate in a teacher incentive program by comparing the average outcomes of schools randomly selected for treatment and the average outcomes of schools randomly selected for control" (p. 390). Additionally, Fryer (2013) used a "treatment-on-treated" effect to measure the impact of the incentive on the schools that chose to participate (p. 390).

Final analysis by Fryer (2013) indicated that the program did not impact teacher attendance or retention and concluded that was because most schools distributed the $3,000 incentive equally between staff members, which was "less than 4.1% of the

average teacher salary in the sample" (p. 402). Fryer (2013) concluded that the incentives did not increase student achievement, except for math scores, and those slightly declined in the last year of the program. Additionally, the math scores were not statistically significant (Fryer, 2013). Additionally, Fryer (2013) indicated that the formula for determining the reward was too complex, and that it was not clear to teachers what they needed to do to positively impact student achievement. A limitation of this study includes using already collected student achievement data from standardized tests rather than collecting new data, and third year data were not available.

Marsh and McCaffrey (2012) also analyzed the SPBP and "interviewed more than 130 individuals in 14 case study schools, interviewed program leaders, and analyzed administrative, individual, and school level student achievement data" (p.52). In interviews with teachers, Marsh and McCaffrey (2012) discovered the incentive was too small for what teachers saw as their best teaching efforts. Additionally, Marsh and McCaffrey (2012) stated that more than one-third of teachers surveyed indicated no awareness of pertinent program structures, such as targets, bonuses, or distributions. Teachers indicated this confusion about the program existed from the beginning of the school year. Furthermore, Marsh and McCaffrey (2012) also concluded that the incentive amount earned after taxes was small and did not impact their classroom teaching practices. A principal, who worked at a school that earned the bonus, reiterated this sentiment: "It should be icing on the cake, because we did the job. It shouldn't be the reason why we get the job done" (Marsh & McCaffrey, 2012, p. 55).

Lastly, based on interview data, Marsh and McCaffrey (2012) concluded that the incentive was ineffective because teachers lacked the knowledge and skills to change

their teaching practices to affect positive outcomes for student achievement. Therefore, teachers continued to teach using the same methods or new teaching methodologies without knowledge of how to produce positive changes for student academic outcomes. Limitations of this study include resource restraints that impacted intended research. Additionally, it was not clear how many teachers and administrators were selected to be interviewed, and how many participated in the interviews, limiting conclusions that can be drawn from the interviews.

The Minnesota Legislature initiated another program, the Minnesota Q-Comp program, in 2005; it was a combined effort of the Minnesota Federation of Teachers, Governor Tim Pawlenty, and the Milliken Family Foundation (Brodsky et al., 2010). It emphasized entire school performance and incentives for school faculty and staff, which included awards based on student achievement and peer evaluations. Teachers received rewards, as incentives for their expertise, when master teachers became instructional leaders and mentors for novice teachers while serving as instructional leaders for the district. The Minnesota Department of Education (MDOE) indicated that traditional public school districts received $260 per student, charter schools and the Perpich Center of the Arts received $254 per student, and intermediate and cooperative districts received $3,000 per teacher in state aid for the program (2016). These funds supported one of the following five areas: Career Ladder/Advancement Options, Job Embedded Professional Development, Teacher Evaluation, Performance Pay, and the Alternative Salary Schedule (MDOE, 2016). For the 2015-2016 school year, 82 school districts and 69 charter schools implemented or were approved to implement Q-Comp (MDOE, 2016).

Hezel Associates (2009) conducted an external review of the Q-Comp program for the MDOE. The group used a "formative-summative approach" to investigate "Q-Comp impact and sustainability" (Hezel Associates, 2009, p. ii). Five research questions guided the study, and data collection and activities—which extended over 5 months—included focus groups, surveys, student data analysis, and group interviews. The final analysis indicated that there was a positive perception of Q-Comp by teachers and administrators as evidenced by the changes in instruction, school culture, communication, administrative practices, and the school improvement process (Hezel Associates, 2009). Additionally, these trends were seen in the ways in which Q-Comp teachers viewed their support systems to improve instruction and grow professionally, especially when compared to teachers who taught in non-Q-Comp schools (Hezel Associates, 2009). Finally, participants did not view money as a motivating factor, and believed that basing the compensation on effort was a fairer practice than basing it on experience (Hezel Associates, 2009).

Limitations of this study included self-reported data, which could skew outcomes to be more positive. Time constraints limited the scope of the research and impacted pre and post testing of surveys, days spent at case study sites, and the number of participants researchers were able to contact. Additionally, some of the participants were involved with Q-Comp for less than 1 year (Hezel Associates, 2009).

Denver's Professional Compensation System for Teachers (ProComp) started as a pilot program in 1999 in 16 schools (Brodsky et al., 2010). Since its inception, the program had union support, but the additional support of the Denver Board of Education in 2004 combined with the local teacher union's backing secured ProComp beyond its

pilot status (Brodsky et al., 2010). The two goals of the program were increased student achievement and attraction and retention of high-quality teachers (Wiley, Spindler, & Subert, 2010). In 2010, researchers from the University of Colorado at Boulder School of Education conducted a formal evaluation of ProComp. Researchers analyzed eight years of student achievement data and then surveyed teachers to look for connections between the achievement data and teacher and administrator responses regarding their attitudes toward ProComp (Wiley et al., 2010). The guiding questions for the evaluation were based on the program's impact; teachers' attitudes, instruction, and practice; and student achievement (Wiley et al., 2010). Researchers used a quantile regression fit growth model to estimate the conditional growth percentiles (CGP), an estimate of a student's percentile achievement among the achievement scores in a reference group. Furthermore, this model was used to estimate teacher effects (Wiley et al., 2010). For this analysis, researchers were limited in their sample (N = 191) because estimates of growth were made in the year prior, so teachers must have been employed in the district in 2007, the year before the study. This sample was surveyed about their beliefs regarding ProComp and the following topics: overall belief about the program; student achievement, traditional salary schedule; and instructional behaviors (Wiley et al., 2010). A Likert scale was used to gauge participants' agreement with the statements.

Study results indicated that student achievement across the district increased 3% to 8% over 8 years. Teachers hired after the ProComp initiation had greater gains in their first year of teaching compared to teachers hired before ProComp was put in place (Wiley et al., 2010). For the survey component, teachers who were involved in ProComp had a more favorable opinion of the program than teachers who were not involved in ProComp.

Additionally, survey results indicated principals had the most favorable outlook of all about ProComp, particularly related to the topic of student achievement (Wiley et al., 2010).

Limitations of this study include the implementation of a statewide student achievement improvement plan called the Denver Plan a few months after the ProComp plan was implemented (Wiley et al., 2010). The ProComp plan, and other state and federal reforms, such as NCLB, were eventually subsumed into the Denver Plan. Researchers, therefore, were not able to directly credit impact on achievement to one plan or the other (Wiley et al., 2010). A second limitation was the impact of the local economy. During implementation of ProComp, Denver saw an increase in population and simultaneous stagnation in employment, impacting the availability of the selection of potential teacher candidates (Wiley et al., 2010). Another limitation was trustworthiness of the data. Researchers indicated that data were accurate, but they were unable to fully verify the connections of students to teachers (Wiley et al., 2010). Finally, student achievement data were from the state mathematics and reading assessments, limiting the information beyond teacher effects for the tested content areas. The effect of the teacher incentive was assessed through a quantile regression model and a VAM. To minimize selection bias, researchers used a score matching strategy based on teacher characteristics, experience, and compensation variables (Wiley et al., 2010).

In an analysis of these programs through the conceptual framework, the Q-Comp program built on the instrumentality and expectancy aspects of motivation theory into the program by utilizing master teachers to help novice teachers. Additionally, the entire faculty and staff were eligible for the incentive, likely increasing the instrumentality

associated with the award. In both the Q-Comp and the ProComp plans, teachers involved in the initiatives had a more favorable outlook compared to teachers not involved in the program, also indicating teachers with knowledge and involvement demonstrating a positive expectancy toward the programs versus teachers not involved in the program. Through the lens of ERG theory, the teachers' needs were met on the physical, social, interpersonal, and creative levels. Conversely, the New York program did not build in valance, as indicated by the teachers in their final evaluation, and as concluded by Fryer (2013) based on the final amount of the incentive once it was distributed. Teachers expressed they were unclear about how to improve student achievement. Additionally, although the teacher's union was brought into the policy portion of creating the program, teachers were not encouraged to work together to increase student achievement. Unlike the Q-Comp program that established teaching partnerships between veteran and new teachers, the New York program was aimed at the teachers as individuals.

As illustrated by the previous examples, teachers' unions were active participants in the creation of teacher incentive programs. Other influences at the federal and state levels have been analyzed in the following section. These are also examples of individual programs generated to improve student achievement. In some cases, student achievement increased, and programs continued (i.e., Minnesota and Denver). In other cases, the programs ceased because student achievement had not been positively impacted (i.e., New York City).

Studies Investigating Economic Incentive Policies

Economic incentive policies for recruitment of teachers in HTS schools have taken many forms. One form was an economic incentive to recruit and retain teachers to work in HTS K-12 schools. A second recruitment tool created a pay for performance reward system for teachers demonstrating student academic growth, and exhibited this through standardized test scores. Miles (2011) suggested that regardless of compensation for teachers, making a long-term commitment to restructuring the current teacher compensation package could save school districts money and increase student achievement. Finding the correct balance when compensating staff for working in a HTS school organization has presented challenges in the school districts where this took place.

Clotfelter et al. (2008) investigated a program in North Carolina. From 2001 to 2004, North Carolina attempted a teacher bonus program to retain teachers with certifications in math, science, and special education in HTS middle and high schools. To be eligible, teachers had to work at the school during the program implementation, and they continued to be eligible as long as they taught math, science, or special education at that school—even if the school itself became ineligible. A school became ineligible by failing to meet one of the following requirements: (a) 80% or more of students qualified for free or reduced lunch, or (b) less than 50% of students meeting grade-level standards on the state standardized test in Algebra I or biology (Clotfelter et al., 2008). Over the course of the 3 years, 65 districts and 146 schools participated in the program and 1,992 teachers received the bonus. This research utilized a difference-in-difference type regression model to attempt an analysis of the relationship of the program's effects on the retention rates of teachers in selected schools. The evaluation of

data revealed that the number of certified teachers in math, science, and special education declined as schools became more disadvantaged and the failure rates of students increased; thus, more failures in those areas indicated fewer certified teachers. Researchers concluded that there were fewer certified teachers in the neediest schools from the start, and that the program failed to impact the recruitment of teachers to these schools (Clotfelter et al., 2008).

The study also revealed problems with the implementation of the program at the district level. Researchers attempted telephone interviews with personnel from 65 districts; personnel from only 49 districts responded. Interviews with district personnel revealed that half of the respondents utilized the program to recruit teachers. Other important findings were that district personnel thought the $1,800 was too small an amount to convince teachers to transfer to eligible schools, and implementation of the program did not occur because of the late dissemination of information. Several others, however, thought the program encouraged the retention of teachers at eligible schools. Ultimately, researchers indicated the program failed to meet the goals of recruiting and retaining science, math, and special education teachers in some of North Carolina's neediest schools. Knowledge about the program was based on principals' experience; 69% of principals with 1 or 2 years' experience were aware of the program compared to principals with 10 or more years of experience, 90% of whom were aware of the program. What the principals learned about the program was insightful for understanding lack of success of the program; 16% of principals stated they did not become aware of the program until after the school was eligible and, therefore, did not inform teachers. Clotfelter et al. (2008) also found that *how* the principals learned about the program

revealed further insights; two-thirds of principals did learn about it through the district, but one-third gained knowledge of the program in other ways (e.g., television news).

The teacher surveys revealed the disjointed nature of the program. For example, 13% of teachers who received bonuses did not have knowledge of their eligibility, and 7% were aware of the timeframe associated with receiving the bonus, calling into question the purpose of teachers receiving an incentive. Researchers also discovered that teachers were not aware of why they were receiving the bonus and half of the new teachers were unaware of the bonus until after they accepted the job, which negated the recruitment purposes of the program (Clotfelter et al., 2008).

Rice et al. (2012) evaluated Prince George's County, Maryland's differential pay scale program called Financial Incentive Rewards for Supervisors and Teachers (FIRST). This initiative was funded by a grant through the TIF. This funding source was important to the program because TIF supported efforts to aid HTS school organizations through a differential pay scale incentive program designed to attract teachers to these types of school organizations. The parameters of TIF included compensation based on student achievement levels, multiple classroom evaluations, and incentives for teachers to assume additional responsibilities and leadership roles (Rice et al., 2012). In addition, the school district assumed more responsibility for program costs as the initiative expanded. Prince George's County paid 75% of the teacher incentive pay supplements by the fifth year with funds other than those provided by TIF.

Prince George's County served 130,000 students as the 18th largest district in the United States. This study obtained partial funds from the school district and conducted studies after the initial year of implementation (Rice et al., 2012). To gain an

understanding about teacher incentive programs, researchers surveyed existing literature and found a gap related to outcomes pertaining to teacher incentive programs:

> The literature on educator incentives demonstrates that programs vary in terms of their overarching goals, their eligibility criteria, the nature of the award, and the measures used to determine the award. While we know about the features of the programs, we know little about their impact in part because the literature based on program effectiveness is limited in scope and plagued by methodological concerns. (Rice et al., 2012, p. 894)

Researchers collected documents, conducted observations, and interviewed 144 individuals who were directly involved in the design and implementation of FIRST or who were close observers of the process. Three case studies were chosen for the study by the school district administrators, based on the following qualities: principal effectiveness, ability to help participants remain anonymous through a high percentage rate of participation, and participants' honesty about their experiences. Using qualitative methodologies, researchers found that the planning and pilot years were critical to achieve successful outcomes. During implementation, organizers were thorough in bringing together stakeholders to garner proper support from the onset. Secondly, to deter criticisms of the program, rules were lenient for participants who fully participated (Rice et al., 2012). These actions impacted outcomes because principals altered evaluation forms, which created a challenge when tracking teacher evaluations at the end of the program. As a result, the committee did not agree about this practice, and those in favor of it highlighted the importance of participants receiving the incentive (Rice et al., 2012). The implications of inconsistent evaluation forms and subsequent tracking of teacher evaluations calls into question the fidelity of the implementation of this program and the validity and reliability of the results.

In analyzing teacher motivation, eligibility for the extra pay was an issue with respondents. Content area teachers were eligible for the incentive; however, counselors, media specialists, instructional specialists, and paraprofessionals were not. In interviews, researchers found this pay structure negatively impacted a school's culture of work from the onset, and eventually "[altered] the view of how the work should be carried out" (Rice et al., 2012, p. 924). In interviews, teachers indicated they were not satisfied with the final payment amount distributed because 40% of the incentive was lost to taxes (Rice et al., 2012). Despite these negative aspects, teachers indicated in survey responses that the incentive did help them put forth "moderate to considerably more effort" (Rice et al., 2012, p.10) into the program.

In addition to analyzing district efforts, researchers also evaluated how teachers are attracted to work in different types of schools. In an investigation to inform policymakers' decision-making that included the recruitment of expert teachers to high-need schools, Amrein-Beardsley (2012) surveyed 207 expert teachers in Arizona. The list of teachers was generated from a list of teachers who were all National Board Certified Teachers (NBCT). Respondents were asked 60 Likert questions and five open-ended questions. The participation rate was 89 teachers, or 40.30%; therefore, the researcher conducted an intergroup, comparative analysis using information from personnel records shared from the Arizona Department of Education, as well as additional information collected from the Department's website about school-level data. This included: "school demographics, indicators of school poverty, and standardized test scores and proficiency level" (Amrein-Beardsley, 2012, p. 6). Information was coded through "calculated factor means, standard deviations, significance levels ($\alpha = 0.05$), and

effect sizes [to] verify if the sample of respondents differed from the sample that declined to participate in the study" (Amrein-Beardsley, 2012, p. 12). Pearson chi-square and ANOVA statistics suggested that the respondents were not statistically significantly different from those who declined to participate based on the following characteristics: expert teacher distinction, gender, schools by state achievement profile, schools by percent of students meeting or exceeding state standards on grade 3, 5, 8, and 10 state tests, schools by percent of colleagues by years of teaching experience, schools by percent of students by race (Amrein-Beardsley, 2012).

The survey instrument focused on the following eight factors deemed important in the formation of policymakers' decisions regarding recruitment of expert teachers into the neediest schools:

- salary;
- community;
- school;
- administration;
- teacher professionalism;
- student achievement; and
- caring (Amrein-Beardsley, 2012).

In addition, open-ended items were "read, coded, re-read, and categorized" (Amrein-Beardsley, 2012, p. 12).

Survey results indicated that two important characteristics of recruiting expert teachers identified by respondents were principal quality and salary. This study limited generalizability because the survey included only expert teachers from Arizona. Additionally, this study posed hypothetical outcomes; it was unknown if these teachers would consider employment in a HTS school if policymakers implemented their suggestions. The researcher admitted these hypotheticals as limitations.

In a study conducted to inform school leaders about retaining quality teachers in HTS schools, Greenlee and Brown (2009) used a convenience sample of 97 teachers enrolled in the educational leadership program at the University of South Florida. Data were gathered through a "Likert-type scale, ranking, and open-ended survey items" (Greenlee & Brown, 2009, p. 100) and reported in percentages based on the frequency of responses. This salary component took several forms: bonus, tuition reimbursement, performance pay, or enhanced retirement benefits (Greenlee & Brown, 2009). Other important factors were autonomy and resources for curriculum enhancement (22%), professional development (12%), and participation in school-based decision-making (9%). The primary limitation of this study was that due to insufficient information regarding methodology, the validity of the study could not be assessed.

Martin (2010) used the SASS data from NCES for 3,500 school districts, including salary compensation information, to conduct a logistic regression for the analysis of relationships between school district characteristics and how likely a district would be to offer incentive pay. Martin (2010) grouped incentives into the following dependent variable categories: shortage (i.e., districts that offered incentives for teaching in a shortage content area or HTS location); excellence (i.e., districts that offered incentives for teachers with National Board Certification or some type of award in teaching); and professional development (i.e., districts that offered incentives for extra time spent attending specific meetings or conferences). The independent variables were based on the following student characteristics: free/reduced lunch; race; crime rate in district community; urban/suburban; proportion of private, religious, and charter schools within the district; number of districts in the county; union membership. Researchers

found that the most common form of incentive was merit pay, but research findings clashed on the effectiveness of this practice. Districts were more likely to offer incentives if the student population was primarily Black or Hispanic, if it encountered racial or socioeconomic segregation, if it possessed larger student populations, or if it had higher crime rates (Martin, 2010). Researchers concluded that pressure was on districts to increase student achievement, and results of the study could help districts modify their salary schedule to offer more incentives to teachers. Although the study took place after the 2008 recession, researchers had not offered an explanation that focused on how the recession and the economic implications influenced the study. This study had the limitation of the data set, which was 10 years old.

In another quantitative study of 189 high school teachers in one southeastern state, Petty et al. (2012) researched the reasons that attracted and detracted teachers from teaching in high-need high schools. Researchers adapted a survey used with elementary school teachers and coded the survey using thematic analysis. The survey contained two parts; the first part focused on demographics and the second part focused on the nine open-ended questions regarding attraction and retention of administrators in high-needs high schools. These nine open-ended questions focused on how teachers perceived their work environment regarding the strategies needed to attract successful teachers to HTS high schools to raise student achievement. Next, researchers coded the results of the open-ended items through thematic analysis. To analyze item responses, they also reported frequencies. Researchers found that having a caring nature was cited by 28% of teachers as being an important characteristic. In addition, classroom management and pedagogical flexibility were named as important characteristics. Teachers indicated that

preservice teachers needed more experience in HTS schools. To attract teachers to these schools, monetary incentives were an important factor, as was administrative support. To retain teachers, money, respect and recognition for student achievement, and teaching resources were notable characteristics. Researchers indicated that this study was limited due to response rate: 35%. Additionally, the study also had self-reported data, which the researchers concede can impact reliability, but the responses provide unique perspectives from teachers (Petty, et al., 2012).

The Motivation Theory Connection

Through the lens of motivation theory, Vroom (1964) theorizes that an individual puts forth effort based on expectancy (effort will increase performance), instrumentality (achievement will result in attainment of the reward), and valence (the reward is valuable). Alderfer's theory is grounded in existence (basic needs), relatedness (connectedness to others), and growth (desire to create meaningful tasks). The studies reviewed fall into two categories—the theory of pay for performance and the practicality of it. The theory of incentive pay is enticing, as analyzed in the studies by Amrein-Beardsley (2012), Greenlee and Brown (2009), and Petty et al. (2012). These studies lack information regarding the practicality. For example, there is no information about teachers beliefs their effort increases student performance (expectancy), or that their achievement will result in attainment of the reward (instrumentality). Additionally, the studies are not grounded in teachers' basic needs (existence), the opportunity to connect with others to reach the goal (relatedness). Teachers in the studies did indicate they foresee a desire to engage in such a program (growth). Conversely, the studies that focused on pay for performance programs addressed the practicality of such programs.

For example, in the study by Clotfelter et al. (2008), the program failed to recruit the teachers it was seeking, suggesting that the value of the reward (valence) was missing. Additionally, teachers were unaware of the program, so motivating factors could not be associated with the reward system. According to Rice et al., (2012) teachers were critical that other staff members could not receive a reward, so their relatedness needs were not met, negating the opportunity for growth. Furthermore, teachers did not view the program as worthwhile because they lost 40% of the incentive to taxes, negating the valence of the reward.

Virginia Performance Pay Incentives

In July 2011, the VDOE introduced the VPPI, a pilot program of the new teacher evaluation model that utilized monetary incentives for increased student academic achievement. In the press release from the VDOE, Governor McDonnell stated:

> There is a growing consensus that public schools must move beyond a compensation model that equally rewards mediocrity and excellence and is rooted in a past when our young people did not face fierce global competition. Bipartisan support – at both the state and federal levels – is allowing Virginia to implement performance pay in some of the Commonwealth's most hard-to-staff rural and urban schools. I look forward to this program getting underway. It will benefit our students, parents and teachers. I congratulate the participating superintendents and school boards for their courage in taking advantage of this opportunity to encourage and reward excellence. The performance-pay pilots will provide a foundation for compensation reform that will reward excellence and attract new talent to the Commonwealth's classrooms. (Commonwealth of Virginia, 2011c, para. 3)

The program included 13 (10%) of the Commonwealth's districts (VDOE, 2011c). This encompasses the qualified districts that applied to participate in the program. To understand the VPPI, an overview of the Virginia school finance structure is presented. This sets the context of Virginia's education system and the ways students have been educated.

According to Salmon (2010), school finance in Virginia had evolved since the inception of the Literacy Fund in 1810. The purpose of the fund was to give poor White children access to schools. Salmon (2010) explained that the Literacy Fund changed in 1918 to a school capital construction fund that provided only one-third of the necessary fiscal resources to local education agencies. From 1930 through the 1950s, localities varied in their abilities to fund public schools. To assist localities, a general sales tax initiative was established in 1966. This gave one cent of every dollar to public schools, and provided localities with an additional one cent for public schools, which all localities accepted (Salmon, 2010). Racial problems and funding issues prompted a 1968 lawsuit that charged Virginia with not doing enough to address disparities between local school districts (Salmon, 2010). The federal district court would not rule that the funding system was unconstitutional (Salmon, 2010). Roch and Howard (2008) indicated that individuals in wealthier areas influenced education finance reform due to their political influence in the state. Despite the negative reaction from the court, the new Virginia Constitution of 1971 specified the responsibilities of the public school system through the Standards of Quality (SOQ).

The SOQ is the minimum educational program a school must provide (Dickey, 2013, slide 1). The locality has been obligated to match the state funds, but may spend more, offer more programs, and employ more staff than the state requires (Dickey, 2013, slide 1). The components that encompass the SOQ have been the salary and benefits, support positions, and support costs (e.g., supplies and utilities). State and local expenditures have influenced cost funding and "most SOQ funding is equalized based on the local ability to pay as determined by the Composite Index of Local Ability-to-Pay"

(Dickey, 2013, slide 20). Each locality's obligation was calculated individually. SOQ funding was calculated yearly through a school district's average daily membership (ADM) as of March 31 of every fiscal year, multiplied by per-pupil expenditure, resulting in program cost. State and local financial obligations were determined through the local composite index (LCI), the index of local taxpaying capacity (Driscoll & Salmon, 2008). LCI tax was accounted for using three measures:

> True valuation of real and public service property, Virginia adjusted gross income, and taxable retail sales and use receipts . . . It yields two shares—one state and one local . . . The local share becomes the local required expenditure . . . the state pays the remaining SOQ costs. (Driscoll & Salmon, 2008, p. 41)

The 2008 recession heightened concerns for researchers who analyzed public school funding in Virginia. As Joyce and Pattison (2010) stated, financial predictions for school funding for the next 10 years were made in prior years, and those decisions impacted states and localities. Disruptions in the current local tax base had implications for future expenditures. According to Driscoll and Salmon (2013), Virginia cut $81.9 million of direct funding to public schools from fiscal year (FY) 2007 to FY 2011. In addition to losing state funding, assessed property values dropped, reducing the funds available for localities to spend on public schools. This trend occurred nationally: state revenues increased by 0.2% during 2010-2011, but states decreased per-pupil spending by 1.6% (NCES, 2016). Additionally, localities carried the burden of meeting the financial obligations set forth by educational standards and mandates. In Virginia, 53.1% of public school spending was through local contributions, 37% through state contributions, and 9.9% through federal contributions (NCES, 2016).

This lack of consistent financial support is concerning because Virginia has been facing changes in its population, which researchers predict will have impacts on the

public education system (Driscoll & Salmon, 2013). Population growth in Virginia has been most concentrated in Northern Virginia, Richmond, and Hampton Roads, together representing 82 % of the state's total population growth (Driscoll & Salmon, 2013). Additionally, population diversity has increased. For the 2011-2012 school year, nonwhite students under the age of 6 accounted for 45.9% of the children in Virginia, and 16 county localities and 18 city localities (i.e., 42% of the total student population) reported a minority-majority student population (Driscoll & Salmon, 2013). This raised concerns among researchers concerned with the poverty levels associated with this change in population because in 2010, 17.2% of Virginia residents were 150% below the federal poverty level (Driscoll & Salmon, 2013). Of these residents, 30% were non-Hispanic Blacks and 25% were of Hispanic origin; and the child poverty rate was 14.6% (Driscoll & Salmon, 2013). More recent data from the Commonwealth of Virginia's performance and accountability group, Virginia Performs, (2016) reported that the poverty rate for Virginians under the age of 18 was 15.9 %. This information is important to have because these students will need educational services from an expert teaching force equipped to educate a diverse student population.

In Virginia, overall funding structures have been traditional, and neither a statewide differential pay scale nor merit pay scale had been implemented prior to the VPPI. Instead, localities were awarded TIF grants. The following districts received those grants: Henrico County Public Schools, Prince William County Public Schools, and Richmond Public Schools.

The Henrico County program was open from 2010 to 2015, involving 323 teachers and principals in eight schools (VDOE, 2011d). Teachers who taught a course

associated with an SOL test or Phonological Awareness Literacy Screening [PALS] assessment were eligible to receive up to $8,000 a year—$3,000 for the classroom learning environment based on the district's teaching standards, and $5,000 for demonstrated student achievement on local and state assessments (VDOE, 2011d). Principals were eligible to receive up to $10,000 a year—$4,000 for supporting teachers' creation of the classroom environment and $6,000 for demonstrated student achievement on state and local assessments (VDOE, 2011d).

The Prince William County Public Schools [PWCS] TIF program was in effect from 2010 to 2015. According to the VDOE (2011d), the program was open to teachers and principals and was designed to:

"Integrate a performance-based compensation system with the new PWCS standards-based evaluation system, Professional Performance Process; provide professional development that focuses on individualized school and educator improvement; and build on growth targets set in the Division's recently adopted Strategic Plan for 2010–11 through 2014–15" (p. 54).

The PWCS system was open to thirty schools identified as having an economically disadvantaged student population. The award system was based on a locally developed performance system made up of six performance standards and demonstrated student academic growth as measured through the SOL test (ED, 2010b). The six standards were: knowledge of students; knowledge of content; planning, delivery, and assessment of instruction; safe effective learning environment; communication and collaboration; professionalism. Calculation of the VAM of academic growth on the SOL test was conducted by taking SOL data which were "converted to normalized scores to show growth from one test to the other" (ED, 2010b, p. 7). To be eligible for the award, schools had to first make the AYP benchmark and achieve full accreditation status with

the VDOE. Two tiers of rewards were granted. The first tier was for principals and teachers of SOL subjects, or grade levels for elementary teachers, as well as special education and Limited English Proficiency (LEP) teachers. All other school staff received a Tier II reward, which was $1,000 (Community Training and Assistance Center, 2016). Tier I teachers received $3,216 and principal awards varied based on school level (elementary, middle, or high). A summary of the rewards if found in Table 4.

Table 4

TIPA Awards

Staff Position	Tier	Amount of Award
SOL Teachers	Tier I	$3,216
Non-SOL Teachers	Tier II	$1,000
Elementary School Principals	Tier I	$6,026
Middle School Principals	Tier I	$6575
High School Principals	Tier I	$6,685
Elementary School Assistant Principals	Tier I	$4585
Middle School Assistant Principals	Tier I	$4747
High School Assistant Principals	Tier I	$5380
Other Administrators	Tier II	$1,000

Note. Adapted from "Getting the Pieces Right: Professional Development, Compensation, and School-Wide Performance. Teacher Incentive Performance Award Prince William County Public Schools" by Community Training and Assistance Center, 2016, p. 15.

The Richmond City Public Schools TIF program focused on the NBPTS (VDOE, 2011d). The program was "designed to help school districts recruit and retain National Board- certified teachers, improve professional development opportunities, expand

mentor training, align teacher and principal evaluations with national student achievement standards, and reward teachers and principals whose students meet targeted standards" (VDOE, 2011d, p. 54). The program included performance based awards, with two levels of student growth measures. Teachers teaching subjects tested through the SOL test were evaluated on SOL outcomes with a value added measure, and teachers outside of SOL tested areas were evaluated on student knowledge based on pre- and post-tests as well as learning objectives aligned with local and state curriculum goals (VDOE, 2011d). The award structures were: $500 per academic year for mentoring new teachers; $1500 for meeting schoolwide goals; $500 per candidate for being a Candidate Support Provider; $1500 for site-based coordinators (ED, 2017b).

A pay for performance initiative in the Commonwealth of Virginia was instituted on January 18, 2011. Then-Governor Robert McDonnell included the VPPI through the second phase of the Commonwealth's 2010-2012 biennial appropriations process. The VPPI was a pilot of a new teacher evaluation system that tied 40% of a teacher's evaluation to student academic growth, and gave teachers a monetary incentive for improved student achievement (RMC Research Corporation, 2012).

The targets of the grants were HTS schools, and participation was by invitation through a grant application process. Schools were identified by data from the 2010-2011 school year. Initially, the grant was open to 169 HTS schools (VDOE, 2011e). Not enough HTS schools applied, so SIG schools were mandated to participate (RMC Research Corporation, 2012). SIG schools received money from the federally funded TIF program, a component of which included differential compensation. In the end, the VPPI included nine HTS schools and 16 SIG schools, for a total of 25 schools across 13

school districts. The schools were heterogeneous in location and characteristics, including each of the eight Superintendent's Regions that encompassed rural and urban schools across elementary, middle, and high school levels.

To qualify for the incentive, participating teachers needed an exemplary rating on their teacher evaluations. The HTS and SIG award amounts differed. Teachers in HTS schools could earn up to $5,000, and teachers in SIG schools could earn up to $3,000. Funding structures were different because the grants allocated different amounts of money. A total of $3 million from the general fund was used for teachers in HTS schools, and matching federal funds from SIG monies paid teachers in SIG schools (RMC Research Corporation, 2012). Additionally, flexibility was built into legislation, allowing school districts to provide varying incentives to different teaching groups (e.g., teachers in an SOL-tested subject area, support area teachers, or teachers of record; RMC Research Corporation, 2012). The VDOE collaborated with Dr. Stronge and a team from The College of William and Mary to conduct training for teachers and principals in the VPPI participating schools about the implementation of the new teacher evaluation system. The RMC Research Corporation directed the final program evaluation.

The final evaluation of the VPPI by the RMC Research Corporation was a mixed-methods study consisting of surveys, document reviews, principal interviews, communication with staff from the VDOE and The College of William and Mary, and an analysis of internal consistency of principal ratings. In addition, a logic model was used to "clarify the relationships between the program components and clearly define expected outcomes" (RMC Research Corporation, 2012, p. 15). The evaluators stated that the program evaluation was intended not to judge the outcomes, but to provide information

about the outcomes related to the new teacher evaluation process and the teacher incentive (RMC Research Corporation, 2012). This information was communicated; however, information regarding student achievement, the foundation for earning the additional compensation, was not reported. Other information in the evaluation, such as interview data, was limited. To help participants maintain their confidentiality, principals and teachers who were interviewed were not identified. It was unclear if interviewed teachers received an incentive. Furthermore, a discrepancy existed between fall teacher survey participation and spring teacher survey participation, making it unclear if respondents were represented in one or both surveys.

This was a pilot program of new teacher evaluation criteria. Therefore, to set the context for the VPPI and how the teachers were evaluated during the pilot, an explanation of the new standards is included in Table 5.

Table 5

Virginia Performance Standards

Standard	Description
Performance Standard 1: Professional Knowledge	The teacher demonstrates an understanding of the curriculum, subject content, and the developmental needs of students by providing relevant learning experiences.
Performance Standard 2: Instructional Planning	The teacher plans using the *Virginia Standards of Learning,* the school's curriculum, effective strategies, resources, and data to meet the needs of all students.
Performance Standard 3: Instructional Delivery	The teacher effectively engages students in learning by using a variety of instructional strategies to meet individual learning needs.
Performance Standard 4: Assessment of and for Student Learning	The teacher systematically gathers, analyzes, and uses all relevant data to measure student academic progress, guide instructional content and delivery methods, and provide timely feedback to both students and parents throughout the school year.
Performance Standard 5: Learning Environment	The teacher uses resources, routines, and procedures to provide a respectful, positive, safe, student-centered environment that is conducive to learning.
Performance Standard 6: Professionalism	The teacher maintains a commitment to professional ethics, communicates effectively and takes responsibility for and participates in professional growth that results in enhanced student learning.
Performance Standard 7: Student Academic Progress	The work of the teacher results in acceptable, measurable, and appropriate student academic progress.

Note. Adapted from "Evaluation Report: Governor's Virginia Performance Pay Incentives Initiative Pilot," by RMC Research, Inc., 2012, p. 84.

The student academic progress standard (Standard 7) was new to the teacher evaluation system. To receive the incentive, teachers had to receive an Exemplary (4) on the final evaluation. It was unclear how many times teacher observations occurred. To ensure validity of the summative evaluation, each principal submitted a data spreadsheet with all participating teachers' evaluation information to the RMC Corporation. The RMC Corporation conducted a review of the performance standards ratings and summative decisions for two teachers from each school. A lead evaluator from RMC Corporation used a document review template to "categorize and triangulate" the information and "cross-referenced with documented evidence across the 16 schools" (RMC Corporation, 2012, p. 26). The findings from this review indicated that 13 schools had strong evidence to support the ratings, 1 school had adequate evidence, and 2 schools had unacceptable evidence (RMC Corporation, 2012). The schools with unacceptable evidence did not submit evidence, which does not indicate that they did not have evidence, but more of an indication that it was not available to evaluators. There is no evidence that the evaluator had access to the data. The RMC Corporation (2012) indicated that personal judgment dictated "whether the evidence supported the ratings of teachers on each of the Standards" (RMC Corporation, 2012, p. 67). It is unknown if teachers met benchmarks, but not given credit because of lack of follow through from their administrators.

Overall results were that teachers and principals highly rated the training provided by The College of William and Mary. Teachers indicated more concrete examples of expectations would have helped because the intent of the teacher evaluation system

design was to be a collaborative process between the principal and teacher. Teachers also indicated they needed additional support from administrators "to answer teacher questions and address anxieties and concerns about the process; clearer guidelines on working with students with disabilities and English Language Learners in terms of goals, standards" (RMC Research Corporation, 2012, p. 41). Teachers also sought additional training and support in data collection and documentation and in writing Specific Measurable Attainable Realistic and Timely (SMART) goals with examples. In the fall teacher survey (N = 388), teachers made similar comments. Additionally, in the fall, the RMC Research Corporation (2012) asked teachers and principals the question, "What is your expectation for the future success of performance-pay improving teaching and learning at your school?" (p. 49). Out of 189 teachers, 52% indicated they had low expectations and out of 23 principals, 39.1% indicated they had moderate expectations of improving teaching and learning at the school (RMC Corporation, 2012). Consequently, The RMC Research Corporation (2012) stated that more quality measures needed to be in place to ensure that teachers had support during program implementation. Suggestions included deeper document reviews and understanding of the connections between evaluations and performance pay. Other than the information about an Exemplary rating, there was no indication in the final evaluation about how teachers received assistance with Standard 7, which addressed student academic growth. Although the incentive was based on student achievement through the SOL scores, there was no information about what student achievement growth was made, leaving questions about the success of a program based on student achievement.

At the end of the VPPI, 782 teachers had participated in the pilot; 340 teachers were from HTS schools and 442 teachers were from SIG schools. Awards for teachers in the HTS schools ranged from $600 to $5,000, and teachers in SIG schools earned an award of $3,000. Performance awards were given to 225 teachers, or 28.77% of participants. HTS teachers made up 125 of the pay recipients, and 100 were from SIG schools. Although student academic growth was the basis of the incentive, this information was not available in the final evaluation, making it unclear about the impact on student achievement. A weakness of this study was that the evaluators missed a unique opportunity to address the connection between teacher incentives and student achievement. An emphasis was on the process of the evaluation, not student achievement outcomes. This gap was the basis for the study.

Motivational Challenges of Implementing Incentive Policies

The concept of teacher motivation has raised questions about the implementation of pay for performance programs. As reviewed in Chapter 1, Vroom's theory of human motivation (1964) indicates that people seek out jobs that will give them positive outcomes. Alderfer's theory of motivation (1969) focuses on three principles: existence, relatedness, and growth (ERG). An extension of Maslow's hierarchy of needs, this theory provides that people have physical needs, social needs, interpersonal needs, and creative needs, but individuals can regress between levels to have their needs met (Alderfer, 1969). Both theories can help explain why people are satisfied with one job and not another.

In analyzing teacher motivation with respect to pay for performance programs, the amount of the award needs to have meaning for the teachers so they will want to

participate in the program; the size of the incentive must have a value (valence) and teachers must believe their personal efforts will help them achieve the academic benchmarks (expectancy). The goals and objectives of the program should be transparent so that teachers understand the expectations of them, and the teachers need to understand the connections between the components of the program and their jobs (growth) so they will achieve the goals (Alderfer, 1969; Kelley, Heneman, & Milanowski, 2002; Rice et al., 2014; Vroom, 1964).

Yuan et al. (2013) analyzed three pay for performance programs, located in Nashville, Tennessee; Round Rock, Texas; and New York City, New York. Researchers analyzed online surveys teachers completed during the years they participated in the pay for performance programs. They then evaluated the survey items, calculated descriptive statistics for each item-level response, and conducted an exploratory factor analysis (Yuan et al., 2013). This evaluation helped researchers create "10 scales measuring key constructs related to teachers' motivation and practices" (Yuan et al., 2013, p. 10). Researchers examined the effects of the programs on teachers' motivation to improve student achievement, instruction, relationships with colleagues, and stress by examining "three randomized controlled trials . . . in which the units of accountability were randomly assigned to either the treatment group (who were eligible to receive the bonus) or the control group (who were not eligible to receive the bonus)" (Yuan et al., 2013, p. 8).

Outcomes of the study were that the expectancy was low across all three programs. First, 50% of the teachers in the treatment group did not believe they could earn the bonus due to outside influences on a student's life, such as family characteristics

(Yuan et al., 2013). For the valence aspect of the study, 77% of teachers from one treatment group and 64% of teachers from another indicated they were interested in the bonus, but only 40% and 20% expressed that the bonus motivated them to improve their teaching. Additionally, 42% and 50% of treatment group teachers revealed that the bonus was not large enough to motivate them to expend the extra effort. In terms of instruction, treatment and control groups did not differ except in test preparation, where treatment group teachers admitted there was more of an emphasis on test preparation (Yuan et al., 2013), but the program did not change their teaching. For both treatment and control groups, teachers worked longer than their contracted work (Yuan et al., 2013). Treatment and control groups did not report significant differences in their collegial relationships, except for one treatment group that expressed more collaboration over a control group.

This study had several limitations. First, teacher data were self-reported, which can indicate bias (Yuan et al., 2013). Second, teacher attrition for the program in Nashville was high, and the response rate for the New York City program was low (Yuan et al., 2013), both limiting information from the full population of teachers who participated in the programs. Finally, it was not possible for researchers to answer clarifying questions or follow up with participants.

In a study of Chicago Public Schools, Finnigan and Gross (2007) examined survey data from a 1997 administration of the Consortium for Chicago School Research's (CSSR) survey of elementary school teachers. Utilizing the lenses of Vroom's expectancy theory of motivation (1964) and Lawler's theory of motivation (1973),

researchers analyzed the impact of the Chicago Public Schools' 1995 School Reform Act on teacher motivation, to find out if accountability policies influence teacher motivation.

Through interviews and focus groups, researchers found that the status of the school influenced the valence and expectancy teachers experienced. Additionally, Finnigan and Gross (2007) found that, as teachers experienced positive gains in student achievement, their motivation increased to "confidence and encouragement" (p. 604). As in the previous study, teachers whose schools did not move off probation had weaker motivation; researchers reported this decreased morale as an unintended consequence of the accountability policy. Despite the weakening of motivation, when schools were on probation, teachers did try new instructional practices and expressed that they increased their efforts. The incentives also influenced teachers in terms of going back to "the missions of serving students" (Finnigan & Gross, 2007, p. 614).

Job status was a concern for two schools on probation. One school was in probation status while high schools in the district were reconstituted, and during this time teacher turnover was elevated due to principal actions. The second school had declining enrollments and changes in the community such as neighborhood gentrification, declining student enrollment, and leadership issues (Finnigan & Gross, 2007). In evaluating the influence of accountability policies on teacher motivation in these two schools, Finnigan and Gross (2007) state that this takes advantage of teachers' intentions for entering education, and that:

> Teachers appear to be particularly sensitive to their status within a professional community and the threat to that status implied probation identification . . . The policy appears to effectively exploit teachers' desires to see their students succeed . . . The material incentives, namely concern for job loss, affect teachers. Thus, teachers' internal motives, stimulated by the policy mechanisms, subsequently

caused teachers to refocus and, in some cases, redefine their goals. (p. 616)

The study was limited because teachers self reported about their instructional practice and the changes were not observed. Additionally, this data set also includes only the perspectives of elementary school teachers, leaving out the middle school and high school teacher perspectives.

Through focus groups and teacher interviews, researchers found three major factors related to teacher expectancy. First, teachers who were new to a school had a higher expectancy than teachers who were veterans at the school. Second, teachers with higher levels of education, at both probation and nonprobation schools, had higher levels of expectancy. Finnigan and Gross (2007) associated advanced education with experiences in high expectancy. Additionally, researchers emphasized that school performance may be improved if school performance accountability policies were aimed at teachers who lacked essential knowledge and skills (Finnegan & Gross, 2007). Third, higher levels of expectancy were related to teachers' race. African American teachers had higher expectancy levels than teachers of other races; this was attributed to the higher population of African American students in the Chicago Public Schools (Finnigan & Gross, 2007). A limitation of this study is that these data are limited to the interviewed teachers in Chicago Public Schools. Additionally, researchers did not test differences.

The study is limited because the survey data were collected after schools were designated as being probation or nonprobation status, so teacher morale based on that information could have affected the survey responses. Finally, the data are only reflective of the elementary level, leaving out unique perspectives of middle school and high school teachers.

Conclusion

Legislative agendas have influenced and continue to influence education policy in the United States. This impact is demonstrated through the actions resulting from the report, *A Nation at Risk: The Imperative for Educational Reform* (1983), the report that established the framework for comparisons of how well students were excelling academically in relation to their peers outside of the United States. Additionally, this report prompted legislative action through the allocation of federal and state resources toward ways to improve student achievement. Some allocations were in the form of teacher incentive programs, reviewed in this chapter.

One influence within teacher pay for performance programs is motivation. As reviewed in Chapter 1, individuals are motivated to work when the work creates meaning, fits with their personality and expectations, and creates growth opportunities (Alderfer, 1969; Vroom, 1964). Chapter 2 presented studies that evaluated teacher motivation within pay for performance programs (Finnegan & Gross, 2007; Rice et al., 2014; Yuan et al., 2013). These analyses suggested that teachers may not be motivated by pay for performance programs due to misunderstandings of program structures (Yuan et al., 2013), and may put forth more effort because of the promise of the incentive (Rice et al., 2014). Issues that can impact teacher motivation are the influence of outside elements, such as a student's familial life (Yuan et al., 2013) and fairness, because the support of a student extends beyond one teacher's classroom (Rice et al., 2014). Furthermore, the final compensation amount does need to be substantial enough after taxes to be satisfactory to teachers (Rice et al., 2014). Additionally, teachers with higher levels of education have higher expectancy levels (Finnegan & Gross, 2007). Moreover,

a school's accreditation status can impact a teacher's level of expectancy and motivation (Finnegan & Gross, 2007).

The impetus to create pay for performance programs is based on challenges of teaching in HTS schools, including high poverty levels of the student population (Greenlee & Brown, 2009). Research exists to support the premise that salary can be an enticement to keep teachers in those schools (Greenlee & Brown, 2009). The promise of attracting and retaining quality teachers has prompted the creation of pay for performance programs. The ideal, sustainable program is yet to come to fruition, although attempts to craft the perfect model have met with mixed outcomes (Clotfelter et al., 2008; Fryer, 2013; Marsh & McCaffrey, 2012; RMC Research Corporation, 2012; Rice et al., 2012). Complexities of the models challenge education policy makers, and there is no one consensus about how much of an incentive is enough, nor is there information about sustainability of incentive. Additionally, who gets the incentive varies as well, and has included teacher of record, content area teachers, teaching teams, administrators, and entire faculty and staff (Clotfelter et al., 2008; Fryer, 2013; Marsh & McCaffrey, 2012; RMC Research Corporation, 2012; Rice et al., 2012).

In the case of the VPPI, the focus of this study, changes in student achievement for the schools that were invited and the schools that participated in the program were not clear in the final evaluation by the RMC Corporation (2012). Thus, a gap emerged when analyzing student achievement on the SOL tests for grades 5 and 8, and the EOC exams. This study, therefore, evaluated gain score pass rates for the entire population of schools invited to participate in the VPPI, to analyze differences between participating and nonparticipating schools.

Chapter 3: Methods

The overall goal of schooling is to help students graduate on time, prepared for college or the workforce, but districts with schools designated as hard to staff (HTS) have difficulty keeping qualified teachers in the schools (Goldhaber et al., 2010; Petty et al., 2012). This results in school districts continuously utilizing resources to replace teachers who leave for schools with less challenging populations or for better salaries and is especially difficult in HTS schools because of "potentially unattractive job characteristics associated with poverty, low student achievement, and racial or ethnic differences" (Kelley et al., 2009, p. 2).

One solution has been to offer some type of monetary compensation to teachers, as reviewed through the literature in Chapter 2. A popular solution has been connecting additional compensation and student achievement (Dickey, 2013; Marsh, & McCaffrey, 2012; Martin, 2010; Miles, 2011; Petty et al., 2012). The monetary incentive has been a primary source of interest because, as stated by Greenlee and Brown (2009), "high poverty schools provide a unique set of challenges" (p. 106). The United States Department of Education (ED) addressed this education policy initiative in 2010 when 62 school districts in 27 states were awarded $442 million for the Teacher Incentive Fund (TIF). This program offered a monetary incentive to attract principals and teachers to HTS schools. Many of these awards have been linked to achievement tests to demonstrate success, prompting researchers to suggest that "the federal government is reshaping local education policy to a greater extent than ever before" (Gratz, 2011, p. 157).

This exploratory study used an ex-post facto research design with data from the entire population of schools invited to participate in the Virginia Performance Pay Incentives Initiative Pilot (VPPI). The objective of this study was to analyze gain scores based on Standards of Learning (SOL) test score pass rates in grades 5 and 8, and on end-of-course high school (EOC) tests to examine if a difference existed between scores of participating and nonparticipating schools. The data included the pass rates in school years 2010-2011 and 2011-2012 (i.e., the year before and the year of the program implementation). The final evaluation of the pilot program, performed by the RMC Research Corporation (2012), did not include information about differences between student test score pass rates, the content areas evaluated, differentiation between participating and nonparticipating teachers, or comparisons of the test score pass rates to those of schools that were invited but did not participate. By these omissions, the evaluator missed a unique opportunity to quantitatively measure those test score differences. According to the RMC Research Corporation (2012), the evaluation was a mixed methods analysis that included interviews, surveys, document reviews, and statistical analysis. Although school districts were asked to submit individual teacher test scores for those who participated in the program, neither that information nor the pass rates were revealed in the final publication. In the final evaluation, the RMC Research Corporation (2012) indicated that the focus of the evaluation was to help provide the Virginia Department of Education (VDOE) with information regarding the VPPI, not to give a definitive answer about the success of the program related to student achievement.

During further investigation into the connection between the program and student achievement, a phone conversation with the Assistant Superintendent for Teacher

Education and Licensure at the VDOE revealed that the names of teachers were unknown to the VDOE because money was distributed in lump sums to school districts and then distributed to teachers at the school level (personal communication, April 2014). Further research revealed a 2013 presentation to the Virginia House Appropriations Committee about the school districts, number of teachers, and grant amounts awarded. That information is available in Table 1.

Although the information about the number of teachers in each district and the amounts awarded was important, the omission of teachers' names limited the opportunity to analyze data at the teacher level, a limitation of the study addressed later in this chapter. Schools in Virginia are accredited based on pass rates on the SOL test for English, mathematics, history, and science (Virginia Department of Education, 2017) and leaving these data out of the evaluation report created a gap in knowledge about the effectiveness of the incentive on improving SOL scores.

This chapter provides the methodology used in this study. It includes the research questions, design of the study, participant selection, and description of the population, validity, reliability, data handling, data analysis, and ethics.

Research Questions

This was an exploratory study with an overarching purpose of examining whether the VPPI had a positive impact on student academic achievement. All schools invited to participate in the VPPI were included in the study. The objectives of the study were pursued by analyzing the difference in gain scores between participating and nonparticipating schools as measured by the Virginia SOL tests in grades 5 and 8, and on

EOC. The years included were 2010-2011 and 2011-2012 (i.e., the year before and the year of the VPPI).

Research Questions:

1. What is the difference in grade 5 Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the incentive?
2. What is the difference in grade 8 Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the incentive?
3. What is the difference in end-of-course high school Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the incentive?

Research Design

This was an exploratory study about the impact of monetary incentives on student academic achievement as measured through standardized test scores. The VPPI took place in the 2011-2012 school year. A total of 169 hard to staff (HTS) schools were invited to participate and apply through a grant application process. A low number of HTS schools applied, so federally funded School Improvement Grant (SIG) schools—schools that received Title I funding from the federal government—were mandated to participate. In the end, 25 schools representing 13 school districts identified as HTS or

SIG schools participated in the program. This group made up 11 elementary schools, 5 middle schools, and 9 high schools. Schools needed to meet four out of eight of the following criteria to participate in the VPPI:

- Accredited with warning
- Average daily attendance rate two percentage points below the state average
- Percent of special education students 150 percent of statewide average
- Percent of limited English proficient (LEP) students exceeds 150 percent of the statewide average
- Percent of teachers with provisional licenses exceeded 150 percent of the statewide average
- Percent of special education teachers with provisional special education licenses exceeded 150 percent of the statewide average
- Percent of inexperienced teachers (0 years of teaching experience) hired compared to the total number of teachers exceeded 150 percent of the statewide average
- The school had one or more inexperienced teachers (0 years of teaching experience) in a critical shortage area. (RMC Corporation, 2012, p. 10)

Information about why other identified HTS schools chose not to apply for the grant was not available. For the purposes of this study, the entire population was included, so parameters were reported. The pass rates for SOL test scores in all core content areas (math, reading, writing, science, and history) for grades 5, grade 8, and EOC at both participating and nonparticipating schools were analyzed to determine if

differences in student achievement existed between participating and nonparticipating schools.

This ex-post facto quantitative study focused on outcomes in the form of gain scores based on pass rates on the Virginia SOL tests. The study used existing data from the VDOE's School Report Card page (VDOE, 2016) to analyze the entire population of schools (HTS and SIG) invited to participate in the VPPI (N = 178) for the years 2010-2011 and 2011-2012. These school years encompassed the year before the incentive and the year the incentive was available. Discontinuation of the VPPI prohibited this study from moving beyond the 2012 school year.

Internal review board permission was not required due to the use of an existing data set; therefore, prior to conducting the study, permission from the Senior Associate Dean of Academic Affairs was sought. Once permission was granted, the names of the HTS schools and SIG schools invited to participate were entered into a Microsoft Excel spreadsheet. Information was obtained from the School Report Card page on the VDOE website. The school level SOL results for grades 5 and 8 and EOC, with the pass rates for math, reading, writing, science, and history for the school years 2010-2011 and 2011-2012, were downloaded and saved on a personal computer. Microsoft Excel was used for calculation of gain scores for each of the schools. Further analysis was conducted using Statistical Software for the Social Sciences (SPSS) software. Mean and median measures of central tendency were determined to summarize the data for the gain scores of each content area at grades 5, 8, and EOC. Measures of dispersion were computed to understand the variability of gain scores for each content area at the analyzed grade levels. In this research design, the unit of analysis was the school.

This study used mean and median to summarize the data for the gain score differences between the participating and nonparticipating schools. Dispersion was computed to understand the variability in the differences of the pass rates of school SOL scores in math, reading, writing, science, and history. Each group's scores has an associated frequency polygon. According to Fraenkel and Wallen (2006), the frequency polygon "will show all the information about each group" and indicate "whether a mean or median is the most appropriate measure of central tendency to use" (p. 248). Boxplots and histograms were used to graphically summarized data. Both types of graphic representations were included to compare between the two distributions and between the nonparticipating versus participating VPPI schools (see Appendices A, B, and C for this information).

A potential weakness of this study was the use of school-level data rather than teacher-level data because the use of school data limited the conclusions about teacher effectiveness and motivation. Test score data encompassed the scores of teaching staff eligible and ineligible for the incentive, limiting analysis of comparisons between teachers who did and did not participate in the VPPI. The limited time available for the initiative, which lasted only one year, was another potential weakness because data trends could not be established and longitudinal data could not be tracked. Randomization and manipulation were not possible because groups were previously formed. Data were analyzed based on the type of school: participating or nonparticipating. A weakness of this ex-post facto design is that "relationships can be identified, but causation cannot be fully established" (Fraenkel & Wallen, 2006, p. 371).

Participants The participants in this study were all the schools eligible and invited to participate in the VPPI. These schools were HTS schools or SIG schools. The VDOE defined HTS schools as meeting four of the following criteria:

- The school was accredited with warning.
- Average daily attendance rate was two percentage points below the statewide average.
- Percentage of special education students exceeded 150% of the statewide average.
- Percentage of teachers with provisional licenses exceeded 150% of the statewide average.
- Percentage of special education teachers with provisional special education licenses exceeded 150% of the statewide average.
- Percentage of inexperienced teachers (i.e., 0 years of teaching experience) hired compared to the total number of teachers exceeded 150% of the statewide average.
- The school had one or more inexperienced teachers (i.e., 0 years of teaching experience) in a critical shortage subject area. (RMC Research Corporation, 2012, p. 10)

Additionally, the VDOE indicated that HTS schools were those unable to adequately fill their teaching vacancies or retain quality teaching personnel due to the nature of the student populations (e.g., students living in poverty) or the location (e.g., rural, urban) of the school, which rendered the school economically disadvantaged or isolated (VDOE, 2011). The final evaluation did not indicate how teachers were chosen

to participate, only that participating teachers had to hold a Virginia teaching license and be endorsed in the grade level and subject associated with the teaching assignment (RMC Corporations, 2012).

SIG schools were schools that received funding through the United States Department of Education (ED). This funding was part of a grant program for state education agencies (SEAs). The grant program allowed the SEAs to create subgrants for low-performing schools to raise student achievement (ED, 2016a, para. 1). In Virginia, these subgrants were appropriated for Title I schools. The SIG schools were mandated to participate in the VPPI, but the final evaluation did not indicate how teachers were chosen to participate.

Additionally, all VPPI school administrators received training and support from Dr. Stronge and a team from The College of William and Mary. The evaluation by the RMC Research Corporation (2012) indicated Dr. Stronge and his team were working with SIG schools prior to the VPPI, and that he and his team were brought to the VPPI to support the HTS and SIG schools as they conducted training for school leadership teams about the process of summative decision-making.

Data Collection and Handling

The data for this study were obtained from the School Report Card section found on the VDOE website. The information was a publicly available data set titled "SOL Test Results/Pass Rates" (VDOE, 2016). Through this site, a Microsoft Excel spreadsheet was available with pass rate scores by school level for each grade and content area for which an SOL test is given. Schools and pass rate scores were filtered for core content areas in grades 5, 8, and EOC for the entire population invited to participate in

the VPPI. In gathering data, information for seven HTS schools (three elementary schools, three middle schools, and one high school) did not have SOL scores, bringing the total population of schools analyzed to 178. Details are included later in the document. The report included the data for the 2010-2011 and 2011-2012 school years.

After the Senior Associate Dean of Academic Affairs granted permission, data were downloaded from the VDOE website, and all pass rates for grades 5, 8, and EOC in the schools were obtained. Computation of gain scores was with Microsoft Excel and SPSS was used to determine the measure of central tendency and dispersion. All data were stored on a personal computer. This was a publicly available data set; data security precautions have been taken by the VDOE to ensure that accurate SOL data information is associated with the appropriate schools, and that content level information regarding SOL data is up-to-date for the designated school year.

The procedures for protecting this publicly available data set are outlined in the VDOE manual, *Procedures for Data Collection and Reporting* (2009), which states that "all major agency data should be stored in the Department's Oracle database" (p. 9). Additionally, Section 3.3 of the manual outlines procedures for data confidentiality:

> 3.3 Ensure that confidential data are protected by following these guidelines:
>
> - Never discuss confidential aspects of the data collection activity with unauthorized persons.
> - Keep copies of records, test scores, and other data in a secure place and deliver promptly to the appropriate person/location.

- During the data collection activities, do not keep any notes or documentation that contain identifying information unless expressly required by the research design.
- Comply with all applicable federal, state, and local laws concerning privacy and confidentiality.
- Destroy all records when data collection requirements are completed. (p. 5)

Instrument

At the time of the study, the instrument used to measure student achievement in core content areas is the SOL. Each test is a multiple-choice test except for the SOL writing test, which includes both a multiple-choice test and an essay. Multiple-choice questions are based on the SOL test blueprint, which groups the SOL into reporting categories. The total scale score indicates a student's progress on the test. Each test has a cut score, which is the proficiency level set by the VDOE based on recommendations from committees of educators (VDOE, 2016, para. 9). According to the VDOE (2016), a group of educators who form a Standard Setting Committee reviews an original form of the SOL test, called the standard setting form. This committee makes recommendations about the number of correct responses needed to reach each of the achievement levels. With the adoption of new curricular standards every six years, a test is developed based on the standards; the Virginia Board of Education must adopt the standard setting form, cut score recommendations, and the scale scores of pass/proficient and pass/advanced.

Reliability and Validity

The VDOE has a contract with Pearson, a proprietary organization that develops the SOL test and does not release specific information about the content of the test. To

ensure that guidelines and processes are in place to protect the integrity of the SOL, the VDOE takes steps to work with Pearson to help produce the most reliable and valid data. Each year, the VDOE selects teachers and leaders from a pool of applicants from across the Commonwealth to participate in test review committees. During the summer months, members of the VDOE leadership for each tested content area, the VDOE assessment department, selected teachers and leaders, and representatives from Pearson review SOL test items for acceptance and further use, or rejection (Virginia Statewide Student Assessments Technical Report, 2012).

In addition to SOL test review committees, the VDOE uses item response theory (IRT) to "develop, calibrate, equate, and scale the Virginia Standards of Learning (SOL) was the Rasch model (Rasch, 1980) and its polytomous extension, the Partial Credit model (PCM) (Masters, 1982)" (VDOE, 2012b, p. 26). To equate test items, a subset of items is used on multiple versions of the SOL test for the same grade and same subject; the VDOE refers to these items as *linking items*. To ensure tests remain consistent in difficulty, linking items are used to compare tests across years. The VDOE takes measures to ensure the SOL test is secure and that the same test is not administered twice, so an internal consistency method is used for reliability. Cronbach's alpha is used for internal consistency and reliability, and values are reported to be above .7 (VDOE, 2012b).

The SOL assessment is written based on a test blueprint, which is used to "ensure that each SOL assessment addresses the Commonwealth's content standards within each subject" (VDOE, 2012b, p. 40). The blueprint indicates the number of items for each reporting category, so the SOL test is consistent in what it assesses each year. The

blueprint is created from the SOL curriculum framework, a document used at the district and school level to help define the knowledge and skills students need to develop at each grade level. According to the VDOE (2012b), the "direct relationship between the SOL curriculum frameworks, the SOL test blueprint, and the SOL assessments lends support to the content validity of the SOL assessments" (p. 40). Every 6 years, curriculum is rewritten and blueprints are updated in a continuous process.

Assessment of data takes place after 3,000 students have taken a form of the SOL test. This includes "an examination of raw response frequencies for each item to check for multiple marks and omit rates (multiple-choice) and to see if the item shows the proper range of scores (open-ended)" (VDOE, 2012b, p. 34). At this point, any anchor items that have "displacements greater than an absolute value of 0.5 are identified as possible outliers and removed from the linking set" (VDOE, 2012b, p. 34).

Construct validity is determined in two ways: statistical correlation and factor analysis. The Virginia Statewide Student Assessment Technical Report (VDOE) indicates that "school pass rates on the SOL tests were previously statistically correlated with national percentile ranks of the *Stanford 9* and/or pass rates on the *Literacy Passport Test (LPT)*" (VDOE, 2012b, p. 41). Additionally, the SOL tests were analyzed separately and then together based on the type of administration of the multiple-choice test, which was either paper-pencil or online (VDOE, 2012b).

For the purposes of this study, existing data were obtained from the VDOE in Microsoft Excel format. Although the VDOE had taken steps to ensure that the data were accurate, this does not mean the data were definitely error-free, and so necessary steps

were taken to guarantee the data were accurate (Smith, 2008). Data were reviewed and analyzed in the Microsoft Excel format by cross-checking information.

Data Analysis

The unit of analysis of this study was the school. The entire population of schools eligible for the VPPI was used. School data related to SOL pass rates for grades 5, 8, and EOC for each core content area in school years 2010-2011 and 2011-2012 were downloaded from the VDOE School Report Card page into an Excel file. Gain scores were computed using Microsoft Excel. Data were then imported into SPSS and verified to ensure the data matched the information from the Excel file. After data were loaded into SPSS, measures of central tendency, dispersion, and frequency polygons were computed. Research questions were answered by analyzing the measures of central tendency (mean and median) and the measure of dispersion. The incentive was determined to have made a difference if the median was one point or higher in a positive direction. Median gain scores were chosen because the population was small and would not be affected by higher scores, depicting a more accurate indication of the test scores (Fraenkel & Wallen, 2006). One point was chosen because one point can be the difference between a school being granted accreditation or denied accreditation (VDOE, 2015a). A one-point gain score was deemed important to the findings because of the way schools in Virginia are accredited. For the school years analyzed in this study, a one-point gain score was important because at that time, schools were to ensure that they earned 100% proficiency by 2014 through demonstration of Adequate Yearly Progress (AYP) on the SOL tests (VDOE, 2011a). For the years of the study, AYP for all students overall was to be more than 86% proficiency in reading and more than 85% proficiency

in math (VDOE, 2011a). In addition, 95% of students must have participated in the reading and math testing, and 95% of students from the following subgroups must have participated in the testing for reading and math: White, Black, Hispanic, students with disabilities, economically disadvantaged students, and Limited English Proficiency (LEP) students (VDOE, 2011a). Due to the stringent testing guidelines, every test score counted for the schools, and one point could have been the difference between being accredited and denied accreditation.

Human Participants and Ethics

The risk to human participants in this study was minimal. The unit of analysis was the school. Identifiable data of teachers and students were not available on the VDOE School Report Card, making it impossible to link school scores to individual teachers, students, and schools. Possible conflict of interest and personal bias involved an administrator in a Virginia school district who was in direct supervision of the middle-school and high-school English curriculum and instruction, and also conducted the study. Bias was minimized because the researcher did not work in the districts using the VPPI and was not associated with parents, students, teachers, or administrators in the school districts using the VPPI. In addition, all research was conducted with integrity and in a truthful and accurate manner during the literature review and research design. This was an objective study that attempted to contribute to the body of knowledge regarding pay for performance initiatives for teachers.

Summary

The purpose of this study was to examine gain scores for the population of schools invited to participate in the VPPI (N = 178) to determine if a difference existed

between participating and nonparticipating VPPI schools during the 2010-2011 and 2012-2012 school years. The focus was grades 5, 8, and EOC. All data were collected from the VDOE Virginia Assessment Build A Table, a publicly available data set available through the VDOE website. Gain scores were used to summarize the data. The median was analyzed because of the small population size (Fraenkel & Wallen, 2006). A one point gain was the determination of whether or not the VPPI made a difference. The purpose of the VPPI was to compensate teachers for positively impacting student achievement in the form of SOL scores, yet that information was not available in the final evaluation conducted by the RMC Research Corporation (2012), warranting this study. The corresponding results are explored in the data analysis section of Chapter 4.

Chapter 4: Results

Overview

The purpose of this study was to analyze the gain scores for all schools invited to participate in the Virginia Performance Pay Incentives Initiatives Pilot (VPPI), to see if a difference existed between Standards of Learning (SOL) test scores of participating and nonparticipating schools. The focus was content area subjects in grades 5, 8, and end-of-course high school (EOC). This was the first time the Commonwealth of Virginia had piloted a teacher performance pay program that linked the incentive to student achievement in all Superintendent's regions. As stated in a January 18, 2011 news release, former Governor Robert McDonnell stated:

> As we've said before, a child's educational opportunities should be determined by their [*sic*] intellect and work ethic, not their [*sic*] zip code. We must ensure that students in every corner of the Commonwealth have access to great schools, excellent teachers and the opportunity to learn in an environment that suits them best. This year we are proposing a plan to provide performance pay incentives of up to $5,000 to teachers in schools identified as hard to staff. Our hope is that these grants will attract good teachers to schools that, for example, may be on the verge of losing their accreditation, have trouble with attendance, or have a large number of limited English proficient students – with a goal that these teachers will help to turn the performance at these schools around and help their students reach higher achievement marks. This was a reform I first outlined during my campaign for Governor and a plan that we will continue to develop over the next couple [of] years as we see the effectiveness of the pilot programs. Performance pay is a bipartisan idea, and I have applauded the President for his leadership on this issue. It is time that we moved forward on bringing this idea to the Commonwealth, to the benefit of our students, teachers and parents. (Virginia Department of Education, 2011c, p. 1)

The overarching goal of the VPPI was to increase student academic achievement as determined through the SOL test (RMC Research Corporation, 2012). To meet this goal, the Virginia Department of Education (VDOE) piloted a new teacher evaluation model that included a standard linking student academic achievement to the teacher

evaluation. To pilot this evaluation model, the VDOE implemented the VPPI in hard to staff (HTS) and school improvement grant (SIG) schools that rewarded teachers with a monetary incentive for an exemplary rating on the evaluation. A reward of up to $5,000 was offered for HTS teachers and $3,000 for SIG teachers.

Participation in the VPPI was by invitation, and when a low number of the invited HTS schools applied for the grant, SIG schools were mandated to participate in the initiative. SIG schools received federal funds through a United States Department of Education (ED) grant program designed for schools in challenging urban or rural communities, or schools with students in challenging economic situations. A total of 169 HTS and 16 SIG schools were invited to participate in this pilot initiative; the school population included 83 elementary schools, 47 middle schools, and 54 high schools. A total of 25 schools (9 HTS and 16 SIG) participated in the initiative. This ex-post facto study was conducted to determine if a gain score difference existed in student SOL pass rates in grades 5, 8, and EOC in tested core content areas (i.e., reading, writing, math, science, and history) between the participating and nonparticipating schools invited to participate in the VPPI, in the year before and the year of the program.

In interpreting effectiveness of the incentives, a difference was meaningful if there was at least a one point gain in the median compared to the previous year. As indicated in Chapter 2, a one point gain was significant because of the VDOE school accreditation policy; one point can be the difference between granting or denying a school accreditation (VDOE, 2015a). During the VPPI, schools were under the No Child Left Behind (NCLB) guidelines, and had a pre-established goal of 100% proficiency by 2014. At the time of the VPPI, schools were to be at 70% proficiency in reading and

70% proficiency in math (VDOE, 2011a). Therefore, a one point gain or loss was the difference between accreditation granted and accreditation denied. In Virginia, schools are accredited annually based on SOL test scores (VDOE, 2017).

The research questions that guided this study were:

1. What is the difference in grade 5 Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the initiative?
2. What is the difference in grade 8 Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the initiative?
3. What is the difference in end-of-course Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the initiative?

The entire population of schools invited to participate in the VPPI was included in this study, and measures of central tendency (mean and median) and dispersion were reported. A median gain of one point was used as the important average because the median would provide a more accurate description of the gain scores (Fraenkel & Wallen, 2006). Extreme scores in a small population can inflate the mean gain scores (Fraenkel & Wallen, 2006). The unit of analysis was the school. Data for each school from the entire population of schools invited to participate (N = 178) was included in the

study. Data were coded as P for participating schools or NP as nonparticipating schools. As referenced in Chapter 3, the VDOE took steps to ensure that the most reliable and valid data were reported by ensuring specific guidelines and processes were in place for the administration of the SOL tests. Selection of schools was by either the VDOE HTS identification or receipt of SIG funds, and all data were publicly available from the VDOE website. The test scores for each subject area for the school years 2010-2011 and 2011-2012 were obtained, and the difference between the two years was calculated using Excel. For convenience of mathematical computations, SPSS was used to calculate the mean and median to summarize the data and standard deviation to show the variability of scores. Histograms, box plots, and frequency polygons were constructed based on the calculations to visually summarize data. In the following section, the population is described, followed by an analysis of each of the research questions.

Description of the Population

The population of interest in this study was composed of all the schools in the Commonwealth of Virginia that were invited to participate in the VPPI (N = 178). In the process of gathering data, it was discovered that SOL scores were not available for one high school. In addition, three elementary schools did not have SOL data because the schools housed Kindergarten through grade 4 only. Also, three middle schools included grades 6 and 7, so grade 8 SOL scores were not available for those schools. One school included both grade 5 and grade 8; SOL scores from both grades were included. No students from participating VPPI schools took the geography EOC test; therefore, a comparison between the participating and nonparticipating schools could not be made in that content area.

In the final data analysis, nine participating schools tested grade 5 reading, writing, science, and math. Seven participating schools tested grade 5 history. Three participating schools tested grade 8 reading, writing, and history. Two participating schools tested grade 8 science. Five participating schools tested grade 8 math. For EOC, eight participating schools tested reading, writing, chemistry, geometry, Algebra II, World History I, World History II, and Virginia/United States History. Seven participating schools tested biology and Algebra I. Six participating schools tested earth science.

Data Analysis

As discussed in Chapter 3, data were obtained from a publicly available dataset found on the VDOE website in the School Report Card section. Data included pass rates for all schools. In both the 2010-2011 and 2011-2012 school years, the nonparticipating VPPI student population was larger at all three levels (grades 5, 8, and EOC) being analyzed for this study. This is important to note, because the RMC Research Corporation's (2012) final evaluation of the VPPI did not explain the reasons why nonparticipants outnumbered participants. It is also important to note that for all three levels, at both participating and nonparticipating schools, the math gain scores were negative. Between the 2010-2011 and 2011-2012 school years, the math SOL standards and curriculum framework changed to reflect more rigorous standards. This impacted the rigor of the math SOL test, resulting in a decline of the math SOL scores ranging from 8 to 25 points across the Commonwealth (VDOE, 2012). Although math scores at participating and nonparticipating schools experienced the same change in measurement,

the expectation was that there would be larger negative gain scores, but no reason to expect there would be a differential impact on the two groups.

As indicated by the RMC Research Corporation (2012) grant application, HTS schools had to meet four of eight specific criteria, and SIG schools were included because of the funds received through federal programming. A visual representation of the specific criteria met by each hard to staff school is included in the tables 6, 7, and 8.

Table 6

Qualifying Criteria of Hard to Staff Elementary Schools Participating in the VPPI

School	Accredited with Warning	Daily Attendance two percentage points below state average	Percent of special education students exceeded 150 percent of statewide average (6.9%)	LEP students exceed 150 percent of statewide average (6.96%)	Provisionally licensed teachers exceed 150 percent statewide average (5.09%)	Provisionally licensed special education teachers exceed 150 percent statewide average (1.58%)	Percent of inexperienced teachers hired compared to total number of teachers exceed statewide average (3.55%)	One or more inexperienced teachers in a critical shortage area
Kegotank Elementary	N/A	N/A	N/A	13.28	11.54	N/A	5.77	2
Pungoteague Elementary	N/A	N/A	N/A	12.43	9.3	N/A	9.3	3
Madison Elementary	N/A	N/A	N/A	N/A	12.12	3.03	9.09	2
Blue Ridge Elementary	N/A	N/A	N/A	N/A	11.11	3.7	14.81	3
Hardin Reynolds Memorial	N/A	N/A	N/A	N/A	17.39	4.35	8.7	1

Note: Adapted from Attachment B, Supts Memo no. 113-11, April 15, 2011. Retrieved from http://www.doe.virginia.gov/administrators/superintendents_memos/2011/113-11b.pdf.

Table 7

Qualifying Criteria of Hard to Staff Middle Schools Participating in the VPPI

School	Accredited with Warning	Daily Attendance two percentage points below state average	Percent of special education students exceeded 150 percent of statewide average (6.9%)	LEP students exceed 150 percent of statewide average	Provisionally licensed teachers exceed 150 percent statewide average (5.09%)	Provisionally licensed special education teachers exceed 150 percent statewide average	Percent of inexperienced teachers hired compared to total number of teachers exceed statewide average (3.55%)	One or more inexperienced teachers in a critical shortage area
Dinwiddie County Middle School	N/A	N/A	N/A	N/A	12.5	5.36	5.36	3
Edward W. Wyatt Middle School	N/A	N/A	N/A	N/A	11.32	3.77	7.55	4

Note: Adapted from Attachment B, Supts Memo no. 113-11, April 15, 2011. Retrieved from http://www.doe.virginia.gov/administrators/superintendents_memos/2011/113-11b.pdf.

Table 8

VPPI Qualifying Criteria of Hard to Staff High Schools Participating in the VPPI

School	Accredited with Warning	Daily Attendance two percentage points below state average	Percent of special education students exceeded 150 percent of statewide average (6.9%)	LEP students exceed 150 percent of statewide average	Provisionally licensed teachers exceed 150 percent statewide average (5.09%)	Provisionally licensed special education teachers exceed 150 percent statewide average	Percent of inexperienced teachers hired compared to total number of teachers exceed statewide average (3.55%)	One or more inexperienced teachers in a critical shortage area
Caroline High School	N/A	N/A	N/A	N/A	9.89	4.4	8.79	1
Patrick Henry High School	N/A	N/A	N/A	N/A	9.59	N/A	7.75	8

Note: Adapted from Attachment B, Supts Memo no. 113-11, April 15, 2011. Retrieved from http://www.doe.virginia.gov/administrators/superintendents_memos/2011/113-11b.pdf.

Analyses of the Research Questions

This study examined whether a difference existed in gain scores between participating and nonparticipating VPPI schools during the year before and the year of the program. Data for the entire population invited to participate in the program were obtained from the VDOE website. Mean, median, and standard deviation were used to analyze the data from a preexisting dataset. Gain scores were computed and the mean, median, and standard deviation were reported. Three research questions were answered, and these questions addressed grades 5, 8, and EOC; the main core content areas assessed by the SOL test were reading, writing, math, history, and science. For EOC, the math, history, and science courses became more specific based on the specialized nature of the content areas; therefore, research questions for EOC tests were specific to the specialized content and reported as such. Data are presented in tables 6, 7, and 8. These tables represent each grade level analyzed. Presentation of results follows the presentation of the data.

Research Questions

In all three grade levels analyzed, there was a large disparity between the numbers of schools in the participating versus nonparticipating groups. In the analysis of scores and comparisons between participating and nonparticipating schools, the incentive was interpreted as effective, and a difference meaningful, by looking at the direction of the difference and analyzing if there was at least a one-point median gain compared to the previous year. The number of schools in both the participating and nonparticipating schools was small, so the median provided a better indication of a gain (Fraenkel & Wallen, 2006). This is important information at the school level because one point can

be the difference between maintaining and losing accreditation status (VDOE, 2015). As noted in Chapter 1, a one point gain score was important to this study because of the implications on student achievement (VDOE, 2011b) and school accreditation (VDOE, 2017b).

Table 9

Grade 5 VPPI SOL Mean, Standard Deviation, and Median Gain Scores

	Participating Group				Nonparticipating Group			
Content Area	*N*	μ	σ	Mdn	*N*	μ	σ	Mdn
Reading	9	5.78	8.77	6.00	73	-0.55	5.56	-1.00
Writing	9	.22	8.99	2.00	73	.86	6.43	1.00
Science	9	2.89	10.97	2.00	73	.85	8.17	-1.00
Math*	9	-13.44	12.38	-12.00	73	-20.18	11.43	-20.00
History	7	-5.57	7.67	-5.00	73	-1.08	7.87	.00

Note: N= Total number of schools with students taking the SOL test.
μ= Mean gain score.
σ= Standard deviation.
Mdn= Median gain score.
* The Math SOL change resulted in negative gain scores for participating and nonparticipating groups.

Table 10

Grade 5 VPPI Effect Size for Participating and Nonparticipating Groups

	Effect Size (*d*)
Reading	.71
Writing	.08
Science	.21
Math*	.56
History	.57

Research question one asked: 1. What is the difference in grade 5 Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the initiative? In grade 5, gains of one point or more were made in reading, writing, and science for the participating schools. In the nonparticipating schools, writing had a gain of 1 point. In reading, the standard deviation of the gain scores for the participating group was greater than that of the nonparticipating group, indicating a greater spread of scores. To analyze the variability that existed between the participating and nonparticipating schools, the effect size was accounted for. Effect size was calculated by dividing the difference between the mean of the participating and nonparticipating groups' gain scores by the standard deviation of the nonparticipating group (Fraenkel & Wallen, 2006). For transparency of data analysis,

effect size calculation used mean while analysis of gain scores from participating and nonparticipating schools was from median.

In math, both the participating and nonparticipating groups had scores in a negative direction. It was during the 2011-2012 school year that the math standards changed to a more rigorous format, and math test scores for schools in Virginia subsequently dropped (VDOE, 2012c). The data showed that the participating group did not have as great a decline in scores as the nonparticipating group after the change to the new standards. In the nonparticipating group, the writing had a meaningful gain score (Mdn = 1.00_g), and the history showed no change (Mdn = 0.00_g).

Table 11

Grade 8 VPPI SOL Mean, Standard Deviation, and Median Gain Scores

	Participating Group				Nonparticipating Group			
Content Area	*N*	μ	σ	Mdn	*N*	μ	σ	Mdn
Reading	3	-6.00	5.29	-4.00	38	-3.50	7.44	-2.00
Writing	3	-1.33	9.50	-1.00	38	-1.61	6.89	-2.00
Science	2	-1.00	5.65	-1.00	39	.46	6.99	1.00
Math*	5	-15.00	28.02	-22.00	38	-31.63	13.15	-30.00
History	3	5.67	16.62	-8.00	35	-.89	16.66	1.00

Note: *N*= Total number of schools with students taking the SOL test.
μ = Mean gain score.
σ= Standard deviation.
Mdn= Median gain score.
* The Math SOL change resulted in negative gain scores for participating and nonparticipating groups.

g Gain score

Table 12

Grade 8 VPPI Effect Size for Participating and Nonparticipating Groups

	Effect Size (*d*)
Reading	.38
Writing	.03
Science	.08
Math*	.75
History	.28

Research question two asked: What is the difference in grade 8 Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the initiative?

Grade 8 had the lowest number of participating schools out of the three levels, with two participating schools represented for the science test, three participating schools represented for the reading, writing, and history tests, and five schools represented for the math test. The VPPI participating schools did not have a meaningful gain in any area. All gain scores for the participating schools were in a negative direction. An examination of the box plot for grade 8 history shows no outliers (see Appendix E for visual representation of the data). A further examination of the frequency polygon for grade 8 history shows more positively skewed gain scores (see Appendix E for visual representation of the data). Consistent with trends in grade 5 math, grade 8 math also had scores in a negative direction for both the participating and nonparticipating groups.

Table 13

EOC VPPI SOL Mean, Standard Deviation, and Median Gain Scores

	Participating Group				Nonparticipating Group			
Content Area								
Reading	8	-.13	3.04	1.00	39	-.46	4.12	.00
Writing	8	2.75	4.83	4.00	39	.59	5.32	.00
Chemistry	8	-.75	5.23	1.50	37	1.68	7.63	.00
Biology	7	1.86	6.46	.00	40	1.80	4.77	2.00
Earth Science	6	3.83	4.33	4.50	39	1.10	5.37	1.00
Geometry*	8	-19.13	5.93	-19.00	39	-17.07	12.37	-15.00
Algebra I*	7	-38.29	12.12	-38.00	40	-31.28	14.76	-30.00
Algebra II*	8	-21.88	16.96	-22.50	38	-29.89	18.95	-31.00
World History I	8	1.00	9.79	2.00	32	6.31	8.78	4.50
World History II	8	3.75	8.54	5.00	35	2.00	8.81	2.00
Virginia/ United States History	8	1.38	6.36	1.50	39	1.95	6.99	2.00
Geography	No Participants				17	7.76	25.04	1.00

Note: N= Total number of schools with students taking the SOL test.
N= Total number of schools taking the SOL test.
μ= Mean gain score.
σ= Standard deviation.
Mdn= Median gain score.
* The Math SOL change resulted in negative gain scores for participating and nonparticipating groups.

Table 14

EOC VPPI Effect Size for Participating and Nonparticipating Groups

	Effect Size (*d*)
Reading	.09
Writing	.42
Chemistry	.14
Biology	.01
Earth Science	.05
Geometry*	.21
Algebra I*	.00
Algebra II*	.44
World History I	.57
World History II	.20
Virginia/ United States History	.08

Research question three asked: What is the difference in end-of-course high school Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the initiative?

As shown in Table 8, in EOC tests, the VPPI participating schools showed meaningful gains in more areas compared to the VPPI participating elementary and middle schools. Additionally, VPPI participating schools made greater gains than nonparticipating schools in reading, writing, chemistry, earth science, and World History II. Table 8 shows nonparticipating schools made greater gains in biology, World History I, Virginia/United States History. The three end-of-course math tests showed gains in a negative direction, which is consistent with the negative direction in math scores at the grade 5 and grade 8 levels.

Summary

This chapter provided an overview and analysis to answer the following research questions:

1. What is the difference in grade 5 Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the initiative?
2. What is the difference in grade 8 Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the initiative?
3. What is the difference in end-of-course high school Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the initiative?

The purpose of this study was to examine whether a difference existed in gain scores for the SOL test scores for grades 5, 8, and EOC between the year before and the year of the VPPI. To examine the gain scores between participating and nonparticipating schools, data were obtained from the School Report Card information on the VDOE website. The main findings of the study were that the VPPI participating schools had greater meaningful gain scores on the following tests: grade 5 reading, writing, and science, and EOC reading, writing, chemistry, earth science, and World History II. None of the grade 8 gain scores in the VPPI schools were in a positive direction or considered meaningful in the context of this study. The non VPPI schools had greater meaningful gain scores in the following areas: grade 8 science and grade 8 history; EOC biology, World History I, and Virginia/United States History. Comparison of scores for end-of-course geography did not occur because VPPI participating schools did not have student representation for the test.

There was an obvious difference between the populations at the participating versus nonparticipating schools. The number of participating schools was small, particularly in grade 8. Additionally, negative gain scores were reported. Based on this study, a positive median gain of one point was considered meaningful. Negative gain scores were important to note because they indicated that the incentive may not have made a difference.

The purpose of the VPPI was to pilot the new teacher evaluation model and offer teachers a monetary incentive for increased test scores. The negative gain scores at the participating schools deserve attention. As noted earlier, math had negative gain scores at all levels, for participating and nonparticipating schools. At the elementary level,

participating schools had gain scores in a negative direction in math, as noted in Table 6. At the middle level, all participating schools had scores in a negative gain scores for all areas, as noted in Table 7. With EOC VPPI participating schools, negative gain scores were in math, and were consistent with data at other levels. Additionally, with the exception of math, the only area without a meaningful gain at VPPI participating schools was biology, as noted in Table 8. What this data does not indicate is information at the teacher level for the year of the evaluation or trend data about teacher test scores. This is one of the challenges of this study's research design. The evaluation did not provide sufficient data to draw conclusions about the success of the VPPI. The lesson learned from this study is that in drafting pay for performance programs, a stronger plan from idea inception to implementation is necessary because, essentially, the data collected did allow measurement.

The purpose of this study was to examine if a difference exists in SOL test scores for grades 5, 8, and EOC during the year before and the year of the VPPI. Interpretations, conclusions, and recommendations for further study are included in the next chapter.

Chapter 5: Interpretations, Conclusions, and Recommendations

Summary of the Problem

Incentives are used to recruit and retain teachers in hard to staff (HTS) schools, especially because school-age populations of children are increasing and the pool of teachers is shrinking (Amrein-Beardsley, 2012; Ingersoll & Merrill, 2010; Kolbe & Strunk, 2012; Petty et al., 2012; Podgursky, 2006). In theory, this concept can be enticing for some teachers because most school districts base teacher salaries on a step or lane system (Stronge, Gareis, & Little, 2006). A step or lane compensation system pays teachers for years of experience and provides a stipend for advanced degrees or certifications, but teachers—regardless of the grade level they teach—are paid the same amount and awarded for longevity rather than expertise (Figlio & Kenny, 2005; Loeb et al., 2009; Stronge et al., 2006).

To justify the spending of funds, additional compensation is attached to some type of criteria, such as teacher evaluation systems. In many cases, the additional compensation is awarded for increased student academic achievement, a controversial plan because of the high stakes attached to such a model (Belfield & Heywood, 2008; Caillier, 2010; Fryer, 2013; Goldhaber, Gross, & Player, 2010). Additionally, monetary incentives in education are one way to entice and motivate teachers to work in (HTS) schools, and education policy develops around the incentive initiatives. Using the theories developed by Vroom (1964) and Alderfer (1969) helped conceptualize the study by analyzing teacher motivation through the evaluation of why people work, what they value in their work, and the expectations associated with work.

Regardless of the incentive policy chosen, different factors hinder the success of pay for performance programs because of staffing issues, teacher inexperience, incompetence of teachers, influence of outside forces, and transiency of teachers within the district (Amrein-Beardsley, 2012; Belfield & Heywood, 2008; Goldhaber, et al., 2010). Through a study of recruiting teachers to HTS schools, Amrein-Beardsley (2012) concluded that HTS schools have been faced with teacher shortages and lack qualified personnel to fill the positions, and "those who leave are most likely to be replaced by the most inexperienced teachers: usually state-certified teachers who are new, long-term substitutes, or people with bachelor's degrees who are permitted to teach with alternative teaching certificates" (p. 4). Additionally, research by Goldhaber et al. (2010) suggested that teachers who move to different schools within a school district place the burden on schools to find qualified candidates because of costs that schools and districts have absorbed for recruitment, hiring, and training of new staff.

This study focused on one model of incentives for teachers, the Virginia Performance Pay Incentives Initiative (VPPI) pilot. This program was the initial examination of a new standard in the teacher evaluation model that based 40% of the evaluation on student academic achievement and a pilot of performance pay. This was an important program to study because it was the first of its kind proposed across the Commonwealth. Additionally, the final evaluation by the RMC Research Corporation (2012) did not include information about student achievement rates after program completion. This study was an analysis of the gain scores of all the schools invited to participate in the initiative to see if a difference existed between the participating and nonparticipating schools the year before and the year of the program. The names of the

teachers were not available for this study; therefore, the unit of analysis was the school. Although information at the teacher level was not available, the results contributed to the current body of knowledge about monetary incentives for teachers, particularly those incentive programs that connect teacher evaluation and student academic achievement. This was the first time an incentive program was initiative in all Superintendent's regions in Virginia. Study design flaws prohibited the measure of student achievement data. For example, data about teacher level data were not available. In an effort to find out information the program structure at the school level, all participating school districts were contacted. When possible, district personnel were asked about teacher recruitment and participation in the VPPI, the grades participating teachers taught, and the content areas participating teachers taught. Accomack County responded with an unofficial response that the teachers taught testing grades 3-5, and every content area was part of the evaluation (personal communication, 2018). The principal at the time of the VPPI at Edward J. Wyatt Middle School indicated that all reading, math, science, and social studies core content teachers in grades 6, 7, and 8 participated. Award of the incentive was incumbent on SOL score outcomes (personal communication, 2018). Attempts at communication with other participating districts went unanswered. This omission warranted the need for the study.

This study confirmed that initiatives need more than one year to document sustainable change, and teachers need to know how to change their teaching practices to influence student achievement. As discussed in Chapter 1, consideration of motivation theory in the analysis of monetary incentives for teachers was an important factor in this

study. Motivation occurs when the outcome helps create success and derive some type of pleasure (Alderfer, 1969).

The following research questions framed the study:

1. What is the difference in grade 5 Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the initiative?
2. What is the difference in grade 8 Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the initiative?
3. What is the difference in end-of-course high school Virginia Standards of Learning gain scores between participating and nonparticipating VPPI schools in the following content areas: (a) reading, (b) writing, (c) science, (d) math, and (e) history the year before and the year of the initiative?

Purpose

The purpose of this chapter is to analyze the findings, implications for future application, and provide suggestions for further research. Although the literature on pay for performance incentives is mixed, the findings are important because of the effects on teaching and learning outcomes and the consequences on the careers of educators. As indicated in the literature review, student achievement data measuring effectiveness of teachers is flawed because assigning students to certain classrooms is not random, and can be based on parent requests, students' academic needs, or students' special education

status (Amrein-Beardsley, 2012, Rice et al., 2012). An additional complication is teaching level. Elementary teachers typically teach a variety of subjects, but may only have expertise in one or two of those subjects, which may not be the subjects in which a judgment about effectiveness is made (Caillier, 2012). At the secondary level, middle and high school teachers typically have expertise in one subject, but may also be adept at recruiting students to sign up for a course; as a result, student assignment to a course may not be random (Caillier, 2012).

Confounding matters further is federal education policy that connects teacher outcomes and student academic achievement. The introduction of the Teacher Incentive Fund (TIF) in 2006 and, later, the Teacher and Leader Incentive Fund (TLIF) in 2011 were monetary incentives originally directed at creating an incentive-based system designed for teachers and school leaders in schools serving poor, minority, and disadvantaged students (Smarick, 2011). Those incentives later became a way to evaluate and reward teachers for positive student achievement outcomes on standardized assessments. In a 2016 evaluation of the third and fourth years of 2011 TLIF grantees, researchers found that as of the third year of the study, school districts reported it was easier to implement the program elements, such as feedback components related to growth measures about student achievement and teacher observations (Wellington et al., 2016). Additionally, in the third year of TLIF implementation, reported that maintaining the program was a challenge, which was down 64% from the previous study (Wellington et al., 2016). In an analysis of student achievement, reading and math gains were one to two percentile points higher than schools that did not offer the bonus, which was equal to

four weeks of learning (Wellington et al., 2016). The structure of the VPPI—the subject of this study—is similar to the TLIF.

Summary of Results

Grade 5 Scores

In grade 5, math, science, social studies, reading, and writing gain scores were analyzed. Effect size was calculated to show how much variability existed between the gain scores of the participating and nonparticipating schools. VPPI participating schools had larger and positive gain scores in reading ($\mu = 5.78_g$, $\sigma = 8.77_g$, Mdn $=6.00_g$), science ($\mu = 2.98_g$, $\sigma = 10.97_g$, Mdn $= 2.00_g$), and writing ($\mu = .22_g$, $\sigma = 8.99_g$, Mdn $=2.00_g$) than did nonparticipating schools.

Grade 8 Scores

In grade 8, the content areas for which gain scores analyzed were reading, math, science, social studies, and writing. The gain scores for grade 8 were larger and positive in nonparticipating schools in science ($\mu = .46_g$, $\sigma = 6.99_g$, Mdn $=1.00_g$) and history ($\mu = -.89_g$, $\sigma = 16.66_g$, Mdn $=1.00_g$) than the corresponding scores in participating schools, which were in a negative direction for science and history. Participating schools had scores in a negative direction, but less of a loss in writing ($\mu = -1.33_g$, $\sigma = 9.50_g$, Mdn $= -1.00_{g)}$ and math ($\mu = -15.00_g$, $\sigma = 28.02_g$, Mdn $= -22.00_g$).

High School End-Of-Course Scores

In high school, end-of-course tests become more specific based on the specialized nature of the content. In the EOC tests, three tests had larger gain scores for the nonparticipating VPPI schools than the participating schools. Nonparticipating schools' EOC gain scores were larger in biology ($\mu = 1.80_g$, $\sigma = 4.77_g$, Mdn $=2.00_g$), World

History I ($\mu = 6.31_g$, $\sigma = 8.78_g$, Mdn = 4.50_g), and Virginia/US History ($\mu = 1.95_g$, $\sigma = 6.99_g$, Mdn =2.00_g). Nonparticipating schools had less of a loss in math with the exception of Algebra II ($\mu = -29.89_g$, $\sigma = 18.95_g$, Mdn = $-31.00_{g)}$. Geography could not be compared because there were no participants from the VPPI-participating schools.

Based on this study, it was found that gain scores were negative for all math courses and at all grade levels. This can be explained by a change in the math standards during the 2011-2012 school year, and SOL math scores across Virginia dropped (VDOE, 2012c). This data reflect the decline for all the schools invited to participate in the VPPI. For the purposes of this study, a meaningful gain was a median difference of one point, and a reflection of that difference is in the score reporting.

Interpretation of the Findings

Based on the results of this study, the VPPI participating schools in grade 5 had higher gain scores than the nonparticipating schools. Gains for participating schools in grade 8 were in a negative direction, and did not show what was expected. In EOC, the participating schools had seven content area scores that represented at least one point difference between the two years compared, and the nonparticipating schools had at least one point difference on five content area scores between the two years compared. Many factors are unknown regarding the differences in gain scores. For example, with the exception of programs in two school divisions, it was unknown which teachers received the incentive or what types of classes the teachers taught. Secondly, it is unknown if this was the same teachers from year 1 to year 2. Additionally, SOL scores

---

$_g$ Gain score

are combined for all teachers at a school in each tested area; therefore, participating schools include the scores of teachers no participating. Finally, it is unknown if teachers were aware that they needed to change their teaching practices to positively impact student achievement, or how to make those changes. For example, in EOC tests, the nonparticipating and participating schools showed gains in earth science, World History I, World History II, and Virginia/United States History. Many factors could have contributed to these gains, such as changes in teaching practices or difficulty level of the test questions, or composition of classes or groups of students. Interventions are only effective when teachers are aware of how to positively improve their teaching practices to positively impact student achievement (Fryer, 2013).

In the case of the VPPI, the RMC Research Corporation's Spring 2012 evaluation indicated that in a survey of teachers (N = 419), 64% strongly agreed or agreed that it was difficult to implement the VPPI. Teachers said that they needed additional support from administrators "to answer teacher questions and address anxieties and concerns about the process; clearer guidelines on working with students with disabilities and English Language Learners in terms of goals, standards" (RMC Research Corporation, 2012, p. 41). Teachers also sought additional training and support in data collection and documentation and in writing Specific Measurable Attainable Realistic and Timely (SMART) goals with examples. In the fall teacher survey (N = 388), teachers made similar comments. Additionally, in the fall, the RMC Research Corporation (2012) asked teachers and principals the question, "What is your expectation for the future success of performance-pay improving teaching and learning at your school?" (p. 49). Responses

were on three levels: Low Expectation (performance pay will not improve teaching and learning at the school); Moderate Expectation (although there are still changes regarding performance pay, the incentive will improve teaching and learning at the school); or High Expectation (teacher buy-in will increase as teaching improves and student achievement improves. Performance pay will make a significant impact on student learning and retaining teachers) (RMC Corporation, 2012). Out of 189 teachers, 52% indicated they had a low expectation of performance-pay improving teaching and learning at the school, and out of 23 principals, 39.1% indicated they had a moderate expectation of improving teaching and learning at the school (RMC Corporation, 2012). In relating this back to Vroom's motivation theory, the expectancy associated with the success of performance pay and the impacts on teaching and learning was low. Based on teachers' comments, they did not associate their performance with the award.

In reviewing the professional development timeline for implementation of the VPPI, initial training about the program and the new teacher evaluation model took place at the end of July and the beginning of August (RMC Research Corporation, 2012). This timeframe for initial training would pose challenges at a school level because of necessary preparations for the school year taking place. Additionally, it is unclear how many teachers attended training, only that a "small number" participated (RMC Research Corporation, 2012, p. 35). Teacher training in setting student achievement goals did not occur until November, and this was a virtual training (RMC Research Corporation, 2012). Finally, schools support came from the team at The College of William & Mary, although the RMC Research Corporation (2012) did indicate that some of the support did

not take place until spring 2012. Although the intent of these trainings was good, it seems there was a lack of consistency in training and support systems.

Using Gain Scores for Teacher Evaluation and as Predictors of Student Success

The information related to the composition of the classes, the assignment of students to classes, and the assignment of teachers to classes is unknown for this study, thus limiting conclusions about the effect of the VPPI. Research suggests that one underlying principle in education policy is that teacher effects, such as "instruction, teacher characteristics, and behavior" (Konstantopoulos, 2014, p. 9) are the factors that influence student achievement outcomes. Educational research on teacher effects is broad, and the conclusions have not fully studied the impact of ever-changing education policy. Policymakers are not waiting for research outcomes before moving on to another policy. This is the case with the VPPI. The program lasted one year, and then shifted to a new program, the Strategic Compensation Grant (SCG). The SCG differed because the local school district determined the program structure for the monetary incentive, not the state authority.

Goldhaber (2009) suggested that the current research about teacher effects has not yet been conclusive about whether teachers become motivated to improve their practice or, instead, leave the teaching profession. Although offering monetary incentives is one possible way to entice educators to work in HTS schools, research suggests that an ideal program design has not emerged, and the programs that offer monetary incentives for teachers varied in outcomes (Yuan et al., 2013), with few variations of differentiated compensation programs researched (Gratz, 2011) and outcomes that have questionable methodology or limited information about effects of the programs (Rice et al., 2012).

Regarding the VPPI, the RMC Research Corporation (2012) noted that teachers interviewed in the spring (N = 417) overwhelmingly indicated that the extra money promised as a part of the VPPI did not motivate them to do a better job (n = 245). One teacher indicated that:

> I give 100 percent to my class each year. A promise of more money for doing what I already do has not been an additional motivator. The paperwork required has made me less effective because I have had to focus my time on duplicating papers for show. (RMC Research Corporation, 2012, p. 44)

Conversely, principals interviewed (N = 17) reported a change in teacher behavior and increased discussions among faculty about student achievement. Some agreed (n = 9) that performance pay evaluation initiatives were an effective way to recruit and retain effective teachers (RMC Research Corporation, 2012).

These divergent views of the VPPI show how administrators and teachers differ regarding pay for performance. This is an important consideration for policy makers interested in pay for performance as part of education reform. Based on the VPPI, the design and implementation of a policy are strong factors to take into account. Timelines are an important factor considering the teacher school year schedule. For example, holding training sessions during the summer may not be an optimal time because in most instances teachers will not contractually obligated to attend. Additionally, implementation of the program is an important factor to analyze, with time built in for teachers and administrators to begin understanding a program and building a system that will result in positive student achievement outcomes and attainment of the incentive for the teachers. The condensed timeline of the VPPI demonstrates the dichotomy between ideology and reality.

Conclusions and Implications for Future Research

When analyzing the gain scores of at least one point between the year before and the year of the incentive for the population of schools invited to participate in the VPPI, it is inconclusive if the VPPI made a difference in student achievement at the school level. Results were inconsistent, and there were negative gain scores for both groups. Additionally, not knowing which teachers participated, the content areas they taught, or the composition of their classes limits the conclusions about test scores at the teacher level. The VPPI was the first initiative that Virginia implemented across the Commonwealth, and the original evaluation did not include data about student test scores. This study was an attempt to add to the evaluation of the program, but more information is needed; for example, information about the number of teachers at each school that participated or why schools did or did not choose to participate. Learning taken from this study is that one year of a program is not enough. Further analysis of motivation theory and adding a qualitative component that includes division leadership is important to consider for future research. Although results are inconclusive, research such as this is important because education policy initiatives are continuously linking student achievement and incentives for teachers. Through the studies, researchers can be a voice for teachers.

Another missing piece in the evaluation of Virginia's teacher incentive programs is whether additional compensation motivates teachers to improve their teaching practices. The conceptual framework of this study is through a motivation paradigm. The assumption is that more money will help spur educators to do better what they intrinsically want to do, which is to prepare a next generation of young people to be

positive contributing members of society. To evaluate incentive policies with a more robust understanding of motivation would help researchers understand the depth and breadth of the rationale behind why educators choose the teacher profession. The valence aspect of the reward system is important to understand because this is the value a teacher associates with the reward. However, digging deeper into the expectancy and instrumentality aspects will enhance an understanding of motivation theory by analyzing teachers' reasoning for the effort he or she associates with performance, if teachers understand how to enhance their performance to help students achieve at higher academic levels, and if teachers believe they superiors will support them to earn awards. As Vroom (1964) states, "There is no way of predicting in advance for a particular organism how much pleasure or pain will be generated with a particular event" (p. 15). Further analysis into the research of Alderfer (1969) will assist researchers into an evaluation of why teachers progress or regress when participating in performance pay programs.

Further in-depth analysis regarding the connection between the motivation of teachers and incentive programs in Virginia would lend insights that could positively impact the implementation of incentive programs. Rather than a continuous reinvention of the same program, positive student achievement gains could be the outcome. A bolder move is to consider the current pay system models. Just as educators are prompted to refine their teaching pedagogy to meet the needs of the 21st century learner, shouldn't our pay system for educators also be refined? What would happen if education policy makers put real power behind overhauling the pay system to reimagine a system that brought pay scales up to the levels of other professions? What types of educators could a system like this attract? Beyond this, would the effects ripple through the surrounding communities?

Could a seamless continuity build between educators and the families they serve? Behind the veil of pay for performance is a call for change, and the courage of education policy makers needs to be fostered to spur this change into action.

Implications for Teaching and Leadership

The final evaluation from the RMC Research Corporation (2012) about the VPPI left open a door for researchers about why teacher and administrator opinions about the VPPI differed. This is an avenue for future researcher because of the dichotomy in opinions about the success of the VPPI; teachers expressed frustration about the program and administrators expressed thoughts about program success. Moreover, further research about what occurs at a school level can be insightful to help shed light on the influences of teaching and learning. As stated above, research in this area can push policy makers to investigate not just pay for performance, but a disruption of the traditional teacher pay scale.

Motivation and Teaching

The current issue of the link between incentive programs and motivation theory was vetted throughout this research study, but further research about the connections is warranted. A greater understanding of how schools as systems operate, from the interpersonal relationships between and among teachers, administrators, and students, to teacher training, to the additional responsibilities teachers carry beyond normal teaching duties can help policy makers create better informed decisions when creating pay for performance programs. To fully understand the environment of a HTS school and explore the motivation behind why teachers choose the teaching profession, an understanding of the cultural identity of the students is necessary. This related to the

expectancy portion of Vroom's framework because a teacher's knowledge, skills, and dispositions connect with expectancy. *Cultural competence* is defined by Hansuvadha and Slater (2012) as the "knowledge, behaviors, dispositions, and policies in a system, organization, or individual to interact effectively with other cultural groups," (p. 174). In their research, Hansuvadha and Slater (2012) conducted a case study of two new administrators and concluded that a strong mentorship program and cross-cultural experiences can help new administrators understand and gain necessary skills required of culturally competent leaders. Alderfer (1969) suggests that individuals need to relate to each other so as to grow as human begins. Although this is a broad conclusion to reach based on a case study of two administrators, it is a call for further research. Cultural competence is a topic pertinent to research regarding incentives and motivation theory because of the factors outside of school that impact student achievement, such as family characteristics and socioeconomic status (Franco & Seidel, 2014). More recently, the release of the 2015 Professional Standards for Educational Leaders, formerly known as the ISLLC Standards, has a new emphasis on student outcomes, and establishes expectations for educational leaders (National Policy Board for Educational Administration, 2015). In addition to the renewed emphasis on using the standards to help shape education policy and expectations for leaders, the standards also include a separate standard that addresses cultural competence. Standard 3 states: "Effective educational leaders strive for equity of educational opportunity and culturally responsive practices to promote each student's academic success and well-being" (National Policy Board for Educational Administration, 2015, p. 11). This is a beginning in the process of understanding the connections between teacher incentives and culturally relevant

teaching and leading. Further research of these components can shed light on using motivation theory in education research to meet the challenge of student achievement in HTS schools. Additionally, it can also uncover the pressures put forth on school personnel to have positive student achievement outcomes.

Shifting Education Policy in Virginia

Immediately following the completion of the VPPI, the policy environment shifted to include more school districts in pay for performance programs in the Commonwealth, and a new program—the SCG—was implemented. The context of achievement of students in HTS schools did not change, but the federal and state resources invested toward the output shifted, thereby impacting the parts of the system. Discontinuation of the SCG after 1 year occurred due to lack of interest from school districts. In the minutes from the April 14, 2014, Virginia Board of Education planning session, the State Superintendent stated that school districts were not interested in the SCG because the amount of the stipend needed to be increased (Commonwealth of Virginia, 2014).

Most recently, a VDOE Superintendent's Memorandum on August 12, 2016, announced a program called the Alternative Teacher Compensation Competitive Planning Grants. The request for proposals outlines that the $5,000 grant is open to school divisions "to explore alternative teacher compensation approaches that move away from tenure-based step increases toward compensation systems based on teacher performance and student progress" and that "priority will be given to schools or school divisions that have not previously explored alternative compensation approaches and that have schools not achieving full accreditation or high numbers of at-risk students needing

qualified teachers in hard to staff subjects" (VDOE, 2016, pp. 3-4). The context of achievement of students in HTS schools has not shifted, but the parts of the system again changed, influencing the system. This new program has its merits because of the efforts to help public school districts find ways to meet the needs of students in the neediest schools. It is difficult to find optimism about this program, however, because of the ways in which it is tying compensation to the initiative. Student achievement is not central to the policy except as an evaluation tool of the teacher, with an outcome of money for improved test scores. These types of initiatives can lead to decreased teacher morale, frustration, and isolation and can build a culture of distrust and competition between teachers and between teachers and administrators; therefore, such a program needs to be carefully designed (Amrein-Beardsley, Pivovarova, & Geiger, 2016). Unfortunately, guidance for such a program is lacking in the Commonwealth's latest attempt to assist districts in recruiting and retaining expert teachers in HTS schools.

Final Thoughts

In remarks to the Hispanic Chamber of Commerce in 2009, President Barack Obama stated:

> From the moment students enter school, the most important factor in their success is not the color of their skin or the income of their parents, it's the person standing at the front of the classroom…America's future depends on its teachers. That is why we are taking steps to prepare teachers for their difficult responsibilities and encouraging them to stay in the profession. That is why we are creating new pathways to teaching and new incentives to bring teachers to schools where they are needed most. (ED, 2011, p. 1)

Public education has continued as it always had since President Obama made his speech in 2009, but the education climate continues to change. There are currently 44 states that have created some type of pay for performance system (Amrein-Beardsley et

al., 2016). Additionally, Every Student Succeeds Act (ESSA) is beginning implementation, and development of new education policy is occurring, leaving researchers with more research to gather and more information to analyze. An additional concern is the fate of students in HTS schools. Will teaching occur by passionate, qualified teachers and led by equally passionate, qualified school leaders? Will public education look the same in four years? On this path to analyze a state incentive program, a realization of the complexity of incentive systems became obvious, as did the shallowness of many programs. The teaching and learning process is multifaceted, with many uncontrollable factors by a teacher or school leader, rendering the fairness of such a program questionable.

The research is rich, but conflicted, about making a connection between student achievement and student test scores. Further analysis about the process of teaching in such programs is warranted to assist with developing an understanding of what motivates teachers to want to work and improve their practice in HTS schools, and how an incentive fits with motivation, namely the expectancy of human motivation, which is the belief that more effort will result in the desired output, increased student performance. This requires further evaluation of teachers' past experiences, personality type, and self-confidence to reveal the complicated layers of what it means to be a teacher. Additionally, analysis of teacher interpersonal relationships at the school level will create an understanding of teacher motivation by revealing the interdependence and value teachers place on those relationships, particularly when analyzing teachers' perceptions of pay for performance policies. Weaving those perspectives in with the study's findings

can create more thoughtful education policy that positively impacts teaching, learning, and leading.

References

Alderfer, C. P. (1969). An empirical test of a new theory of human needs. *Organizational Behavior and Human Performance, 4*(2), 142-175. doi:10.1016/0030-5073(69)90004-x

American Educational Research Association (2015). *AERA Issues Statement on the Use of Value added Models in Evaluation of Educators and Educator Preparation Programs* [Press release]. Retrieved from http://www.aera.net/Newsroom/News-Releases-and-Statements/AERA-Issues-Statement-on-the-Use-of-Value added-Models-in-Evaluation-of-Educators-and-Educator-Preparation-Programs

Amrein-Beardsley, A. (2012). Recruiting expert teachers into high-needs schools: Leadership, money, and colleagues. *Education Policy Analysis Archives*, *20*(27), 1-26.

Amrein-Beardsley, A., Pivovarova, M., & Geiger, T. J. (2016). Value added models: What the experts say. *Phi Delta Kappan, 98*(2), 35-40. doi:10.1177/0031721716671904

Ballou, D. & Springer, M. (2015). Using student test scores to measure teacher performance: Some problems in the design and implementation of evaluation systems. *Educational Researcher, 4* (2), 77–86. doi: 10.3102/0013189X15574904

Barela, E. (2008). Title I achieving schools study. *American Journal of Evaluation*, *29*(4), 531-533. doi: 10.1177/1098214008328405

Belfield, C., & Heywood, J. (2008). Performance pay for teachers: Determinants and consequences. *Economics of Education Review*, 243-252. doi: 10.1016/j.econedurev.2008.01.002

Berry, B. (2004). Recruiting and retaining "highly qualified teachers" for hard to staff schools. *National Association of Secondary School Principals. NASSP Bulletin*88, (638) (03): 5-27, http://proxygw.wrlc.org/login?url=http://search.proquest.com/docview/216039418?accountid=11243 (accessed April 21, 2016).

Brodsky, A., DeCesare, D., & Kramer-Wine, J. (2010). Design and implementation considerations for alternative teacher compensation programs. *Theory Into Practice,* 49, 213-222. doi: 10.1080/00405841.2010.487757

Burke, P., & Lind, K. G. (1987). Performance assessment techniques for teacher career ladder advancement. NASSP Bulletin, 71(500), 27-35. doi:10.1177/019263658707150005

Butin, D. W. (2010). *The education dissertation: A guide for practitioner scholars*. Thousand Oaks, CA: Corwin.

Caillier, J. (2010). Paying teachers according to student achievement: Questions regarding pay-for performance models in public education. *The Clearing House*, 89, 58-61. Doi: 10.1080/00098650903386451

Clotfelter, C., Glennie, E., Ladd, H., & Vigdor, J. (2008). Teacher bonuses and teacher retention in low-performing schools: Evidence from the north carolina $1,800 bonus program. *Public Finance Review*, *36*(1), 63-87. doi: 10.1177//1091142106291662

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

Cohen, D., & Moffitt, S. (2009). *The ordeal of equality: Did federal regulation fix the schools?* Cambridge: President and Fellows of Harvard College.

Commonwealth of Virginia. (2016, December 21). *Virginia performs: Measuring poverty in virginia*. Retrieved January 3, 2018, from http://vaperforms.virginia.gov/indicators/economy/poverty.php

Community Training and Assistance Center. (2016). Getting the pieces right: Professional development, compensation, and school-wide performance. Teacher incentive performance award prince william county public schools. Retrieved January 3, 2018, from http://www.ctacusa.com/wp-content/uploads/2016/12/Getting-the-Pieces-Right-2016.pdf

Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational Researcher, 44(2),* 132-137. doi: 10.3102/0013189X15575346

Dickey, K. (2013). *Overview of standards of funding process* [PowerPoint slides]. Retrieved from www.doe.virginia.gov/school_finance/.../soq_funding_presentation.ppt

Driscoll, L., & Salmon, R. (2013). Challenges confronting public elementary and secondary education in the commonwealth of virginia. *Journal of Education Finance, 38*(3), 230-254. Retrieved from http://www.jstor.org.proxygw.wrlc.org/stable/23354865

Driscoll, L., & Salmon, R. (2008). How increased state equalization aid resulted in greater disparities: An unexpected consequence for the Commonwealth of Virginia. *Journal of Education Finance*, *33*(3), 238-261.

Edley, C. & Mariano-Florentino, C. (2013). U.S. Department of Education, The Equity and Excellence Commission. *For each and every child: A strategy for education equity and excellence—A strategy for education equity and excellence.* Washington, D.C.

Every Student Succeeds Act of 2015, 114 U.S.C. § 1177 (2015a). Retrieved from https://www.gpo.gov/fdsys/pkg/BILLS-114s1177enr/pdf/BILLS-114s1177enr.pdf

Every Student Succeeds Act of 2015, Final Consensus-Based Language, 114 U.S.C. §200.2 (2015b). Retrieved from http://www2.ed.gov/policy/elsec/leg/essa/title1a-assessment-consensus-regulatory-lang.pdf

Fact Sheet: Congress acts to fix no child left behind. (2015). Retrieved May 20, 2016, from https://www.whitehouse.gov/the-press-office/2015/12/03/fact-sheet-congress-acts-fix-no-child-left-behind

Figlio, D. & Kenny, L. (2007). Individual teacher incentives and student performance. *Journal of Public Economics. 91,* 901-914. doi: 10.1016/j.pubeco.2006.10.001

Finnigan, K., & Gross, B. (2007). Do accountability policy sanctions influence teacher motivation? Lessons from chicago's low-performing schools. *American Educational Research Journal*, *44*(3), 594-629. doi: 10:3102/0002831207306767

Fraenkel, J. R., & Wallen, N. E. (2006). *How to design and evaluate research in education*. Boston: McGraw-Hill.

Francis, A. (2017, November 25). Compare and Contrast Herzberg's Two Factor Theory with Vroom's Expectancy Theory of Motivation. Retrieved March 1, 2018, from https://www.mbaknol.com/management-principles/herzbergs-two-factor-theory-vs-vrooms-expectancy-theory-of-motivation/

Franco, M. S., & Seidel, K. (2014). Evidence for the Need to More Closely Examine School Effects in Value added Modeling and Related Accountability Policies. *Education and Urban Society, 46*(1), 30-58. doi:10.1177/0013124511432306

Fryer, R. (2013). Teacher incentives and student achievement: Evidence from new york city public schools. *Journal of Labor Economics*, *31*(2), 373-407.

Gardner, D. P. U.S. Department of Education, National Commission on Excellence in Education. (1983). *A Nation at Risk: The imperative for educational reform. An open letter to the American people. A letter to the nation and the secretary of education.* (065-000-00177-2). Washington, DC: Superintendent of Documents: Government Printing Office.

Goldhaber, D. (2009). The politics of teacher pay reform. M. G. Springer (Editor), *Performance incentives: Their growing impact on american K-12 education* (pp. 25-42). Washington, D.C.: The Brookings Institution.

Goldhaber, D., Gross, B., & Player, D. (2010). Teacher career paths, teacher quality, and persistence in the classroom: Are public schools keeping their best? *Journal of Policy Analysis and Management, 30(157-87), doi: 10.1002/pam.20549*

Gratz, D. (2011). Performance pay path to improvement. *Kappa Delta Pi Record, 47*(4), 156-161. doi: 10.1080/00228958.2011.10516582

Greenlee, B., & Brown, J. (2009). Retaining teachers in challenging schools. *Education, 130*(1), 96-109.

Hansuvadha, N. & Slater, C. (2012). Culturally competent school leaders: The individual and the system. *The Educational Forum, 76*, 174-189. doi: 10.1080/00131725.2012.653094

Hezel Associates. (2009). *Quality compensation for teachers summative evaluation.* Syracuse, NY: Author.

Hogge, Susan L. (2013, November 20). Strategic compensation grants (scg) initiative update (House Appropriations Committee). Retrieved from http://hac.virginia.gov/committee/files/2013/11-20-13/II_a--Strategic_Compensation_Grants.pdf)

Houck, E., & Kurtz, A. (2010). Resource distribution and graduation rates in sreb states: An overview. *Peabody Journal of Education*, *85*, 32-48. doi: 10.1080/01619560903523698

Ingersoll, R.M. & Merrill, E. (2011). The status of teaching as a profession. In J. Ballantine and J. Spade (Eds.), *Schools and Society: A Sociological Approach to Education*. (p. 185-189) 4th Ed. CA: Pine Forge Press/Sage Publications.

Ingersoll, R., & Merrill, L. (2010). Who's teaching our children? *Educational Leadership*, *6*(8), 14-20.

Joyce, P., & Pattison, S. (2010). Public budgeting in 2020: Return to equilibrium or continued mismatch between demands and resources? *Public Administration Review*, *70*(s1), 524-532.

Kelley, C., Heneman, H., & Milanowski, A. (2002). Teacher motivation and school-based performance awards. *Educational Administration Quarterly*, *38*(3), 372-402. doi: 10.1177/0013161X02383004

Ko, Y.J., Rhee, Y.C., Walker, M., & Lee, J.H. (2013). What motivates donors to athletic programs: A new model of donor behavior. *Nonprofit and Voluntary Sector Quarterly 43(3)*, 523-546. doi: 10.1177/0899764012472065

Kolbe, T., & Strunk, K. (2012). Economic incentives as a strategy for responding to teacher staffing problems: A typology of policies and practices. *Educational Administration Quarterly, 48*(5), 779-813. doi: 10.1177/0013161X12441011

Konstantopoulos, S. (2014). Teacher effects, value added models, and accountability. *Teachers College Record, 116*, 1-21.

Levačić, R. (2009). Teacher incentives and performance: An application of principal-agent theory. *Oxford Development Studies, 37*(1), 33-46. doi: 10.1080/1360081002660844

Liu, G. (2008). Improving title I funding equity across states, districts, and schools. *Iowa Law Review, 93*(3), 973-1013.

Loeb, S., Miller, L. C., & Strunk, K. O. (2009). The state role in teacher compensation. *Education Finance and Policy, 4*(1), 89-114. Retrieved February 29, 2016, from https://cepa.stanford.edu/sites/default/files/stateRole.pdf.

Marsh, J., & McCaffrey, D. (2012). What are achievement gains worth to teachers? *Phi Delta Kappan*, *93*(4), 57-59.

Martin, S. (2010). The determinants of school district salary incentives: An empirical analysis of, where and why. *Economics of Education Review, 29*, 1143-1153. doi: 10.1026/j.econedurev.2010.06.007

Miles, K. (2011). Transformation or decline? Using tough times to create higher performing schools. *Phi Delta Kappan, 93*(2), 42-46.

Minnesota Department of Education. (2016). *Q-comp.* Retrieved from http://education.state.mn.us/MDE/SchSup/QComp/

National Center for Education Statistics. (2015). *Teaching vacancies and difficult-to-staff teaching positions in public schools*. Retrieved from https://nces.ed.gov/pubs2015/2015065.pdf

National Center for Education Statistics. (2012). *Schools and staffing survey.* Retrieved from https://nces.ed.gov/surveys/sass/

National Center for Education Evaluation and Regional Assistance (2016). *Impact evaluation of the teacher incentive fund* (2016). Retrieved from https://ies.ed.gov/ncee/projects/evaluation/tq_incentive.asp

No Child Left Behind (NCLB) Act of 2001, 20 U.S.C.A. § 6301 *et seq.* (West 2003).

Odden, A. (2001). Enhancing teacher quality through knowledge- and skills-based pay. In: CPRE Policy Briefs. November 2001. [Online]. [Cit. 04-20-2016]. Available from: Consortium for Policy Research in Education Web site: <http://www.cpre.org/images/stories/cpre_pdfs/rb34.pdf>.

Odden, A., Picus, L., & Goetz, M. (2010). A 50-state strategy to achieve school finance adequacy. *Educational Policy*, *24*(4), 628-654. doi: 10.1177/0895904809335107

Opfer, D. (2011). Defining and identifying hard to staff schools: The role of school demographics and conditions. *Educational Administration Quarterly, 47(4),* 582-619. doi:10.1177/0013161x11400598

Paufler, N. A. & Amrein-Beardsley, A. (2014). The random assignment of students into elementary classrooms: Implications for value added analyses and interpretations. *American Educational Research Journal, 51(2),* 328-362. doi: 10.3102/0002831213508229

Petty, T., Fitchett, P., & O'Connor, K. (2012). Attracting and keeping teachers in high-need schools. *American Secondary Education*, *40*(2), 67-88.

Podgursky, M. (2006). Is there a "qualified teacher" shortage? *Education Next 6, (2),* Retrieved from http://proxygw.wrlc.org/login?url=http://search.proquest.com/docview/1237806681?accountid=11243 (accessed April 20, 2016).

Pyle, C. P. (2013). Governor mcdonnell awards $4.5 million in strategic teacher compensation grants. [Press release]. Retrieved from http://www.doe.virginia.gov/news/news_releases/2013/aug09_gov.shtml.

Rice, Malen, Jackson, & Hoyer. (2014). Time to pay up: Analyzing the motivational potential of financial awards in a tif program. *Educational Evaluation and Policy Analysis,* 1-21. doi: 10.3102/0162373714524622

Rice, J., Malen, B., Baumann, P., Chen, E., Doughtery, A., Hyde, L., Jackson, C., Jacobson, R., & McKithen, C. (2012). The persistent problems and confounding challenges of educator incentives: The case of tif in prince george's county, maryland. *Educational Policy*, *26*(6), 892-933. doi: 10.1177/0895904812465708

RMC Research Corporation. (2012). *Evaluation report: Governor's virginia performance pay incentives initiative pilot.* Arlington, VA: Author.

Roch, C., & Howard, R. (2008). State policy innovation in perspective: Courts, legislatures, and education finance reform. *Political Research Quarterly*, *61*(2), 333-344.

Rosenholtz, S. J., & Smylie, M. A. (1984). Teacher compensation and career ladders. *The Elementary School Journal, 85(2),* 149–166. Retrieved from http://www.jstor.org.proxygw.wrlc.org/stable/1001643

Salmon, R. (2010). The evolution of virginia public school finance: From the beginnings to today's difficulties. *Journal of Education Finance*, *36*(2), 143-161.

Smarick, A. (2011). Diplomatic mission. *Education Next*, *11*(1), 56-62.

Smith, E. (2008). *Using secondary data in educational and social research*. Maidenhead: Open University Press.

Springer, M. G., Swain, W. A., & Rodriguez, L. A. (2014). Effective teacher retention bonuses: Evidence from tennessee. *Educational Evaluation and Policy Analysis, 38*(2), 199-221. doi:10.3102/0162373715609687

Streams, M., Butler, J., Cowen, J., Fowles, J., & Toma, E. (2011). School finance reform: Do equalized expenditures imply equalized teacher salaries? *Education Finance and Policy*, *6*(4), 508-536.

Stronge, J., Gareis, C., & Little, C. (2006). *Teacher pay and teacher quality: Attracting, developing, and retaining the best teachers*. Thousand Oaks: Corwin Press.

Strunk, K. O. & Zeehandelaar, D. (2011). Differentiated compensation: How california school districts use economic incentives to target teachers. *Journal of Education Finance 36(3),* 268-293. University of Illinois Press. Retrieved April 21, 2016,

U.S. Department of Education (2017a.) *Every student succeeds act final regulations: Accountability, state plans, and data reporting.* Retrieved March 8, 2017, from https://www.ed.gov/essa?src=rn

U.S. Department of Education (2017b). *Awards -- Teacher Incentive Fund.* Retrieved

September 20, 2017, from https://www2.ed.gov/programs/teacherincentive/awards.html

U.S. Department of Education (2016a, December 29). *School improvement grants.* Retrieved February 22, 2017, from https://www2.ed.gov/programs/sif/index.html

U.S. Department of Education. (2016b, December 21). *Teacher incentive fund*. Retrieved March 02, 2017, from https://ed.gov/programs/teacherincentive/index.html

U.S. Department of Education (2011). *Our future, our teachers: The obama administration's plan for teacher education reform and improvement*. Retrieved September 8, 2016, from http://www.2ed.gov/inits/ed/index/html

U.S. Department of Education (2010a, June 22). *Race to the top—game-changing reforms.* Retrieved from http://www.ed.gov/open/plan/race-top-game-changing-reforms.

U.S. Department of Education (2010b). *Application for New Grants Under the Teacher Incentive Fund Program.* Retrieved from https://www2.ed.gov/programs/teacherincentive/apps/a100094.pdf

U.S. Department of Education. (2006, November 14). *Teacher incentive fund: Notice inviting applications for new awards for fiscal year (FY) 2006, fed. reg. E6-19193.* Retrieved from https://www2.ed.gov/legislation/FedRegister/announcements/2006-4/111406a.html

U.S. Department of Education. (2004, September 15). *Title I — improving the academic achievement of the disadvantaged*. Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/pg1.html

U.S. Department of Labor Bureau of Labor Statistics. (2015, January 23). *Union membership—annual news release.* Retrieved from http://www.bls.gov/news.release/union2.htm

Virginia Department of Education (n.d.). School Report Card. Retrieved July 08, 2016, from http://www.doe.virginia.gov/statistics_reports/school_report_card/

Virginia Department of Education (2017a). *Accreditation and federal reports*. Retrieved May 15, 2017, from http://www.doe.virginia.gov/statistics_reports/accreditation_federal_reports/

Virginia Department of Education (2017b). *School accreditation ratings*. Retrieved February 27, 2017, from http://www.doe.virginia.gov/statistics_reports/accreditation_federal_reports/accreditation/index.shtml#descriptions

Virginia Department of Education (2016a). *Teaching in virginia.* Retrieved July 10, 2016, from http://www.doe.virginia.gov/teaching/

Virginia Department of Education (2016b). *Description of cut scores posted on the virginia department of education website.* Retrieved from http://www.doe.virginia.gov/testing/scoring/explanation_for_cut%20scores.pdf

Virginia Department of Education (2015a). *Accountability in virginia public schools*. Retrieved from

http://doe.virginia.gov/statistics_reports/school_report_card/accountability_guide.pdf

Virginia Department of Education (2015b). *Frequently asked questions: Revised standards of accreditation*. Retrieved from http://www.doe.virginia.gov/statistics_reports/accreditation_federal_reports/accreditation/soa_faq.pdf

Virginia Department of Education. (2013). *Strategic compensation grants (scg) initiative competitive grants.* Retrieved from http://hac.virginia.gov/committee/files/2013/11-20-13/II_a--Strategic_Compensation_Grants.pdf

Virginia Department of Education. (2012a). *Graduation requirements*. Retrieved from http://www.doe.virginia.gov/instruction/graduation/index.shtml

Virginia Department of Education. (2012b). *Technical report 2011-2012 administration cycle*. Retrieved from http://www.doe.virginia.gov/testing/test_administration/technical_reports/sol_technical_report_2011-12_administration_cycle.pdf

Virginia Department of Education. (2012c). *2011-2012 mathematics standards of learning testing: Start of a new trend line in student achievement* [Press Release]. Retrieved from http://www.doe.virginia.gov/news/news_releases/2012/feb22.shtml

Virginia Department of Education, Division of Teacher Education and Licensure. (2011-2012). *Virginia performance pay incentives (vppi) in hard to staff schools*

competitive grant application packet. Richmond, VA: Virginia Department of Education.

Virginia Department of Education (2011a). *Adequate yearly progress (ayp) reports*. Retrieved from http://www.doe.virginia.gov/statistics_reports/accreditation_federal_reports/ayp/

Virginia Department of Education (2011b). *Explanation for sol cut scores.* Retrieved from http://www.doe.virginia.gov/testing/scoring/index.shtml#cutscores

Virginia Department of Education (2011c). *Governor mcdonnell announces K-12 education initiatives for 2011 general assembly session* [Press release]. Retrieved from http://www.doe.virginia.gov/news/news_releases/2011/jan18_gov.shtml

Virginia Department of Education (2011d). *Guidelines for creating a comprehensive pay system.* Retrieved from http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwj8hY7_9IvVAhWK8CYKHVbPBRIQFggmMAA&url=http%3A%2F%2Fwww.doe.virginia.gov%2Fteaching%2Fperformance_evaluation%2Fteacher%2Ftraining_phase1%2Fmaterials%2Fappendix_b_compensation_models.doc&usg=AFQjCNFOA83kPsY5zzKZR4htL4qpk3Ly8w

Virginia Department of Education (2011e). *Virginia performance pay incentives (vppi) in hard to staff schools competitive grant application packet.* Retrieved from http://www.doe.virginia.gov/administrators/superintendents_memos/2011/113-11a.pdf

Virginia Department of Education (2011f). *Virginia sol assessment build-a-table.* Retrieved from http://bi.virginia.gov/BuildATab/rdPage.aspx

Vroom, V. (1964). *Work and motivation.* San Francisco, California: Jossey-Bass.

Wellington, A., Chiang, H., Hallgren K., Speroni C, Herrmann M., Burkander P. (2016). *Evaluation of the teacher incentive fund: Implementation and impacts of pay -for-performance after three years: Executive summary.* (NCEE 2016- 4005). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from https://ies.ed.gov/ncee/pubs/20164004/pdf/20164005.pdf

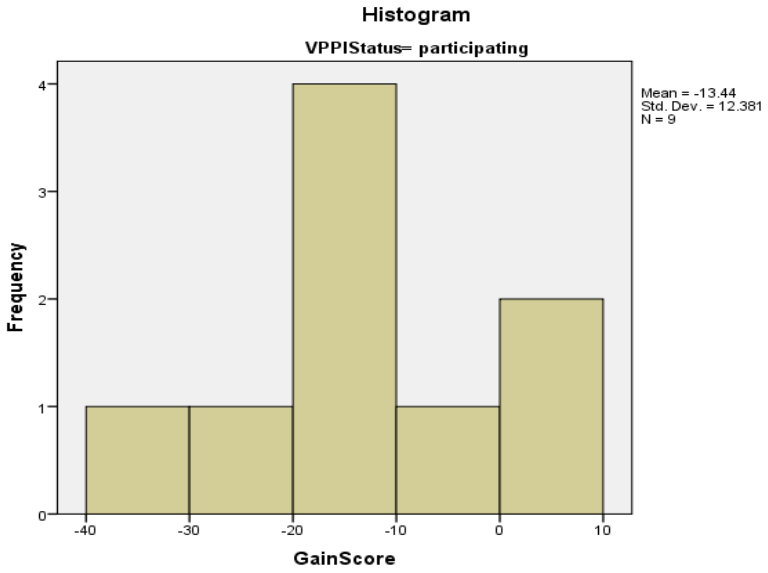
Wiley, E., Spindler, E. & Subert, A. (2010). *Denver procomp: An outcomes evaluation of denver's alternative teacher compensation system.* University of Colorado at Boulder School of Education: Author.

Yuan, K., Le, V., McCaffrey, D., Marsh, J., & Hamilton, L., Stecher, B., & Springer, M. (2013). Incentive pay programs do not affect teacher motivation or reported practices: Results from three randomized studies. *Educational Evaluation and Policy Analysis, 35*(1), 3-22.

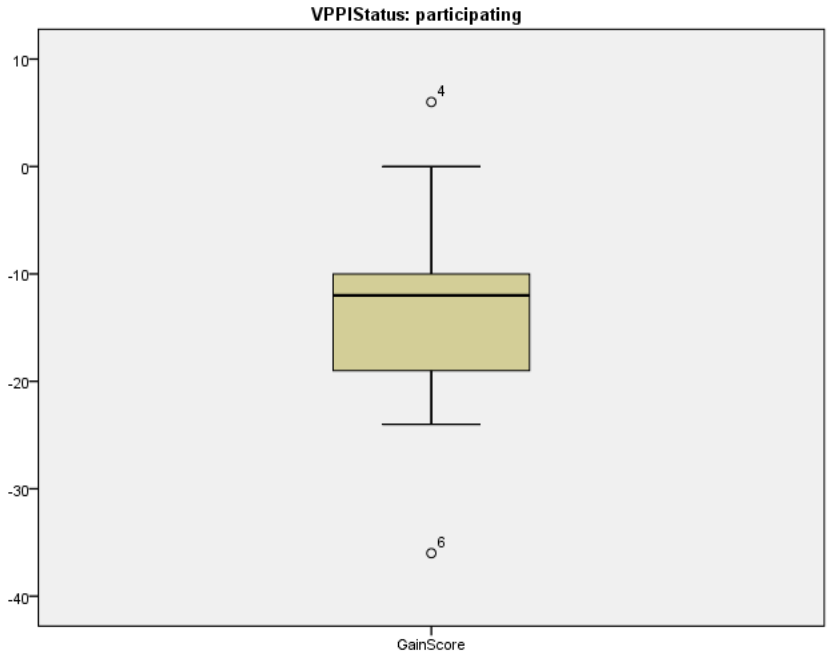
Appendix A

Grade 5 Histogram, Box Plots, and Frequency Polygons

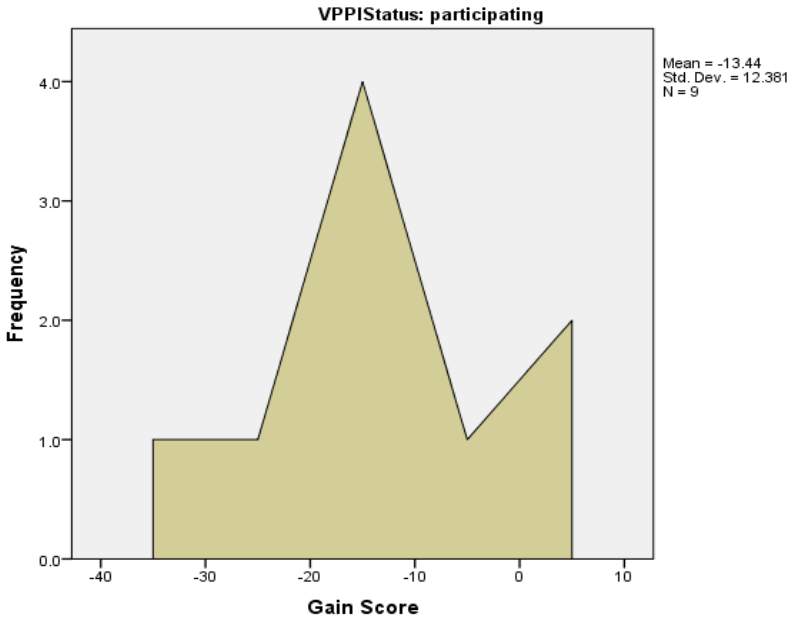
Grade 5 Math Histogram VPPI Participating



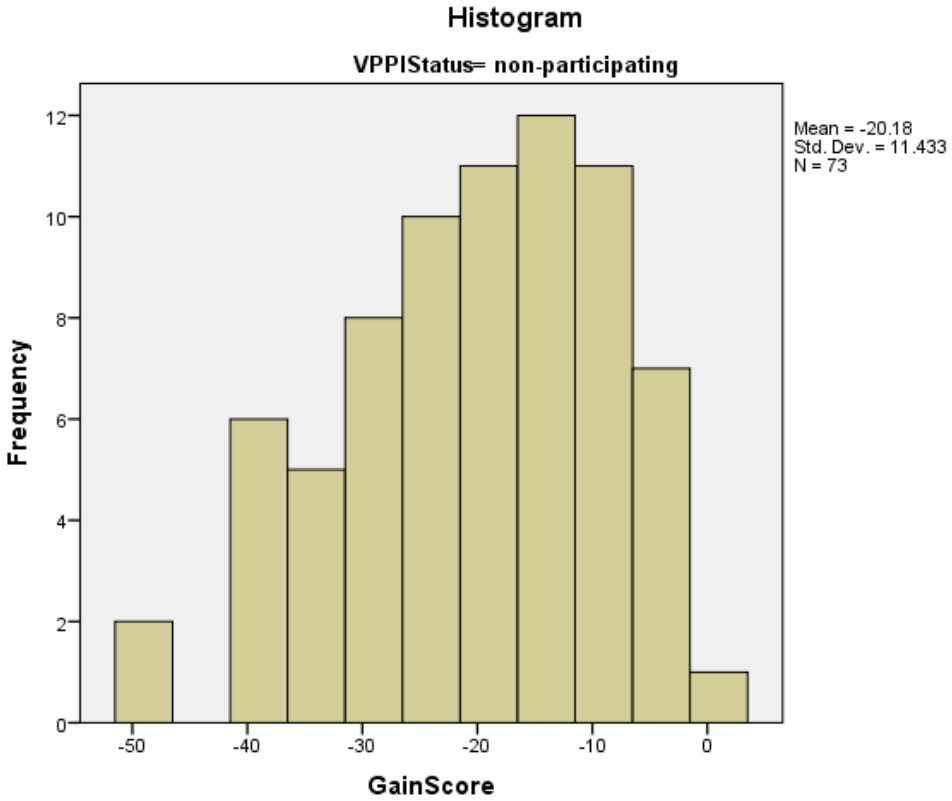
Grade 5 Math Box Plot VPPI Participating



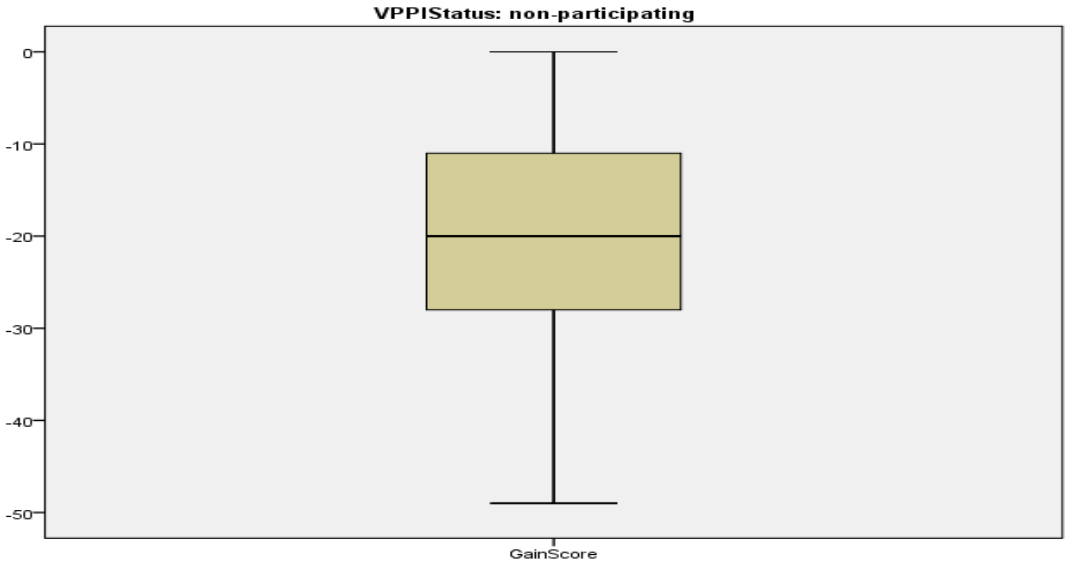
Grade 5 Math Frequency Polygon VPPI Participating



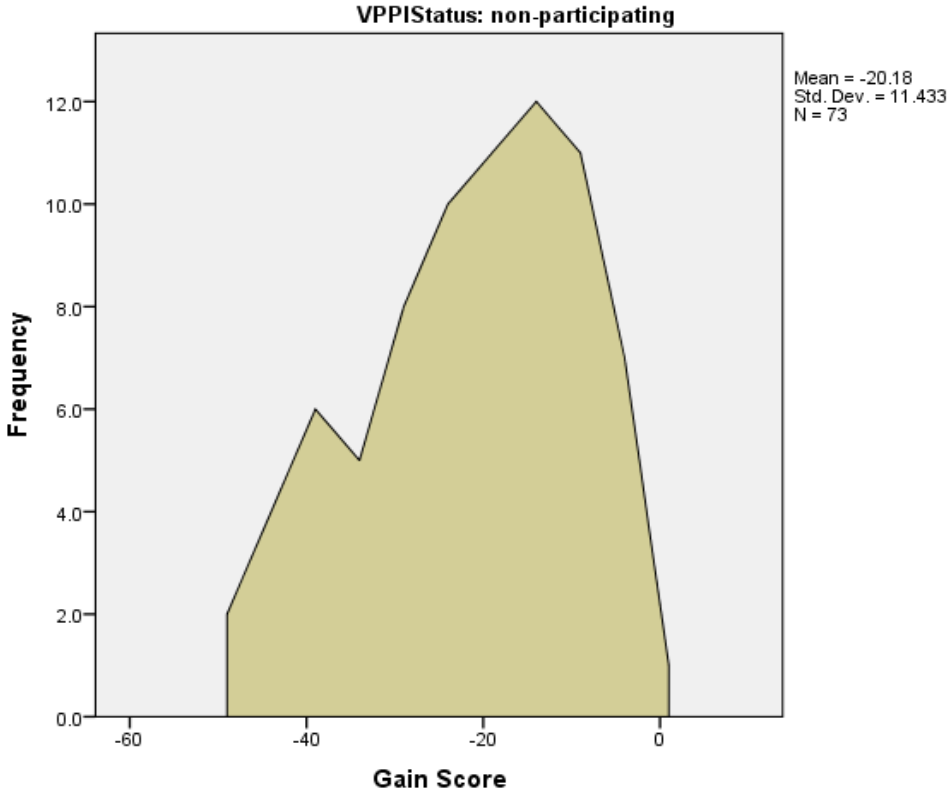
Grade 5 Math Histogram VPPI Nonparticipating



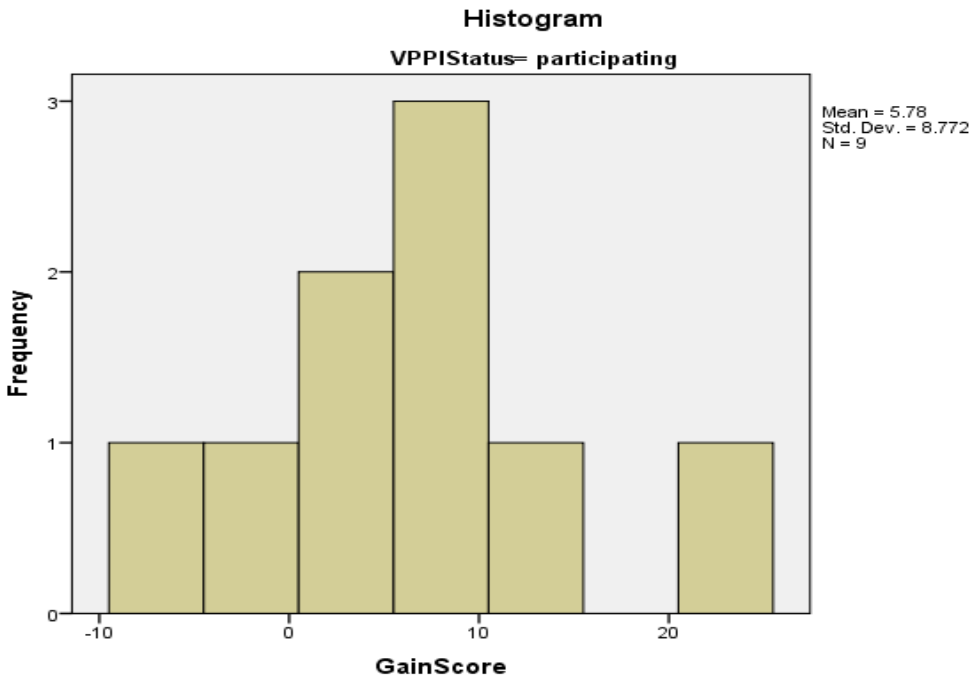
Grade 5 Math Box Plot VPPI Nonparticipating



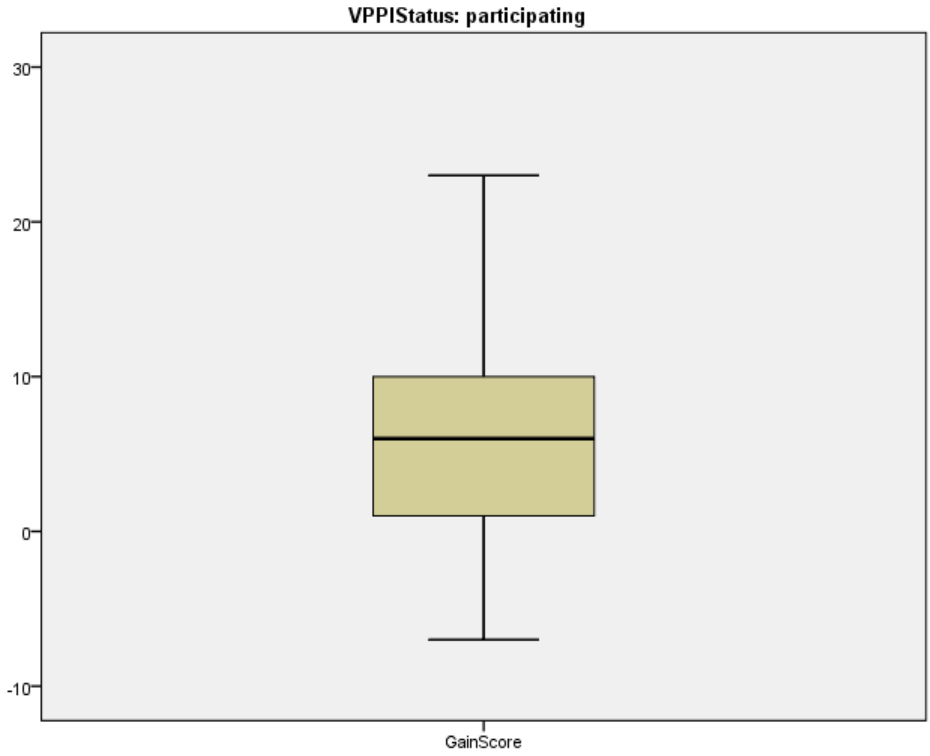
Grade 5 Math Frequency Polygon VPPI Nonparticipating



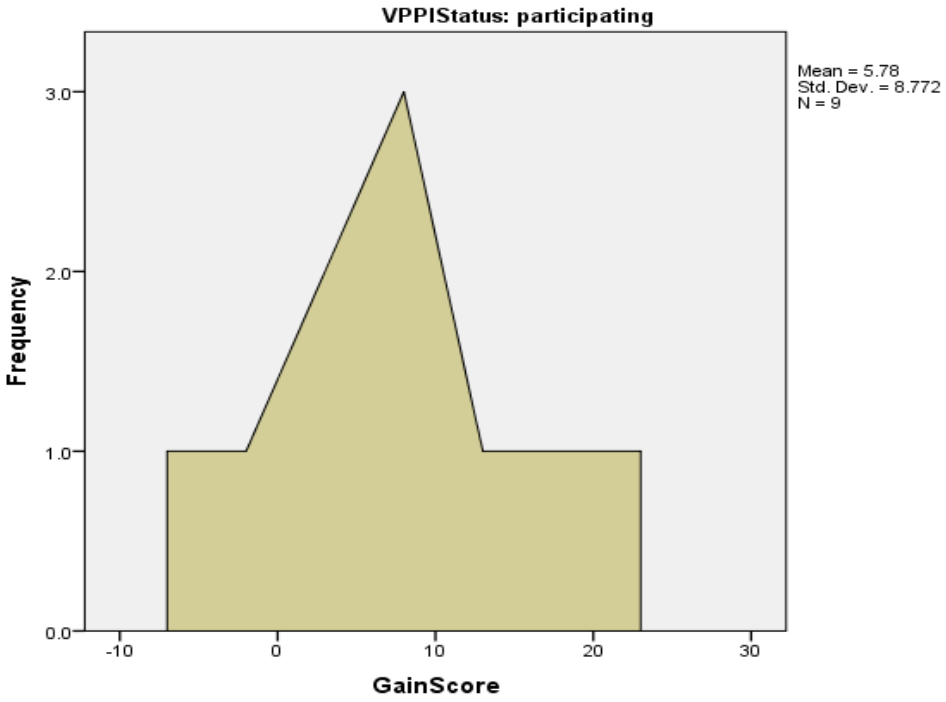
Grade 5 Reading Histogram VPPI Participating



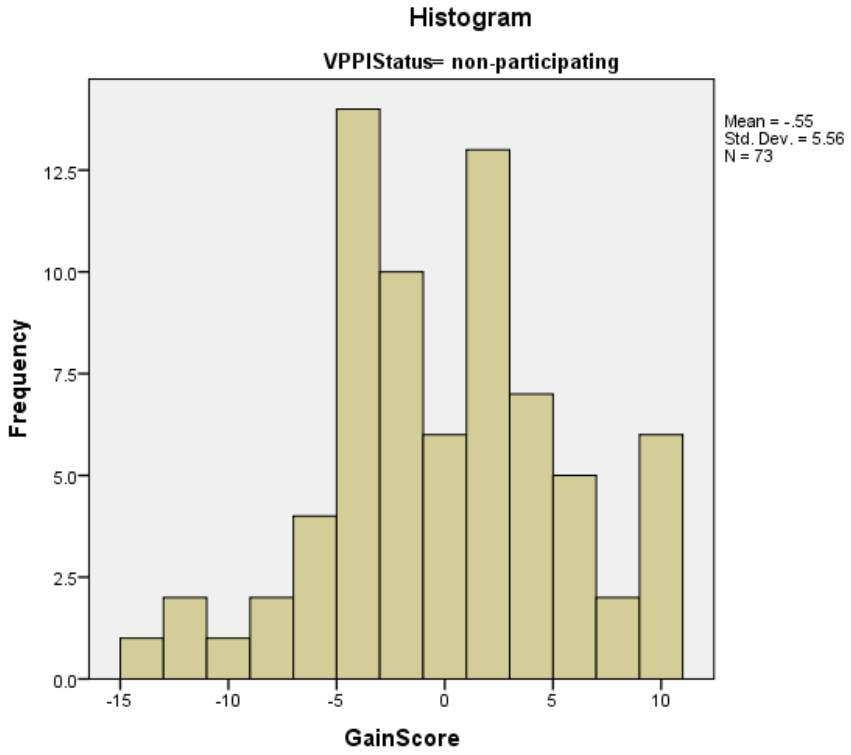
Grade 5 Reading Box Plot VPPI Participating



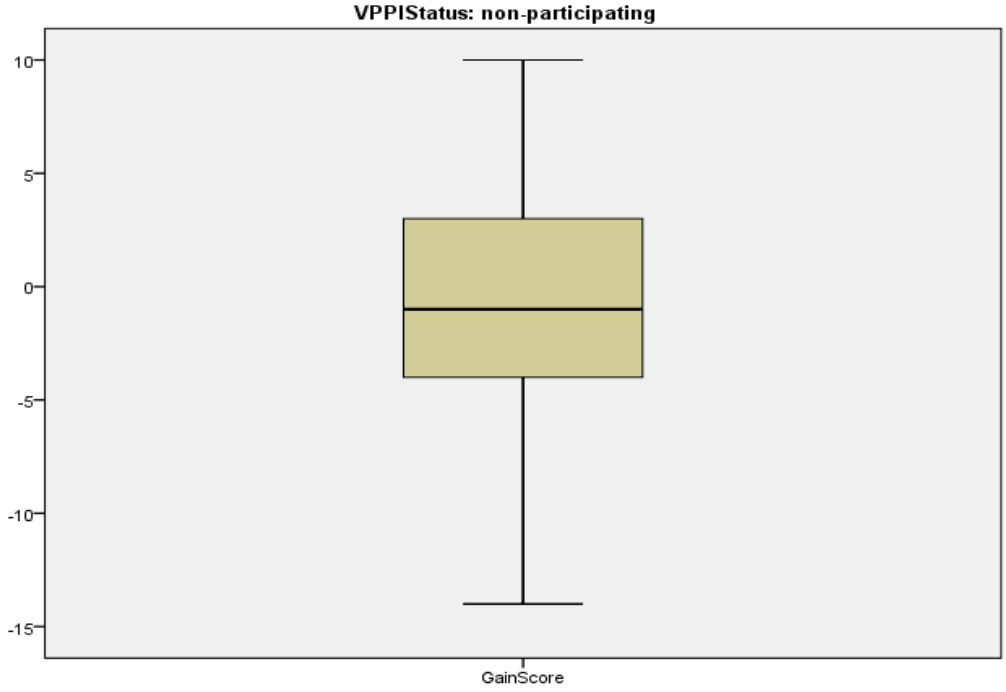
Grade 5 Reading Frequency Polygon VPPI Participating



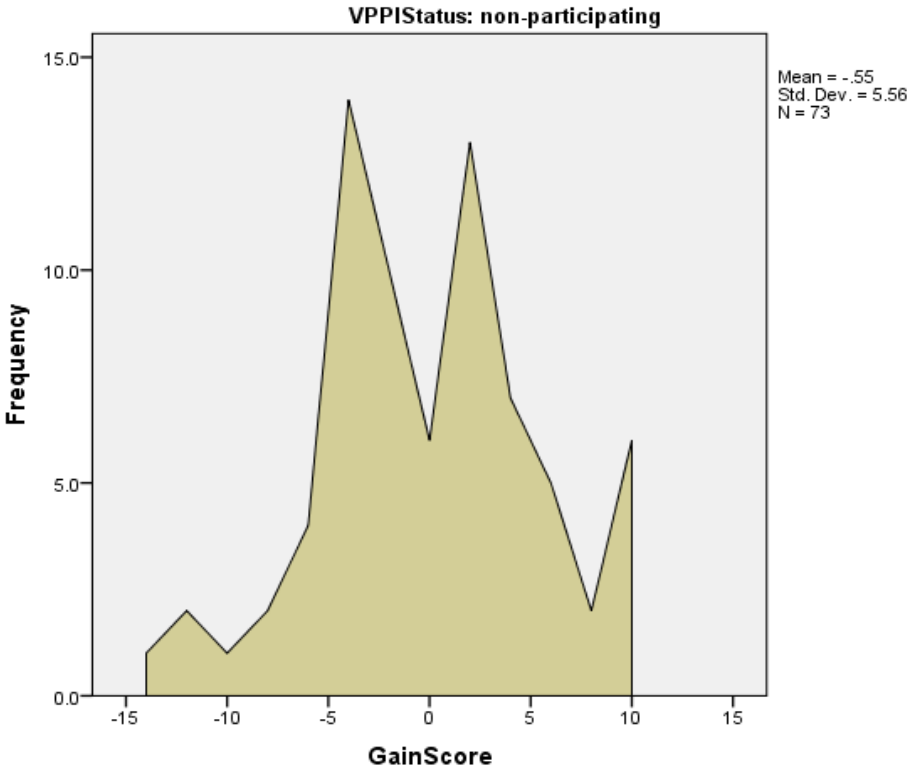
Grade 5 Reading Histogram VPPI Nonparticipating



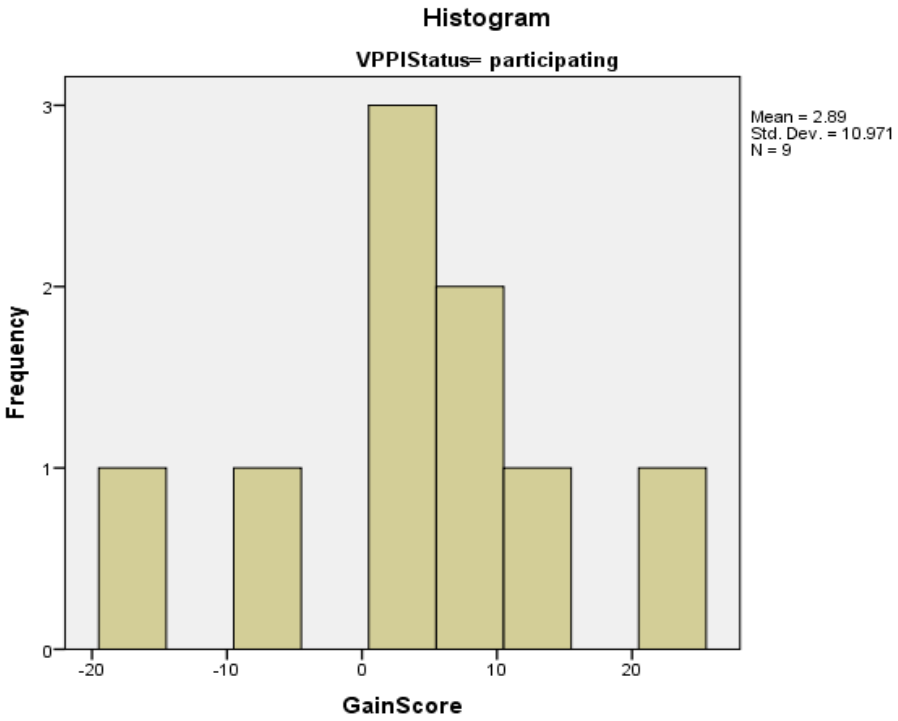
Grade 5 Reading Box Plot VPPI Nonparticipating



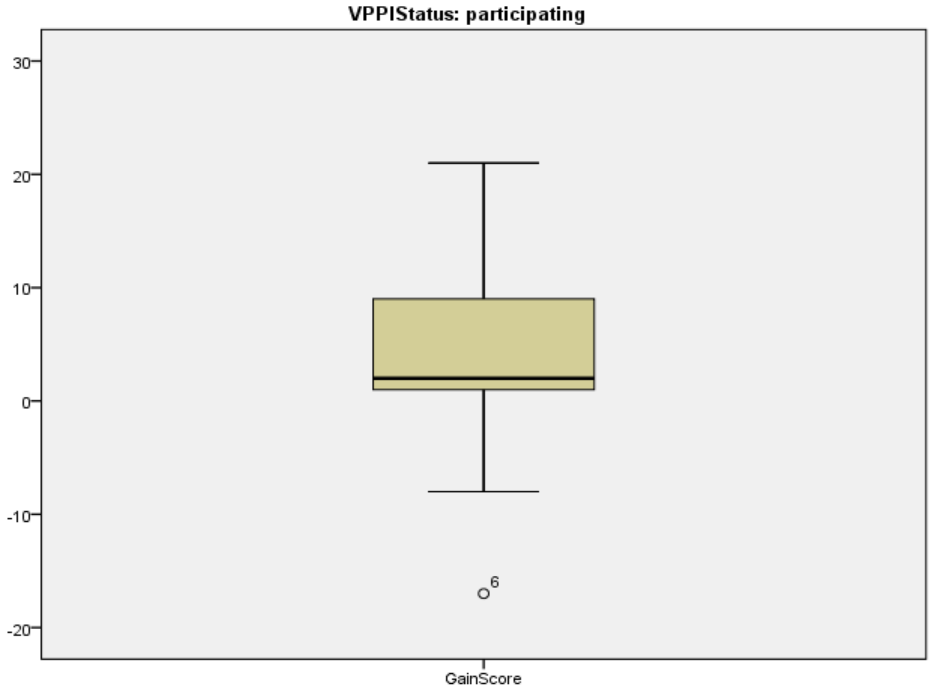
Grade 5 Reading Frequency Polygon VPPI Nonparticipating



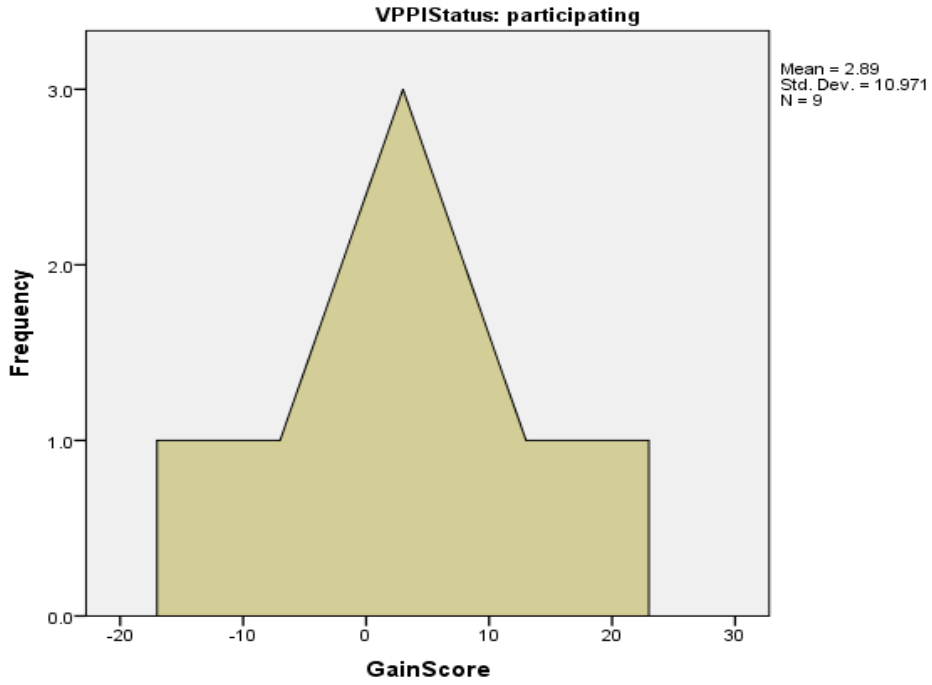
Grade 5 Science Histogram VPPI Participating



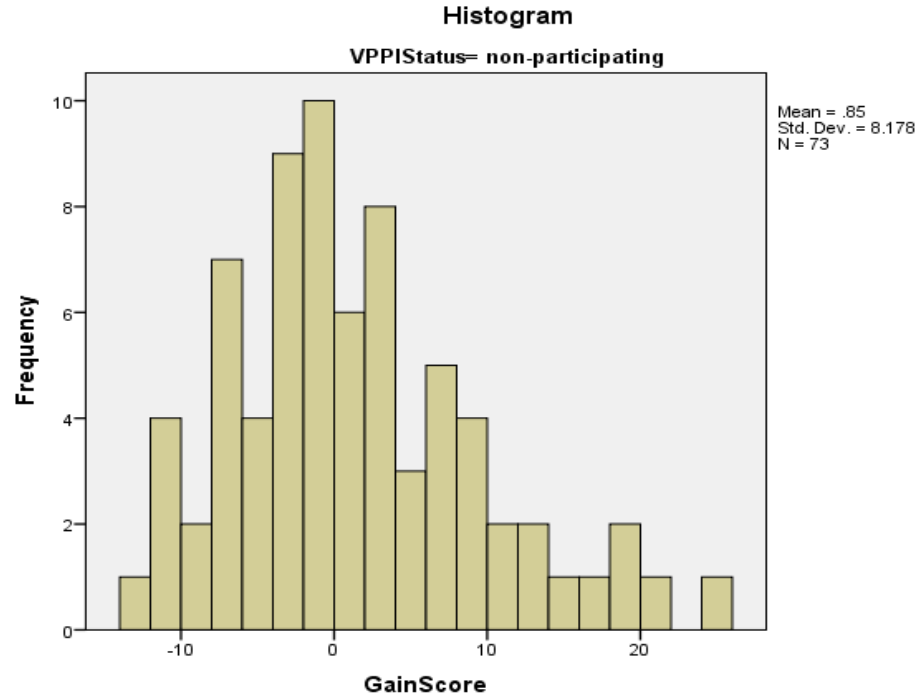
Grade 5 Science Box Plot VPPI Participating



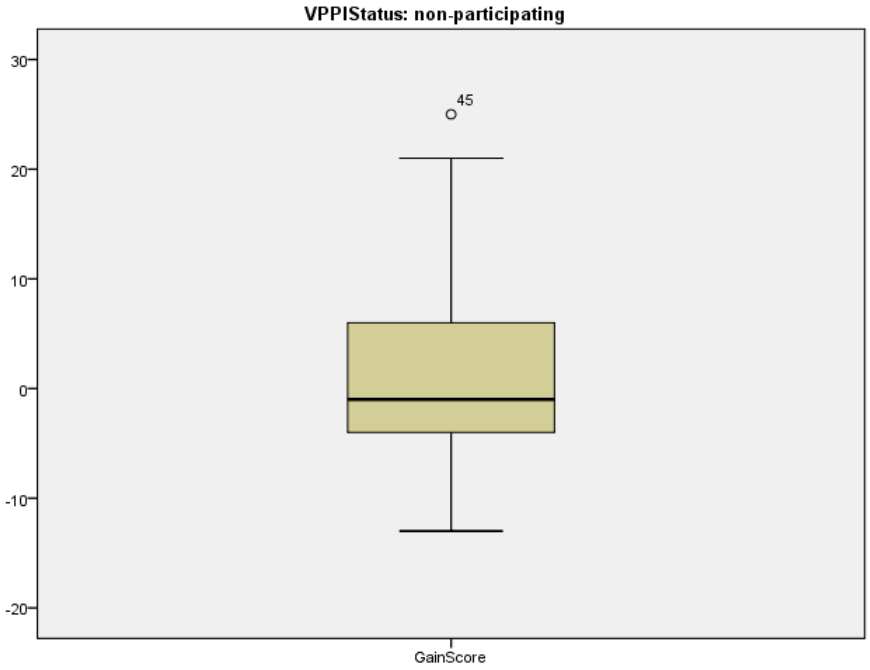
Grade 5 Science Frequency Polygon VPPI Participating



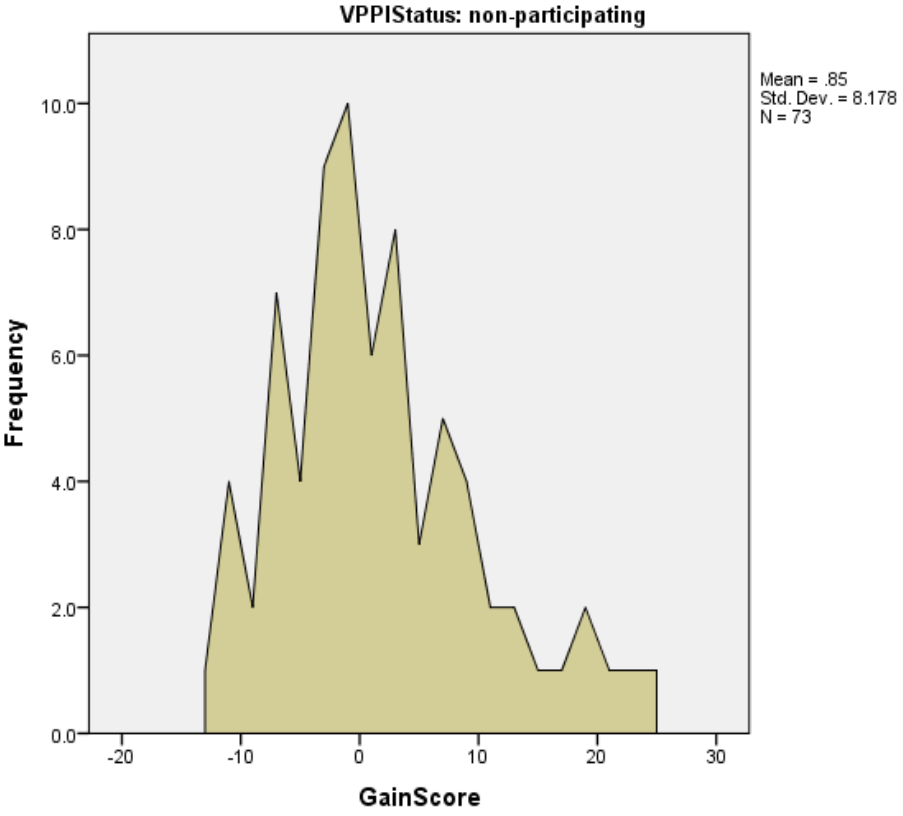
Grade 5 Science Histogram VPPI Nonparticipating



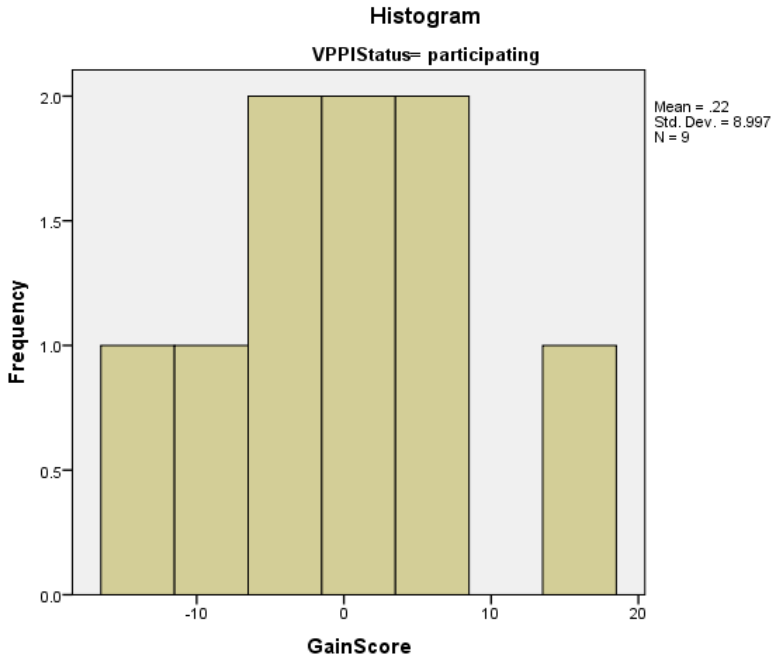
Grade 5 Science Box Plot VPPI Nonparticipating



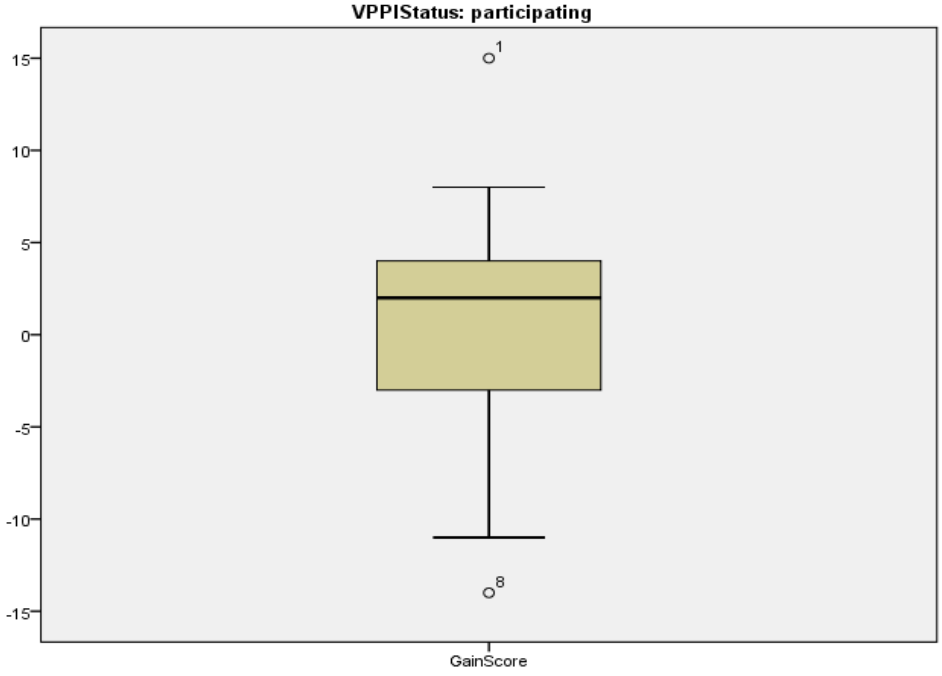
Grade 5 Science Frequency Polygon VPPI Nonparticipating



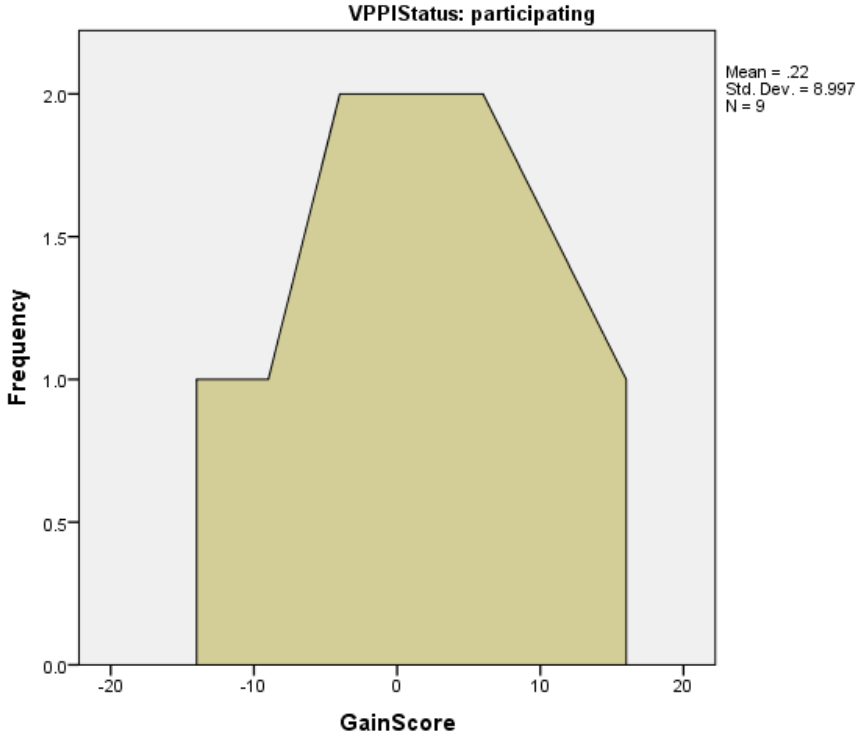
Grade 5 Writing Histogram VPPI Participating



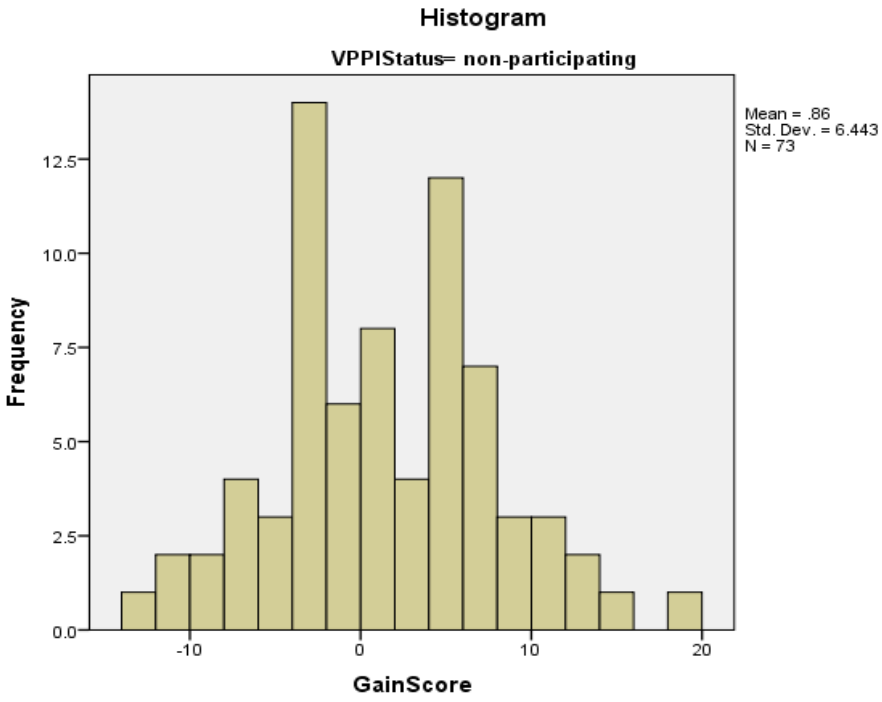
Grade 5 Writing Box Plot VPPI Participating



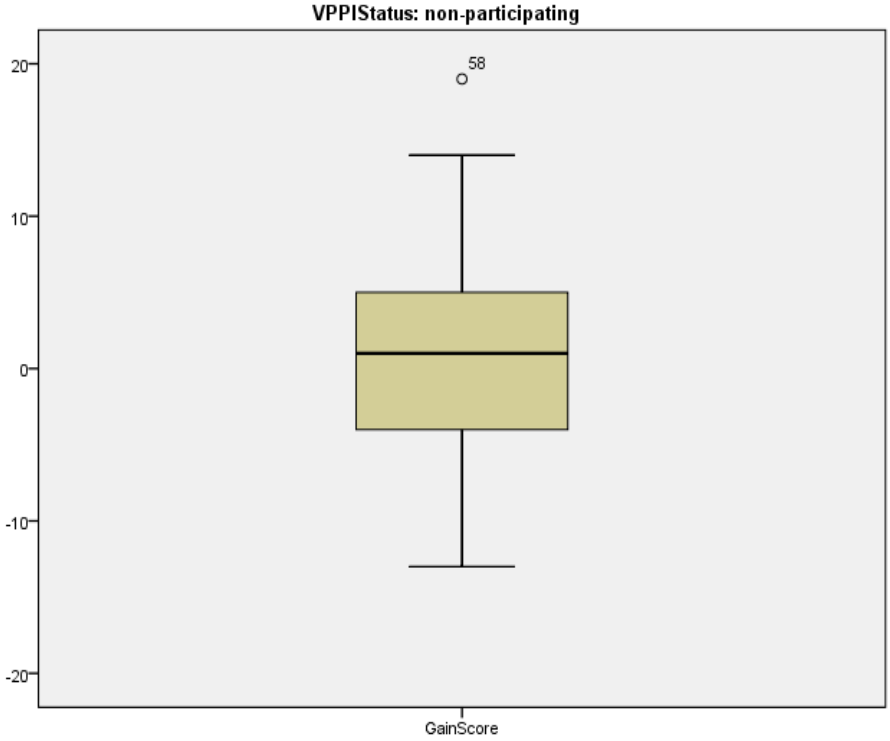
Grade 5 Writing Frequency Polygon VPPI Participating



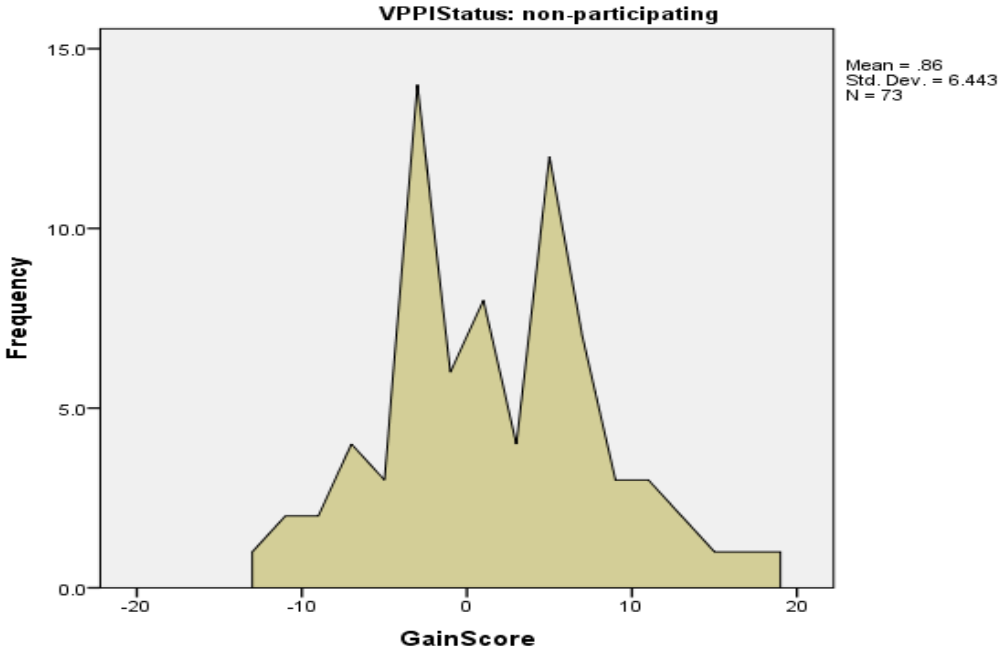
Grade 5 Writing Histogram VPPI Nonparticipating



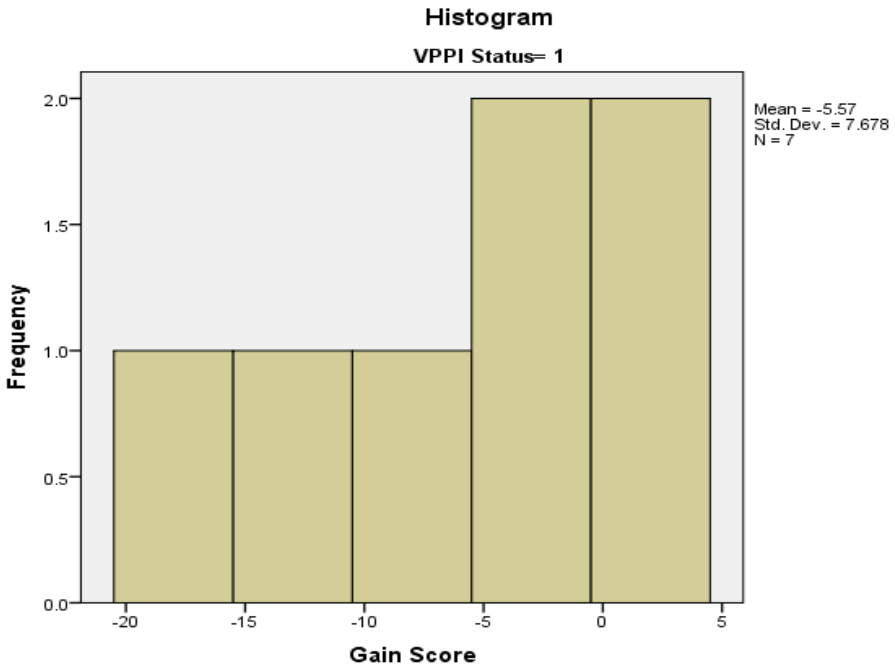
Grade 5 Writing Box Plot VPPI Nonparticipating



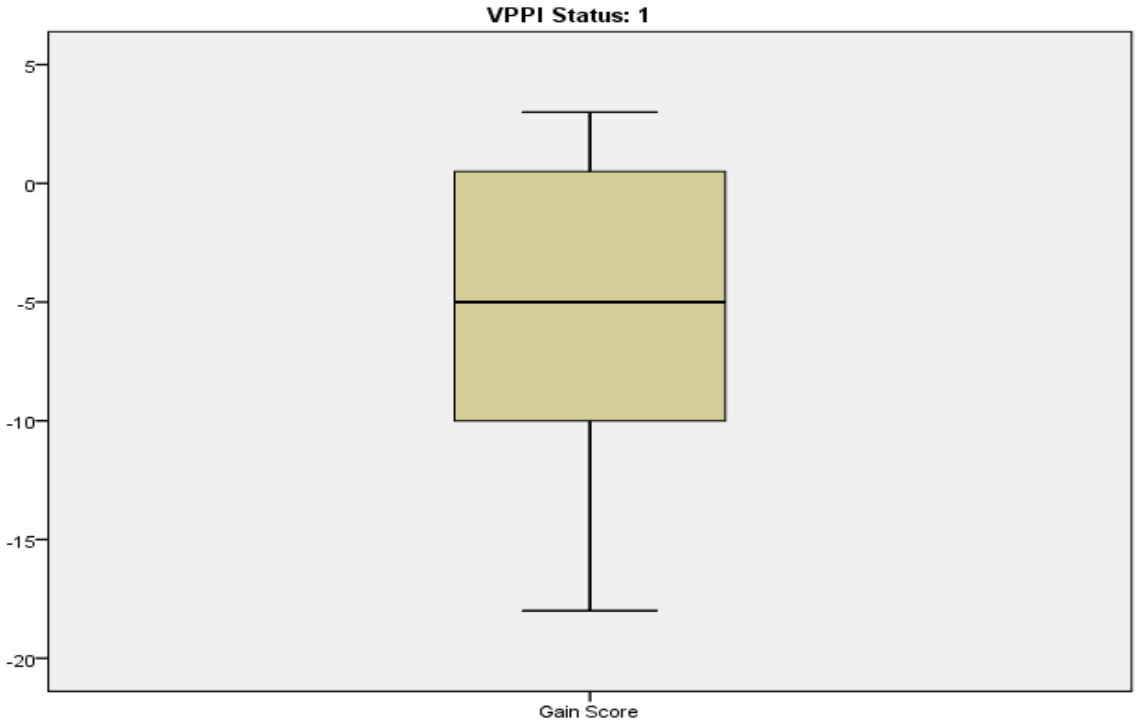
Grade 5 Writing Frequency Polygon VPPI Nonparticipating



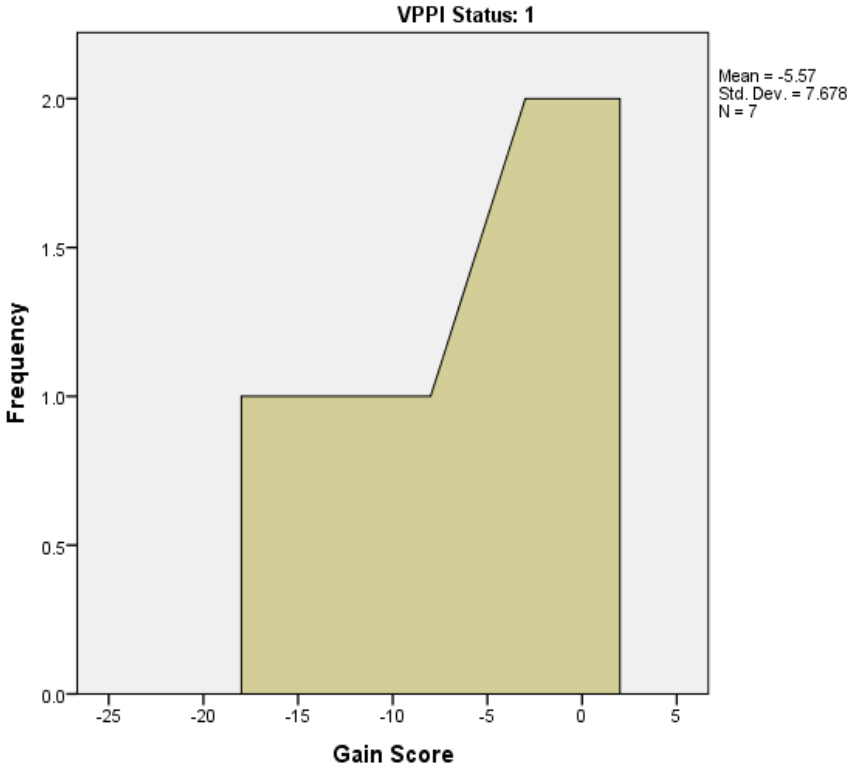
Grade 5 History Histogram VPPI Participating



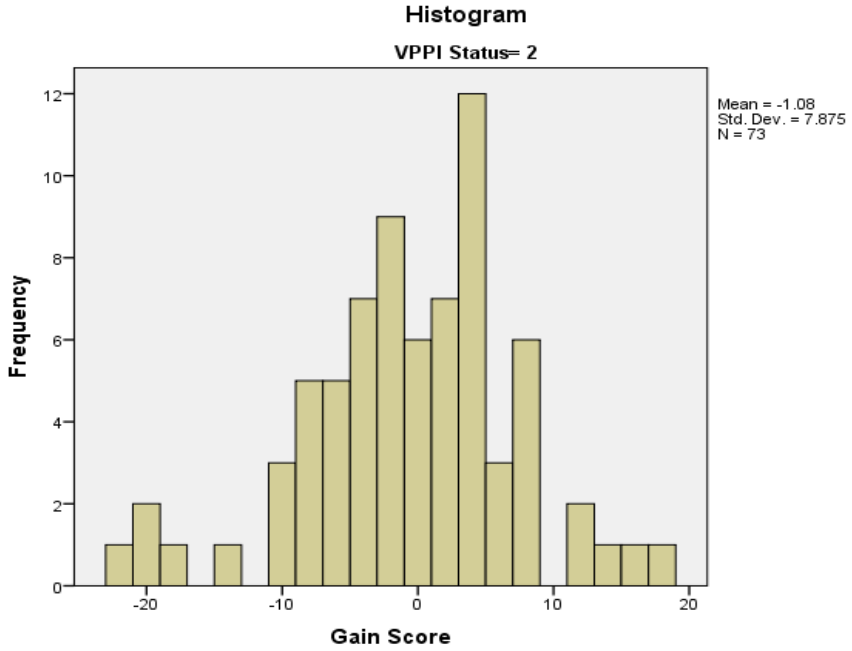
Grade 5 History Box Plot VPPI Participating



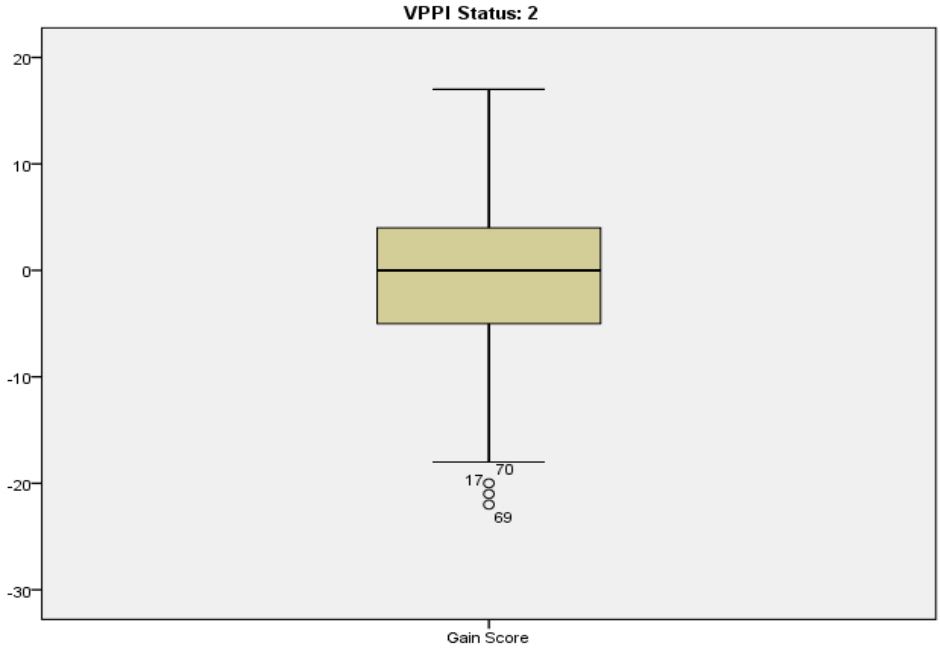
Grade 5 History Frequency Polygon VPPI Participating



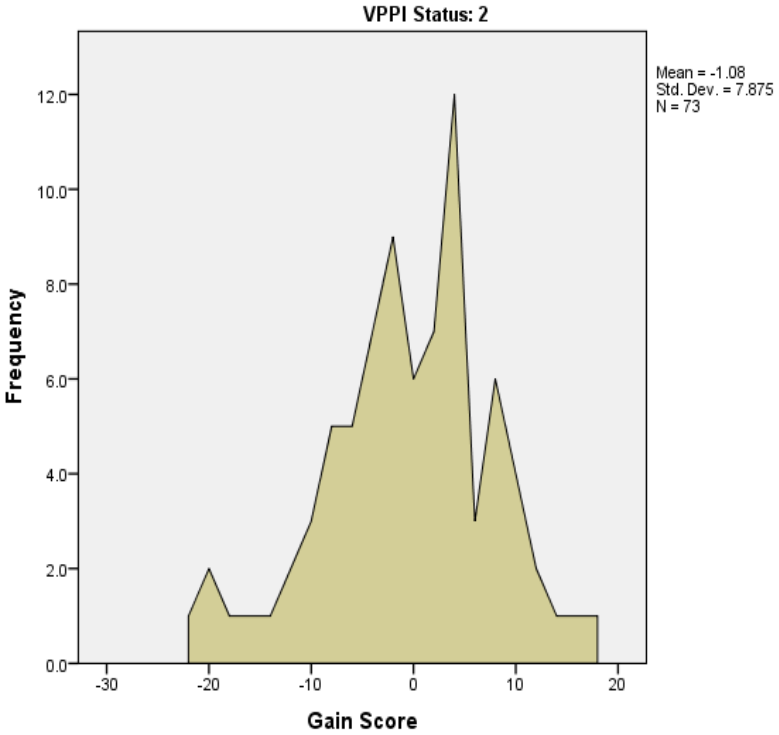
Grade 5 History Histogram VPPI Nonparticipating



Grade 5 History Box Plot VPPI Nonparticipating



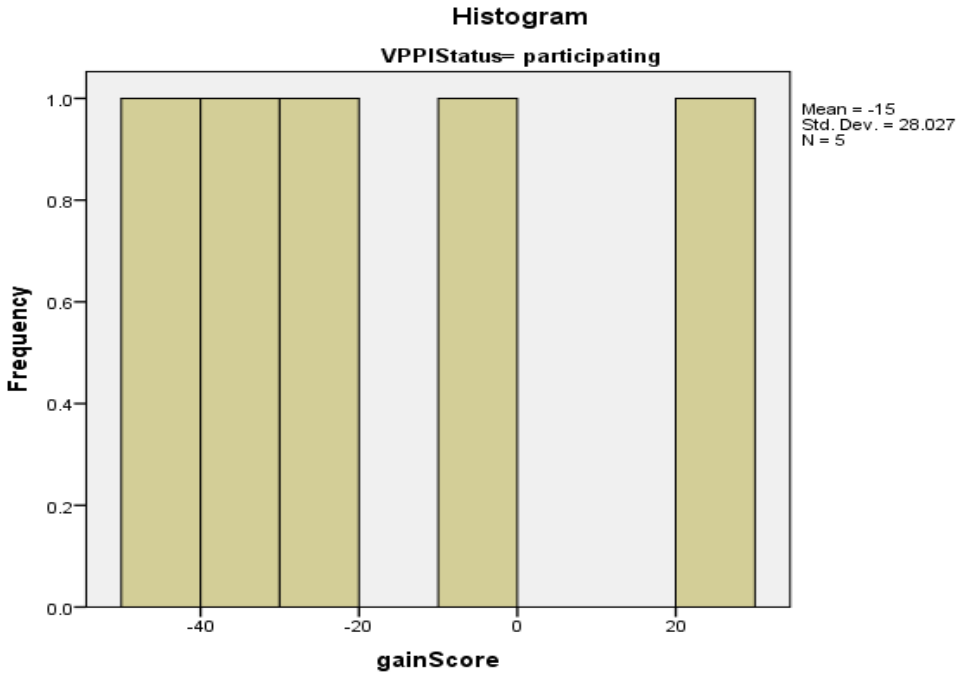
Grade 5 History Frequency Polygon VPPI Nonparticipating



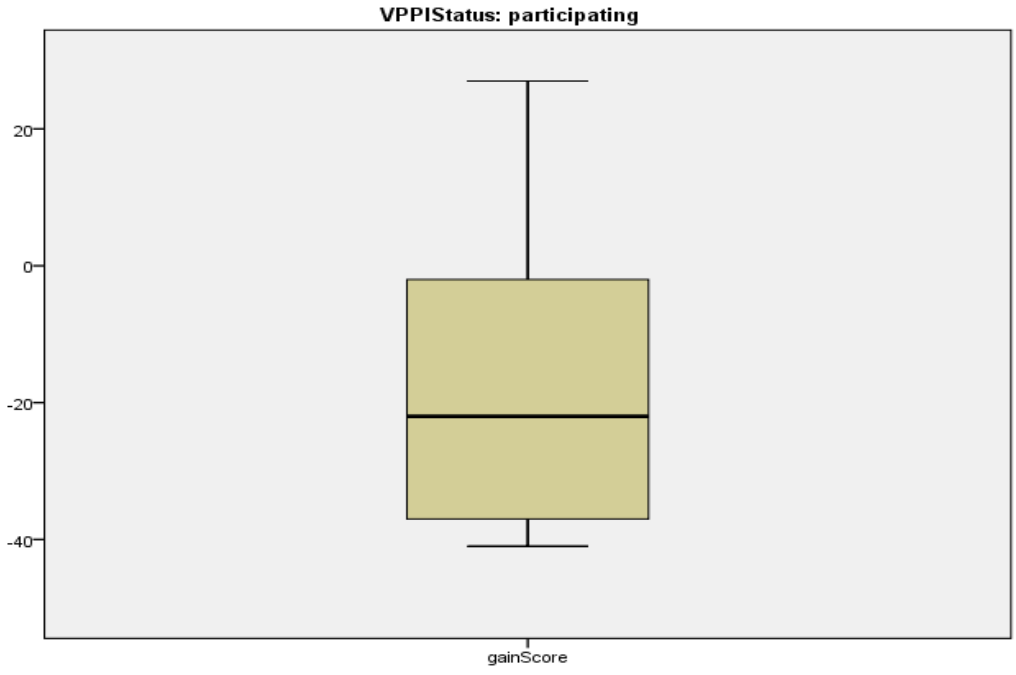
Appendix B

Grade 8 Histogram, Box Plots, and Frequency Polygons

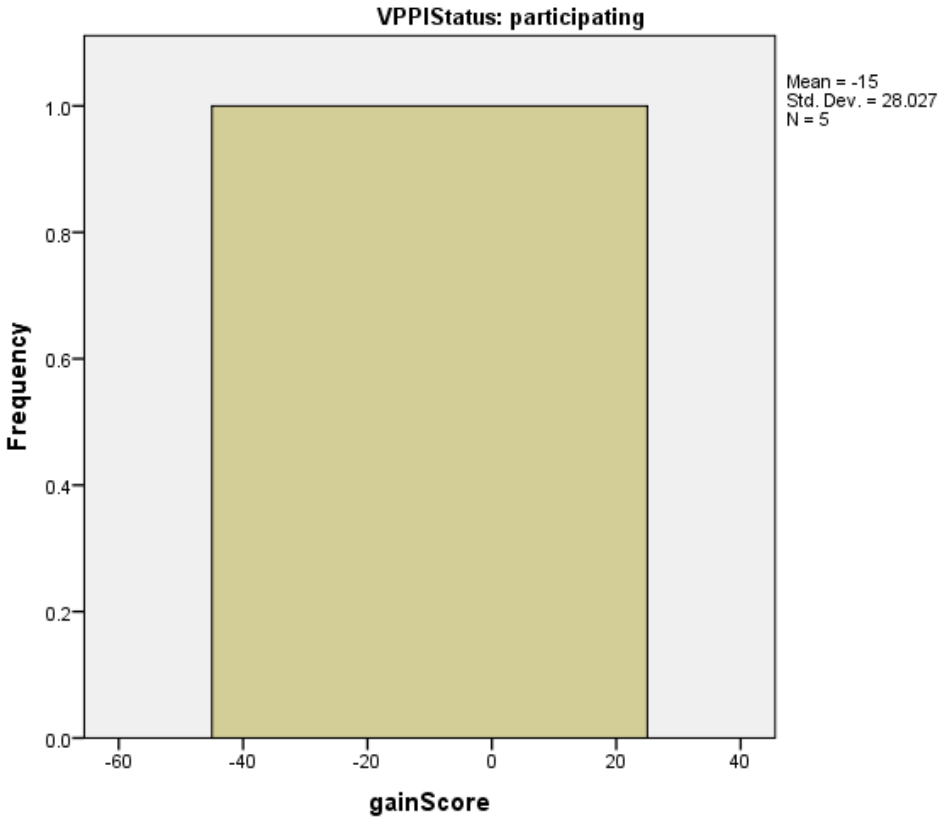
Grade 8 Math Histogram VPPI Participating



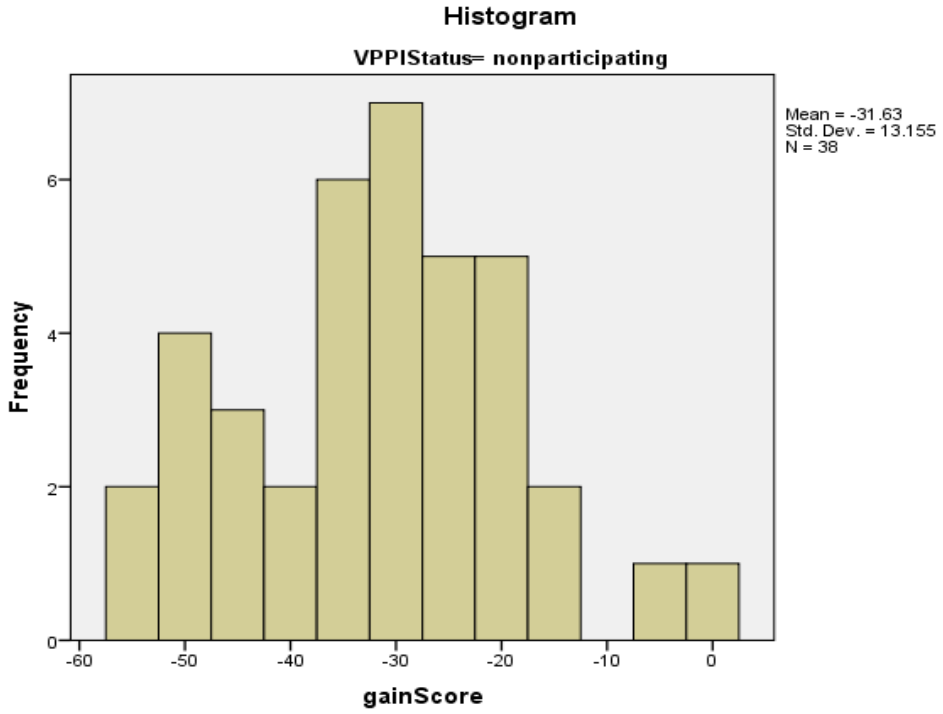
Grade 8 Math Box Plot VPPI Participating



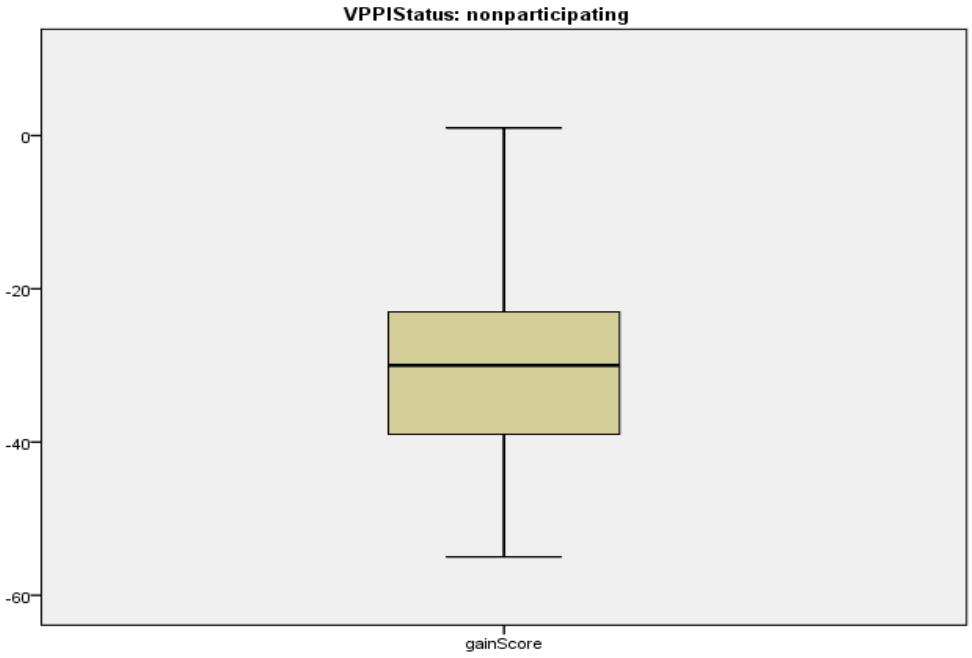
Grade 8 Math Frequency Polygon VPPI Participating



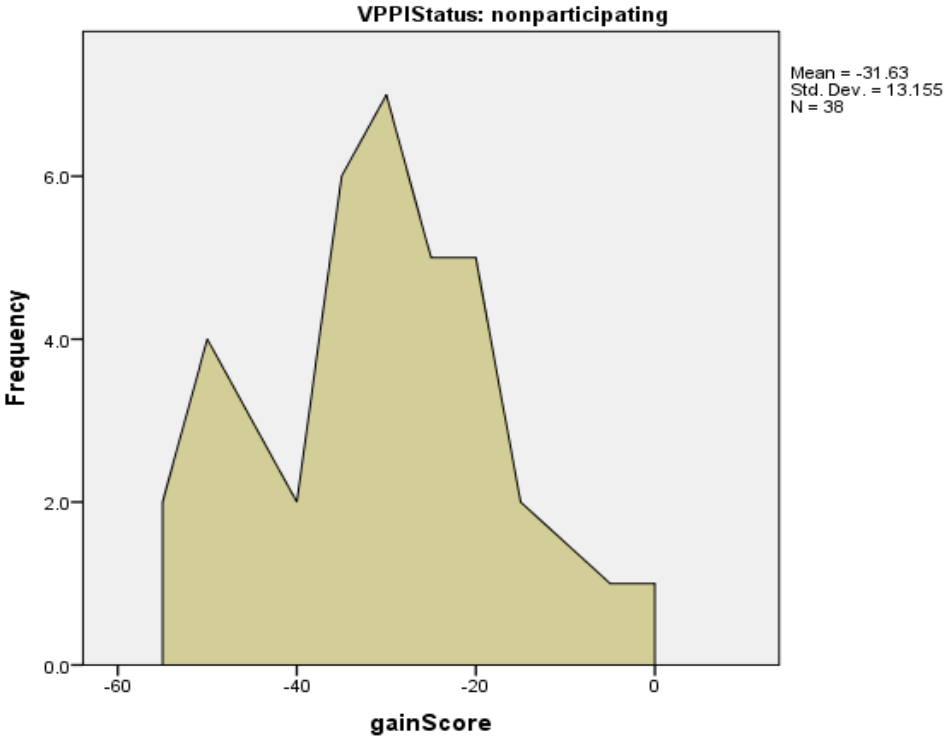
Grade 8 Math Histogram VPPI Nonparticipating



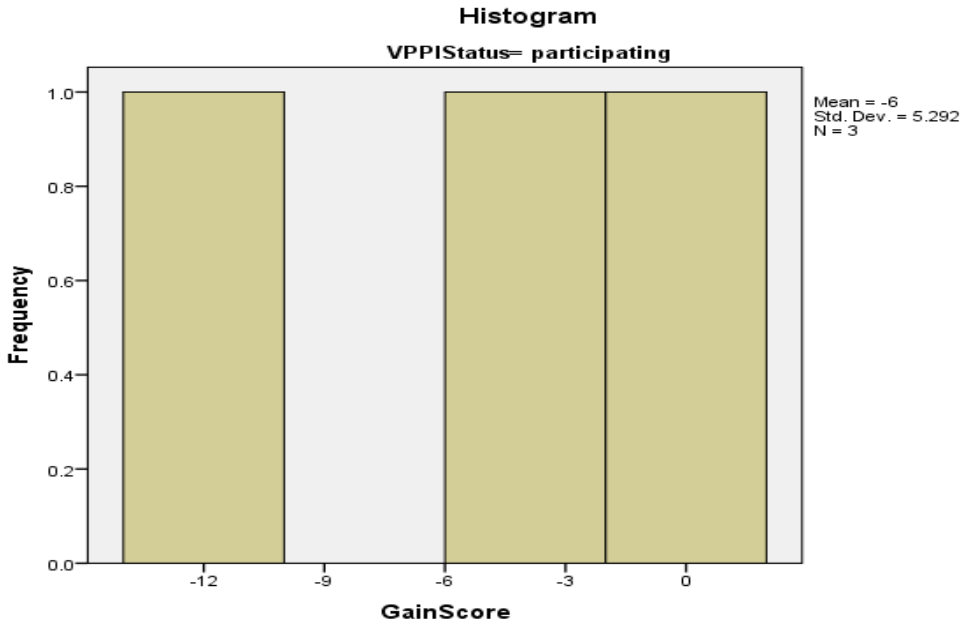
Grade 8 Math Box Plot VPPI Nonparticipating



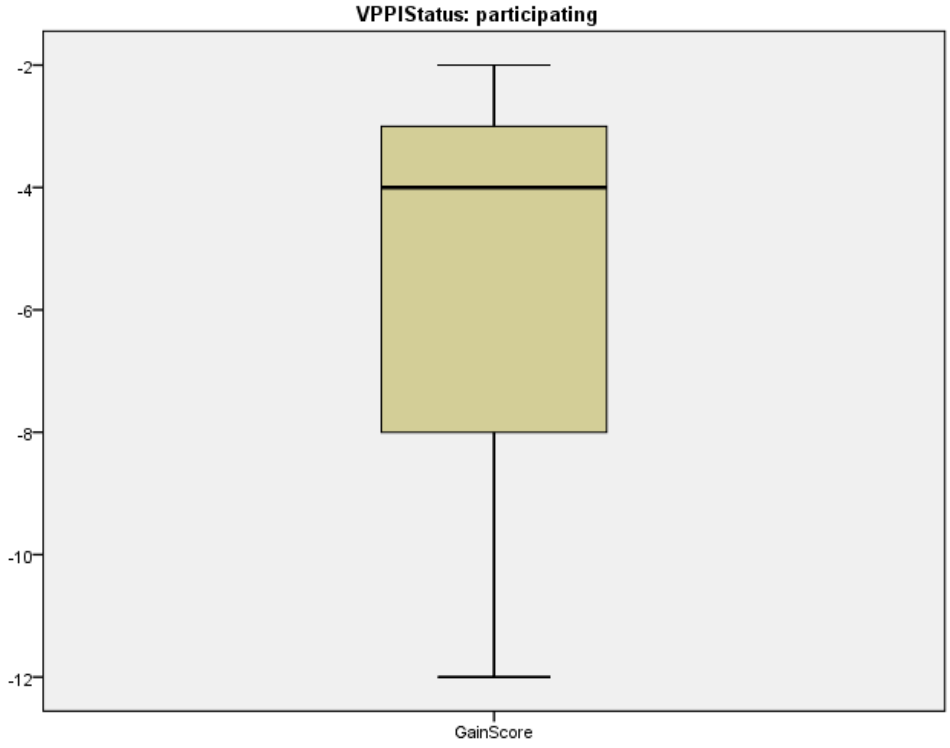
Grade 8 Math Frequency Polygon VPPI Nonparticipating



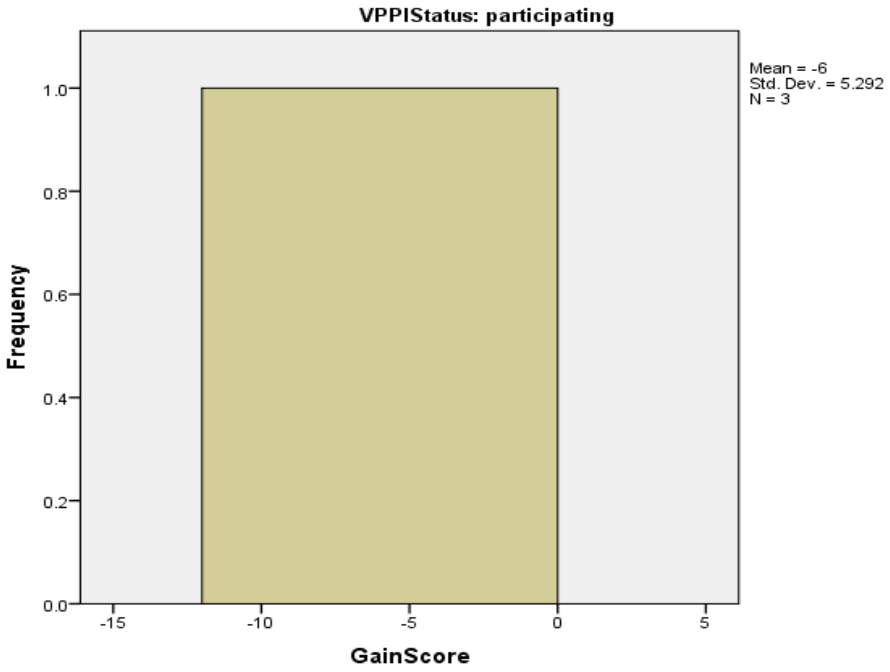
Grade 8 Reading Histogram VPPI Participating



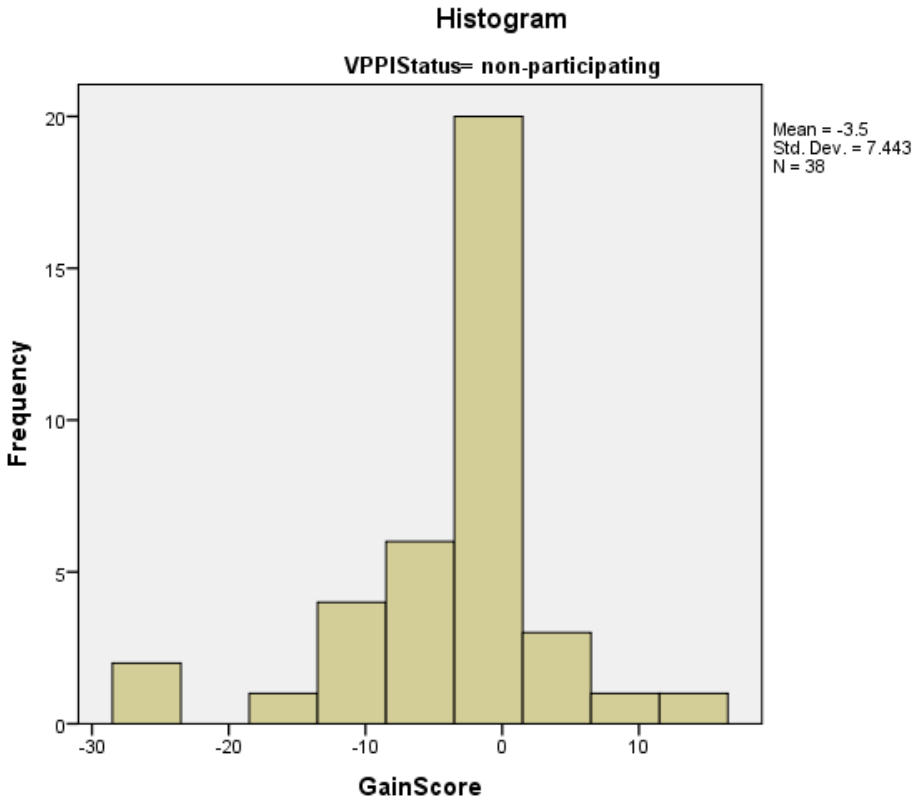
Grade 8 Reading Box Plot VPPI Participating



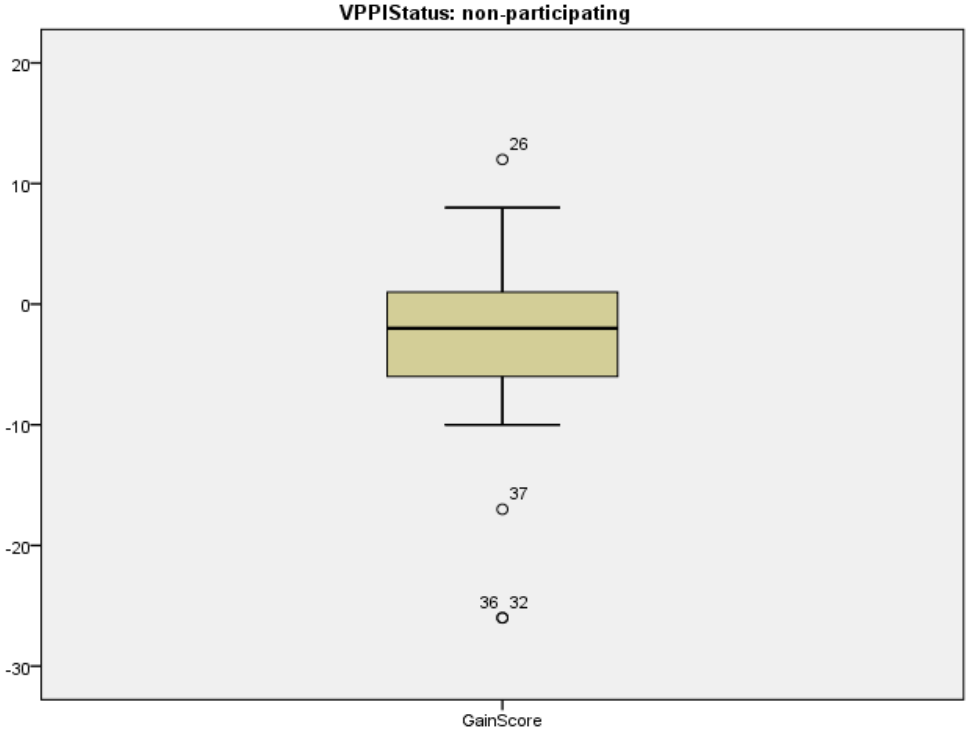
Grade 8 Reading Frequency Polygon VPPI Participating



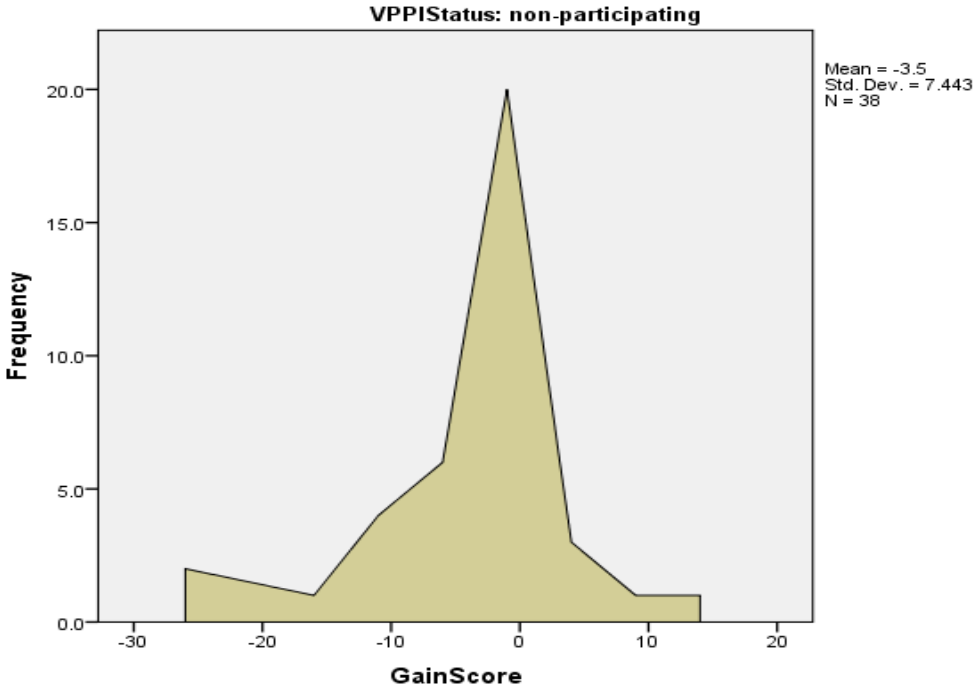
Grade 8 Reading Histogram VPPI Nonparticipating



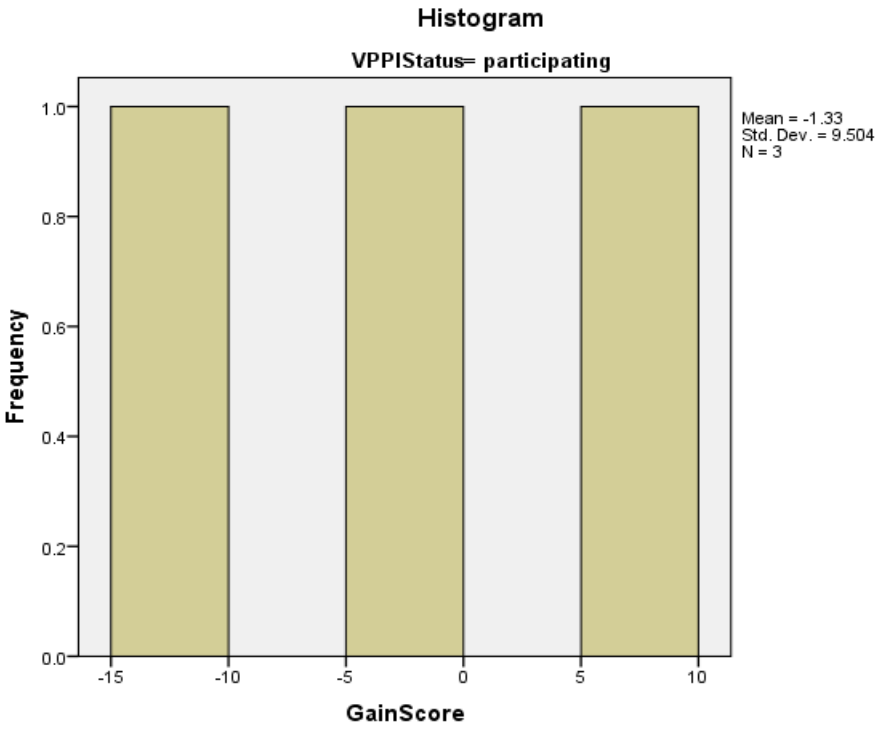
Grade 8 Reading Box Plot VPPI Nonparticipating



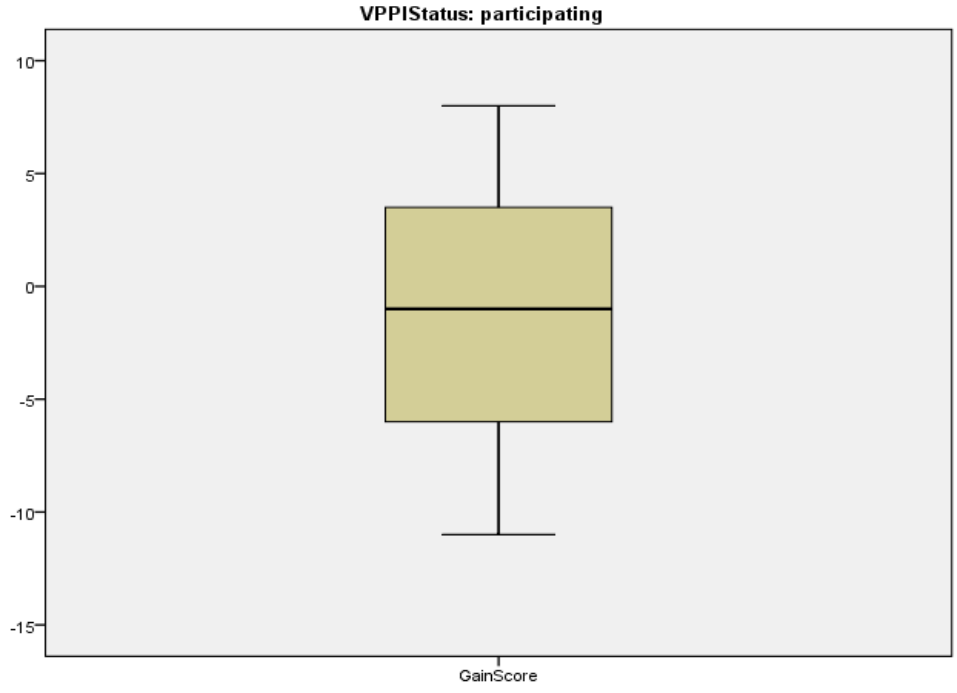
Grade 8 Reading Frequency Polygon VPPI Nonparticipating



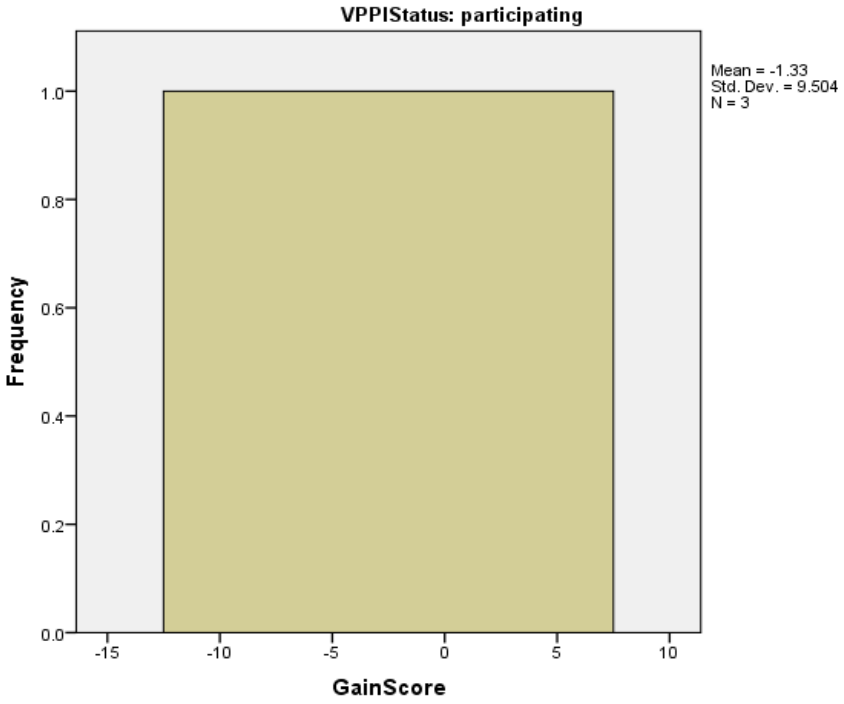
Grade 8 Writing Histogram VPPI Participating



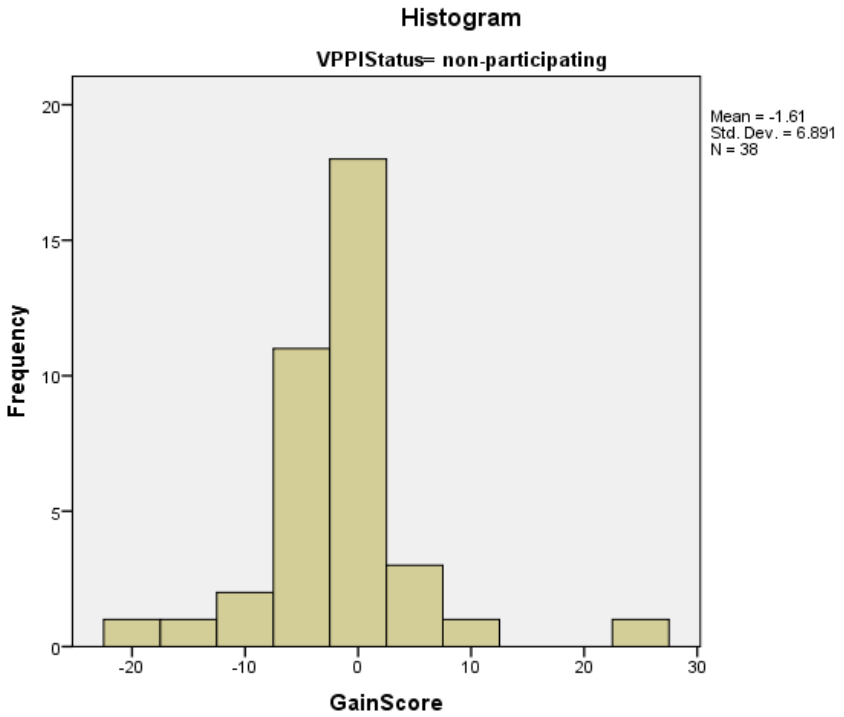
Grade 8 Writing Box Plot VPPI Participating



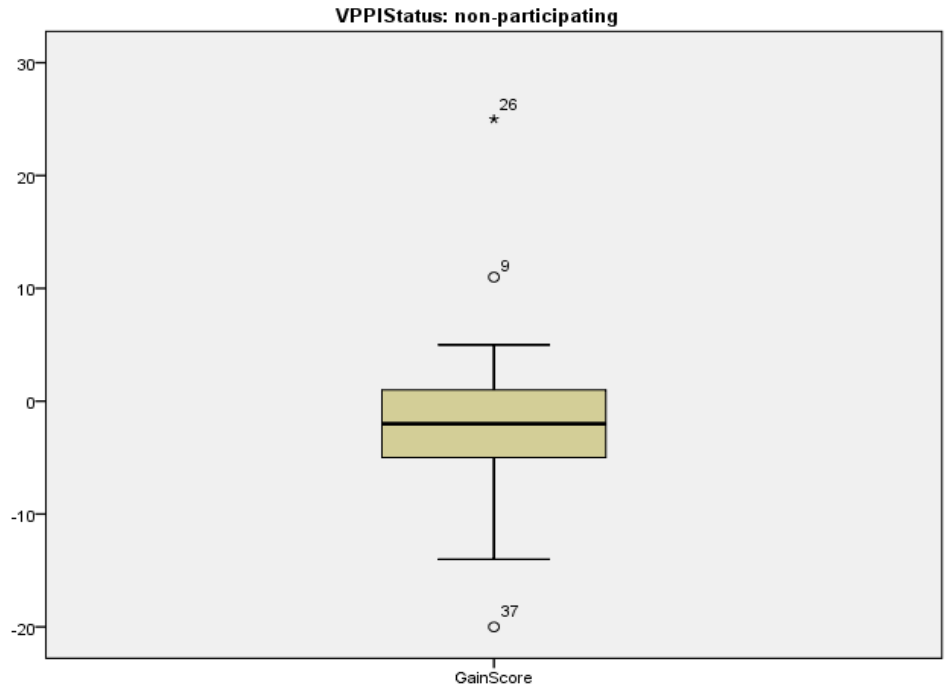
Grade 8 Writing Frequency Polygon VPPI Participating



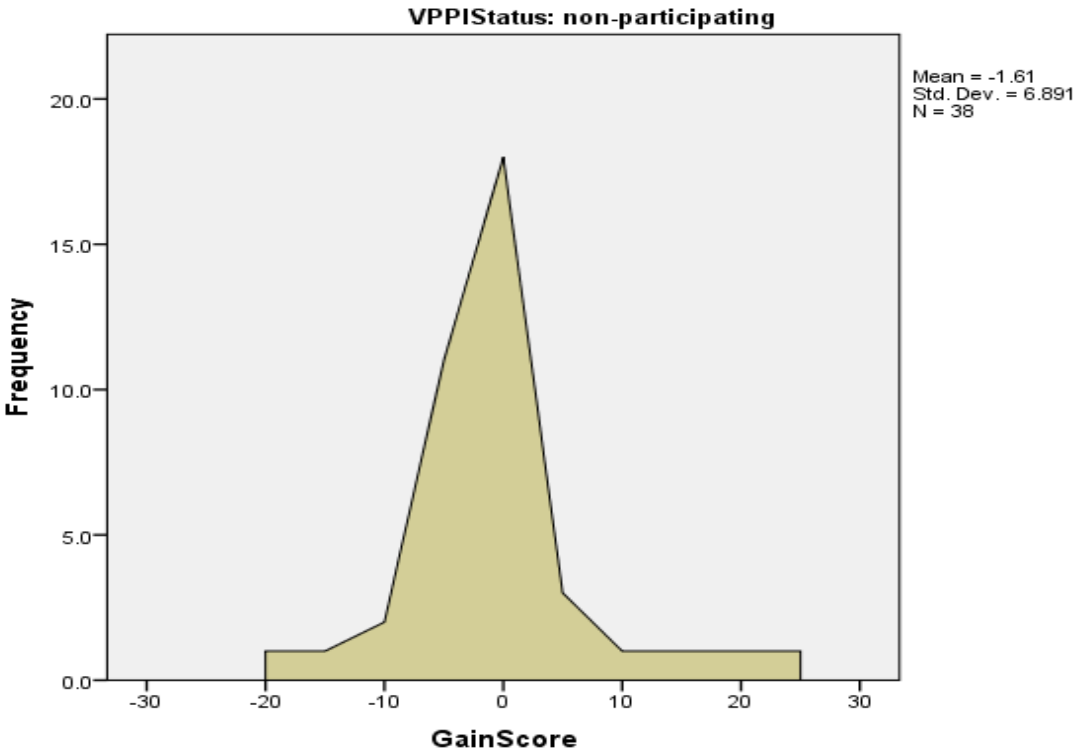
Grade 8 Writing Histogram VPPI Nonparticipating



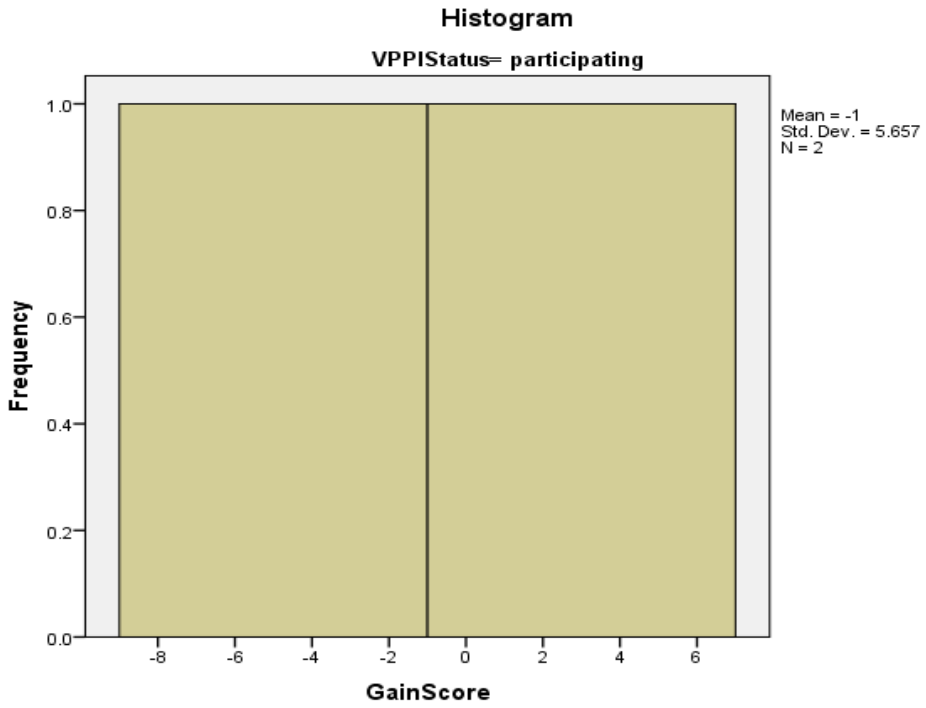
Grade 8 Writing Box Plot VPPI Nonparticipating



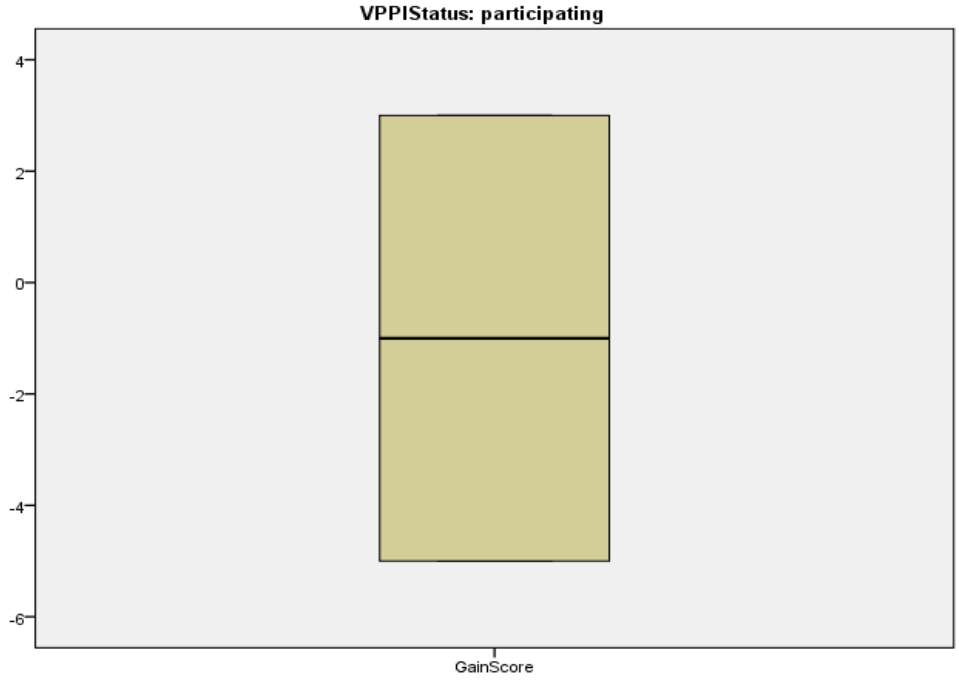
Grade 8 Writing Frequency Polygon VPPI Nonparticipating



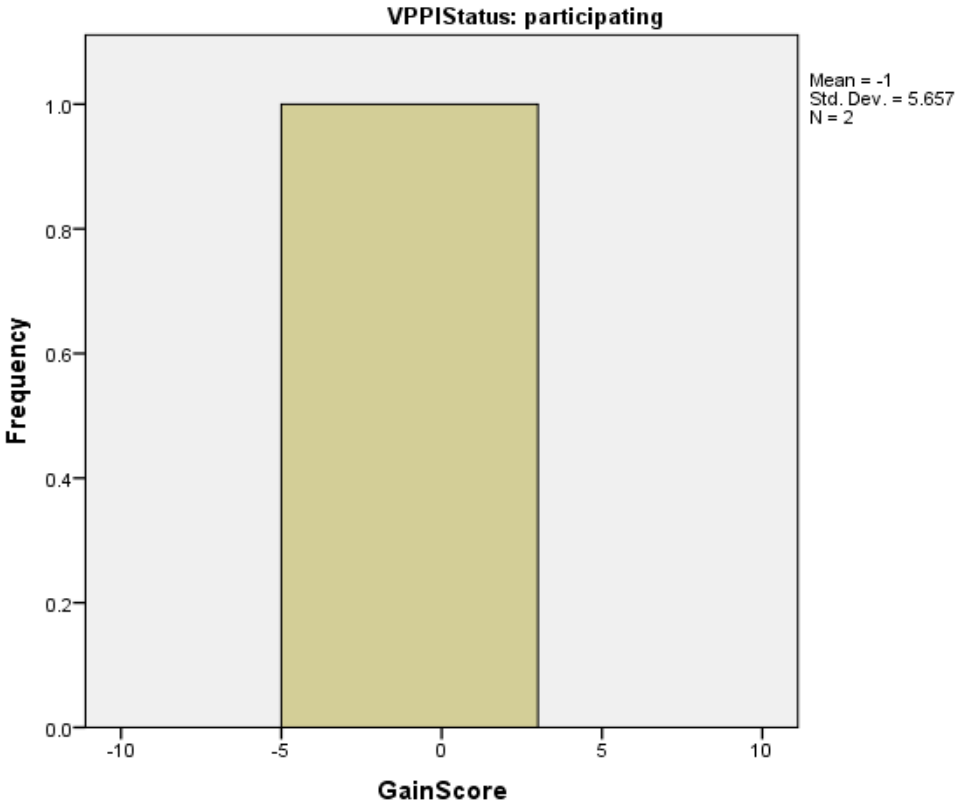
Grade 8 Science Histogram VPPI Participating



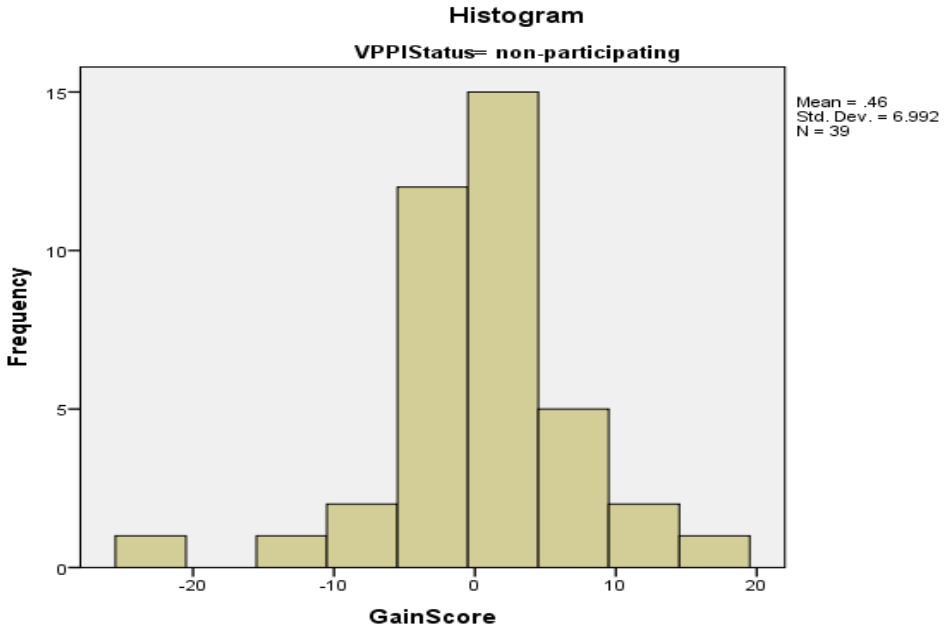
Grade 8 Science Box Plot VPPI Participating



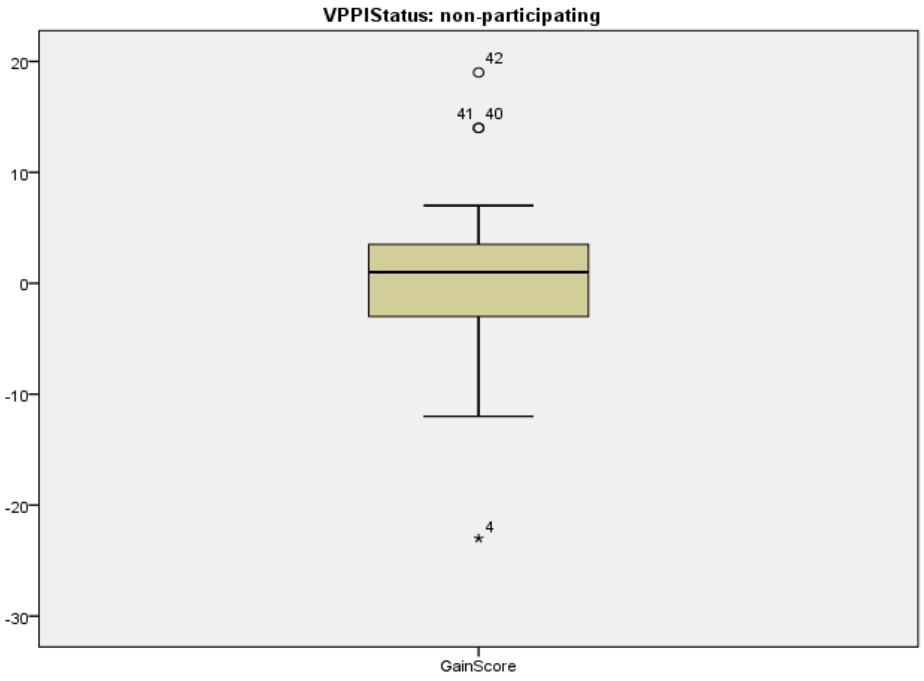
Grade 8 Science Frequency Polygon VPPI Participating



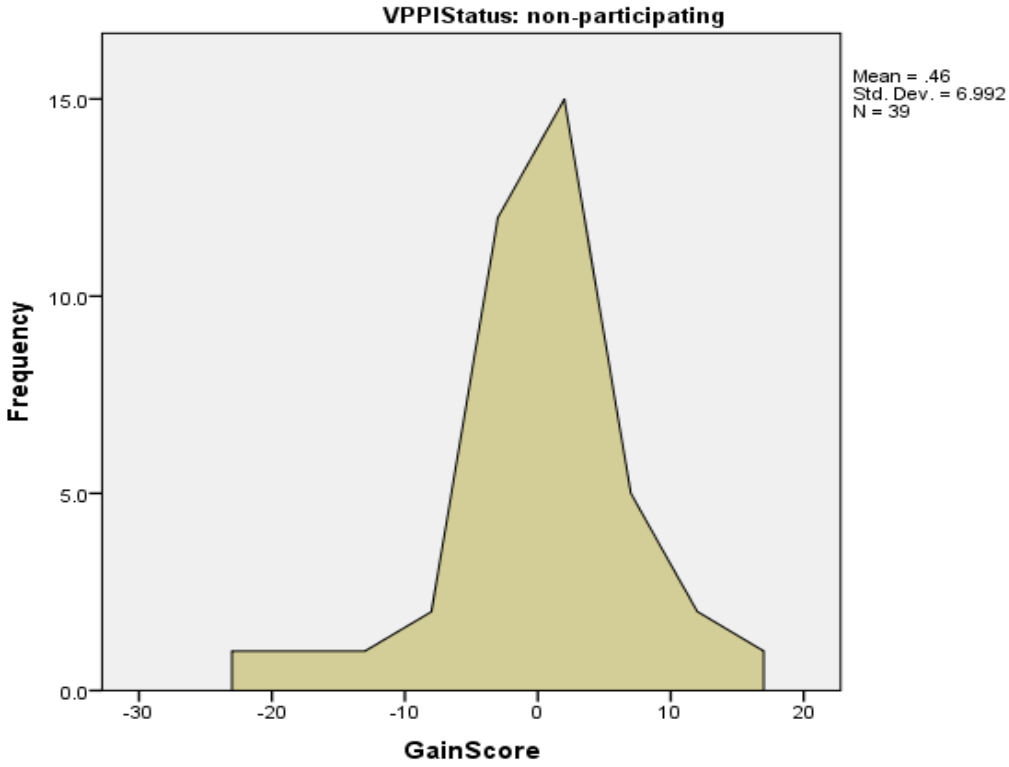
Grade 8 Science Histogram VPPI Nonparticipating



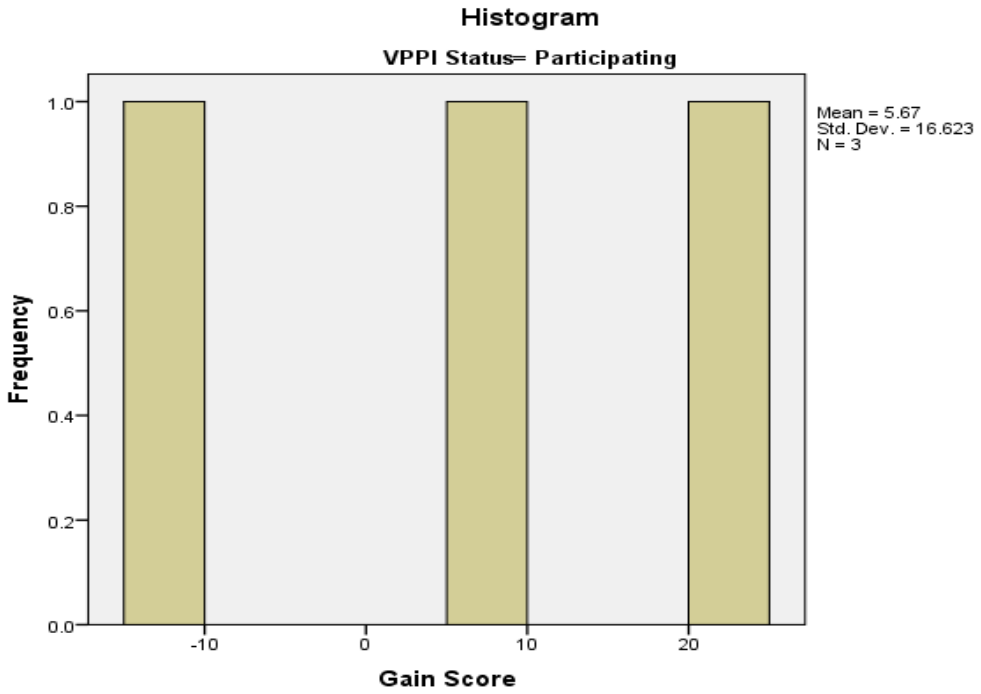
Grade 8 Science Box Plot VPPI Nonparticipating



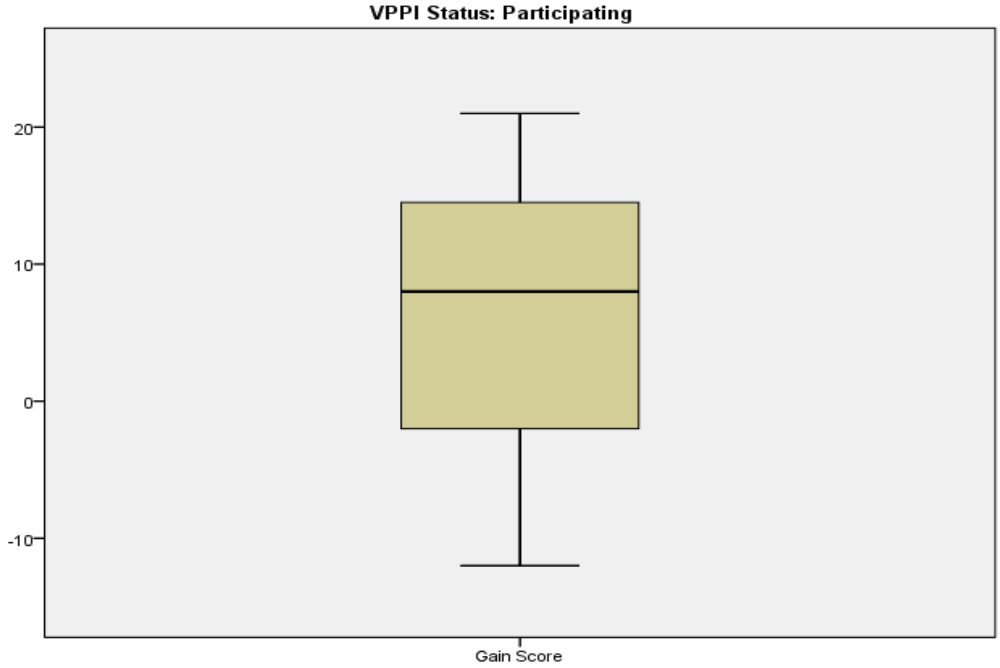
Grade 8 Science Frequency Polygon VPPI Nonparticipating



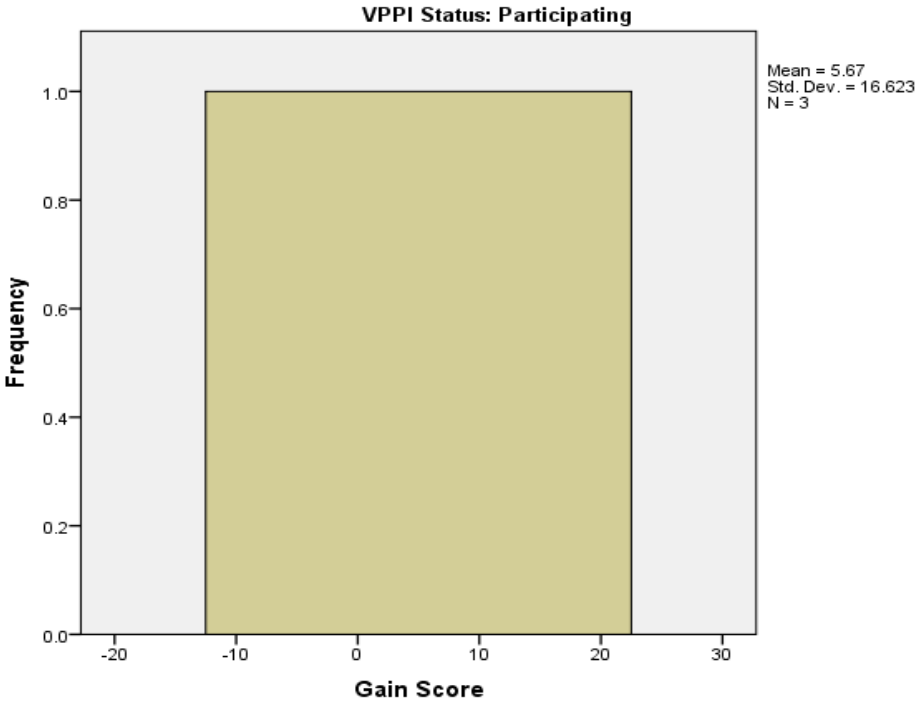
Grade 8 History Histogram VPPI Participating



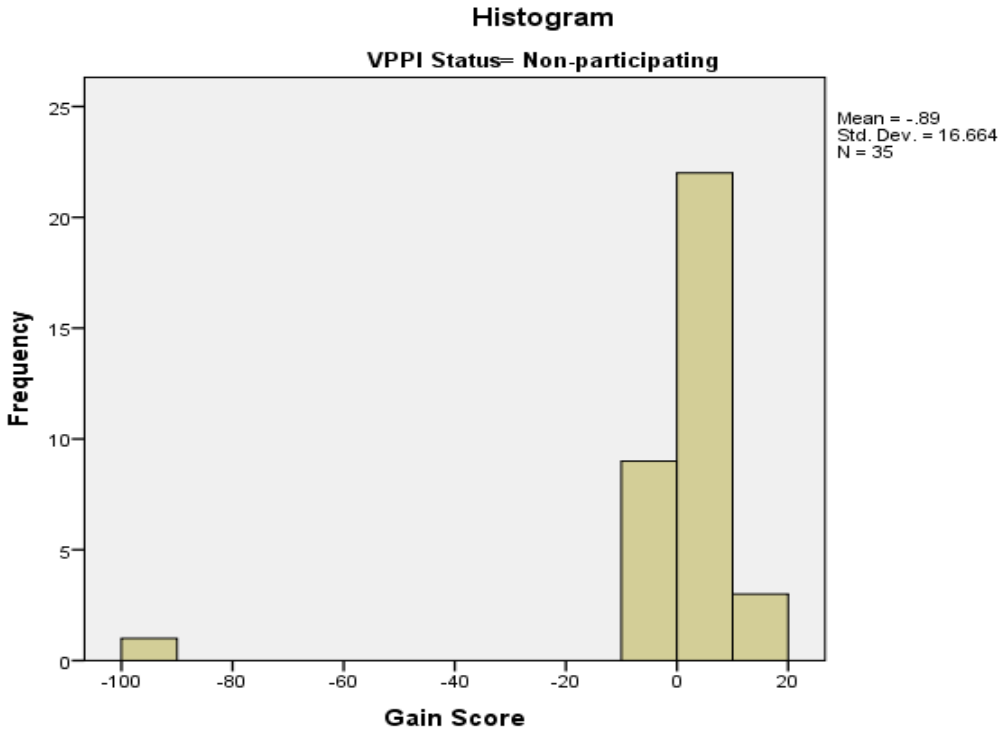
Grade 8 History Box Plot VPPI Participating



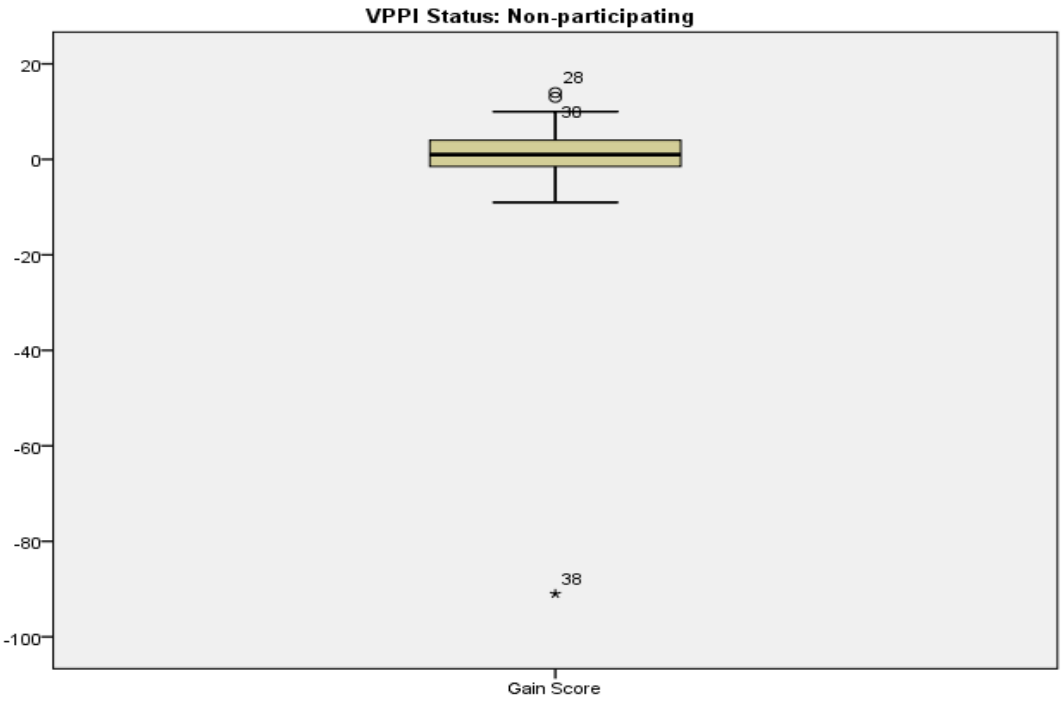
Grade 8 History Frequency Polygon VPPI Participating



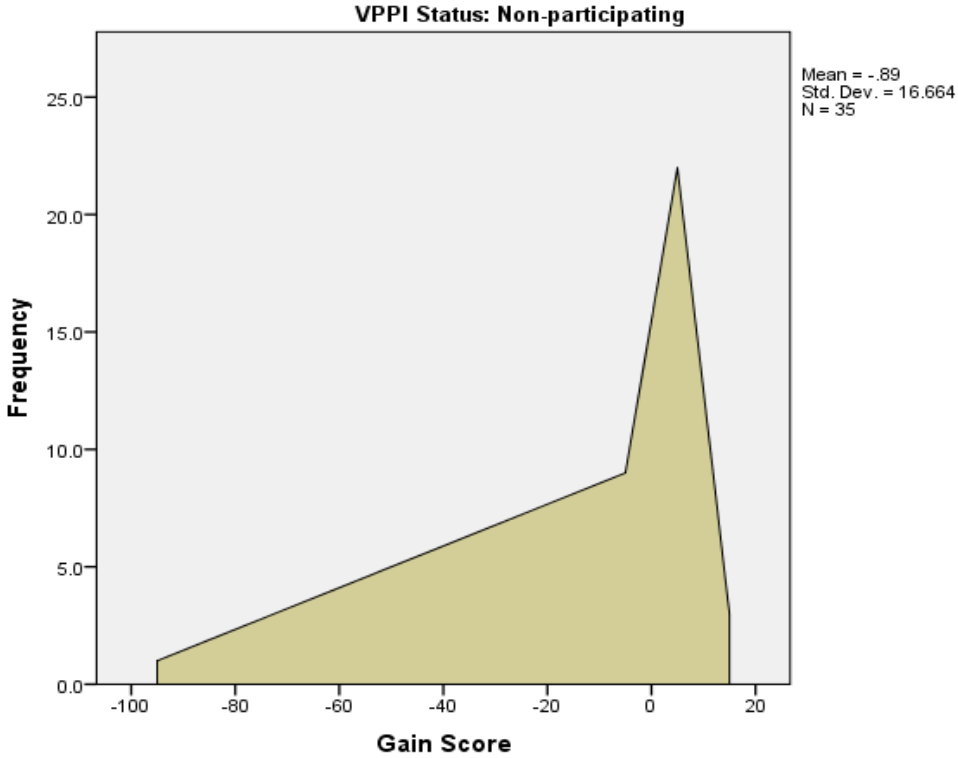
Grade 8 History Histogram VPPI Nonparticipating



Grade 8 History Box Plot VPPI Nonparticipating



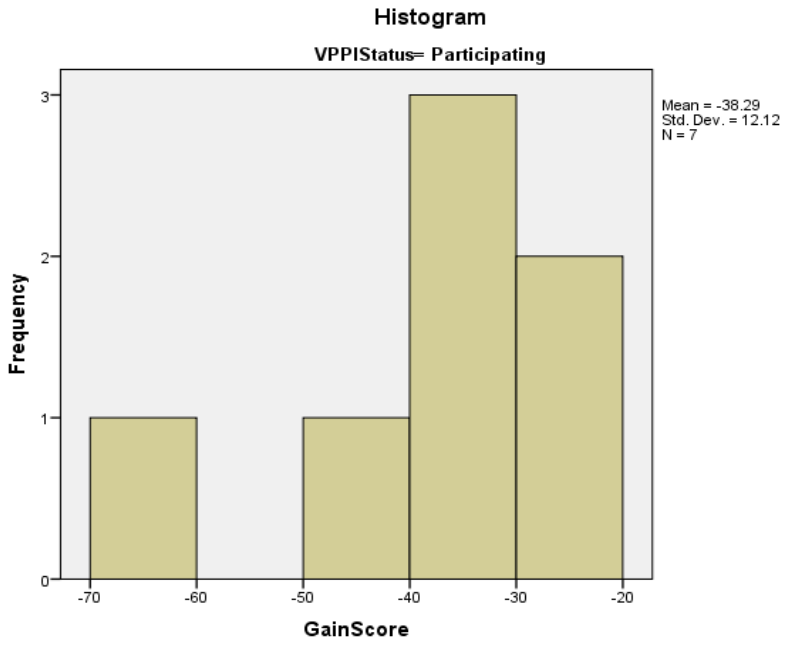
Grade 8 History Frequency Polygon VPPI Nonparticipating



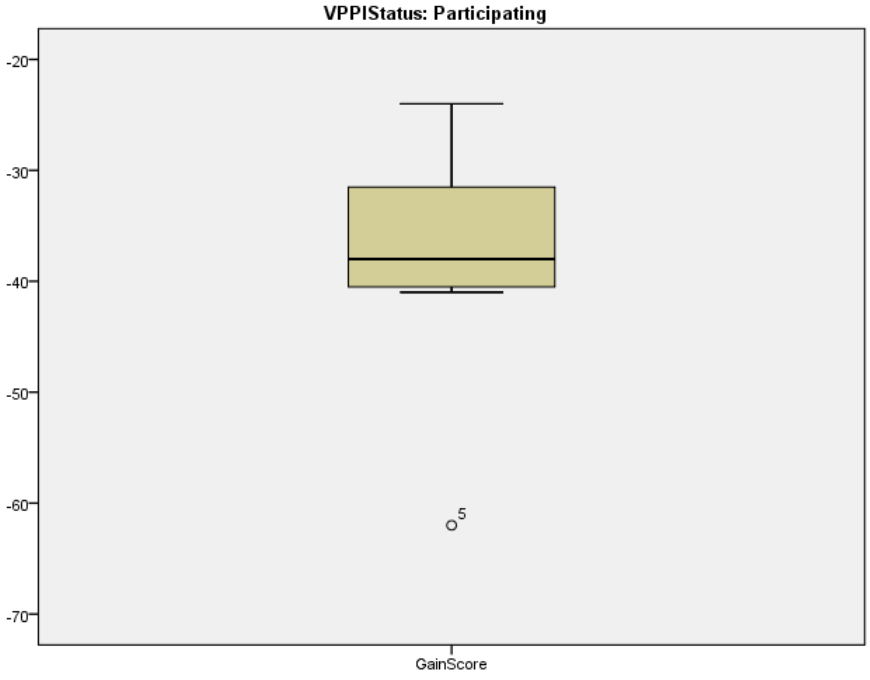
Appendix C

End-of-Course Histogram, Box Plots, and Frequency Polygons

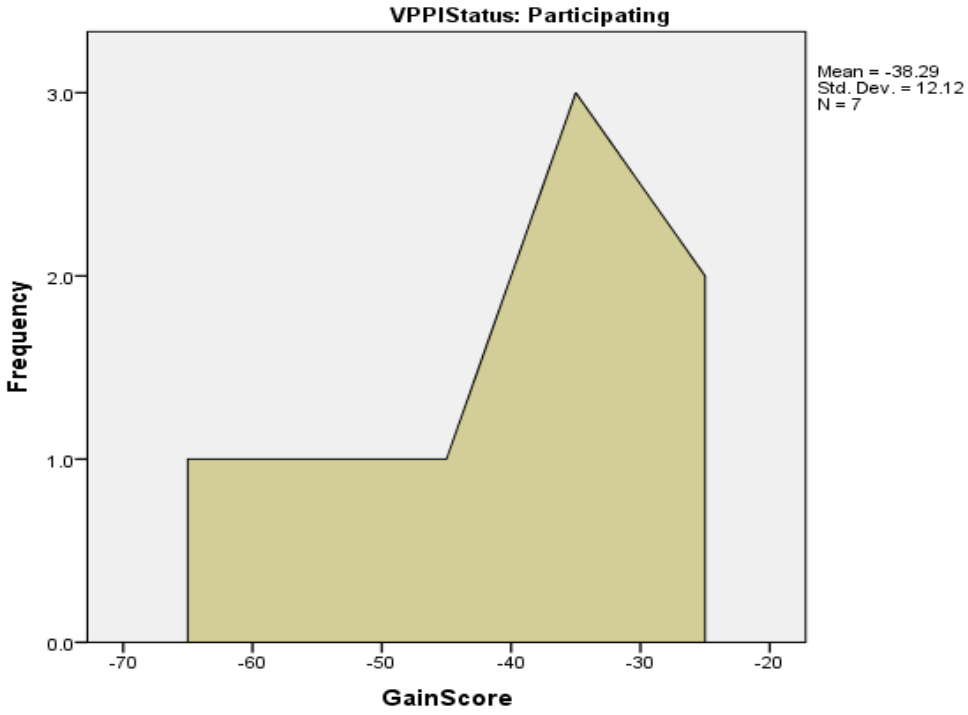
High School End-of-Course Algebra I Histogram VPPI Participating



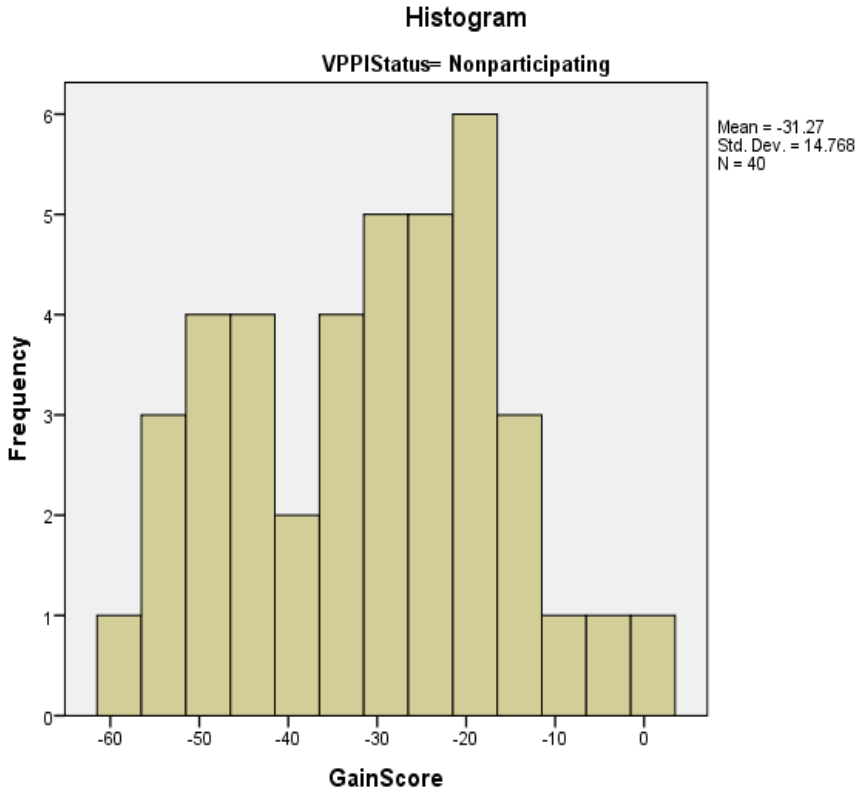
High School End-of-Course Algebra I Box Plot VPPI Participating



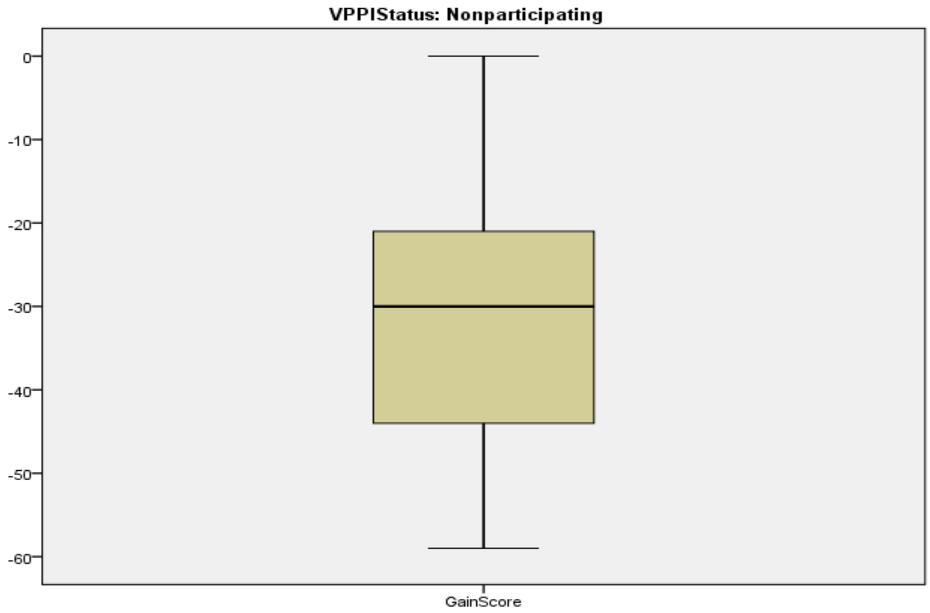
High School End-of-Course Algebra I Frequency Polygon VPPI Participating



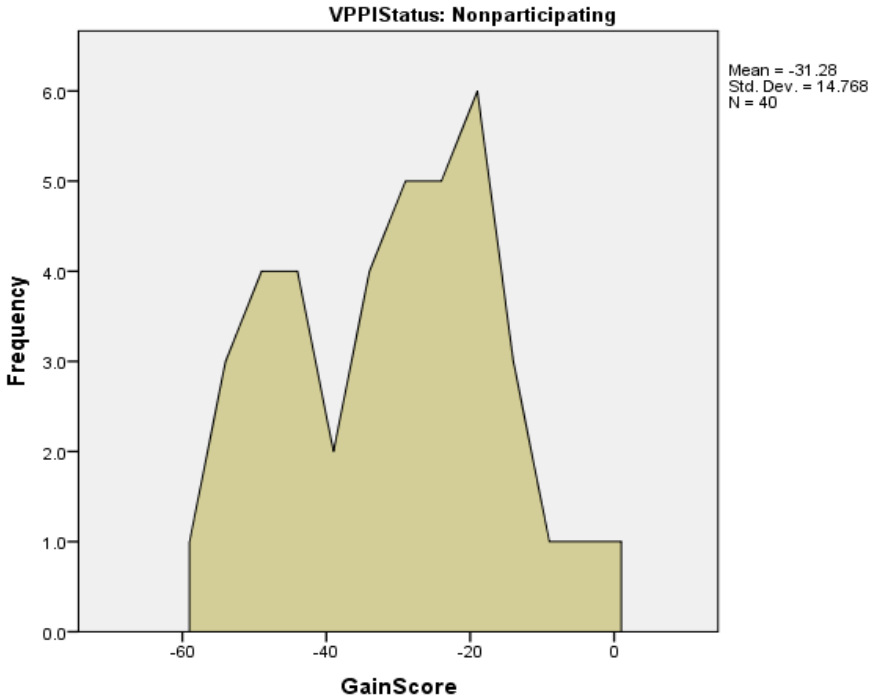
High School End-of-Course Algebra I Histogram VPPI Nonparticipating



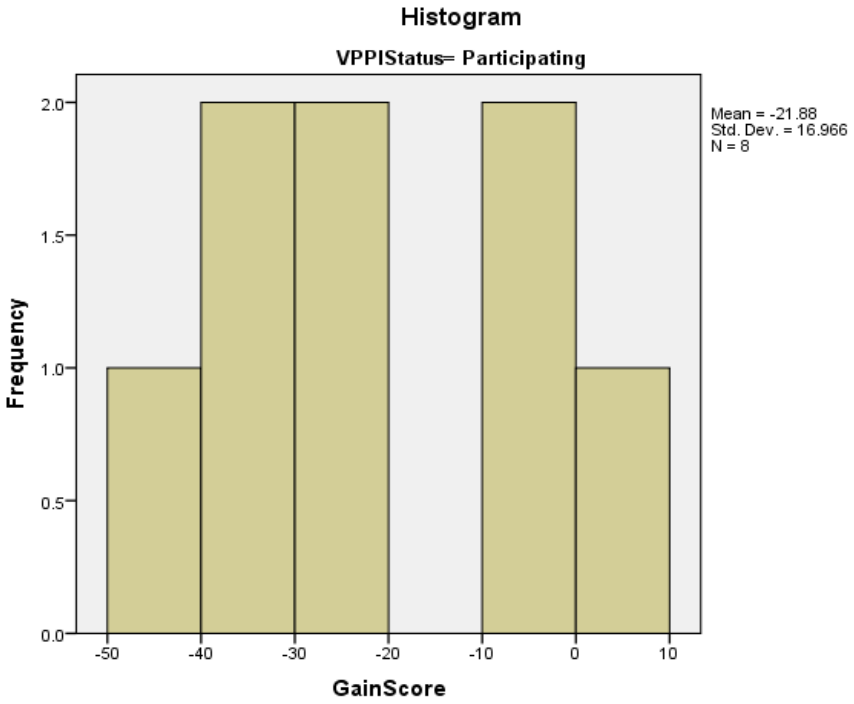
High School End-of-Course Algebra I Box Plot VPPI Nonparticipating



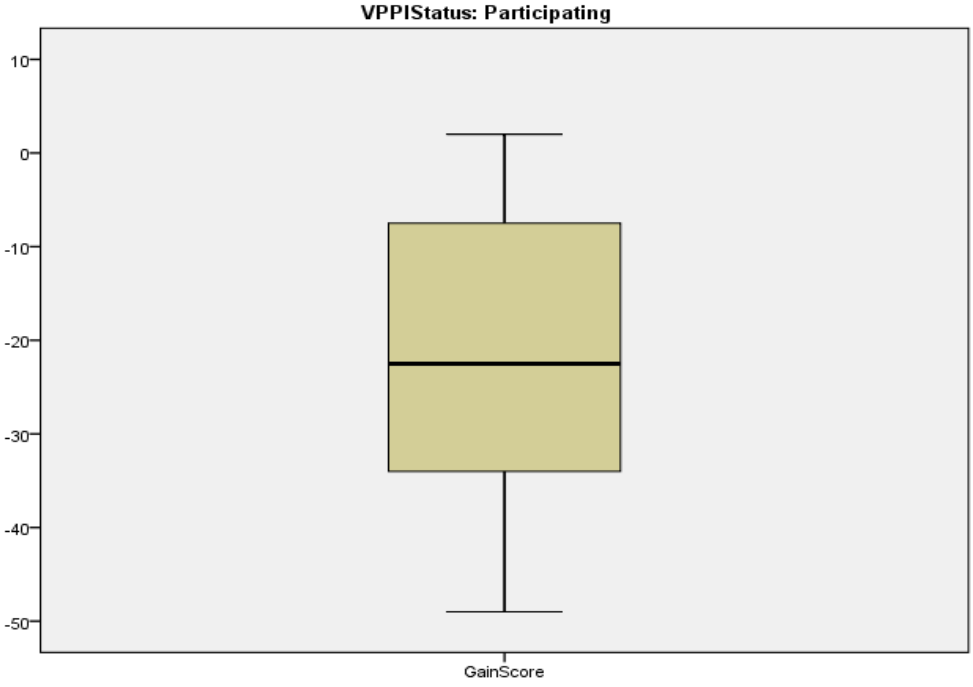
High School End-of-Course Algebra I Frequency Polygon VPPI Nonparticipating



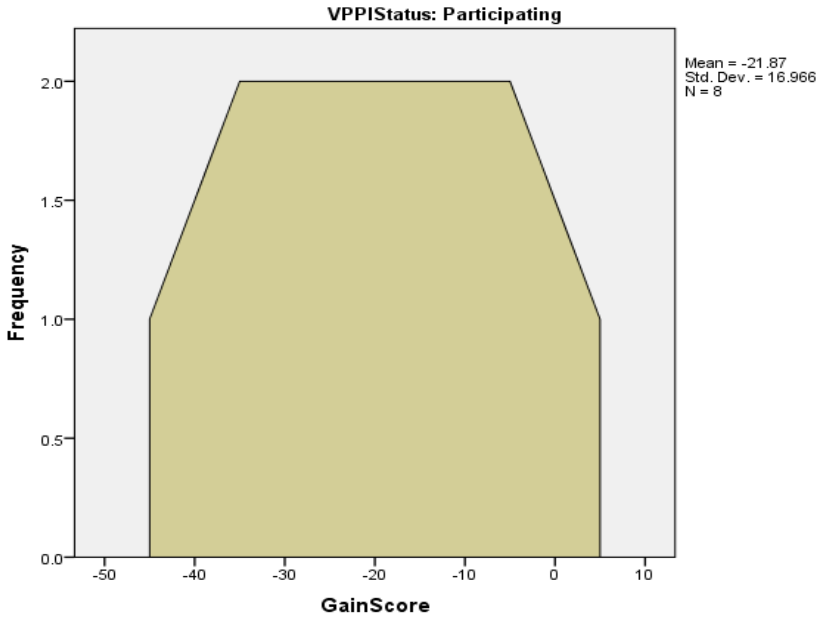
High School End-of-Course Algebra II Histogram VPPI Participating



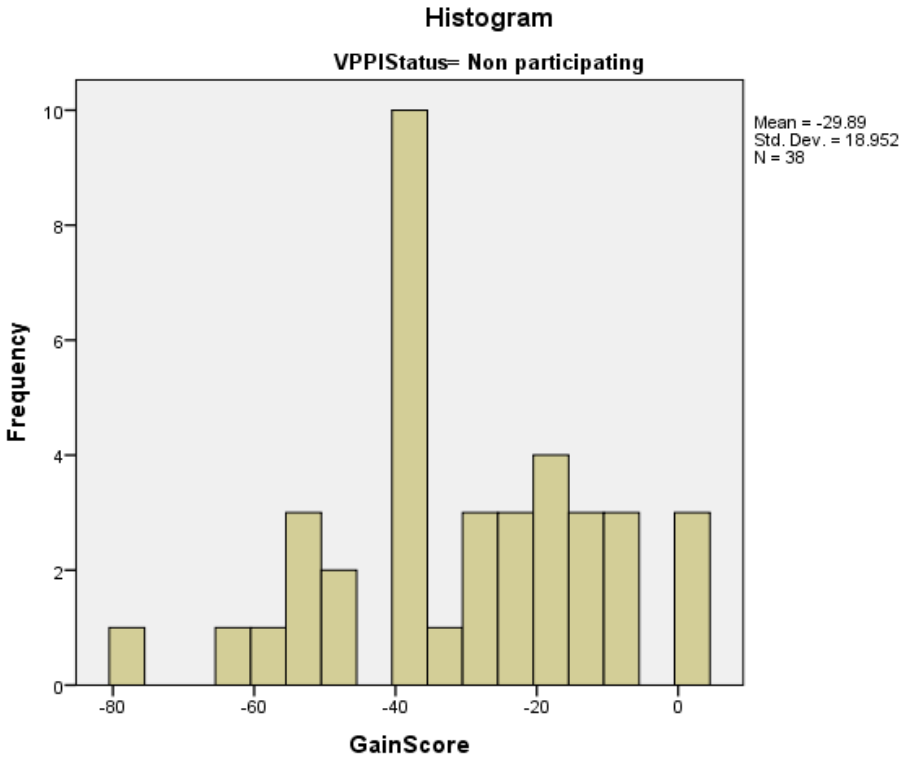
High School End-of-Course Algebra II Box Plot VPPI Participating



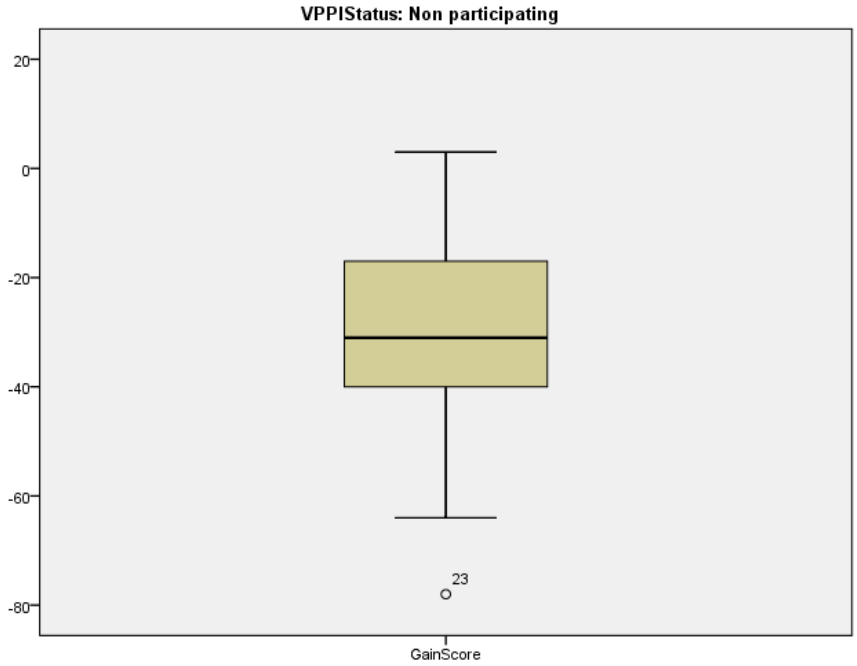
High School End-of-Course Algebra II Frequency Polygon VPPI Participating



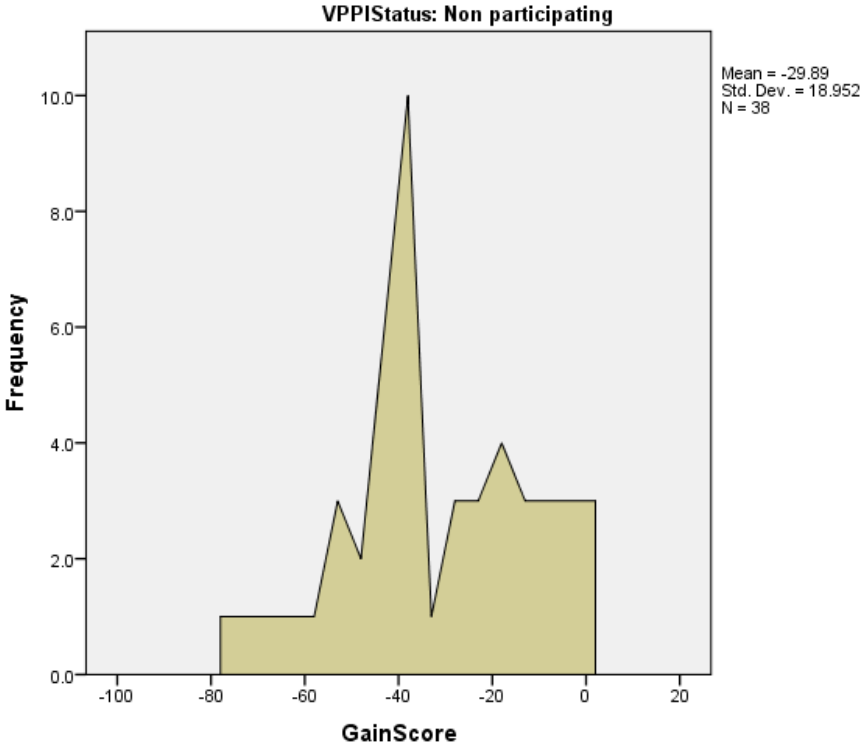
High School End-of-Course Algebra II Histogram VPPI Nonparticipating



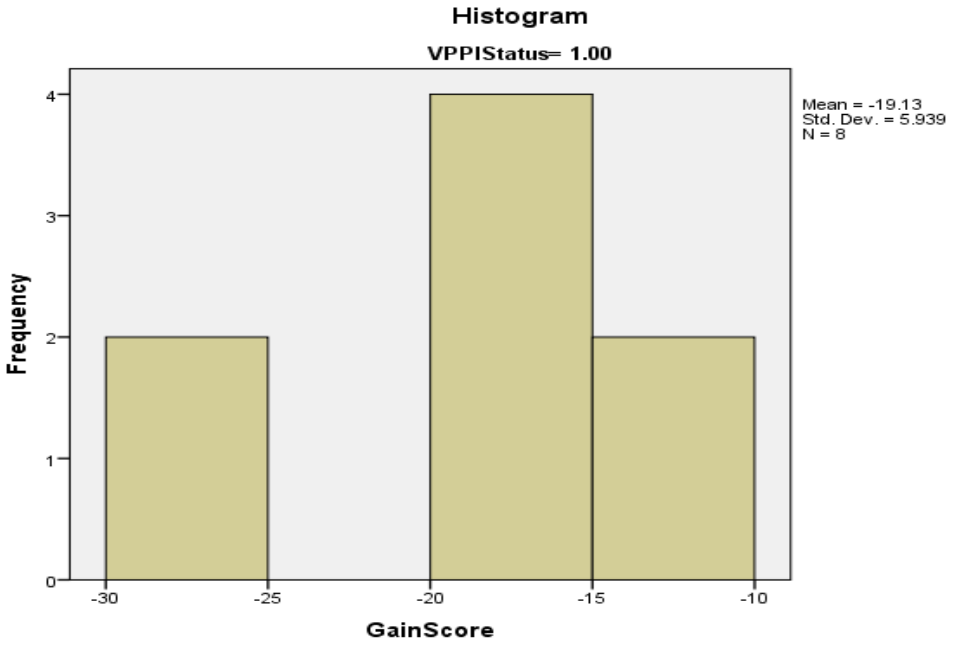
High School End-of-Course Algebra II Box Plot VPPI Nonparticipating



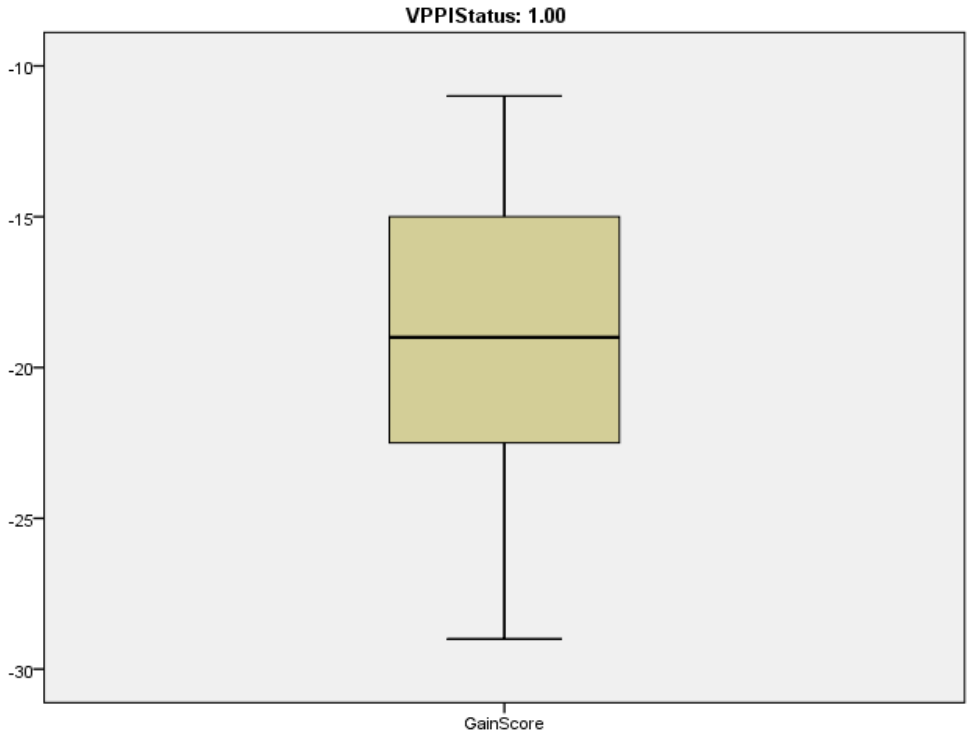
High School End-of-Course Algebra II Frequency Polygon VPPI Nonparticipating



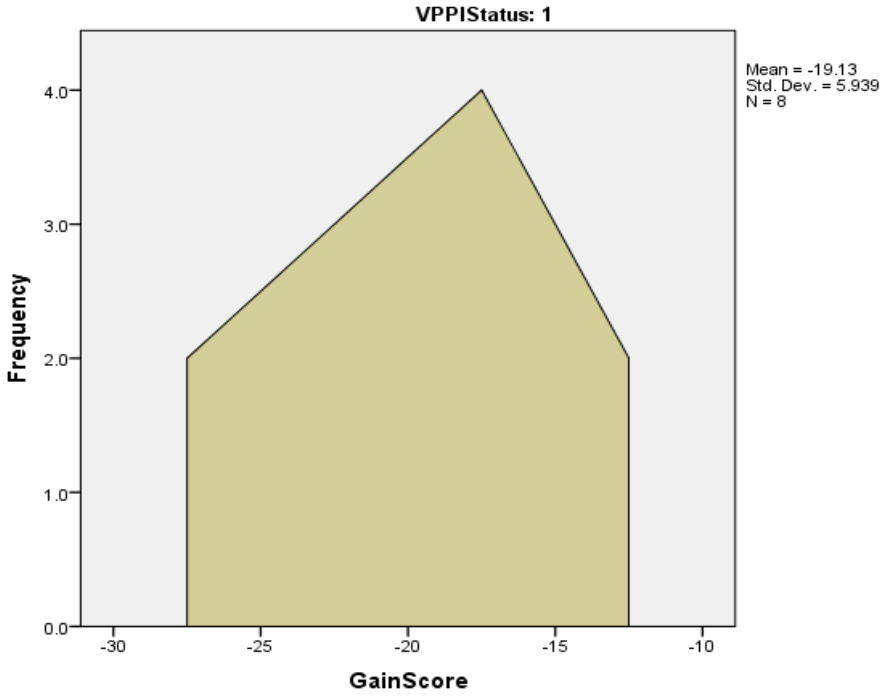
High School End-of-Course Geometry Histogram VPPI Participating



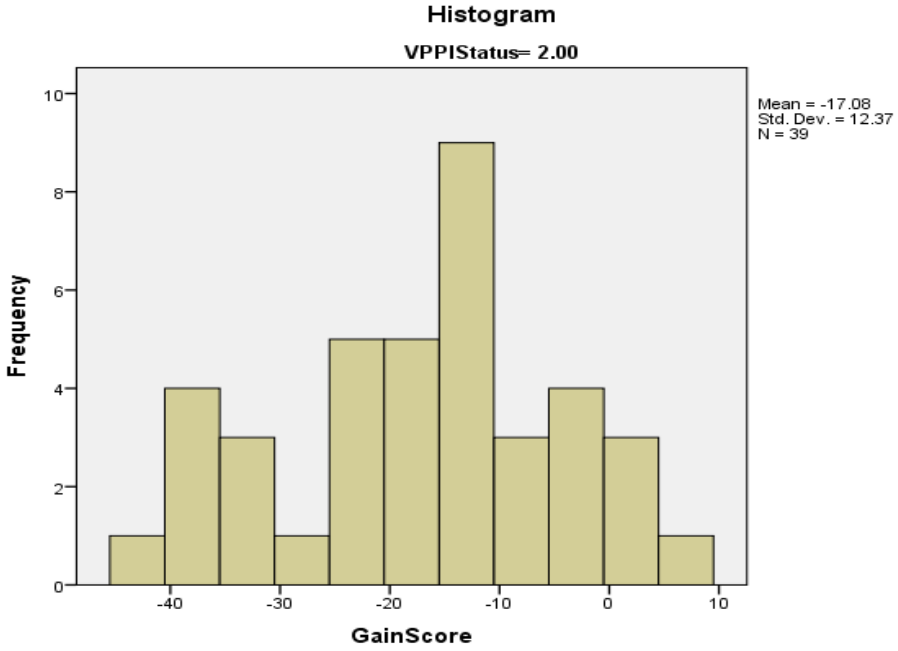
High School End-of-Course Geometry Box Plot VPPI Participating



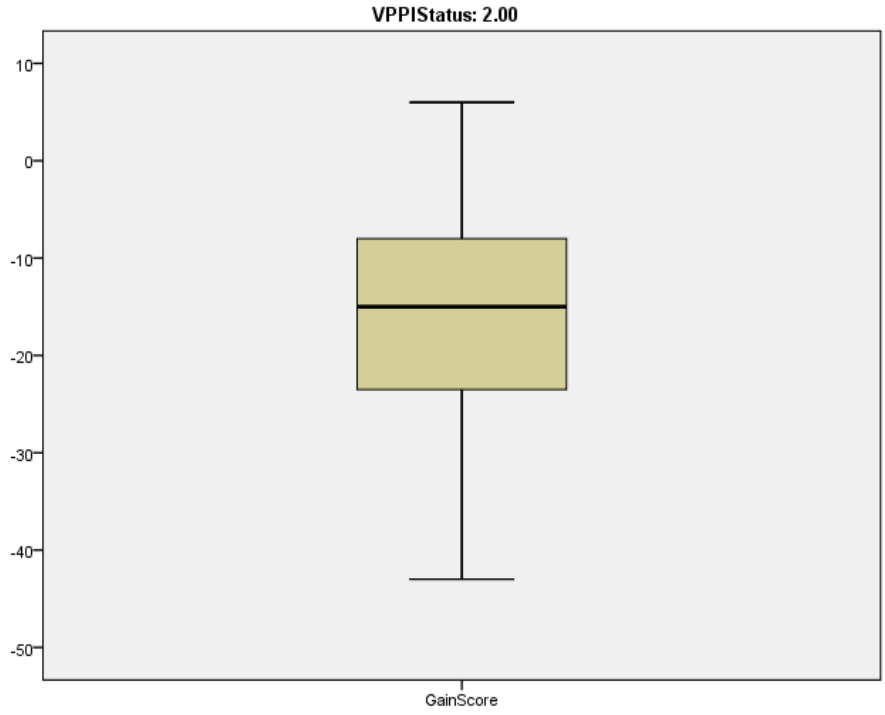
High School End-of-Course Geometry Frequency Polygon VPPI Participating



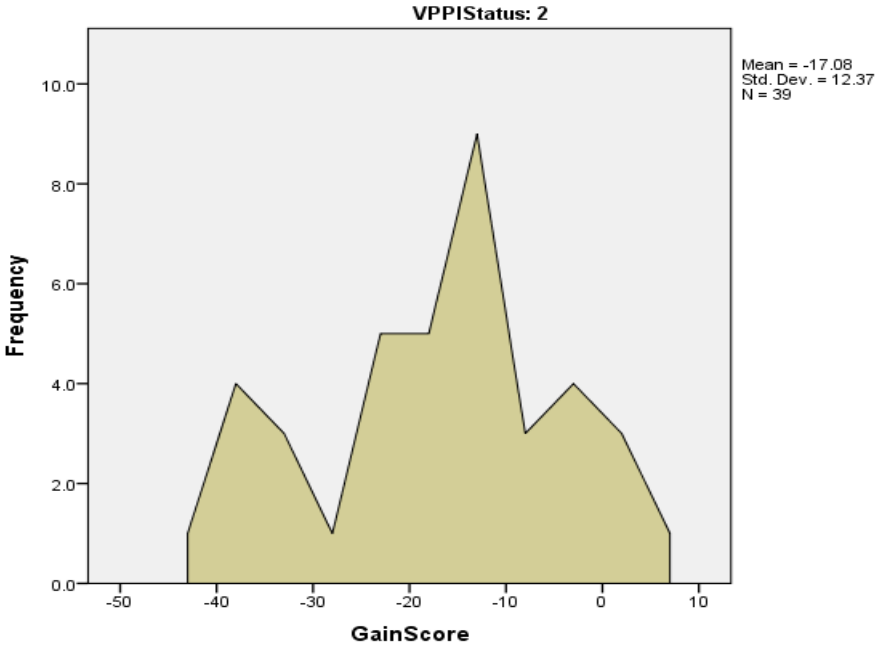
High School End-of-Course Geometry Histogram VPPI Nonparticipating



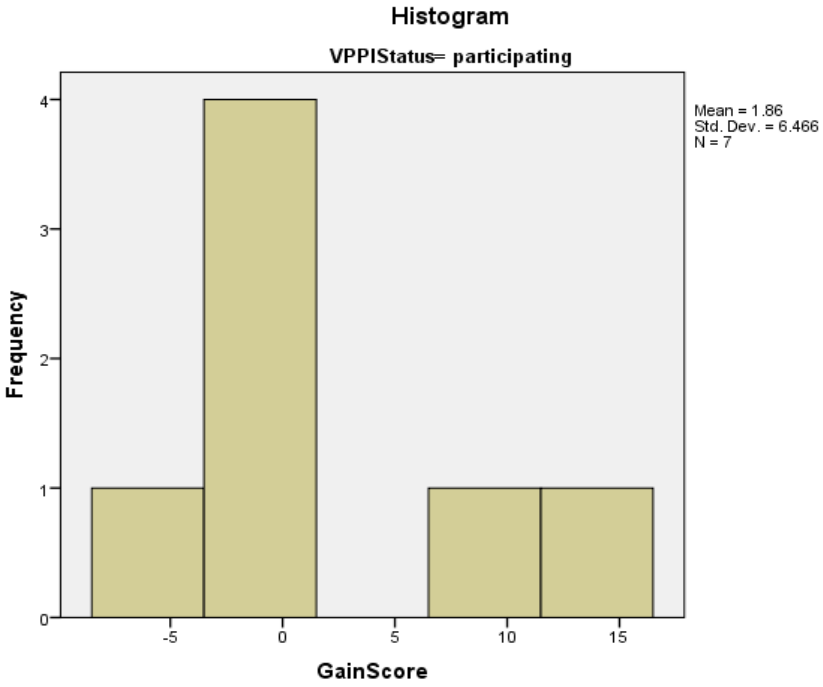
High School End-of-Course Geometry Box Plot VPPI Nonparticipating



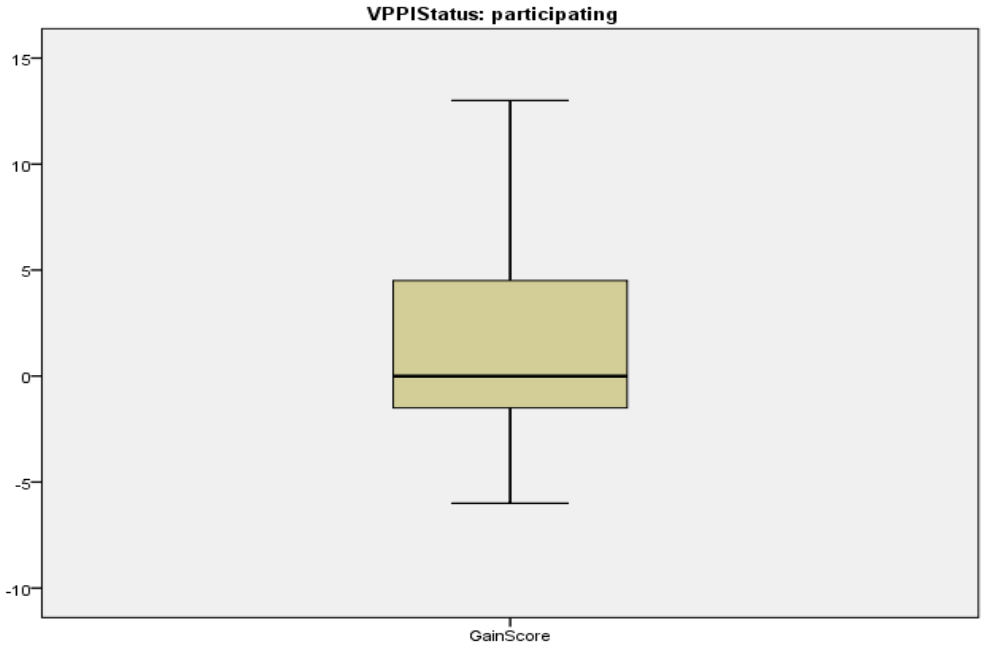
High School End-of-Course Geometry Frequency Polygon VPPI Nonparticipating



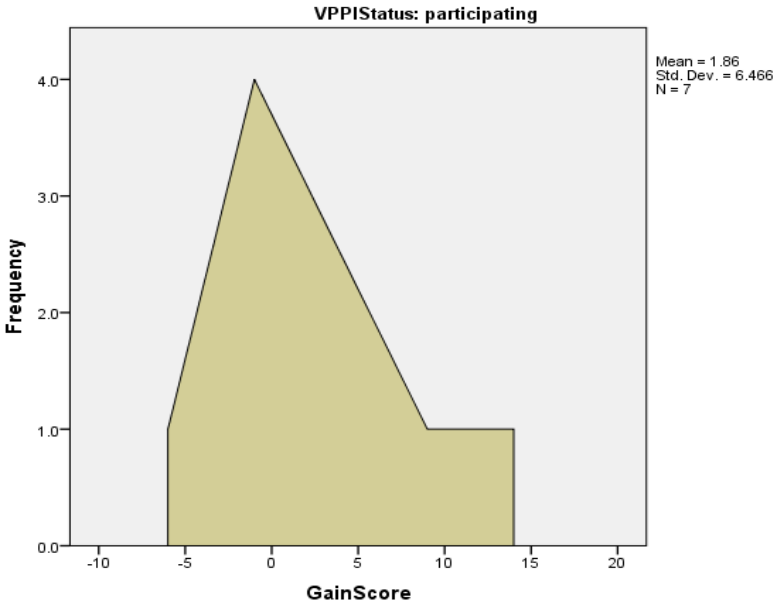
High School End-of-Course Biology Histogram VPPI Participating



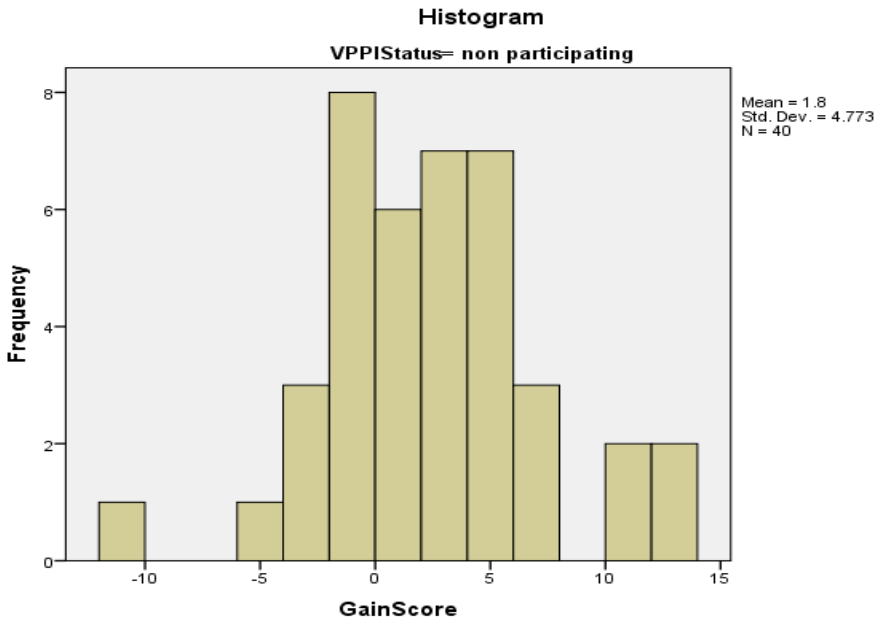
High School End-of-Course Biology Box Plot VPPI Participating



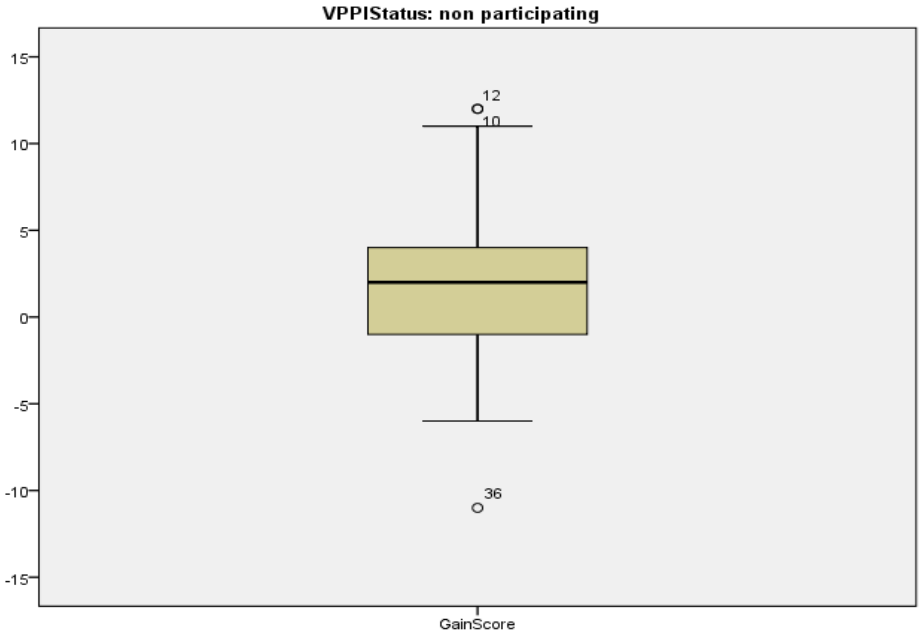
High School End-of-Course Biology Frequency Polygon VPPI Participating



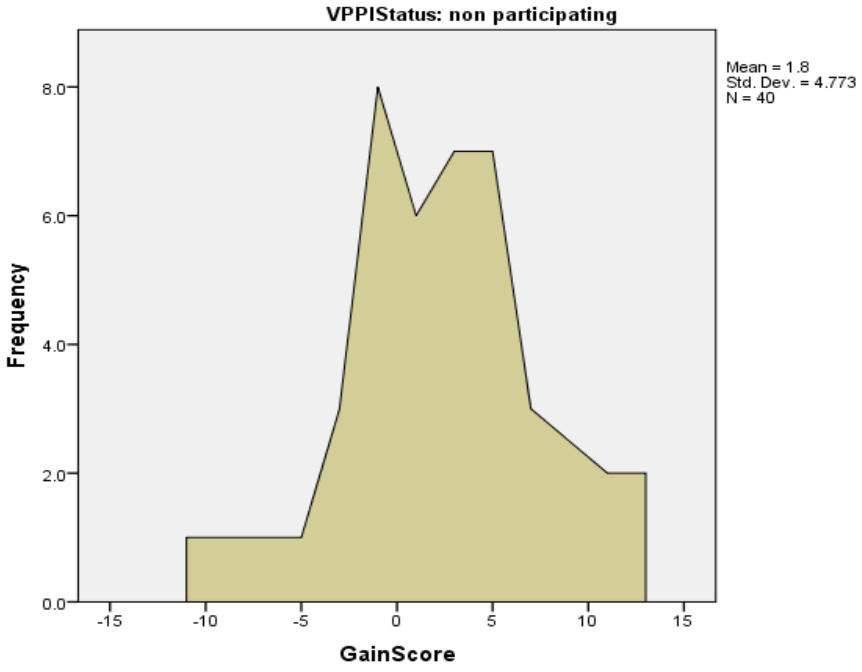
High School End-of-Course Biology Histogram VPPI Nonparticipating



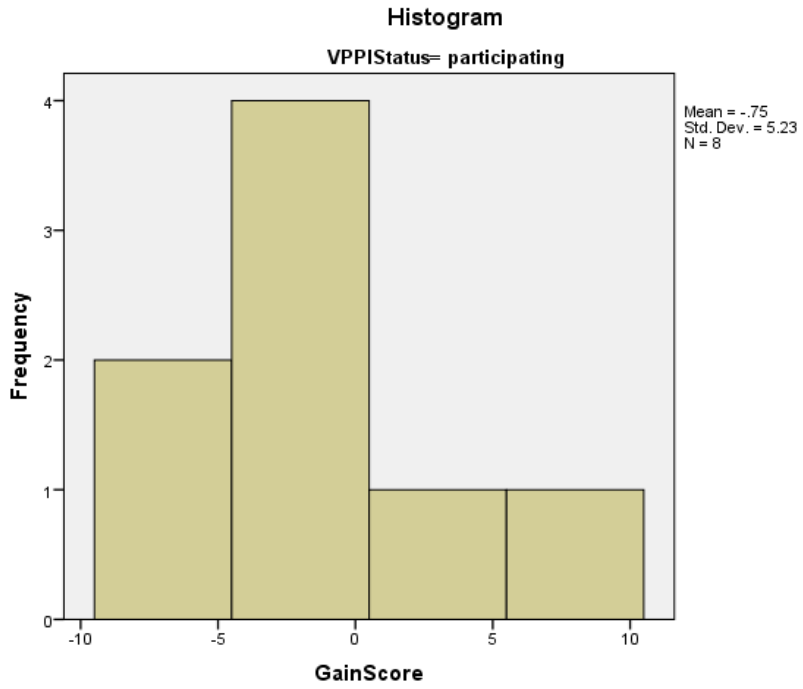
High School End-of-Course Biology Box Plot VPPI Nonparticipating



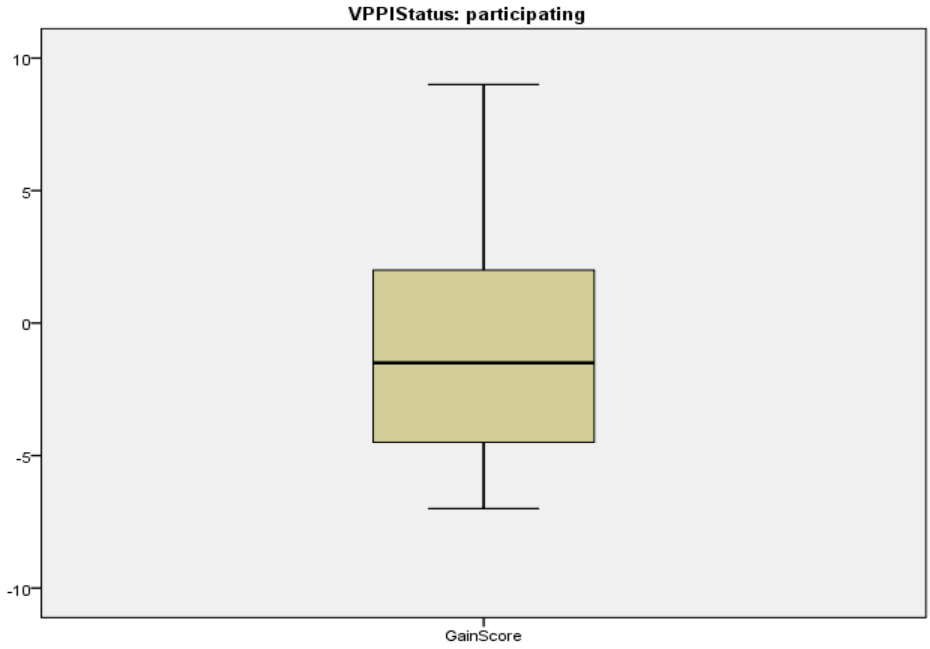
High School End-of-Course Biology Frequency Polygon VPPI Nonparticipating



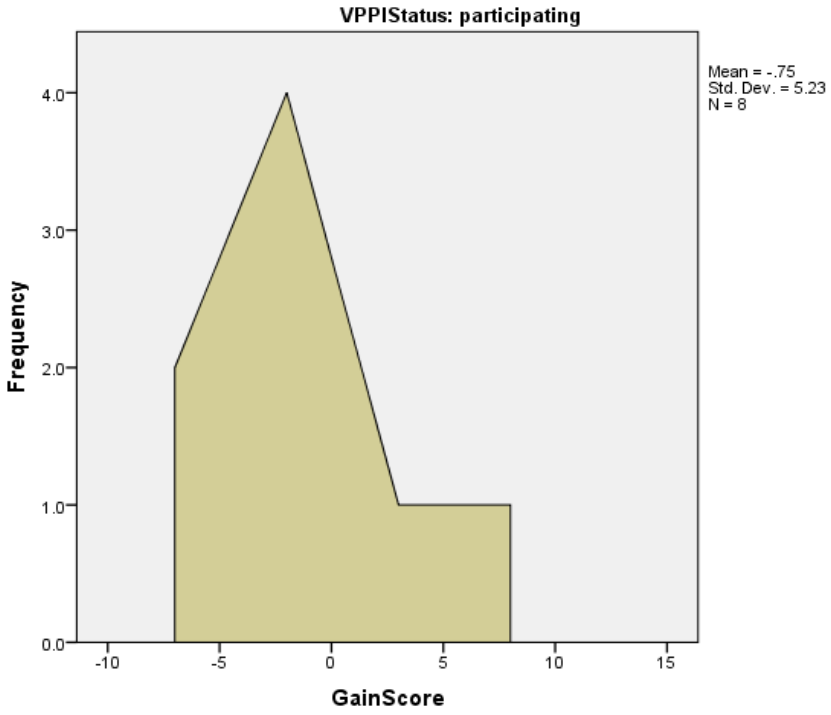
High School End-of-Course Chemistry Histogram VPPI Participating



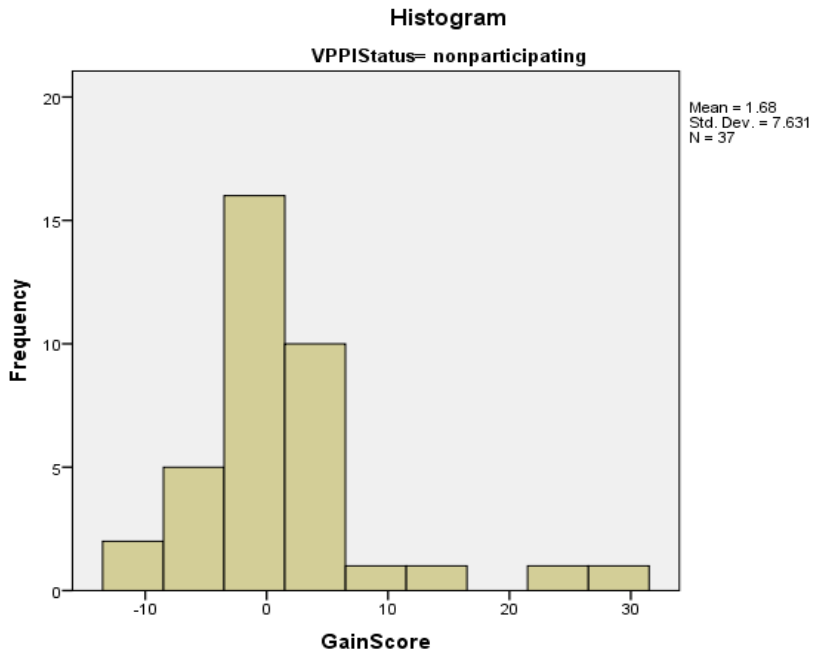
High School End-of-Course Chemistry Box Plot VPPI Participating



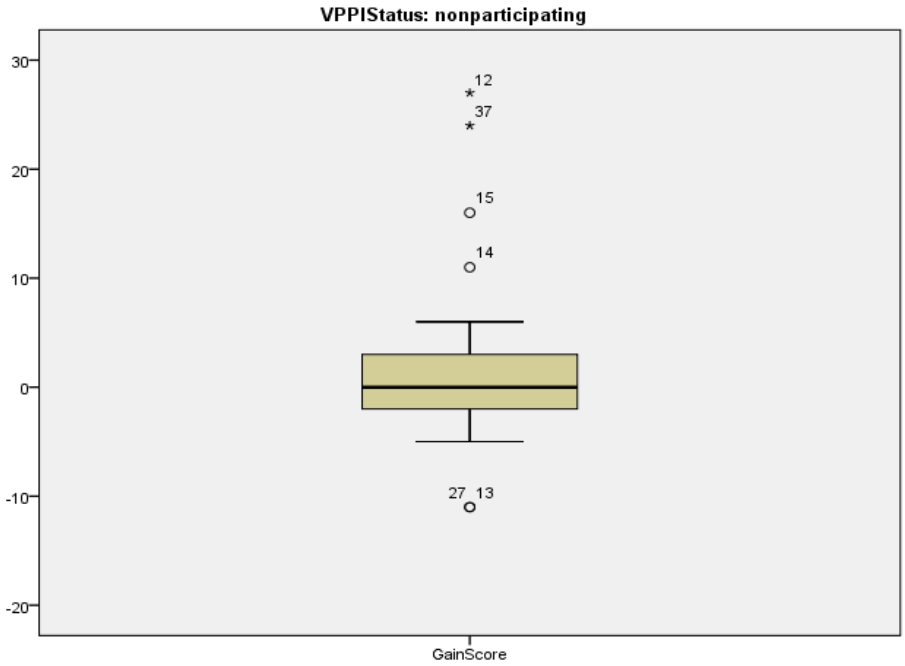
High School End-of-Course Chemistry Frequency Polygon VPPI Participating



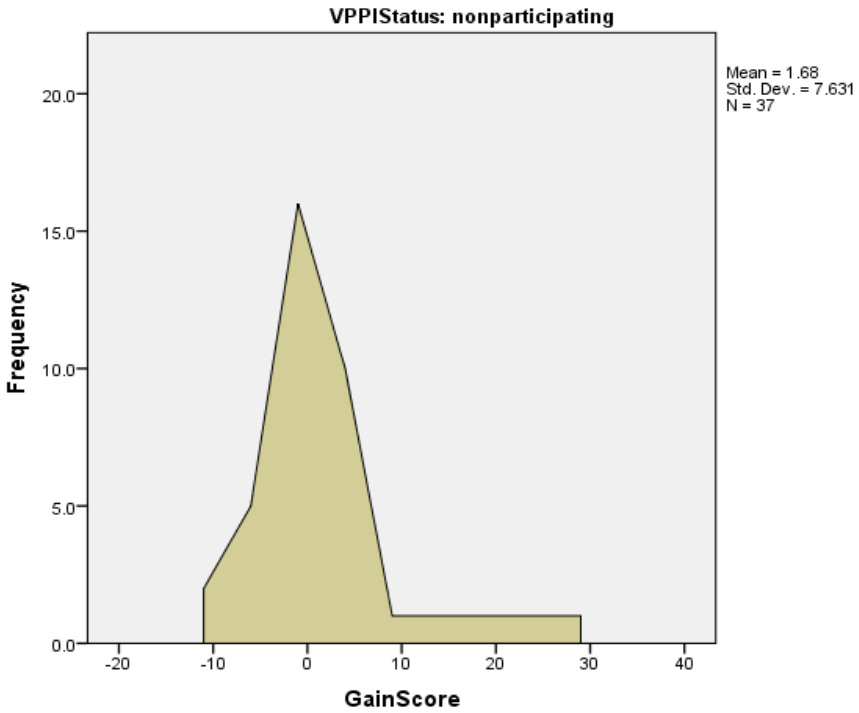
High School End-of-Course Chemistry Histogram VPPI Nonparticipating



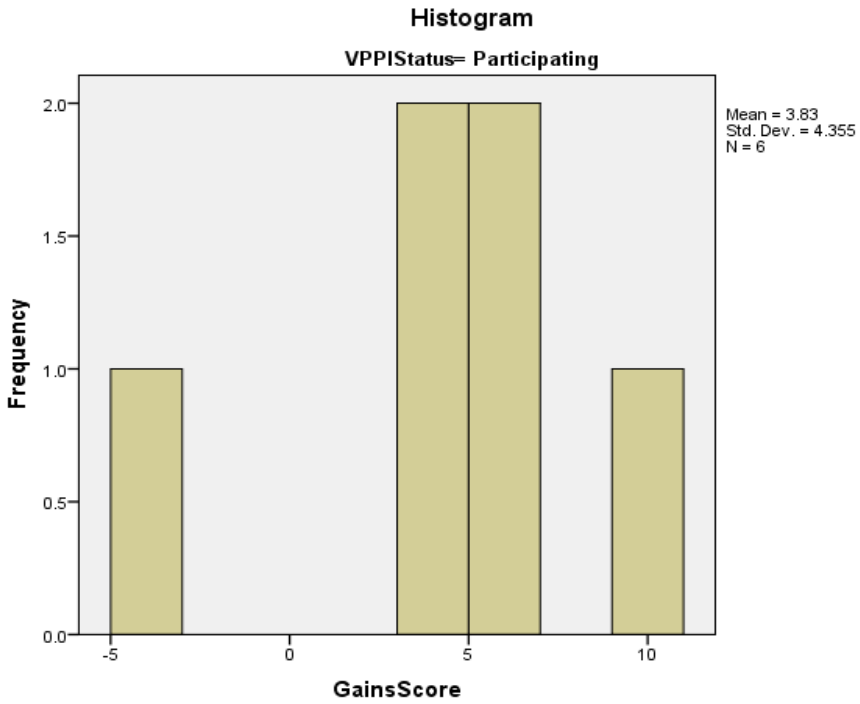
High School End-of-Course Chemistry Box Plot VPPI Nonparticipating



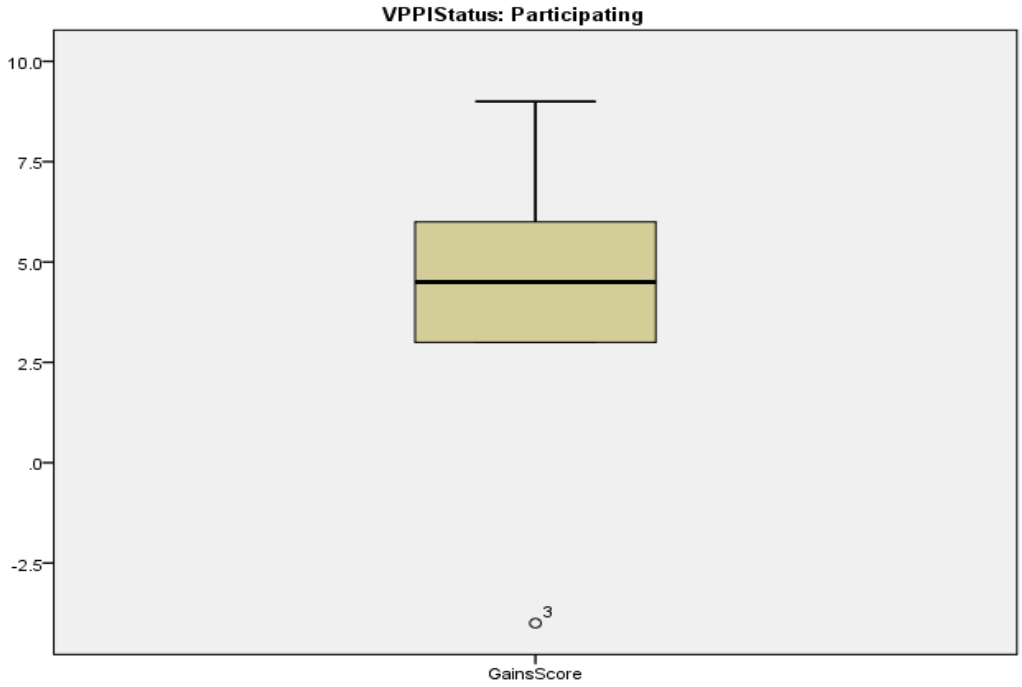
High School End-of-Course Chemistry Frequency Polygon VPPI Nonparticipating



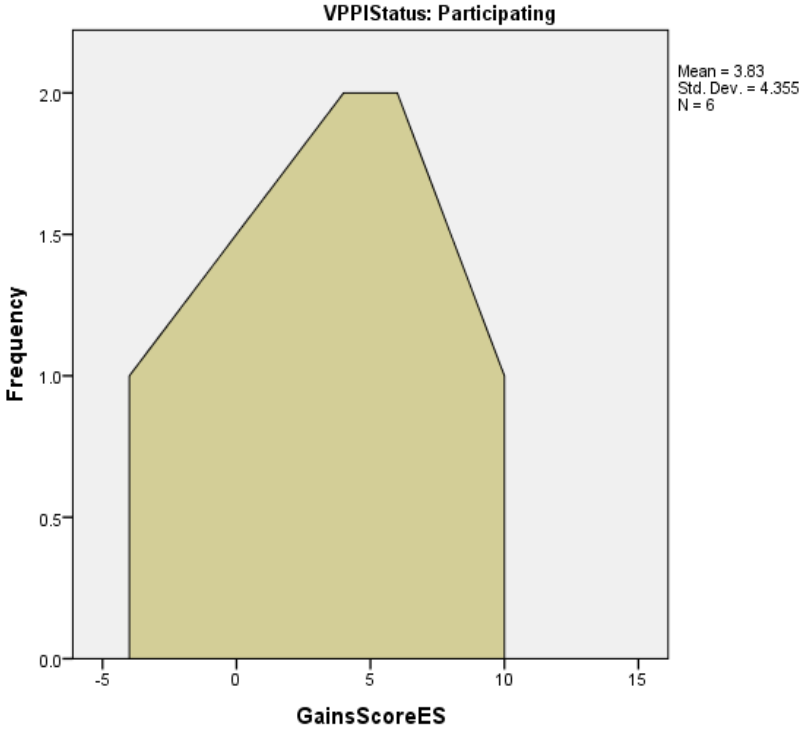
High School End-of-Course Earth Science Histogram VPPI Participating



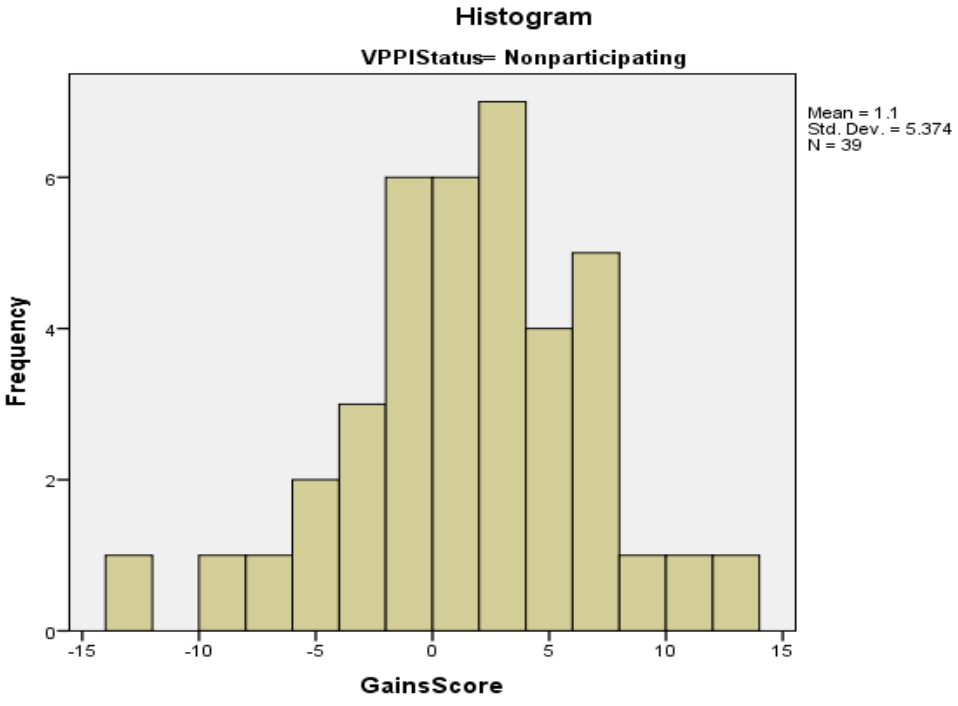
High School End-of-Course Earth Science Box Plot VPPI Participating



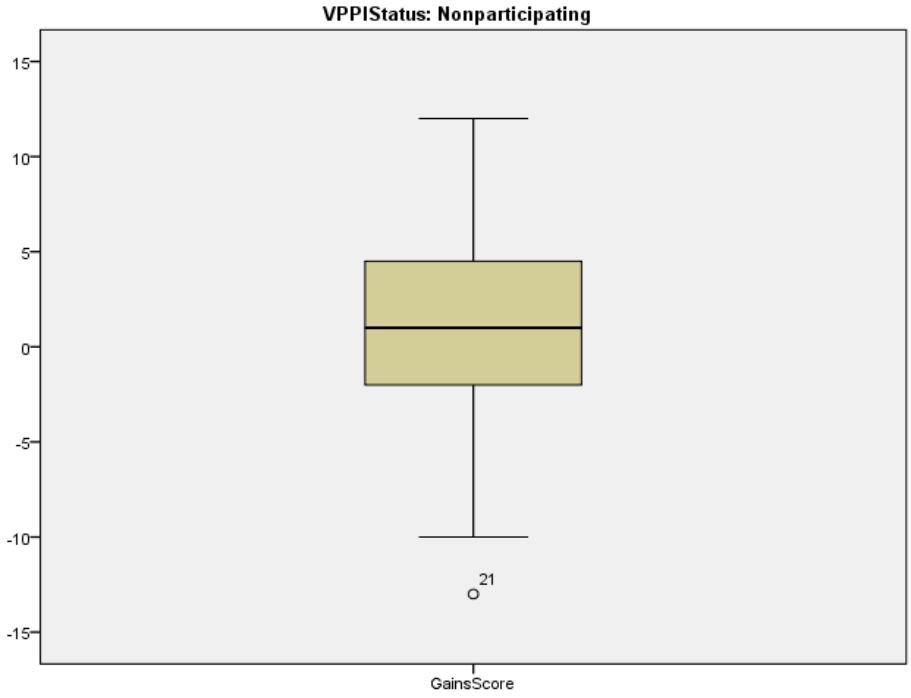
High School End-of-Course Earth Science Frequency Polygon VPPI Participating



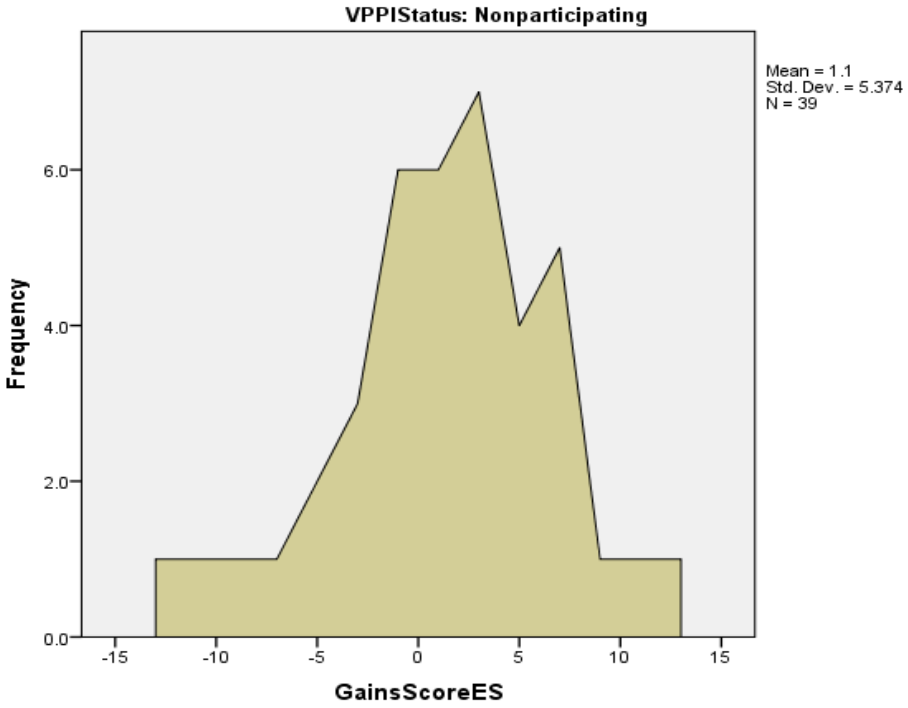
High School End-of-Course Earth Science Histogram VPPI Nonparticipating



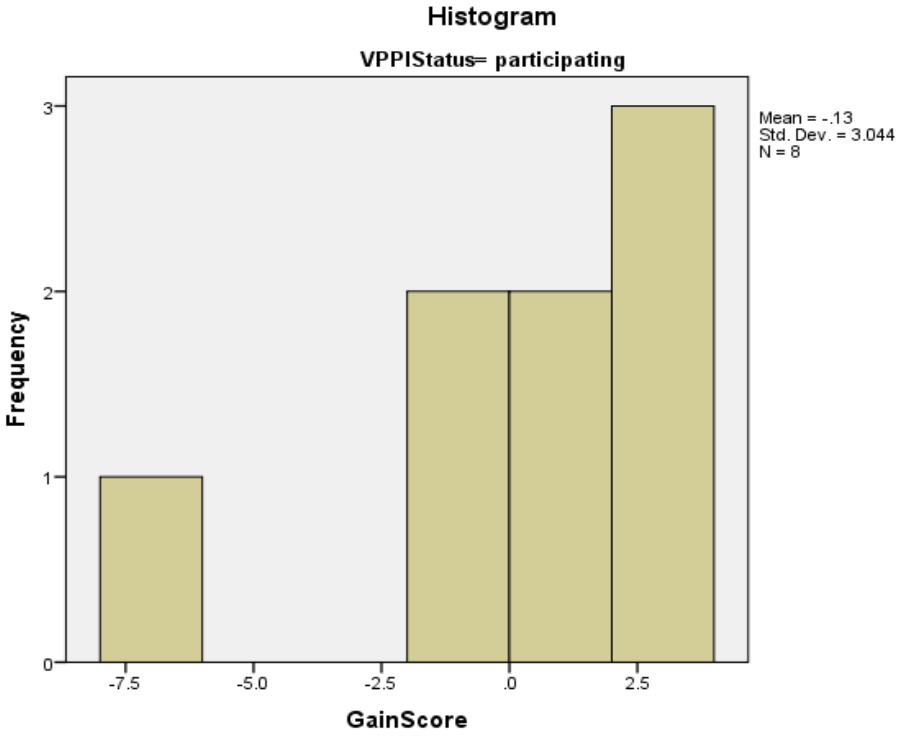
High School End-of-Course Earth Science Box Plot VPPI Nonparticipating



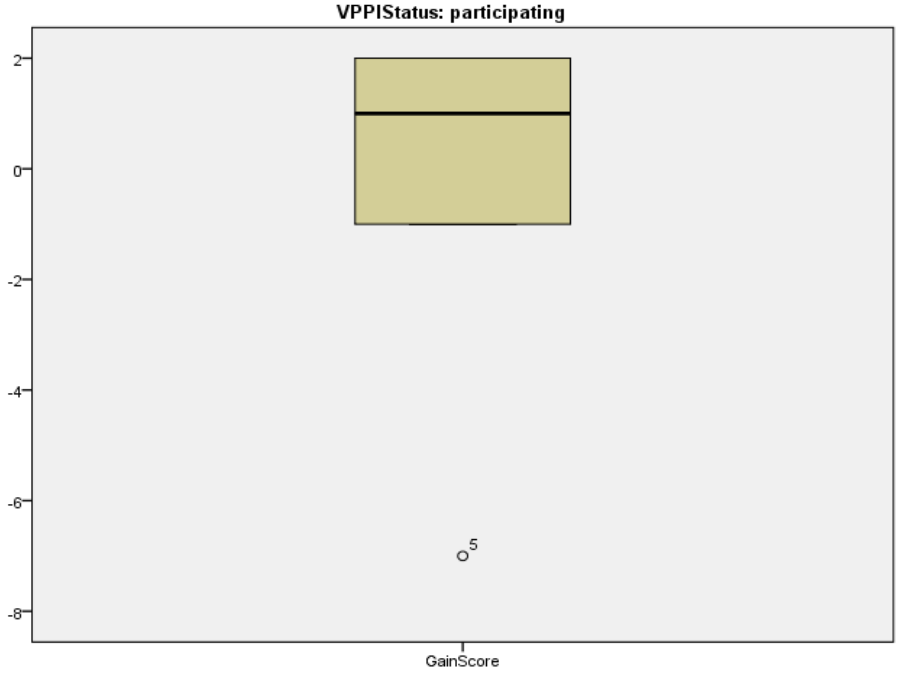
High School End-of-Course Earth Science Frequency Polygon Histogram VPPI Nonparticipating


VPPIStatus: Nonparticipating
Mean = 1.1
Std. Dev. = 5.374
N = 39
Frequency
6.0
4.0
2.0
0.0
-15
-10
-5
0
5
10
15
GainsScoreES

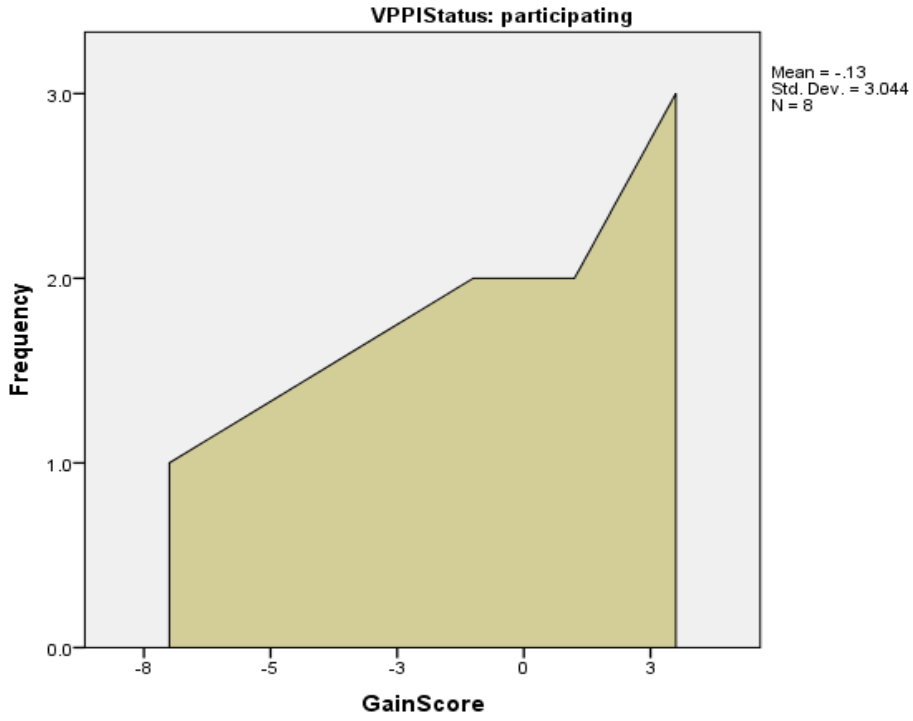
High School End-of-Course Reading Histogram VPPI Participating



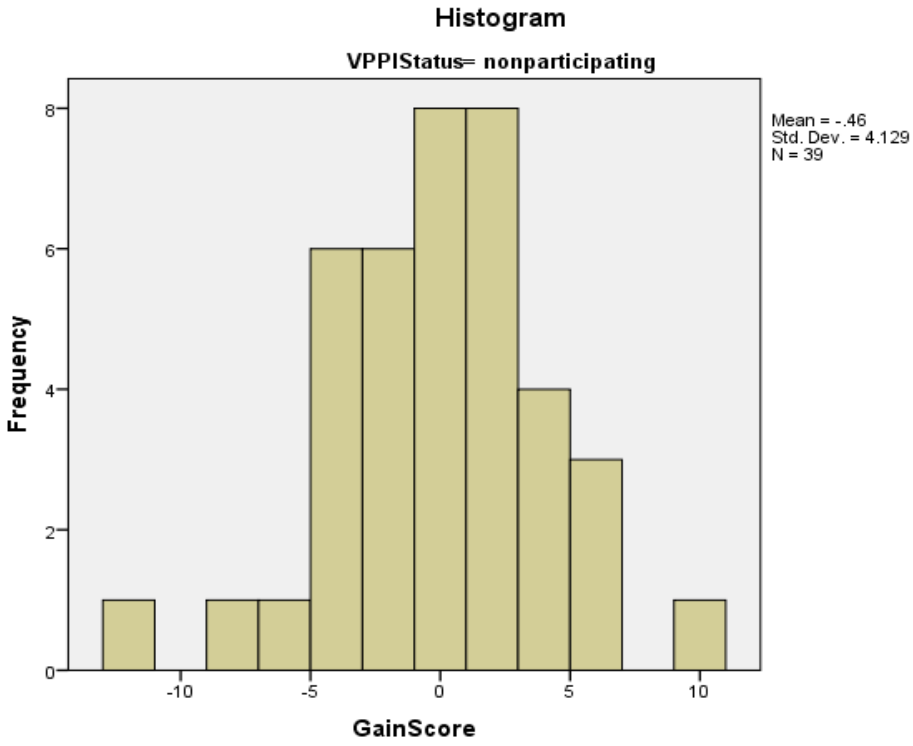
High School End-of-Course Reading Box Plot VPPI Participating



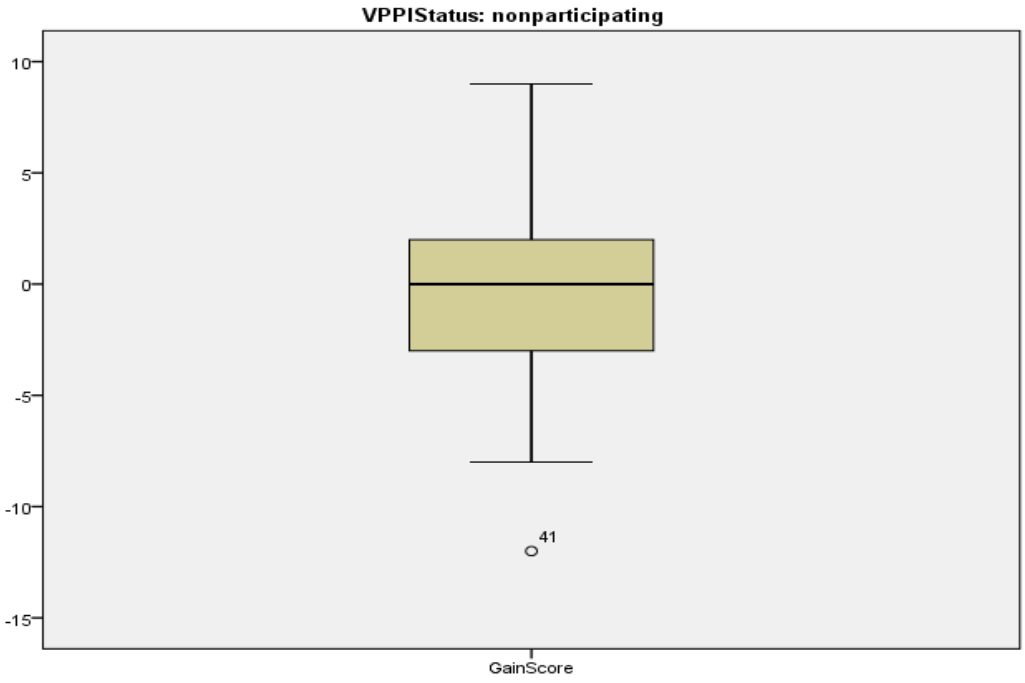
High School End-of-Course Reading Frequency Polygon VPPI Participating



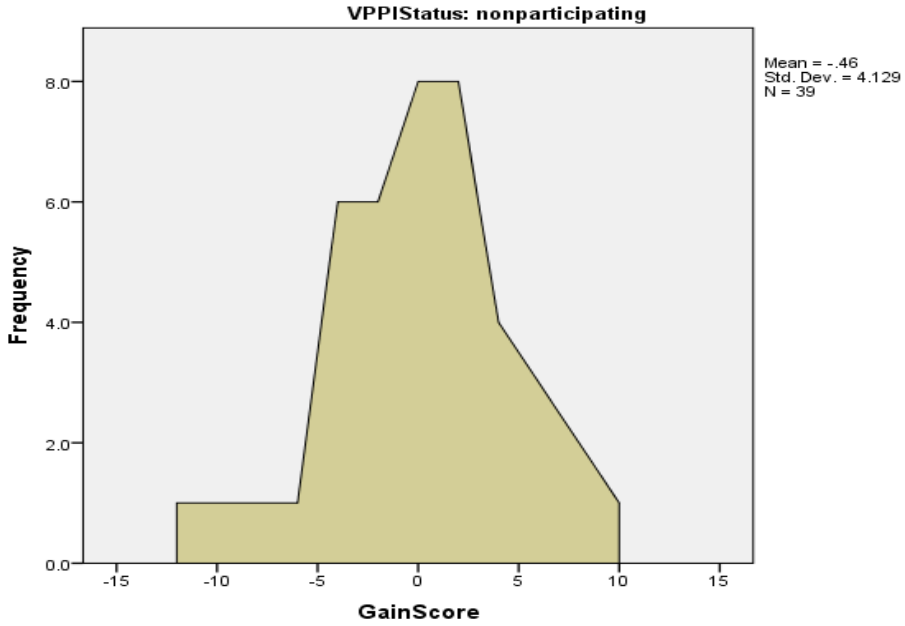
High School End-of-Course Reading Histogram VPPI Nonparticipating



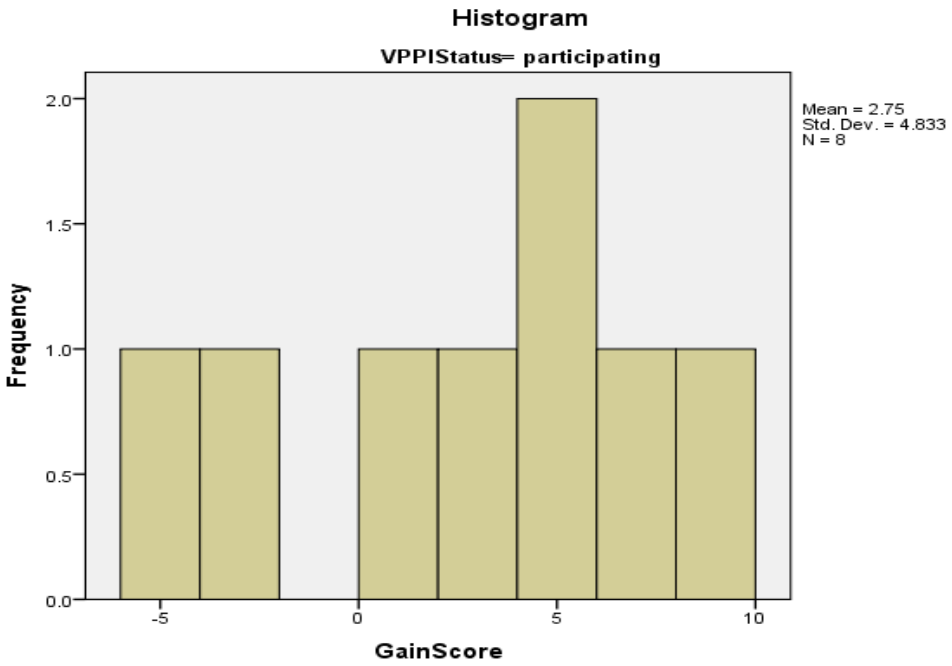
High School End-of-Course Reading Box Plot VPPI nonparticipating



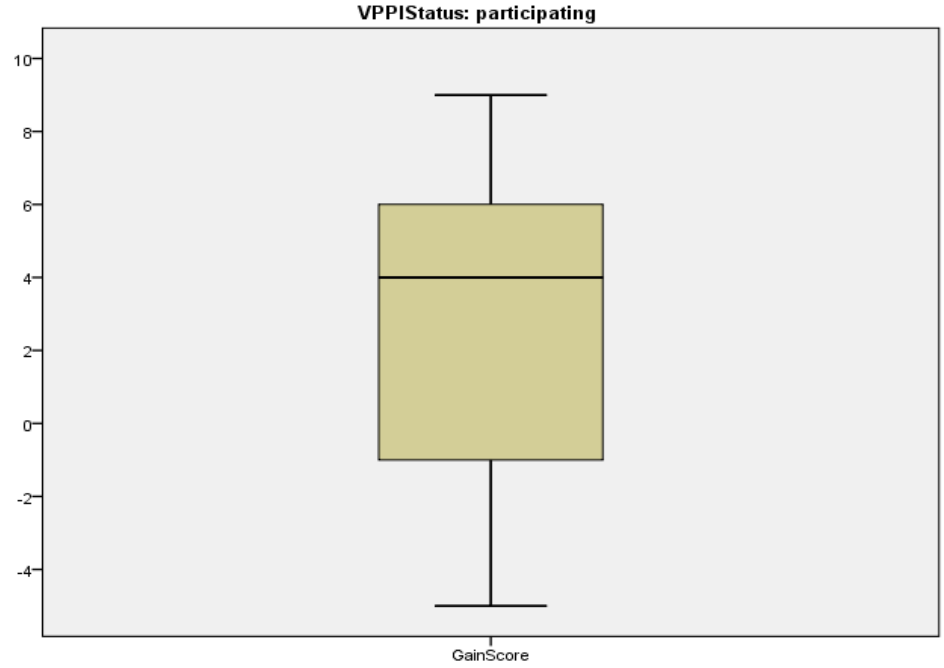
High School End-of-Course Reading Frequency Polygon VPPI Nonparticipating



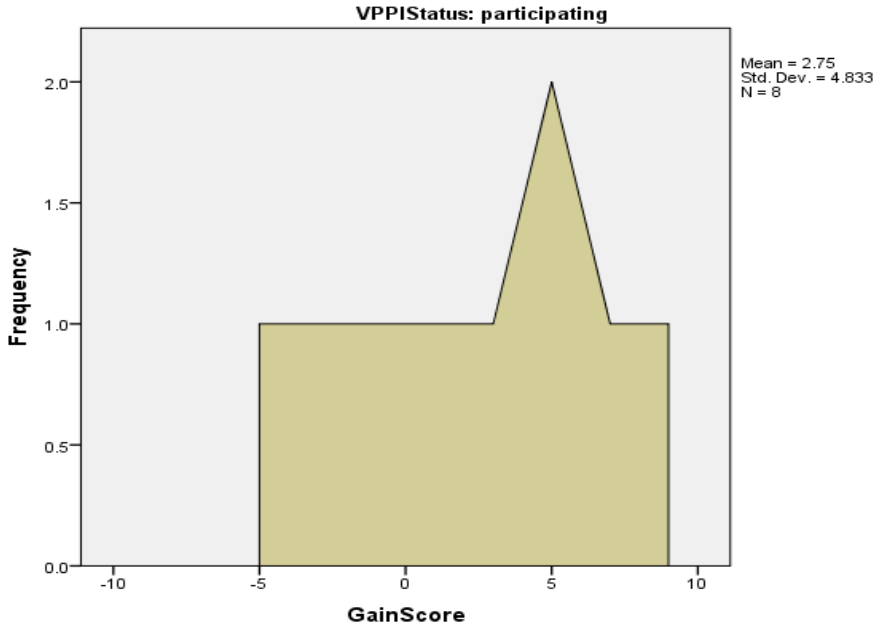
High School End-of-Course Writing Histogram VPPI Participating



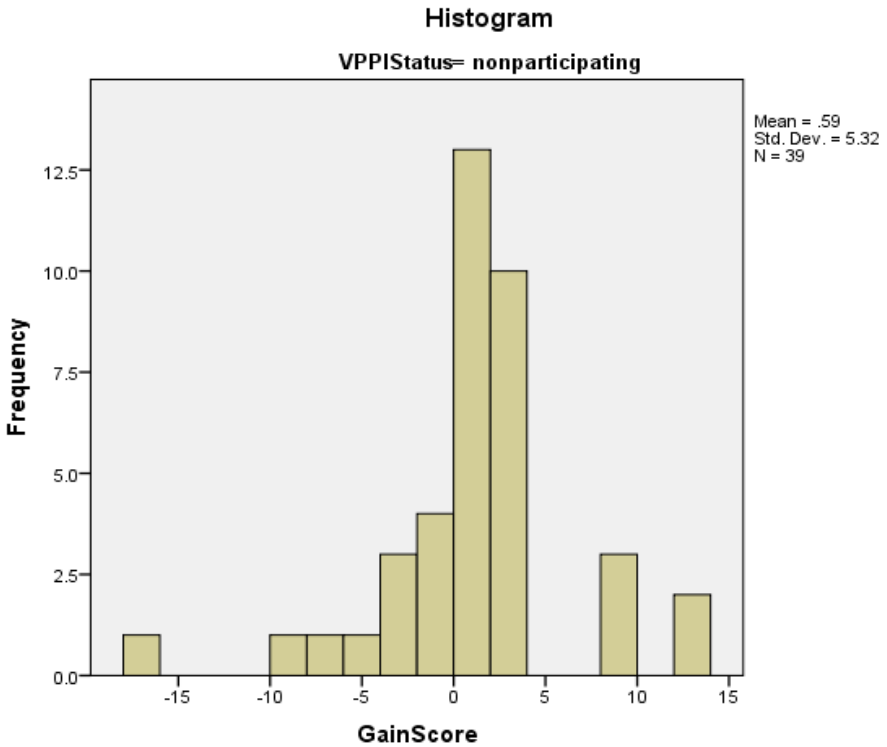
High School End-of-Course Writing Box Plot VPPI Participating



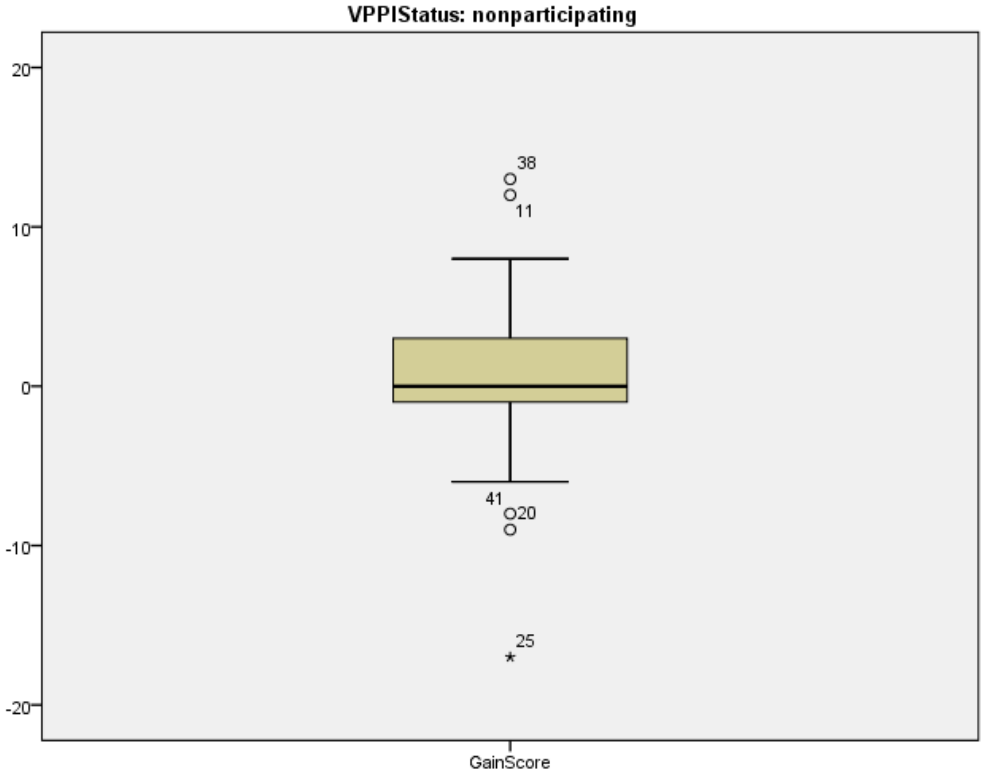
High School End-of-Course Writing Frequency Polygon VPPI Participating



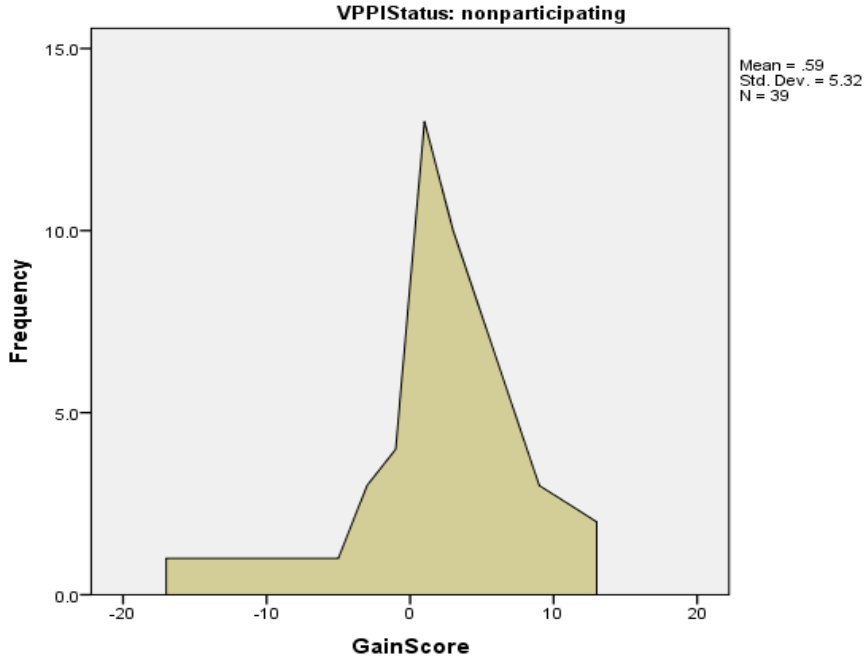
High School End-of-Course Writing Histogram VPPI Nonparticipating



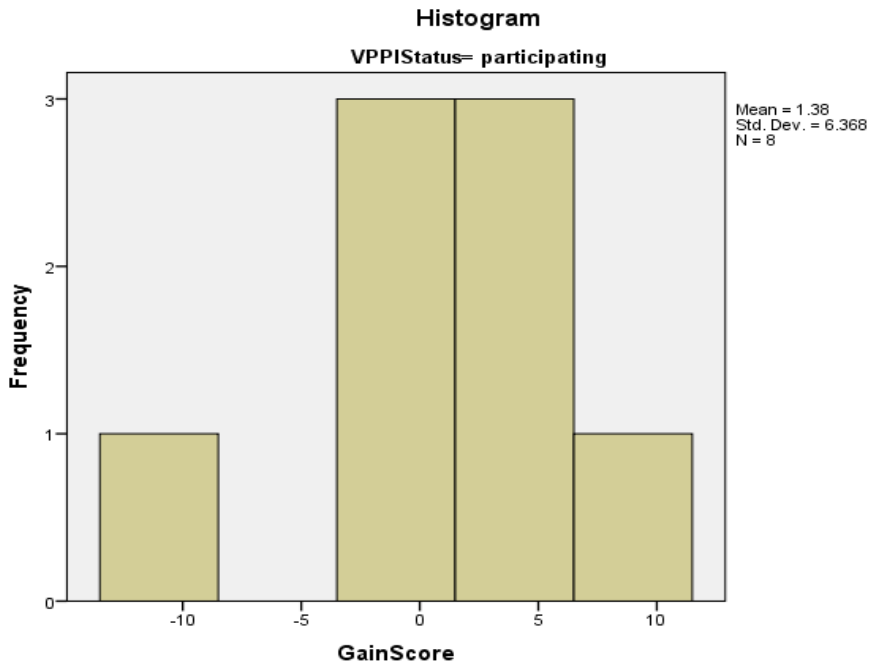
High School End-of-Course Writing Box Plot VPPI Nonparticipating



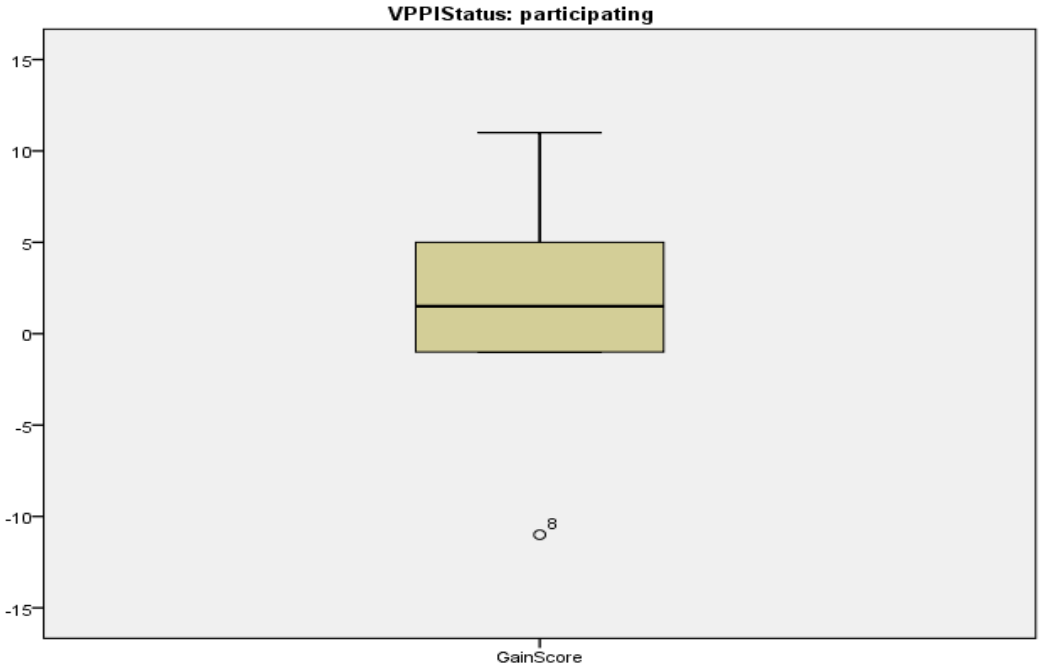
High School End-of-Course Writing Frequency Polygon VPPI Nonparticipating



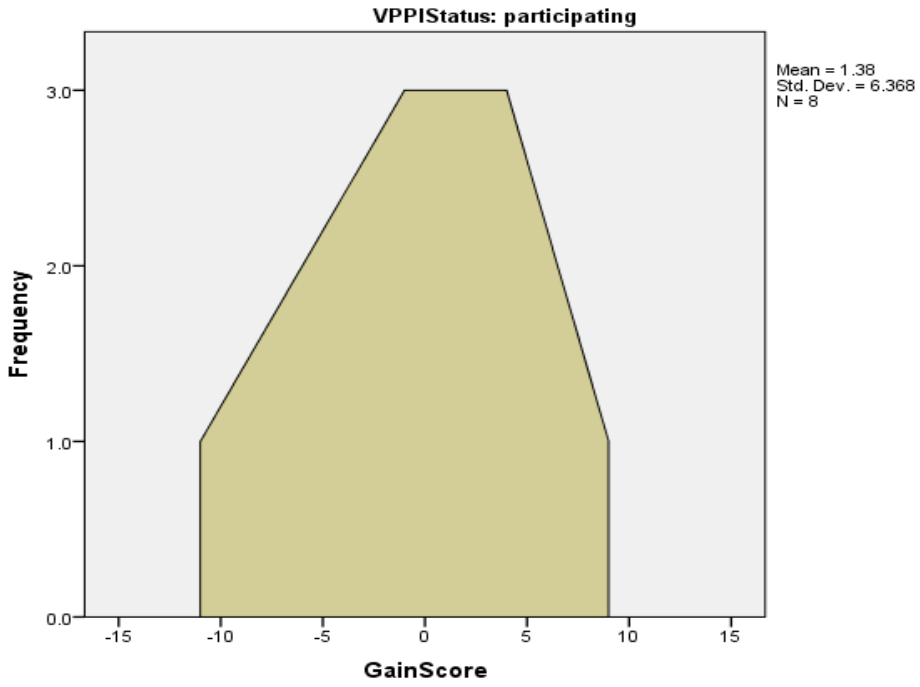
High School End-of-Course VA/US History Histogram VPPI Participating



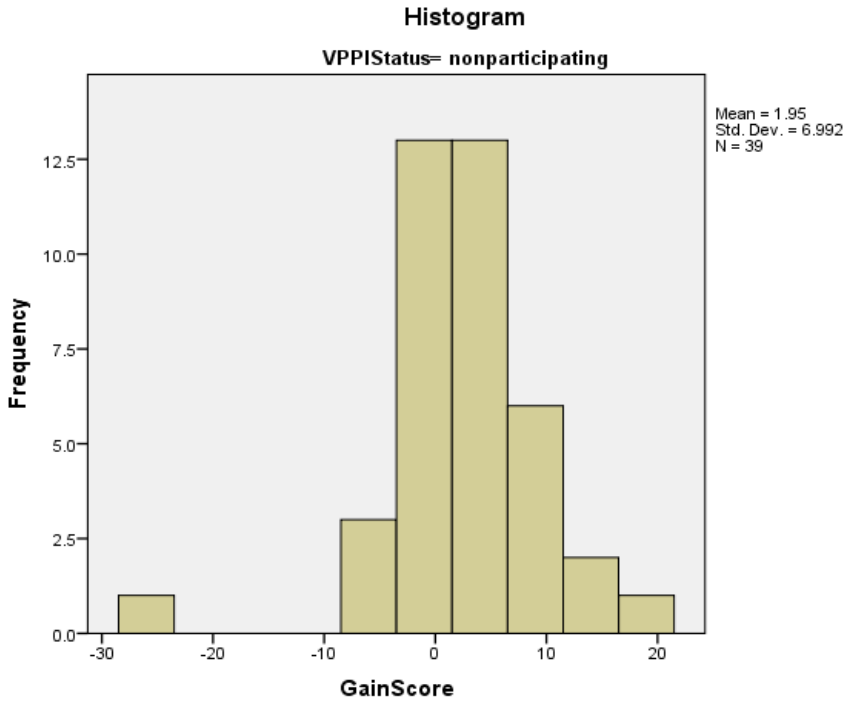
High School End-of-Course VA/US History Box Plot VPPI Participating



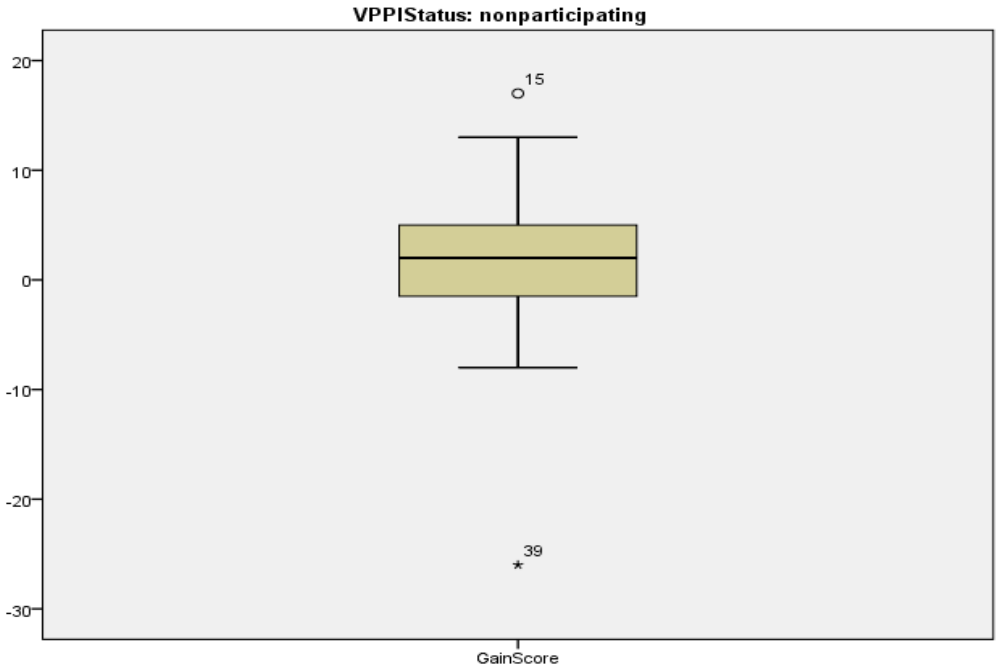
High School End-of-Course VA/US History Frequency Polygon VPPI Participating



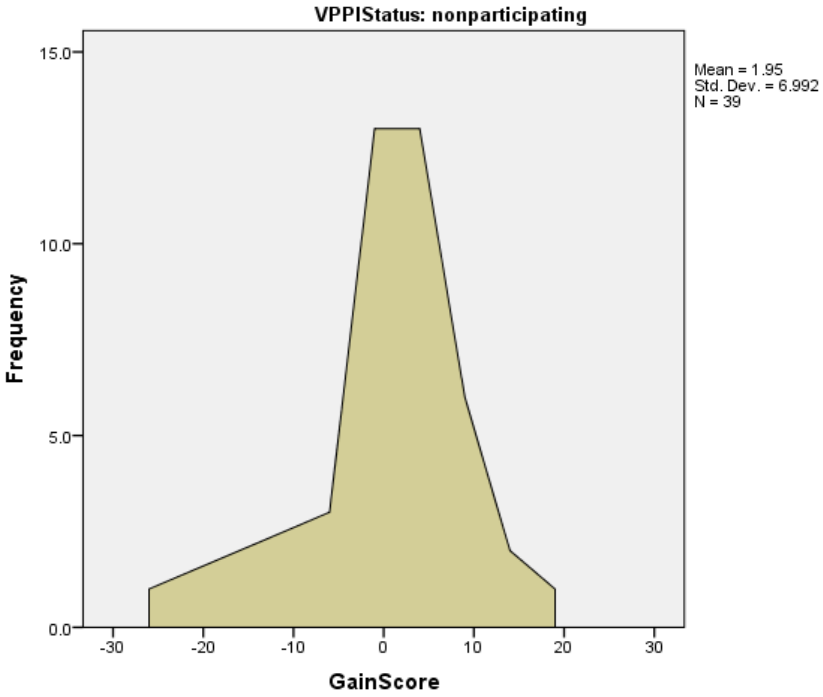
High School End-of-Course VA/US History Histogram VPPI Nonparticipating



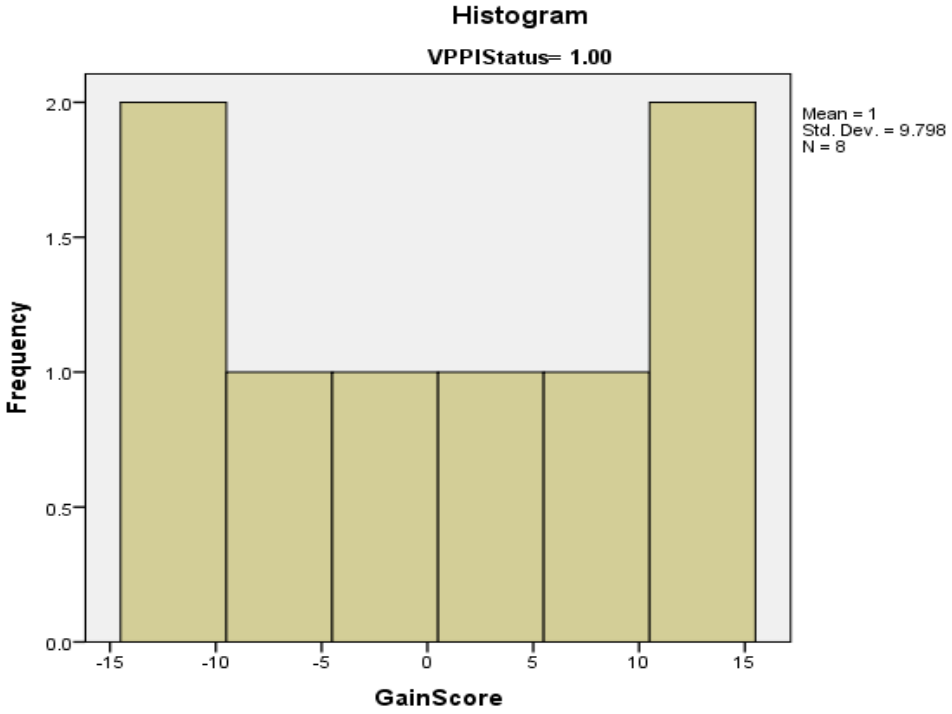
High School End-of-Course VA/US History Box Plot VPPI Nonparticipating



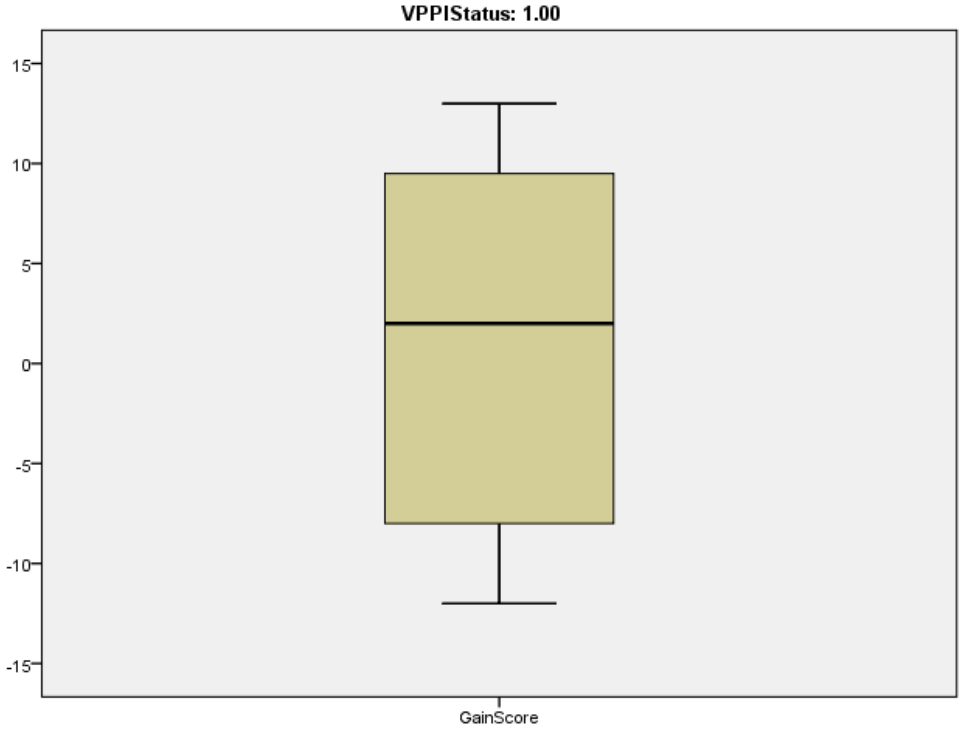
High School End-of-Course VA/US History Frequency Polygon VPPI Nonparticipating



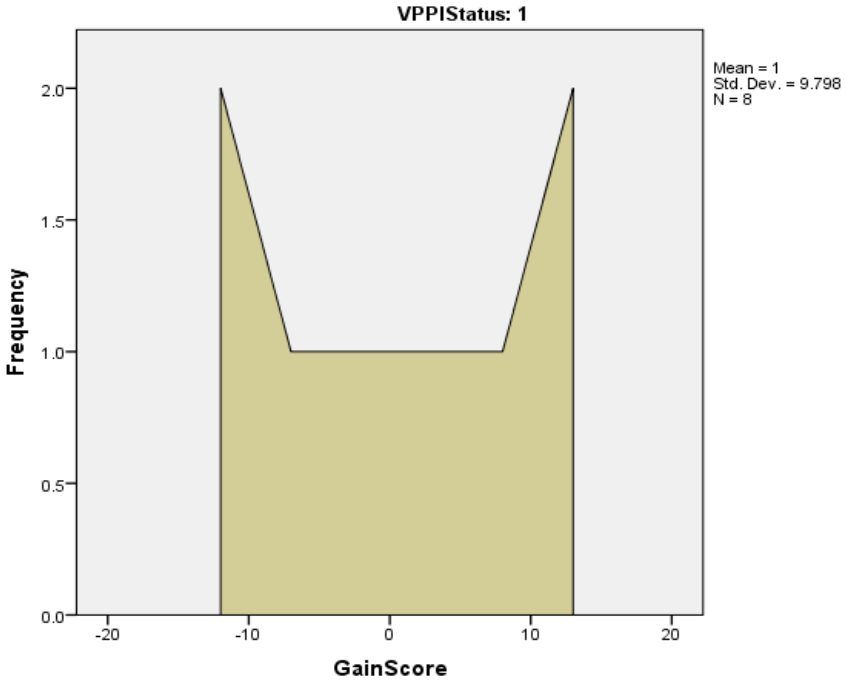
High School End-of-Course World History I Histogram VPPI Participating



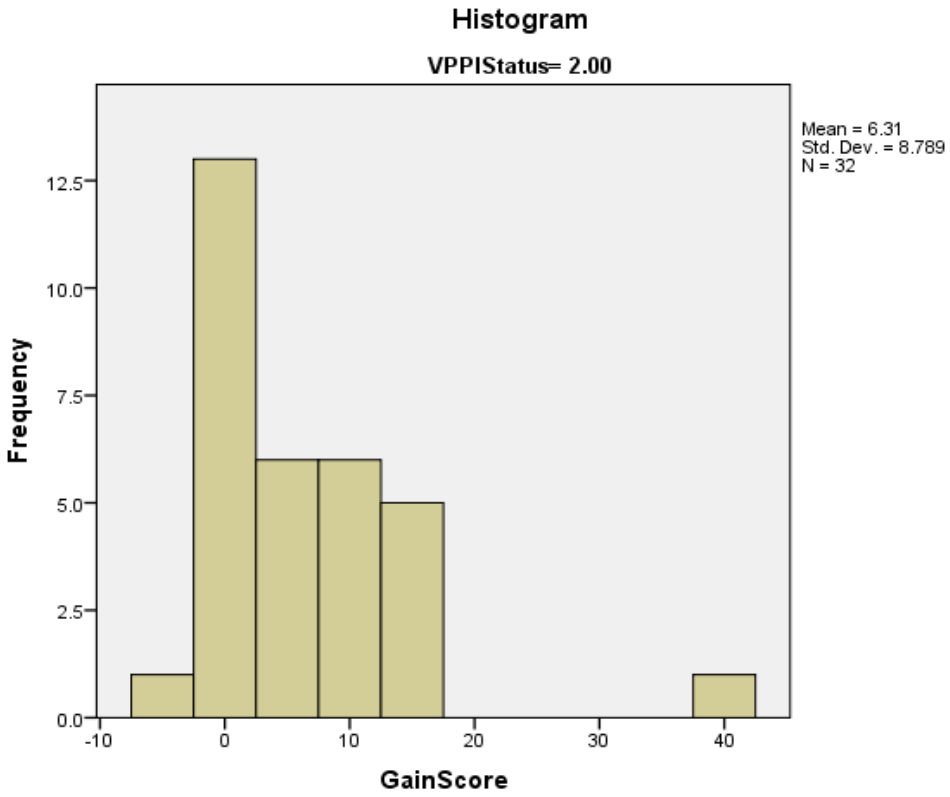
High School End-of-Course World History I Box Plot VPPI Participating



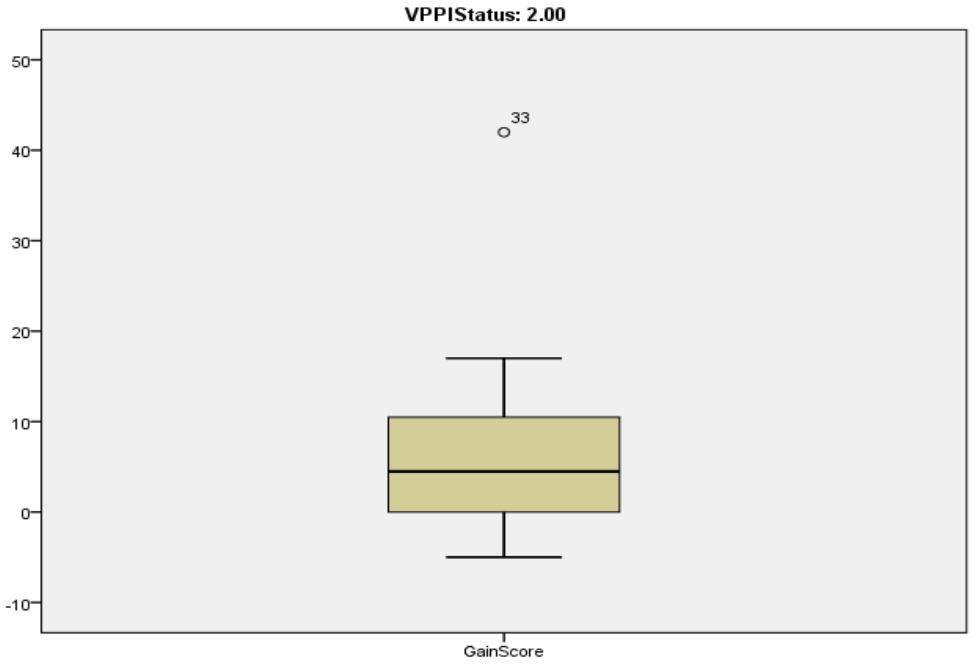
High School End-of-Course World History I Frequency Polygon VPPI Participating



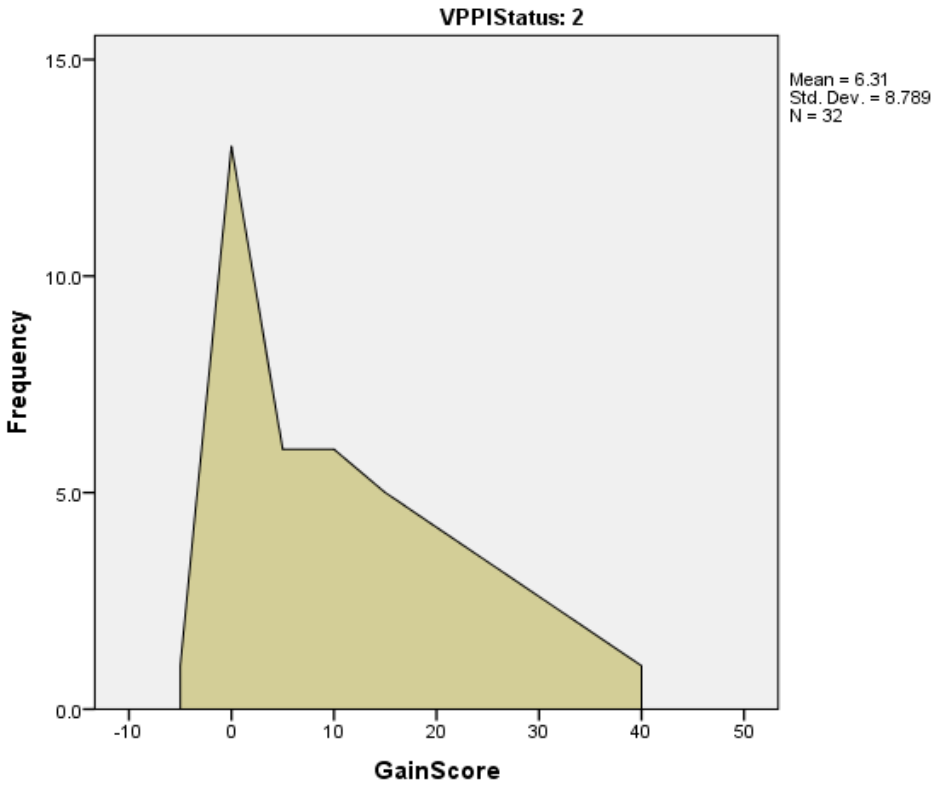
High School End-of-Course World History I Histogram VPPI Nonparticipating



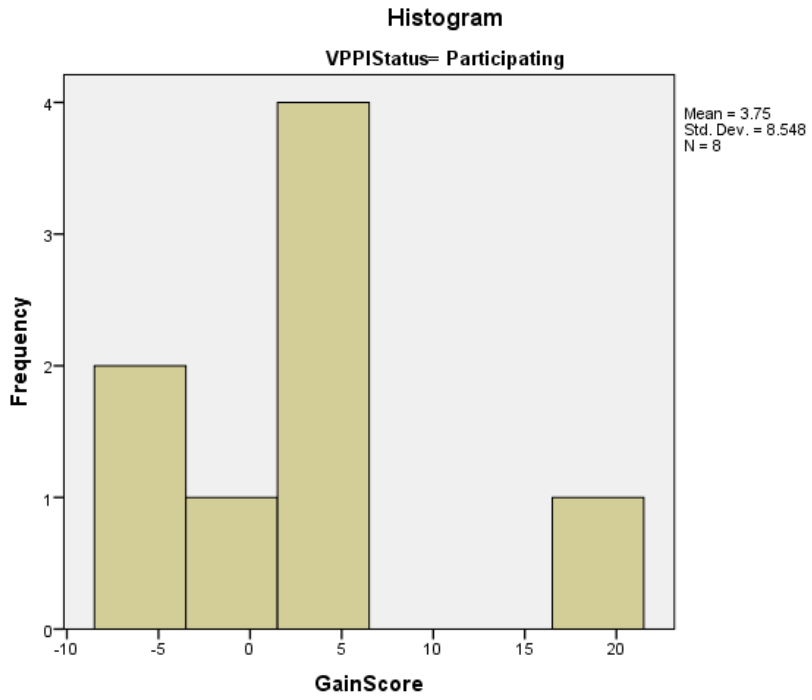
High School End-of-Course World History I Box Plot VPPI Nonparticipating



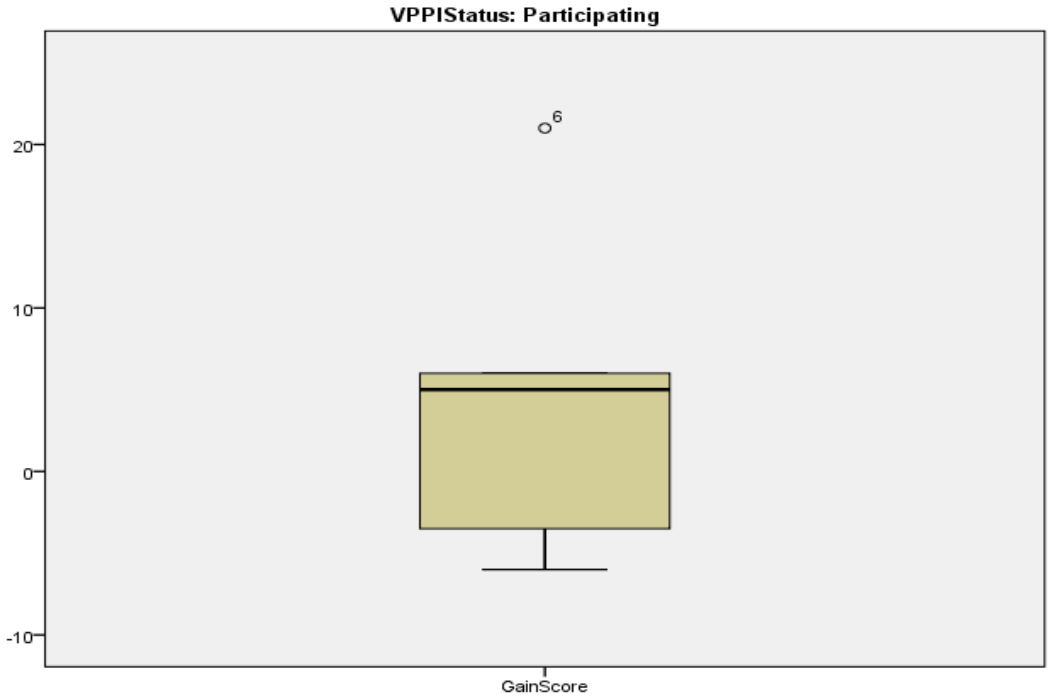
High School End-of-Course World History I Frequency Polygon VPPI Nonparticipating



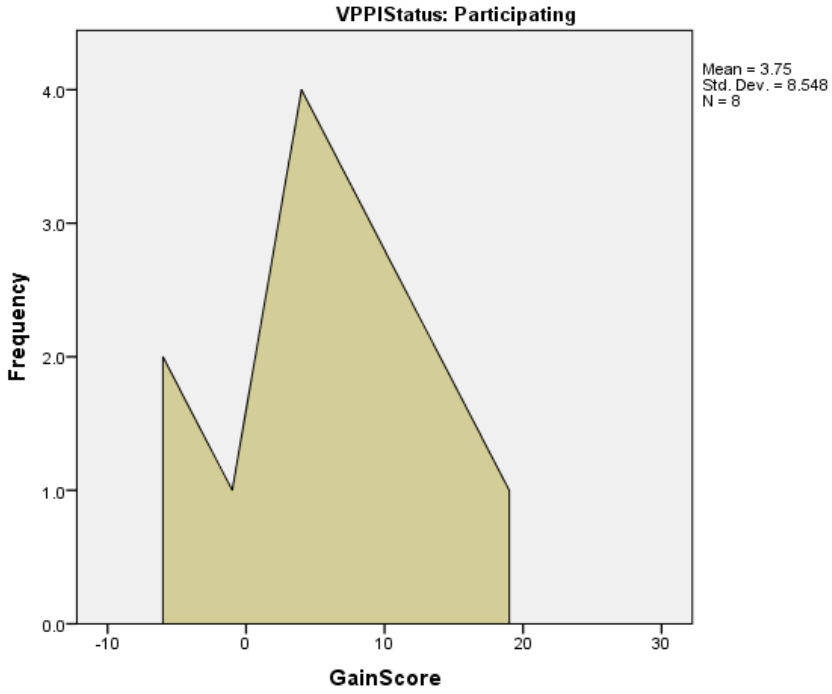
High School End-of-Course World History II Histogram VPPI Participating



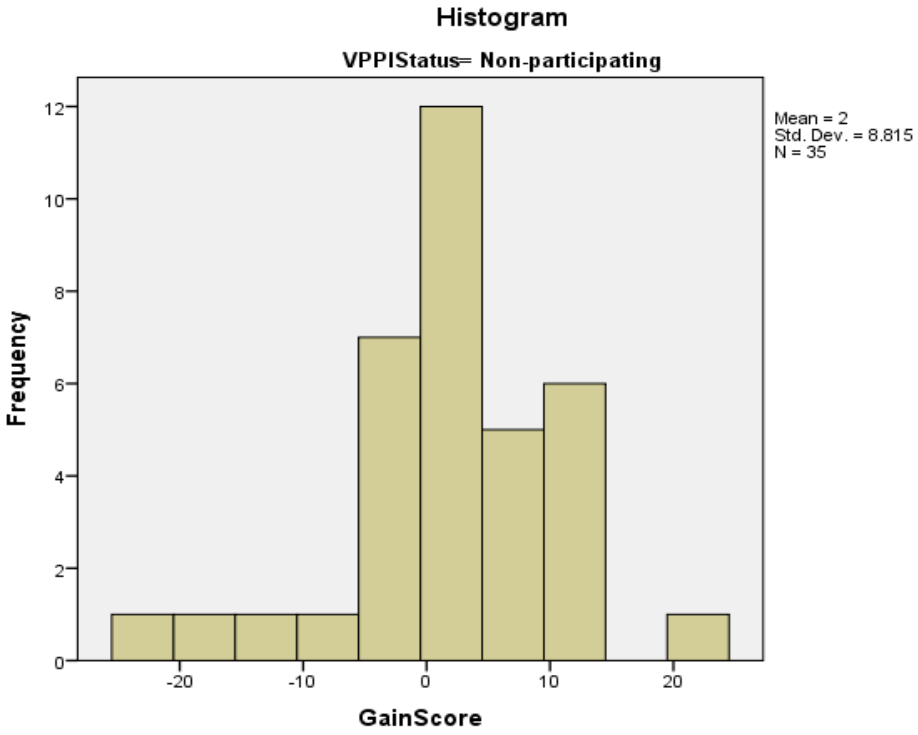
High School End-of-Course World History II Box Plot VPPI Participating



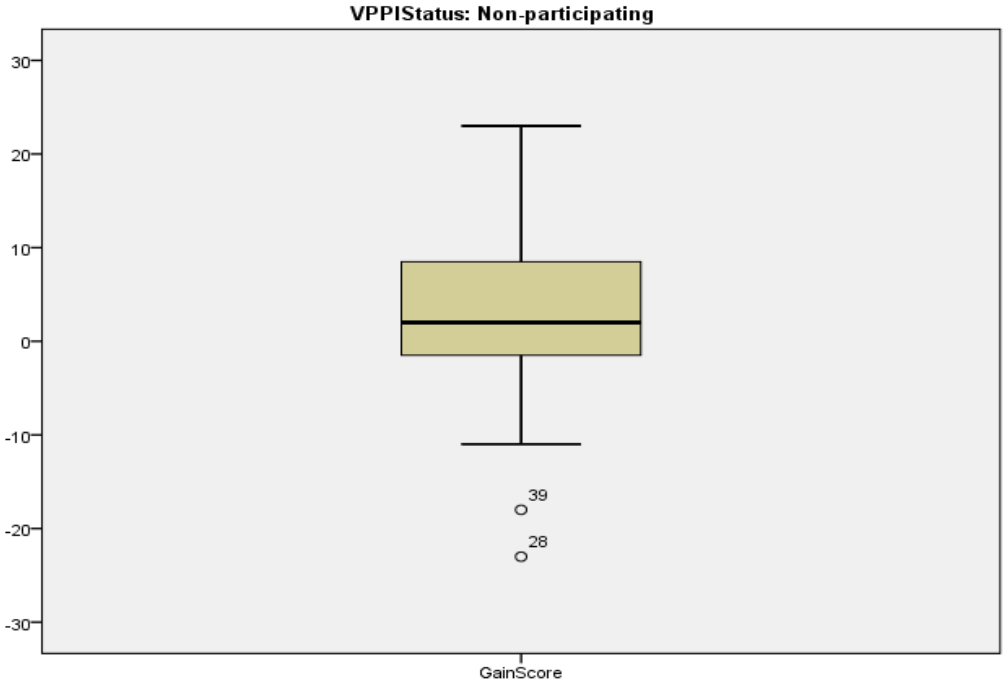
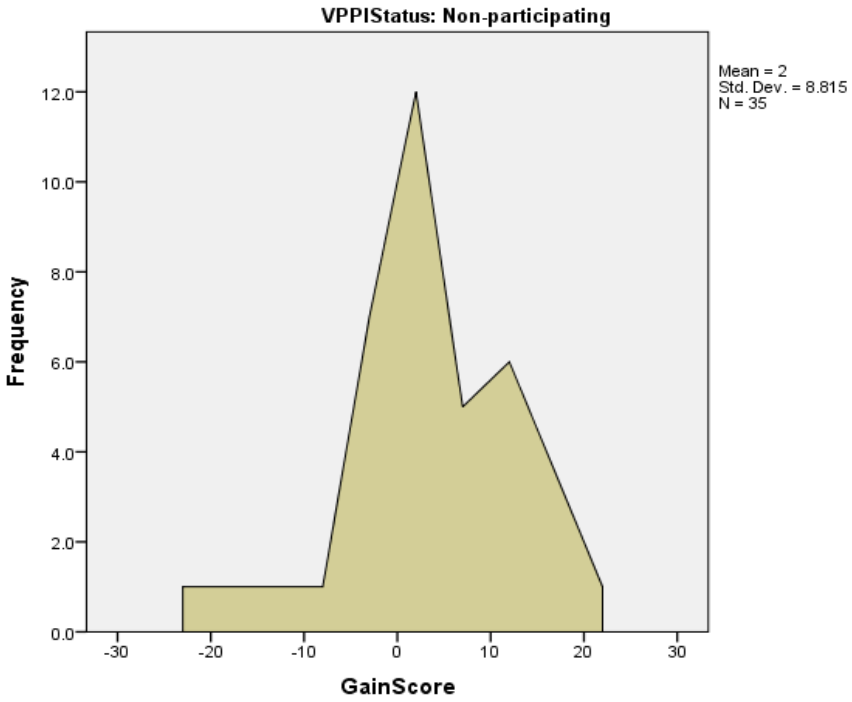
High School End-of-Course World History II Frequency Polygon VPPI Participating



High School End-of-Course World History II Histogram VPPI Nonparticipating



High School End-of-Course World History II Box Plot VPPI Nonparticipating



High School End-of-Course World History II Frequency Polygon VPPI Nonparticipating